THE VOWELS OF SOUTH AFRICAN ENGLISH

Ian Bekker

Thesis submitted for the degree Doctor of Philosophy in English at the Potchefstroom Campus of the North-West University

Promoter: Prof A.J. van Rooy

Summary and Key Terms

This thesis provides a comparative analysis of vowel quality in South African English (SAE) using the following data: firstly, the existing impressionistic literature on SAE and other relevant accents of English, the former of which is subject to a critical review; secondly, acoustic data from a similar range of accents, including new SAE data, collected and instrumentally analyzed specifically for the purposes of this research. These various data are used to position, on both a descriptive and theoretical level, the SAE vowel system.

In addition, and in the service of providing a careful reconstruction of the linguistic history of this variety, it offers a three-stage koinéization model which helps, in many respects, to illuminate the respective roles played by endogenous and exogenous factors in SAE's development.

More generally, the analysis is focussed on rendering explicit the extent to which the synchronic status and diachronic development of SAE more generally, and SAE vowel quality more particularly, provides support for a number of descriptive and theoretical frameworks, including those provided in Labov (1994), Torgersen and Kerswill (2004), Trudgill (2004) and Schneider (2003; 2007). With respect to these frameworks, and based on the results of the analysis, it proposes an extension to Schneider's (2007) Dynamic Model, shows Trudgill's (2004) model of new-dialect formation to be inadequate in accounting for some of the SAE data, provides evidence that SAE is a possibly imminent but 'conservative' member of Torgersen and Kerswill's (2004) SECS-Shift and uses SAE data to question the applicability of the SECS-Shift to FOOT-Fronting.

Furthermore, this thesis provides evidence that SAE has undergone an indexically-driven arrestment of the Diphthong and Southern Shifts and a subsequent and related diffusion of GenSAE values at the expense of BrSAE ones. Similarly, it shows that SAE's possible participation in the SECS-Shift constitutes an effective chain-shift reversal ‘from above’. It stresses that, in order to understand such phenomena, recourse needs to be made to a theory of indexicality that takes into account the unique sociohistorical development of SAE and its speakers.

Lastly, the adoption of the three-stage koinéization model mentioned above highlights the merits of considering both endogenous and exogenous factors in the historical reconstruction of new-dialect formation and, for research into SAE in particular, strengthens the case for further investigation into the possible effects of 19th-century Afrikaans/Dutch, Yiddish and north-of-English dialects on the formation of modern SAE.

Key Terms: South African English, SAE, vowels, vowel quality, English pronunciation, English accents, sociophonetics, phonetics, acoustic phonetics, acoustic analysis, new-dialect formation, koinéization, Southern Hemisphere Englishes, Diphthong Shift, Southern Shift, indexicality.

Acknowledgements

I would like to acknowledge the contributions made by a number of individuals.

First and foremost, thanks go to my supervisor, Bertus van Rooy, who provided timely and pertinent guidance and who bravely and uncomplainingly undertook a close reading of my somewhat turgid earlier drafts. The fact that what follows is at least to some degree digestible is in large part thanks to his intervention. Any remaining ‘turgidity’, as well as the other inevitable faults and omissions remain solely my responsibility.

Secondly, thanks are due to my colleagues at the English Language and Linguistics Department at Rhodes University: for their encouragement, support and, perhaps more importantly, their good humour. In particular, I would like to thank Sally Hunt and Louise Todd for possessing ‘ears’ far better than mine and correcting some of my grosser errors of observation. On many levels, I suspect that they would have done a better job of describing contemporary SAEP than me. Mark de Vos also deserves special mention, mainly for his willingness to engage in discussion on a number of phonetically and phonologically-related topics which, while they interested me, hold, if truth be told, only a tangential relationship to his core academic concerns.

I would also like to thank Andrew van der Spuy, an ex-colleague of mine at Rhodes and current Head of Department at the University of the Witwatersrand, both for his friendship and for his unbending commitment to accuracy and truth in linguistics and academic work in general. I can only hope that this thesis provides some evidence that his fine example has had some positive effect on my own work. More generally, I would like to acknowledge the contribution of delegates at a number of LSSA[1] and UKLVC[2] conferences; individuals who

[1] Linguistics Society of South Africa.
[2] United Kingdom Language Variation and Change.

have provided insightful and penetrating commentary on some of the results reported on in this thesis. Names that come immediately to mind include Charles Boberg, Paul Foulkes, Paul Kerswill, Ron Simango, Rajend Mesthrie and Arista da Silva, but there are many more.

Last, but certainly not least, I would like to thank my family for their support, encouragement and love. To my parents, Leon and Doreen, for instilling in me a love of thought and argumentation. To my wife, Adelaide, for reminding me to view thought and argumentation (and academia more generally) in their proper perspective. And to my darling son, Carlito, for *forcing* me to adopt said perspective.

This thesis is dedicated to my daughter, Dayna.

Contents

List of Figures

List of Tables

Typographical Conventions and Phonetic Symbols

- Where words are being directly referred to, these are placed in italics e.g. "…the vowel in *park* is [ɑː]".

- Words written in capital letters, such as PRICE, are Wells' (1982) keywords for standard lexical sets which "refer concisely to large groups of words which tend to share the same vowel, *and to the vowel which they share*" (Wells 1982:xviii; my emphasis).

- The default phonetic symbol set used in this thesis is "the version revised by a Convention of the International Phonetic Association held in Kiel in 1989, subject to a subsequent set of minor modifications and corrections approved by the Council of the Association" (International Phonetic Association 1999:3). In many cases of transcriptions drawing on a symbol set prior to, or not adoptive of, the above-mentioned Kiel conventions, these transcriptions have, where necessary, been translated into the modern alternatives and been included in brace-brackets, i.e. { … }, after the original. Pullum and Ladusaw's (1996) excellent *Phonetic Symbol Guide* has made this task of translation much less onerous than it would otherwise have been. Transcriptions based on American conventions (e.g. those used by Labov in various publications) have, in general, been left untranslated.

Waar enigsins moontlik, moet die studie van fonetiek beginne met die klanke van die moedertaal. Dis alleen dàn wat 'n mens volkome seker is van die uitspraak van 'n klank en dit kan herhaal so dikwels as nodig mag wees, twee onontbeerlike vereistes ten einde sekuur na te gaan wat die artikulasie van 'n klank is en wat die indruk is wat dit op die oor maak. By vreemde klanke daarenteë gebeur dit maar alte dikwels dat iemand nie seker is van die uitspraak van 'n klank nie, of, wat nog erger is, meen dat hy die korrekte uitspraak gee, terwyl dit volstrek nie die geval is nie. Waar iemand 'n dialek praat, kan dit in seker gevalle nodig wees om selfs daarmee te beginne.

(Le Roux and De Villiers Pienaar 1927:3).

In the case of language . . . the sensuous is reduced to a mere instrument and is thus annulled. If a man spoke in such a way that one heard the movement of his tongue, he would speak badly; if he heard so that he heard the air vibrations instead of the words, he would hear badly . . . Language becomes the perfect medium just at the moment when everything sensuous in it is negatived.

(Kierkegaard 1959:66).

CHAPTER 1

Introduction

> "A peculiar blend of general and particular interests is involved in the pursuit of science; the particular is studied in the hope that it may throw light upon the general" (Russell 1928:40).

1.1 Rationale and Scope

Compared to other accents of English, South African English Pronunciation (SAEP) is under-described and unintegrated into current model- and theory-formation. Lass (1990:272), for example, characterizes it as "enormously complex and grossly under-described":

> "While some of the major variation patterns and segmental particulars . . . have been set out clearly . . . except for the very early Hopwood (1928) and the recent Wells (1982:§8.3), there seems to be nothing available describing whole systems (and even these are conflated 'general' pictures)"(Lass 1990:272).

For all intents and purposes, the gap of under-description remains; one which, I believe, is linked partly to the fact that, apart from the all-but-forgotten Webb (1983), no examples of research employing instrumental (particularly acoustic) methods of analysis exist. In the extant literature, decisions regarding the nature and value of particular phonetic features have relied solely on the subjective, though often informed, judgements of the investigators concerned. Although typically guided by a high level of phonetic training and no doubt perfectly accurate in many respects, the lack of any objective acoustic analysis in these studies means that the validity of the relevant descriptions are hard to assess and, as a reader, one has to simply accept them on faith. The application of acoustic methods of analysis might thus assist in solving this problem and, more specifically, in resolving certain points of disagreement in the literature, especially those possibly derivative of the commonly attested phenomenon of different (even highly-trained) phoneticians coming to different conclusions about the quality of a particular (especially vowel) sound[1]. So, to provide just one example, the exact values of the various allophones of the KIT vowel in SAE have been a 'bone of contention' in the literature, as can be seen by a quick comparison of Lanham (1967:70–4;80) with Lass and Wright (1985:148–54).

The implications of an impressionistic method for, among other things, the study of language change are of particular concern, since it is often impossible to guarantee that one is 'hearing' the data in exactly the same manner as a researcher did 20 or 30 years ago: differences between one's own results and those of earlier descriptions are thus attributable to either language change or differences in judgement.

A related concern is the impossibility of replicating research performed on this basis, due to the obvious constraint of requiring the same researcher (i.e. 'set of ears') to perform it. Again, for obvious reasons, this is not always possible. Foulkes and Docherty (1999:23), for one, emphasize the importance of accountability in any form of empirical research, one important aspect of which is the "replication of results".

Boberg (2005:126) summarises this state-of-affairs in the following manner:

> "A reliance on impressionistic transcription can sometimes introduce various sources of error, ranging from the imprecision of the analytical categories used in impressionistic transcription to problems of intertoken and intercoder reliability and objectivity . . . the inherent limitations of auditory-impressionistic analysis suggest that our knowledge . . . could be advanced by further study using large-scale acoustic analysis".

The above quotation is from an article reporting on research on the so-called "Canadian Shift" in Montreal. Much the same could, however, be said of SAE.

Another drawback of impressionistic methods of analysis is that they are often rather conservative, being strongly influenced by the expectations of the the analyst, and thus, arguably, unsuited to capturing new developments in a dialect. As emphasized by Labov (1991:3), "understatement in impressionistic transcription is inevitable . . . for an accurate view of dialect differences, instrumental analysis is needed".

The current study goes *some* way towards resolving the above-mentioned problems with respect to SAE. It achieves this by, among other things, isolating and thoroughly analysing,

[1] See Ladefoged (1967:50–142) for an excellent example and discussion of this and related phenomena.

using modern (acoustic) instrumental methods, the vowels of one sub-variety of SAE i.e. General South African English (GenSAE). As such, this research is different to standard sociolinguistic research, e.g. Labov (1966), in that, in the acoustic analysis, it attempts to control social variables instead of factoring them directly into the research design. What the current research *does* share with most Labovian-type sociolinguistic research, though, is an interest in the nature and direction of linguistic change: although the focus of the acoustic analysis *per se* is on the isolation, description and analysis of the vowel system of one variety of SAE at one point in time, for one age group and in one style, a perhaps more important part of the broader analysis is the contextualization of the acoustic data within the framework of a range of other data and within the framework of a number of attempts to model and theorize change in the English vowel-system more broadly.

This thesis begins with two quotes, one from the Danish philosopher and theologian Søren Kierkergaard. The quote in question would appear, on first reading, to be at odds with an approach that emphasizes the use of acoustic methods of phonetic analysis; indeed at odds with phonetic analysis per se. Its inclusion was, however, meant as a reminder that the analysis of the new acoustic data collected for this thesis should never devolve into nothing more than a simple description of the acoustic facts. Firstly, and as already intimated above, it compares the outcome of the acoustic analysis with a range of other data sources in order to provide reasonably motivated arguments regarding the development of the General SAE (and SAE more generally) vowel system over the last two centuries as well as its current status *vis-à-vis* other accents of English. The data sources include:

1. A thorough analysis of the available literature on SAE vowels. Most of this literature is impressionistic, with the exception of Webb (1983), whose acoustic data provides a useful basis for comparison.

2. Acoustic and impressionistic data drawn from other relevant accents of English.

The relationship between SAE and other accents of English is given particular prominence in the final analysis. In particular, an attempt is made to analyze the acoustic data in the light of work on Labov's (1991) so-called 'Southern Shift,' Torgersen and Kerswill's (2004) study of the English of the southeast of England, as well as Trudgill's (2004) more recent work on Southern Hemisphere Englishes.

Secondly, the acoustic data is, in this thesis, consistently viewed and conceptualized as the material basis for a linguistic and indexical *system* and only truly of interest in this light. As such, and where possible, an attempt is made to view the data from a systemic, phonologically-orientated perspective. With an eye to future research, possible links between the phonetic and phonological levels are highlighted. The establishment of such putative links is, however, dependent on certain basic assumptions concerning the relationship between the two mentioned levels; these are sketched briefly in §1.4.

A concern in this thesis is also shown for the indexical value of varieties and variants of SAEP, mainly given that any broader concern for the diachronic status of this variety can hardly, I believe, overlook this perspective. A brief excursus into some of the issues relating to the indexical value of linguistic variants is provided in §1.3.3.

A broader rationale for the above-mentioned attempts to integrate a description of the SAEP vowel-system into current theoretical frameworks relates to a general trend in South African linguistics, "away from microlinguistics / descriptive linguistics towards certain macrolinguistic and particularly applied linguistic domains" (Van Rooy and Pienaar 2006:201). This thesis constitutes one attempt to counteract this imbalance as well as to, hopefully, engender further microlinguistic and theoretically-framed research in SAE. On the last point, it is acknowledged that the current thesis is still primarily descriptive in nature, with an emphasis on descriptive as opposed to explanatory adequacy; primarily, however, as a result of the fact that, with respect to SAE, as with many other South African varieties, "so much still needs to be done in terms of . . . description and understanding" (Van Rooy and Pienaar 2006:201). It is hoped, however, that this thesis will go some way toward 'clearing the ground' for future theoretically-motivated work on (or using) SAE data. This speaks directly to the emphasis placed in this thesis on embedding the analysis into a number of current theoretical frameworks and models.

In the opinion of the author, one further obstacle to the development of research on SAEP as well as its integration into existing theory, is the lack, aside from occasional broad announcements regarding the merits or demerits of any particular work, of any real attempt to *critically* engage with the existing literature. This thesis provides just such a critical review and, as a consequence, takes the slightly unusual approach of incorporating much of the review of the SAEP literature under the various results chapters[2]. As such the critical review is viewed as a 'result' of this research. One positive side-effect of this approach is that it 'deposits' all the technical detail for any particular vowel in one 'place', a design feature which will, hopefully, benefit the reader.

Turning to issues of focus, it should be clear from the title of this thesis that the consonantal features of SAEP are not given much attention. A decision to limit this research to vowel quality was based on a number of considerations:

- While it is a truism that different accents of English differ in the nature and realisation

[2] Chapters 6 to 9.

of their consonantal inventories, it is also true that most accents of English are more readily distinguishable from each other in terms of their vowels.

- Much theory and model-construction, particularly with respect to the diachronic development of English, has focussed on the vocalic sub-system; in addition, Labovian-type research into change-in-progress often focusses on vowel-shifts and patterns and is also often characterized by the use of acoustic data.

- The simple issue of scale has precluded the expansion of an already lengthy thesis into further territory.

Still, where these have arisen as an unforseen, but welcome, side-effect of the research process, consonantal features derived particularly from the new acoustic data have been commented on and earmarked for later research. Another exception to the otherwise exclusive focus on vowel-quality has been the inclusion of an analysis of /l/ in SAE, particularly when in post-vocalic position. As will become clear in later sections, the rationale for focussing on this positional variant of /l/ is its dramatic effect on preceding vowel quality.

Lastly, since the focus of the acoustic analysis contained in this research is on only one, tightly constrained, variety of SAE, the results obtained from this analysis are, *in isolation*, not of course generalisable to the SAE speech community as a whole. Placed in the *context* of the various data, models and frameworks alluded to above, reasonable conclusions can, however, be drawn regarding the synchronic and diachronic status of the SAE vowel system.

Moreover, the unrepresentativeness of a sample (or rather the lack of any real sample at all) has not prevented previous commentators from making statements about SAE more generally. So, for example, Lass and Wright (1985:148–54) reanalyze the KIT vowel, and reject Lanham's (1967) previous analysis in important respects, on the basis of recordings of two speakers. This study makes use of twenty-seven.

1.2 The SAE Vowel System in Context

As repeatedly stressed in the previous section, one of the main foci of the current research is to 'place' the SAEP vowel system in the context of a number of relatively recent, general theoretical frameworks which, in turn, attempt, among other things, to explain the emergence of similar developmental patterns across different (regionally separated) varieties of English. Groups of dialects have similar features and seem to move in similar directions with respect to structural change, and work by Labov (1991) and, more recently, Trudgill (2004) has sought to explain these similarities. Although Labov (1991:34) does not attempt to provide "an exhaustive or definitive taxonomy of English dialects," he certainly attempts to encompass most of them into a tripartite division of English accents, one division of which he labels the 'Southern Shift' and which includes dialects from the southeast of England, the southern states of the USA and the southern hemisphere. Trudgill (2004) looks specifically at the Southern Hemisphere Englishes and attempts to explain their similarities by making recourse to a number of more general principles of linguistic change, while Torgersen and Kerswill (2004) report on a chain-shift of the short vowels of the English of the southeast of England; research which has interesting implications for the current and future status of SAE and other 'southern' Englishes.

Schneider (2003; 2007), on the other hand, provides a more macro-level account of the development of New Englishes, or what he refers to as Postcolonial Englishes, under which he includes colonial L1-varieties such as American English and those of the southern hemisphere. In essence, he posits that similar sociohistorical developments applying across the different colonial settings into which English was transplanted, mean that the different varieties of English undergo a similar, staged sociolinguistic development.

The above frameworks, as well as other more minor bases of comparison, are given further attention in latter sections, both in terms of their explication and in terms of determining the extent to which the data collected and analyzed in later sections supports or challenges these frameworks.

1.3 Sociolinguistic considerations

Perhaps the best way to classify the current research would be to say that it falls within the field of 'accent studies' as defined by Foulkes and Docherty (1999:4–6). These authors use the term to refer to a broad field of interest "which intersects (at least) dialectology, sociolinguistics, phonetics and phonology" (Foulkes and Docherty 1999:6). Furthermore, while the current research is not an example of a Labovian-type study of change-in-progress and is, perhaps, not even classifiable as sociolinguistic, it does draw, at various points, on the work of Labov and his fellow variationists; and, more generally, it attempts to draw connections between various forms of data and broader trends in SAE and English accents in general, the trends being both of an internal (structural) nature and those involving what might be called the 'social life' of a dialect i.e. the various factors involved in and consequent upon the interplay between language and society.

More specifically, because the results of the current research are meant to be viewed as, among other things, an effective 'launching-pad' for Labovian-type studies of change-in-progress in the various urban centers of South Africa, a brief excursus into the main foci, methodologies and results of modern variationist research, particularly that of a phonetic nature, is provided in §1.3.1 below.

The description arising out of the acoustic analysis that constitutes a core source of data for the present research is that of a particular sub-variety of SAE i.e. General South African English (GenSAE), conceived of as the modern standard for the vast majority of SAE speakers. While a full motivation for choosing a standard variety is provided in Chapter 4, some brief remarks about the notion of a standard and its particular use in the South African context are required. These are contained in §1.3.2.

Closely related to the notion of a standard is the notion of prestige and of the indexicality of linguistic choice in general. Since indexicality is a central issue in much, particularly recent, sociolinguistic research it was felt necessary to embed the later review of SAE within a relatively clear view of the nature and internal structure of indexicality. The main outlines of such a view are provided in §1.3.3.

1.3.1 Sociophonetics

Interest in phonetic variability, particularly in the form of dialect geography, has a long history. The term ‘sociophonetics’ is, however, relatively recent: according to Foulkes (2005:495), “among the first to define her work as ‘sociophonetic’ is Deshaies-LaFontaine (1975), in a study of Canadian-French”. According to Foulkes and Docherty (2006:410), the term has often been used by phoneticians to refer to work focussed on providing “descriptive accounts of speech production across different dialects, speaker groups or speech styles”. The term is also closely linked to the growth of Labovian variationist sociolinguistics and its focus on “social, regional and stylistic variation in speech production” (Foulkes 2005:495). Work in this field is characterised theoretically by an attempt to merge the fields of “dialect geography, sociolinguistics, phonetics, and historical linguistics”(Labov 1994:25) and, in particular, on the connection between synchronic variability and language change. The latter is also characterised by a number of methodological developments which, according to Labov (1994:25), include computational advances in the instrumental analysis of acoustic data as well as the application of various kinds of mathematical and statistical models to the analysis of sociolinguistic data in general.

A common focus in modern sociophonetics has been on capturing change-in-progress in urban centres, mostly in North America and the United Kingdom. This shift from the traditional study of rural varieties, as evidenced in earlier studies in dialect geography, to the study of change-in-progress in urban settings is defended by Labov (1994:22–3) in the following manner:

> "It is … commonly reported by dialectologists that local dialects are disappearing, and that we have entered a new period of linguistic convergence instead of divergence. But research in urban areas shows the opposite. Since 1972, I and others have been reporting evidence of continued sound change in the dialects of the major English-speaking cities. In every large speech community studied in the United States, Canada and Australia, we observe the vigorous development of the local vernacular …".

This is not to, however, underestimate forces operating in the opposite direction i.e. those linked to "migration … mobility, and expansion in the tertiary economy and labour-market flexibility" (Torgersen and Kerswill 2004:25), all of which have the potential to lead to a degree of dialect levelling and dialect convergence, terms which are defined by Torgersen and Kerswill (2004:24; my parenthesis) in the following manner:

> "dialect levelling … can be defined as the reduction in the number of realisations of linguistic units found in a defined area, usually through the loss of geographically and demographically restricted, or 'marked', variants … the closely related notion of dialect convergence … [can be defined in terms of a process] … by which two or more varieties [become] more alike through convergent changes. These are both seen as the outcome of various, mainly contact-based, scenarios".

Much of the research on dialect levelling and related areas has been conducted on British varieties, which are numerous and regionally diverse in an obvious way, or on early colonial Englishes, which were mostly characterised by a variety of input e.g. early New Zealand English (Trudgill 2004). From this perspective (i.e. as a colonial English), SAE has gone through its process of dialect levelling. Still, the same forces that created a relatively stable SAE variety, might have (and possibly still do) mitigate against subsequent regional and other divergence within SAE. With respect to SAE, there is certainly a lack of clarity concerning these issues, and while a focus on particular urban centers and regional variety in general is not a focus of the current research, it requires some preliminary attention, given that the subjects come mainly from two separate areas. This issue is taken further in §3.3.3.

1.3.2 GenSAE as a Standard

Lanham and Traill (1962:5) have the following comments to make regarding the aim and scope of their study on SAE pronunciation:

> The research on which this study is based has, as its ultimate objective, a description of pronunciation systems relating to all social, regional, etc. dialects of English in

> South Africa. This is, however, a long-term project requiring considerable resources in trained personnel, time and money. It seemed to us, however, that there was an important intermediate objective capable of achievement in a much shorter time, viz. the description of ‘educated English’ - the ‘prestige dialect’ as far as the English community in South Africa is concerned.

In the current research, the choice of GenSAE over less-prestigious forms of SAE with respect to the collection of the new acoustic data has, similarly, been motivated by a number of considerations, both purely practical and otherwise. The purely practical considerations, as in the case of Lanham and Traill (1962), did have a central role to play in the choice of subjects; the most prominent of these was the relative ease of access the author has had to speakers of GenSAE. While the less practical, more methodologically and theoretically-driven considerations are dealt with in some detail in Chapter 4, and, in particular, in §4.2.3, it is worthwhile stressing at this point that one of the most basic motivations for this research was the establishment of a linguistic reference point or ‘baseline’ from which full-scale (instrumental) sociophonetic studies of the various urban centers of South Africa could proceed. Labov (1994:63), for one, advocates the use of such ‘conservative’ reference points in the study of language change. Furthermore, linguistic communities are often not definable in terms of the language used by their speakers, mainly due to the great degree of variation present in many speech communities, particularly the large urban ones. Such communities are, however, often definable in terms of the sociolinguistic norms that their members share. Labov (2003:244) puts this bluntly: “a uniform set of norms defines a speech community”. The accurate description and analysis of the normative, prestigious variety of a speech community is one way of coming close towards determining the nature of those norms that bind a speech community together.

It should be stressed, however, that while the new acoustic data, as collected for this research project, is focussed on the more formal end of the sociolinguistic spectrum, the inclusion of other forms of data (both acoustic and impressionistic) derived from the existing literature on SAE, provides for an important basis of comparison and control in the overall analysis of vowel quality in SAE. Webb’s (1983) acoustic data is particularly useful in this regard. These and other issues will be taken up at greater length later in the thesis, particularly in Chapter 4.

1.3.3 A Brief Note on Prestige and Indexicality

In §1.3.2 reference is made to prestige. Although the establishment of the indexical value of particular variants and varieties of SAE is not a main focus of this thesis, reference to indexical value (e.g. prestige, in whatever form, or its lack) is unavoidable in reviewing variance in SAE and in terms of speculating on its history and possible phonetic and phonological future. As such it is important to position this thesis *vis-à-vis* the notion and theory of indexicality, particularly as employed in research in the Labovian tradition. In particular, while it is accepted that broad connections between class, gender etc. and particular phonetic variants can (and should) be made, recent work illustrating and exploring the complex and 'grounded' nature of indexicality in language also needs to be recognized. Zhang (2008:202), for example, identifies the following:

> "a broader move in sociolinguistics, particularly variationist sociolinguistics, to adopt practice-based approaches to the study of linguistic variation . . . compared with earlier variationist studies, this strand of research pays more attention to speaker agency and the ways in which social meanings are constructed through deployment of linguistic and other semiotic resources".

This constitutes a change of emphasis which leads Bucholtz and Hall (2008) to motivate for the adoption of the term 'sociocultural linguistics' in order to capture this new emphasis on the context of language use and thus on the links between language, culture and society as well as the use of concepts and methodologies closely connected with other social sciences such as anthropology and sociology.

Thus, by way of illustration, a study by Dodsworth (2008) shows how GOAT-Fronting in Worthington, Ohio "largely resists traditional sociolinguistic explanations . . . [while] . . . a close analysis of four speaker's mental representations of the local tensions surrounding urban sprawl reveals significant differences which are argued to account for their variable use of fronted /o/" (Dodsworth 2008:34; my parenthesis). In a similar vein, Johnstone and Kiesling (2008) show how MOUTH-Monophthongisation in Pittsburgh, "clearly indexes local identity for some speakers . . . [but] . . . these are not, on the whole, the speakers who actually use it" (Johnstone and Kiesling 2008:6; my parenthesis). Within the context of a sociolinguistic investigation of SAE, Da Silva (2007:42; my parenthesis) outlines the problems (and assumptions) underlying the practice of simply rating variants of a variable on a scale of lowest to highest prestige i.e. it assumes "that the ranking of prestigious and stigmatised forms is uniform by [sic] all members of that speech community . . . this approach is often implicit in sociolinguistic research since Labov's first studies".

While this thesis cannot (and is not meant) to attempt sophisticated 'on-site' accounts such as those provided in Zhang (2008), Dodsworth (2008), Johnstone and Kiesling (2008) and Da Silva (2007), it can at least embrace a degree of sophistication in this regard by recognizing various *orders* of indexical meaning as, for example, outlined in Johnstone and Kiesling (2008:8–9):

first-order indexicality: similar to Labov's (1972) notion of an 'indicator'. Below the level of consciousness, indicative of a particular demographic group (e.g. class or region) and not subject to stylistic variation in the speech of the individual.

second-order indexicality: similar to Labov's (1972) notion of a 'marker'. Shows stylistic variation; "speakers use different variants in different contexts, because the use of one variant or another is socially meaningful. Speakers are not necessarily aware of the variables or their social meanings" (Johnstone and Kiesling 2008:8).

third-order indexicality: similar to Labov's (1972) notion of a 'stereotype'. A variable form which is commonly the topic of overt comment; it is often not actually used in the vernacular and is inclined to eventually recede from the later.

Furthermore, an understanding of indexicality that goes at least some way beyond the oversimplistic needs to be sensitive, firstly, to the possibility that indexical value is 'in the eye of the beholder' and, secondly, to the social history of the variety (and variable) in question. Ultimately, "the social meanings of variables are contextually bound and always in flux" (Bucholtz and Hall 2008:409) and, as stressed by Foulkes and Docherty (2006:432), the meaning of a variant can change over time and thus be "redeployed in the same environment with different indexical consequences". Within the specific context of SAE, Lanham and Macdonald (1979:69) summarize this state-of-affairs in the following manner:

> "Prestige … lies in the eye of the beholder and social values as ultimate motivation are to be examined as a product of an individual social experience and commitment as well as group norms. Prestige as a property assigned to a social type or group requires different specification in different societies".

We will also see later, in §3.3, that, more generally, the indexicality of SAE varieties and variables provides important challenges that will be kept firmly in mind in the actual analysis of the SAE vowel system as provided in later chapters.

An agnostic position is, however, adopted in this thesis with respect to the 'locality' of such indexical information in the overall cognitive framework of the individual speaker i.e. whether it forms part of a more broadly-conceived phonology, as advanced in the exemplar-based model of phonological knowledge provided in Foulkes and Docherty (2006), or whether it forms part of performance conceived as separate from grammatical competence[3].

[3] See (Yang 2008) for an eloquent defence of this position.

Certain assumptions about the link between phonetics, phonology and language change have been embraced however, as outlined briefly in the next section.

Lastly, the role of indexicality (in the form of identity, prestige etc.) in language change and, in particular, new-dialect formation, needs to be considered in the light of statements such as those of Trudgill (2004:148), who claims that "patterns of interaction ... should always be consulted for possible explanations before one jumps to conclusions about identity and prestige". From this angle, linguistic diffusion etc. is the result of the inevitable accommodation that takes place in face-to-face encounters and the changes are, accordingly, mechanical, predictable and minimally associated with social evaluation etc. Trudgill (2004) draws explicitly on this mechanical, deterministic notion of language change and new-dialect formation, in his attempt to explain the similarities that exist between the various Southern Hemisphere Englishes. In short, "if you bake cakes ... from roughly the same ingredients in roughly the same proportions in roughly similar conditions for roughly the same length of time, you will get roughly similar cakes" (Trudgill 2004:20). From this point of view it is, thus, at least theoretically possible, given enough information regarding the speakers of the various dialects and languages transplanted to a new geographical setting, to determine what the eventual outcome will be. In particular, the outcome will not be "as a result of social factors such as status. In determining who accommodates to who – and therefore which forms are retained and which lost – demographic factors involving proportions of different dialect speakers present will be vital" (Trudgill 2004:85). In short, if a particular variant is more common than another the first variant will be retained while the other will not[4]. According to Trudgill (2004:150), this reliance on a purely mechanistic 'method' of feature-selection is of particular relevance to "new-dialect formation ... associated with tabula rasa colonial situations", where by 'tabula rasa colonial situations' Trudgill (2004:26) means "those in which there is no prior-existing population speaking the language in question, either in the location or nearby". Trudgill (2004:26) also adds that by the use of this term he is "not making claims ... about new-town koinés" in which a small but local population already exists before the process of urbanization. Important is that in such cases "there will not necessarily ... be a complete break in contact with any normative traditions already established in the wider community" (Trudgill 2004:29) i.e. issues of prestige etc. might indeed play a role in feature selection. We will see later, however, that the three main South African settlements (Eastern Cape, Natal and Johannesburg) all qualify, effectively, as tabula rasa 'situations'[5].

[4]There are certain provisos provided by Trudgill (2004), such as the role of markedness, but the basic point remains.

[5]I confess that an obvious gap in this list of 'settlements' is Cape Town. I know of no existing attempt, however, to link 'Cape Town English' explicitly with the early development and genesis of SAE. This aspect

Perhaps naturally enough, the emphasis in Trudgill (2004; 2008) on the relative unimportance of indexicality in the development of new-dialects in tabula rasa colonial settings has not remained unchallenged. Schneider (2003; 2007), for one, proposes a model for the development of new Englishes in colonial settings which places issues of identity and indexicality at its core; this model is discussed in greater detail in §3.3.3. More recently, Schneider (2008:262–3) has attributed to Trudgill a "narrow concept of identity" as well as a false dichotomy between accommodation and identity, the former viewed as mechanical, the latter "implying an understanding of the effects of identity as deliberate and conscious". In contrast, Schneider (2008) stresses the link between accommodation and identity, defining the first as "a process of linguistic approximation with the social goal of signaling solidarity by diminishing symbolic distance" and pointing out that "accommodation is one of the mechanisms of expressing one's identity choices". The author also, rightly, intimates that the opposite of accommodation i.e. divergence, where speakers wish "to emphasise their distinctiveness or increase their social distance" (Mesthrie, Swann, Deumert and Leap 2000:151), is difficult to reconcile with Trudgill's rather mechanistic conception of accommodation as linked to "an apparently biologically given drive to behave as ones peers do" (Trudgill 2004:28).

1.4 The Phonologization Approach

As already mentioned in §1.1, at various points in later chapters some speculation will be indulged in with respect to the possible phonological status of certain features of SAE, both with reference to the review of the existing impressionistic literature and the results of this study. Given the phonetic slant of this thesis, the adoption of a phonetically-grounded approach to phonology, and, in particular, to the phonetics-phonology interface, seemed most appropriate.

The existence of less phonetically-driven phonological theories should, of course, be recognized. Thus, by way of example, Chomsky and Halle (1968) have the following to say about the stress contours of English:

> "We do not doubt that the stress contours and other phonetic facts that are recorded by careful phoneticians … constitute some sort of perceptual reality for those who know the language in question. A person who knows the language should "hear" the predicted phonetic shapes. In particular, the careful and sophisticated impressionistic phonetician who knows the language should be able to bring this perceptual reality

deserves further research.

> to the level of awareness, and there is ample evidence that phoneticians are capable of doing this. We take for granted, then, that phonetic representations describe a perceptual reality . . . *Notice, however, that there is nothing to suggest that these phonetic representations also describe a physical or acoustic reality in any detail*. For example, there is little reason to suppose that the perceived stress contour must represent some physical property of the utterance in a point-by-point fashion . . . In fact, there is no evidence from experimental phonetics to suggest that these contours are actually present as physical properties of utterances in anything like the detail with which they are perceived. Accordingly, there seems to be no reason to suppose that well-trained phoneticians could detect such contours with any reliability of precision in a language that he does not know" (Chomsky and Halle 1968:25; my emphasis).

For stress patterns in particular, this perspective is echoed in latter work, such as Ladd (1980:41), who explicitly claims that "perceptions of prominence are indeed an illusion, a very powerful and consistent one". Even more recently, Gussenhoven (2004:20) argues that "a . . . useful, response has been to point out that stress is ultimately a location in phonological structure". The issue at stake is, of course, broader than simply stress; and if Chomsky and Halle (1968) are correct it casts some doubt on the use of acoustic data in phonetic/phonological studies, since the legitimate objects of study, from this perspective, are not the objective facts but rather the internal perception of these facts, guided as they are by the grammar of the language concerned. The starting point is, therefore, the intuitions of the mother-tongue speaker of the language (or dialect) concerned. From a less mentalist perspective, though, the above quotation should be viewed rather as containing a warning against the effect that an individual's expectations about a language or dialect can have on his impressionistic analysis, particularly if the language and/or dialect are his own. Recourse to acoustic data is one way to overcome these predispositions.

More generally, however, it is not the task of this thesis to argue for one phonological approach over another. As such, the underlying, well-motivated (although certainly not universally accepted) assumption of this thesis will be that in terms of the link between phonetics and phonology, the latter is *primarily* driven by and grounded in the former (particularly during the language acquisition process) and that "a primary route from phonetics to phonology runs through sound change, and in particular through the mechanism of phonologization" (Barnes 2006:2).

Given the adoption of this assumption, a distinction needs to be made between the notions of phonologization and phonemicization. While the latter relates to phonemic contrast, the former deals with the reanalysis of a gradient (and possibly variant) phonetic pattern as a categorical, phonological one. By way of example, Nguyen and Fagyal (2008:1), in their study of vowel harmony (VH) in French, introduce the topic in the following manner:

> "It has been suggested that ... VH ..., a morpho-phonological process in many languages requiring all vowels in a word to share ... one or several features, could be understood as the "grammaticalized" end product of "an earlier phonetic process involving vowel-to-vowel assimilation" (Ohala, 1994). According to this hypothesis, VH arose gradually through sound change due to a dissociation parsing error: ... vowel-to-vowel assimilation ... became so strong as to cause listeners to parse the harmonizing vowel as independent of the triggering vowel. Once dissociated from its trigger, the quality of the harmonizing vowel could vary independently from its initial conditioning environment, and thus be assigned new functions in the grammar".

In the actual study conducted by Nguyen and Fagyal (2008), things turn out to be somewhat more complex and the conclusion (with certain provisos) is that VH in French is, in fact, still phonetic and has *not* undergone a process of phonologization. This is, however, irrelevant to the point of the current discussion which is to make it clear that when phonologization *does* occur, what was once a mechanical, phonetic process becomes assimilated into the grammar through misperceptions on the part of new language learners (i.e. children) and in the process gains independence from its earlier mechanical constraints.

As such, and as will be seen later in §2.2.4.2, the distinction between unphonologized and phonologized processes is usefully recast, synchronically, as a distinction between accomodatory and non-accomodatory allophony (Wells 1982:41–4). As pointed out by Cho and McQueen (2008:240), in cases of one sound assimilating to another, the difference between these two kinds of allophony is where, in the first instance, the listener relies on "acoustic remnants of the speaker's intention that might still be present in the speech signal," while in the second instance the speaker has to draw on phonological or lexical knowledge.

One possible link between phonologization and sound-change becomes apparent when we accept that the former requires the incorporation into the linguistic system of, in the case of vowels, a new vowel 'target' with its own susceptibility to coarticulatory forces. Thus, by way of example, once the retracted value of a vowel before final /l/ has become phonologized by the new language-learner, the articulation of this new target by the speaker will, in all likelihood, be under coarticulatory pressure from the self-same conditioning environment (i.e. final /l/); only in this case the retraction will be more pronounced than before; and in itself open to further phonologization.

A related (and perhaps more basic) process similarly involves "the listener's failure to compensate for coarticulation" (Harrington et al. 2008:2825), but instead of the explicit incorporation of the allophony into the grammar (with the retention of an original target/unmarked allophone which is linked to the new target/marked allophone through, for example, an allophonic rule), the original vowel target is simply shifted along i.e. to a position closer to the coarticulatory source. While the former process would seem most likely to lead to systemic changes such as phonemic splits (particularly when the conditioning environment is lost, for example), the second process involves changes in the phonetic realization of vowels, as, for example, commonly found in vowel chain-shifts (see §2.3.2 for more on this and related matters). Thus, by way of example, Harrington et al. (2008) link just such a process to the fronting of /uː/ in RP over the last couple of decades.

Both these processes can be subsumed under the rubric "hypoarticulation-induced sound changes – that is those that can be related to contextual influences synchronically . . . central to this idea is that listeners may incorrectly parse phonetic events from one phoneme that are temporally distributed and interwoven with those of another" (Harrington et al. 2008:2825–6).

It should be pointed out that the above analysis does accept a level of abstraction i.e. phonetics and phonology do constitute different levels. Following van Rooy (1997:386), it will be assumed, however, that the starting point (or null hypothesis) of any analysis should view the distance between underlying phonological and phonetic representations as minimal as possible. It seems likely, for example, that many of the complex derivations described in Chomsky and Halle (1968) reflect historical fact more than synchronic competence. Unlike van Rooy (1997), however, the usefulness of hypothesizing at least a shallow underlying level is embraced for two reasons; firstly, as a useful descriptive tool for describing phonological change and, secondly, as a means of ensuring the integrity of the phoneme in the face of allophonic variation. However, the emphasis lies mostly on the first, with nothing much riding on the second. With reference to van Rooy (1997:386–8) again, it is accepted that one way of accounting for (non-accommodatory) allophony is not by positing a unitary underlying representation linked by phonological rules to a surface representation, but rather by hypothesizing the use of, as-it-were *post-hoc* generalizations (or surface analogies), over separately stored members of an alternation in order, for example, to incorporate new items into the lexicon. Accommodatory allophony is, on the other hand, taken care of by lower-level (semi-)automatic processes.

Turning to language-change in general it is important, though, following Blevins (2006), to distinguish between 'natural' phonetically-driven, language-internal instances of change such as those discussed above, and externally-motivated changes (or inhibitions to change) resulting from, among other things, language-contact, prescriptivism and literacy. This is particularly relevant to SAE, given the existence of a fair degree of debate regarding the origins of some of its characteristic features e.g. whether the centralization of certain allophones of the KIT vowel has its origins in an endogenous, structural change or is, rather, the result of intensive language-contact between English and Afrikaans in South Africa (or both). In this regard, one has to take into account that, according to Torgersen and Kerswill (2004:24) at least, there has been a serious underestimation (in Labovian-type studies at least) of the role played by "dialect contact in change". Thus these authors provide convincing evidence of dialect leveling in the southeast of England, a process intimately connected with dialect contact and involving dialect convergence. This emphasis on the role of contact in linguistic change has recently received added impetus in work such as Mufwene (2008).

Even granted the truth of this (still controversial) perspective, however, it is clear that compared, for example, with English English (or even American English), the role of leveling in *modern* SAE remains unclear. While there is no doubt than a degree of leveling took place during SAE's 'early' years, and while the spread of GenSAE at the expense of other regional and social varieties has also played its part, the relevance of leveling with respect to the containment of incipient diversification in modern SAE remains unclear. The issue of leveling in early SAE as well as regional variation in 'pre'-modern and modern SAE is taken up again in later chapters, particularly in §3.3.3, where the application of Schneider's (2007) Dynamic Model to SAE is reviewed.

It is perhaps also important to recognize that, from a strictly Labovian point of view, the speech errors, misparsings or reanalyses of children do not constitute linguistic change *per se*. For Labov (1994:47; footnote), the issue is not one of "a change in individual habits, but rather the diffusion of new individual forms into the wider community ... the speech errors of children have no privileged relation to the specific language changes that affect a given community". While it is unclear to me what other sources Labov might have in mind with respect to 'internal' changes (as opposed to externally-motivated changes as described above) he is surely correct in asserting that the reanalyses of a few children do not constitute a sufficient condition for the adoption of these new forms into the wider speech community. For such conditions one has to, perhaps, look beyond the linguistic system towards the broader social matrix.

Lastly, while the afore-mentioned conceptualization of the link between the phonetic and phonological levels provides a neat way of reconciling the gradiance of the former and the (generally assumed) categoricality of the latter, it makes no attempt to accommodate for the other form of gradiance so characteristic of language i.e. the fact that in a language "variation may be observed such that a given form is used statistically more by one social group than another, or more in one speech style than another" (Foulkes and Docherty

2006:411). Accounting for such variation means, in the final analysis, accounting for the effects of indexicality, an issue that has already been dealt with in §1.3.3.

1.5 Formal Problem Statements

The following constitutes a brief (formalised) summary of the problems (issues) outlined in the previous sections:

1. SAE is still an under-researched dialect cluster and would benefit from more research generally and from the application of a different methodology in particular. More specifically, research into SAE has, except for Webb (1983), never employed instrumental acoustic techniques of analysis. This renders the research unobjective, uninterpretable and unreplicable in certain important respects. It also implies that existing impressionistic studies of SAE might have failed to capture recent developments in the accent.

2. Disagreement in the literature is conceivably related to systematic differences in the impressionistic (though expert) judgement of different researchers. If so, further appeal to such judgment is unlikely to resolve the disagreement.

3. The existing literature on the nature and history of SAEP has not, in the main, been subject to a critical appraisal.

4. Not enough has been done to attempt an integration of SAE into certain recent theoretical frameworks describing the synchronic status and diachronic development (both structurally and sociolinguistically) of certain varieties of English, specifically those frameworks represented in Labov (1991; 1994), Torgersen and Kerswill (2004), Trudgill (2004) and Schneider (2003; 2007).

1.6 Formal Statement of Objectives

The following provides a brief outline of the objectives of the research. The stated objectives are directly related to the four problems outlined in the previous section.

1. To isolate, describe and analyze, using instrumental acoustic techniques, the vowel-system of one sub-variety of SAEP i.e. GenSAEP; and, in so doing, provide a research design that is replicable as well as results that are objective and directly interpretable in certain important respects. The application of the above-mentioned techniques might also unearth some recent trends in SAE.

2. To apply the above-mentioned techniques as a potential means of resolving points of disagreement in the extant literature.

3. To provide a critical appraisal of the existing literature on the vowel-system of SAEP.

4. To attempt an integration of an overall comparative analysis based on

   (a) the results of the acoustic analysis described in (1);

   (b) the critical appraisal described in (3);

   (c) impressionistic and acoustic data derived from studies on other relevant English accents; with

   the frameworks and models provided in Labov (1991; 1994), Torgersen and Kerswill (2004), Trudgill (2004) and Schneider (2003; 2007); and, as a consequence, 'test' these frameworks against this overall analysis.

1.7 Brief Notes on Methodology

As far as objective (1) in section 1.6 above is concerned, the research qualifies as descriptive in the sense that its overall objective will be to describe a sub-variety of SAE using a particular (largely unused) methodology. A descriptive approach to analysis is well-attested in the linguistic literature, and of course descriptive adequacy is paramount before any attempt can be made to embed discussion of the language or dialect in question into a broader theoretical framework, the latter being the focus of (4) above. Goals (2) and (3) have descriptive adequacy as their focus as well, while the main focus of (4) is, of course, the integration of the various data forms and, more specifically, the vowel system of SAEP into current theory and model-building.

In summary, the methodological approach to be employed in this research can be broadly defined as, firstly, instrumental and acoustic, as opposed to traditional phonetic research which is not i.e. whereas traditional phonetic description relies on the impressionistic (though well-honed) judgements of trained observers, this research relies (at least partly) on more direct measurements of the acoustic signal produced by subjects. With certain (important) provisos it could, therefore, be described as a more objective approach as well. More importantly, perhaps, is that it is far more replicable.

Furthermore, while the research is, on the aforementioned basis, describable as descriptive, data-driven and empirical, it, most importantly, attempts to integrate observation into a number of broader theoretical frameworks, showing how the data supports or, in some cases, challenges these frameworks. The current research is, therefore, far more than an un-integrated collection of facts; or to put it more colloquially, it isn't an example of 'butterfly collecting'.

1.8 Structure of Thesis

Chapter 2 provides the theoretical basis for the acoustic and comparative analysis of the vowels of SAE. It begins with a focus on the acoustic dimensions of vowel quality, in the process highlighting the relationship between acoustic and more traditional representations of vowel quality. Under the rubric 'the problem of non-invariance' the chapter then focusses on reviewing theory relating to two issues that will be taken up later in Chapter 4, which deals with methodology: i.e. the normalization problem and the problem of coarticulation and the specification of vowel targets. Next the chapter undertakes a review of the basic principles and patterns of vowel-shifting as outlined mainly in Labov (1991; 1994). This review is designed as a precursor to the theme of GenSAE's and, more generally, SAE's place in the so-called 'Southern Shift'; a theme that will be picked up intermittently in following chapters. The chapter also briefly outlines a useful framework for sub-dividing and thus describing the vowels of English, i.e. that provided by Wells (1982) and, in addition, briefly reviews the work conducted by Torgersen and Kerswill (2004) on the accents of the southeast of England, the relevance of which will become apparent as the thesis progresses.

Chapter 3 begins with the resolution of some terminological issues, before providing a broad sociohistorical picture of SAE, focussing, respectively, on its sociohistorical development, its division into three sociolects as well as prestige variants within SAE. Most importantly, and with a view towards providing a convincing reconstruction of the history of the SAE vowel system, the chapter provides a three-stage koinéization model, the utility of which is to gain clarity concerning the exogenous inputs at play in the development of SAE. An attempt is then made to integrate the sociohistorical 'picture' provided in earlier sections with Schneider's (2007) Dynamic Model of the evolution of postcolonial Englishes, a process which leads to a proposed adjustment to Schneider's (2007) model. This is followed by a preliminary attempt to determine SAE's 'place' among other accents of English, with the focus falling alternately on its status as an Extraterritorial English (ETE), its participation in Labov's (1994) Southern Shift as well as other chain-shift patterns and, lastly, on its status as a Southern Hemisphere ETE, in the process describing and highlighting Trudgill's (2004) recent work on these varieties.

Chapter 4 then turns to methodological issues. It begins with a close look at the various social variables that have been controlled as part of the *acoustic* analysis (gender, social class, age and style) and then provides additional data on the subjects. The chapter then focusses on, respectively, the processes of data collection, measurement and analysis. Lastly, it 'situates' the acoustic analysis within the broader methodological framework of the thesis, providing at the same time details of the additional acoustic data drawn upon i.e. that of Webb (1983) and that derived from research on other varieties of English.

Chapter 5, a short chapter, provides a brief review of the main work done on SAEP. It has a chronological structure, beginning with Hopwood (1928) and ending with recent work done by, for example, Bowerman (2004) and Da Silva (2007). Its focus, however, is not on the actual, mainly impressionistically-derived, values of the various SAE vowels, but rather on making explicit the different assumptions, terminologies, foci and methods of the various contributors to the existing literature on SAEP.

Chapter 6 to 9 then provide a vowel-by-vowel comparative analysis of vowel-quality in SAE. Each sub-section of each chapter is divided into four main sections:

1. a section overviewing the main features of the respective vowels in other accents of English;

2. a section providing a critical appraisal of the SAEP impressionistic literature as it pertains to the vowels in question;

3. a section comparing the acoustic data analyzed as part of the current research with acoustic data derived from Webb (1983) as well as a number of other accents of English; and

4. lastly, a section which summarizes the main conclusions to be derived from the analyses contained in the previous three sections.

Chapter 6 focusses on the three vowels involved in Labov's (1994) Pattern 1 extension (i.e. PRICE, MOUTH and CHOICE) as well as on most of the vowels involved in his Pattern 4 chain-shift (TRAP, FACE, DRESS, FLEECE)[6]. The KIT vowel is dealt with in a separate chapter, Chapter 7, given the inordinate amount of attention that it has received in the impressionistic literature on SAE. Chapter 8 then shifts the focus to those vowels involved in Labov's (1994) Pattern 3 chain-shift (BATH, THOUGHT and GOOSE) as well as GOAT and CURE. Lastly attention is paid to the remaining short vowels (STRUT, LOT and FOOT). Finally, Chapter 9 deals with the remaining vowels (NURSE, NEAR and SQUARE as well as the weak vowels of English) and with the nature and effect of final /l/.

Chapter 10.2 provides a summary of the overall analysis and provides potentially profitable avenues for further research.

[6]See §2.3.2.1 for more on these various patterns.

1.9 Conclusion

As an under-researched accent of English, SAEP will benefit from the application of modern acoustic techniques of analysis as well as from an attempt to more firmly integrate it into existing theoretical frameworks which attempt to explain, mainly from a diachronic perspective, the similarities and differences between the various accents of English, both on a structural and sociohistorical level.

This introduction has, firstly, provided a framework which is hopefully clear enough to allow the reader to understand the reasons for the conclusions drawn in the previous paragraph, and, secondly, an outline of a number of themes that receive further elaboration in the ensuing chapters. Lastly, the problems that the thesis wishes to go someway towards solving, the objectives it wishes to meet in order to solve said problems, as well as the broad structure of the thesis have been provided.

CHAPTER 2

Vowel Quality and Vowel Systems

2.1 Introduction

This chapter is focussed on providing an account of various dimensions of vowel quality and vowel systems relevant to an understanding of the work undertaken in later chapters. The chapter is consequently divided into two broad sections, the first (§2.2) devoted to vowel quality *per se*, and the second (§2.3) focussed on reviewing two frameworks which divide the vowels of English into sub-systems, both providing useful descriptive frameworks that will be employed in later chapters.

2.2 Vowel Quality

The acoustics of vowel quality is the central concern of this section. Where necessary it also deals with the articulatory production of vowel quality as well as more traditional approaches towards its representation i.e what Catford (1981:19) refers to broadly as the "'Bell-Sweet' model of vowel production and system of classification that underlies most modern description of vowels; including the system of Cardinal Vowels developed by Daniel Jones and used by the International Phonetic Association".

This section begins, in §2.2.1, with a brief overview of some of the basic principles of vowel quality, focussing on the general (though not unanimous) conclusion that vowel quality, as represented in the traditional vowel chart, more closely represents acoustic-auditory rather than articulatory fact. Section 2.2.2 is then focussed on providing a (selective) review of the acoustic correlates of some of the basic dimensions of vowel quality. It begins in §2.2.2.1 with an attempt to situate the acoustic nature of vowel quality within the broader framework of the source-filter model of speech production (Fant 1960). The articulatory basis of vowel quality is touched on, the independence of source and filter is emphasized and the notion of ‘formant’ is defined and explained. Sections §2.2.2.2 to §2.2.2.6 then look at various specific dimensions of vowel quality, including vowel height, the front-back dimension, lip-rounding, vowel length and the tense vs. lax distinction.

Section 2.2.3 focuses briefly on the nature and utility of tonotopic scales. Section 2.2.4 is, on the other hand, focussed on reviewing two aspects of a far broader issue, the existence of invariance on a symbolic, phonemic level in the face of the persistence of massive variation at the phonetic (and particularly acoustic) level. Section 2.2.4.1 focusses on the problem of intraspeaker differences with respect to vowel quality, the source of which is usually conceptualized in terms of differences in the size and shape of the vocal tracts of individuals, and reviews the various techniques that have been designed to account for (and ‘smooth out’) the acoustic differences that arise as a consequence of the relevant physiological differences, while preserving the effect of other sources of variation that may be of interest to the researcher (e.g. allophonic or, more commonly, sociolinguistic variation). Next, §2.2.4.2 focusses on the effects that intrasyllabic context has on vowel quality; in essence the variation that results from the presence of consonantal segments before and/or after the vocalic nucleus.

2.2.1 Articulatory vs. Acoustic Approaches to Vowel Quality

From an articulatory perspective, the quality of a vowel is generally determined by the shape and size of the vocal tract, except in the case of nasal vowels in which case the linked nasal cavity adds an extra source of resonance. The size and shape of the vocal tract is, in turn, mainly determined by the position and shape of the tongue and lips. The traditional vowel chart[1], is commonly meant to represent the various articulatory possibilities of these two organs. Thus the horizontal axis is often used to signify place of articulation while “the vertical dimension is divided into four areas, in effect sub-dividing the degree of stricture” (Laver 1994:276). Lip-rounding is, on the other hand, accommodated for by including two symbols for each main tongue position, the one on the left signalling the unrounded option, while the one on the right represents the rounded variant. Thus, for example, at the extreme top left-hand corner of the chart we find both [i] and [y] – both representing the

[1]The latest version of which is included, as part of the IPA, in Appendix A.

same degree of advancement of the tongue and degree of stricture but with [y] carrying the extra dimension of lip-rounding.

The extreme points of the traditional vowel chart represent the so-called "cardinal vowels" which Jones (1934:28) defines as "a set of fixed vowel-sounds having known acoustic qualities and known tongue and lip positions," the first eight of which are generally used to define the limits of the vowel space as represented by the vowel chart.

However, though it is common practice in, for example, introductory phonetics courses, to define the two axes of the traditional vowel chart in terms of place of articulation and degree of stricture, it has long been recognized that the traditional vowel chart more accurately represents an acoustic-auditory vowel space rather than an articulatory one. It was, in particular, a result of the development of the sound spectrograph and spectrogram (Potter, Kopp and Kopp 1966) and as early as Joos (1948), that the relationship between the axes of the traditional vowel chart and the *formants* of vowels was recognized. In §2.2.2.1 the notion of 'formant' will, among other things, be given a full treatment; for now the dissonance between articulation and the traditional vowel chart can be clearly appreciated in Figure 2.1, which, according to Ladefoged (2006:215), shows "the highest points of the tongue as shown in a . . . set of x-rays of cardinal vowels":

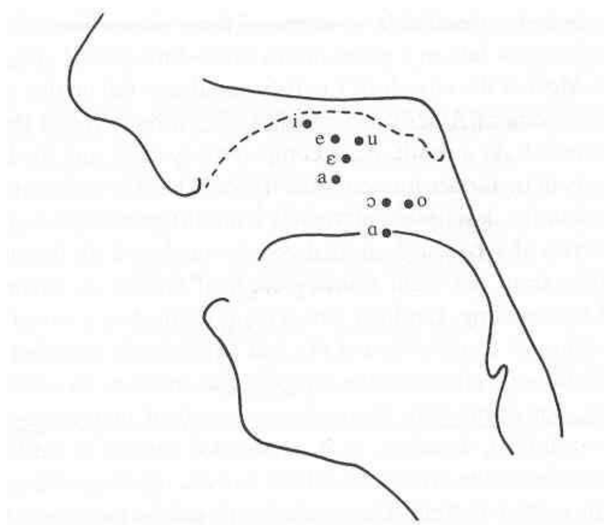


Figure 2.1: Cardinal Vowels: The Highest Points of the Tongue; taken from Ladefoged (2006:215)

Ladefoged (2001:115) summarizes this state-of-affairs in the following (somewhat condescending) manner:

> "The early phoneticians were much like astronomers before Galileo . . . These astronomers

> were certain they were describing how the stars and planets went round the earth. But they were not. The same is true of phoneticians. They thought they were describing the highest point of the tongue, but they were not. They were actually describing formant frequencies".

Still, a number of dissenting voices remain. Catford (1981), for one, takes Ladefoged, among others, to task for making reference to the highest point of the tongue as a basis for determining (traditional) vowel quality. This author claims that traditional phoneticians based their classifications on "the general positioning of the tongue as a whole, not of the highest point of the tongue" and that "the traditional proprioceptive system of vowel classification . . . still works remarkably well in practice, and . . . can be brought into harmony with modern methods of acoustic analysis" (Catford 1981:30–1). An added argument against positions such as Catford's (1981), however, is the well-attested phenomenon of different articulations producing the same acoustic result. As summarized in Nearey (1997:3242), "articulatory targets for the same vowel are quite varied across speakers. Since the corresponding acoustic output is more nearly invariant . . . linguistically relevant properties of vowels are acoustic or auditory rather than articulatory".

2.2.2 Acoustic Dimensions of Vowel Quality

This section provides an overview of the various dimensions of vowel quality, particularly (although not exclusively) as they relate to English vowels and with the added goal of situating these correlates within a distinctive-feature framework. It begins however, in §2.2.2.1 with a summary of general acoustic considerations necessary to understand the more specific facts regarding the various dimensions of vowel quality. Following subsections deal, respectively, with vowel height, the front–back dimension, the effects of lip-rounding, vowel length and the tense vs. lax distinction. In many of these sections a basic distinctive-feature framework will be utilized in order to highlight the link between the phonetic and phonological levels. The features employed are meant to be merely heuristic and in no way imply an acceptance of any particular feature-theory. They are, in most respects, based on those provided in Odden (2005).

2.2.2.1 General Considerations

Most acoustic models of vowel production are conceptualized within the more general source-filter model of speech production (Fant 1960). The model is outlined as follows in Fant (1970:191):

> "The speech wave is a product of a source and vocal tract filter or transfer function. The basic function . . . implies that somewhere in the vocal tract there originates a source, constituting the raw material of the sound whilst . . . the vocal tract provides a filtering, a shaping of the raw material resulting in the speech wave".

According to Fujimura and Erickson (1997:72), one of the cornerstones of this model is the relative independence of the two component parts i.e. while the source of the speech sound provides the 'raw material', the quality of the eventual sound is mainly determined by the filter concerned, with a small, final role being played by so-called lip-radiation (see below).

There are various different sources and filters, depending on the speech sound in question. In the case of voiceless fricatives, for example, the source is the turbulent airflow generated by the close proximity of the relevant active and passive articulators (e.g. the tip/blade of the tongue and alveolar ridge in the case of [s]), while the filter is the shape and size of the cavity anterior to the source and, in the case of sibilants, the effect that the teeth-as-obstacle have on the eventual wave-form produced.

Vocal-fold vibration forms the articulatory basis for the glottal pulse characteristic of all voiced sounds, including vowels. It is this glottal pulse that constitutes the source of the relevant speech sound. In the special case of voiceless vowels the source is the low-intensity frication created by laminar airflow through the vocal tract. In most accents of English, word-initial /h/, as in *hit*, can, acoustically, be regarded as a voiceless vowel. Thus, in the case of *hit*, the vocal-tract configuration for the initial /h/ is the same as for the following vowel, but there is no vocal-fold vibration. The filter is thus the same, but the source different.

Returning to the glottal pulse, the source of the vibration which creates the glottal pulse is the presence of positive subglottal air-pressure which "exerts a lateral force on the [underneath] surfaces of the vocal folds, causing these surfaces to move outward" (Stevens 1998:58; my parenthesis) and eventually open, leading in turn to airflow through the now open glottis. Once airflow has begun, the pressure beneath the vocal folds decreases, which in turn leads to their more rapid adduction i.e. adduction (closing) takes less time to complete than abduction (opening). On adduction of the vocal folds, the cycle begins again with a build-up of subglottal pressure. For adult speakers the rate of vibration of the vocal folds is between 100 and 300Hz, or more simply understood as 100 to 300 vibrations per second i.e. each full cycle of the vibration takes between 0.01 seconds (10 milliseconds) or 0.003 seconds (3 milliseconds) to complete.

As a result of glottal fold vibration, a similar cycle of increase, decrease and lack of airflow is created. This modulation of airflow constitutes the sound source which is subsequently filtered by the rest of the vocal tract. As mentioned the fundamental frequency (F_0) of this sound source or glottal pulse is between 100 and 300Hz. A typical male (modal-voice) F_0 would be 126 Hz and a typical female (modal-voice) F_0 would be 223Hz. These are average figures obtained by Holmberg et al. (1988) "for a large number of adult male and female speakers" (cited in Stevens, 1998:70–1).

Stiffening of the vocal folds (as a result of an increase in their length) leads, however, to a decrease in the time taken for one cycle of vibration to be completed and a consequent increase in F_0. A shortening of the vocal folds with a resultant reduction in stiffness leads to a decrease in F_0. Although an increase in the amount of subglottal pressure also leads to a slight increase in frequency, its main effect is to increase the amplitude (or loudness) in decibels (dB) of the sound wave.

Whereas the glottal pulse created by the mechanisms just described has a fundamental frequency (the F_0 referred to above), it is nonetheless a complex waveform comprised of a series of harmonics all of which are multiples of the fundamental i.e. if the fundamental frequency is 100Hz the first harmonic will be 200Hz, the second 300Hz etc. Some commentators, for example Ladefoged (1996:38), refer to the fundamental frequency as the first harmonic. This will not be the case here, mostly because it is possible for the fundamental frequency of a complex wave-form to *not* constitute a component part of the wave, but only its greatest common denominator; unlike harmonics, which are always constituents of the complex wave (Johnson 1997:10). Technical issues aside, it is, however, the harmonic series that, essentially, provides the 'raw material' for the subsequent filtering process. A spectral envelope of such a harmonic series, as contained in the glottal pulse and 'set-up' by the vocal folds, is depicted in the first box of Figure 2.2 on the next page.

As already mentioned above, the sounds produced during human speech production are the product of the source, the filtering effect of the cavities of the vocal tract and radiation from the lips. In more technical terms, "the output spectrum is the product of a source spectrum . . . , a transfer function . . . , and a radiation characteristic" (Stevens 1998:129). It is, in particular, the transfer function that determines the ultimate quality of the sound produced; in articulatory terms, the configuration of the vocal tract. For any particular configuration of the vocal tract there are a set of natural frequencies known as formants. When the transfer function is superimposed upon the source spectrum (i.e. the harmonic series), the frequency components corresponding to these natural frequencies receive a degree of amplification. These frequency components are, of course, the harmonics of F_0 i.e. the 'raw material'. This combination of source, filter and radiation can best be illustrated graphically.

Figure 2.2 is taken from Figure 3.1 in Stevens (1998:129), which shows the combination of these various influences as well as the final output spectrum of a vowel sound.

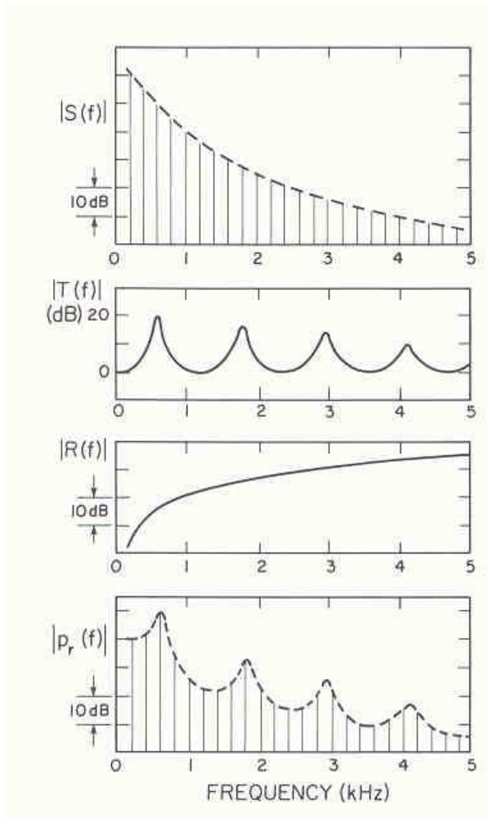


Figure 2.2: The effect of the configuration of the vocal tract on the glottal source

In this figure, S(f) equals the source spectrum, T(f) the transfer function (the natural frequencies are represented by the peaks in the spectrum), R(f) the radiation characteristic and p_r(f) the output spectrum. From this diagram one can clearly see the main effect of the configuration of the vocal tract on the final output, with the effect of the glottal source mainly being the slightly downward slope from lower to higher frequencies. The peaks in the output spectrum are also known as formants. Thus the term 'formant' is ambiguous, referring either to the peaks of the transfer function of the vocal-tract filter or to the peaks of the spectral-envelope that are, mainly, the outcome of the filtering of the source spectrum by this transfer function. Formants are generally labeled as F_1, F_2, F_3 etc. from lowest frequency upwards. Their position in terms of frequency (i.e. the x-axis on the spectrum) is intimately related to vocal tract configuration and thus vowel quality. So, for example, [iː] has a relatively low F_1 and high F_2 whilst [ɑː] has a much higher F_1 but a comparatively low F_2.

The frequencies of F_1 and F_2 are commonly used in the literature to plot the acoustic quality of vowels. Although, as will be seen later, a number of transformations are often added to this basic picture, the relationship between F_1 and traditional vowel height and F_2 and degree of 'frontness' or 'backness' forms the starting point of any acoustic analysis. As mentioned above the relationship between these two formants and the traditional vowel chart was pointed out as early as Joos (1948). According to this author, "in distinguishing among vocalic qualities, it is the lowest two resonance bands [i.e. formants] that are most significant" (Joos 1948:46; my parenthesis) and "it appears that the classical or IPA tongue-position quadrilateral rather more closely resembles the acoustic vowel quadrilateral than it resembles the . . . X-ray tongue-position quadrilateral" (Joos 1948:54). A common misconception is, however, that the first two formants relate directly to the two 'tubes' in front of and behind the stricture created by the highest point of the tongue. Thus Joos (1948:57–9) himself points out the relationship between vowel height and F_1 and frontness/backness and F_2 and tries to explain the articulatory basis of these relationships in terms of a basic two-tube model of the vocal-tract i.e. F_1 is viewed as throat resonance and F_2 as mouth resonance. This perspective has, according to Rosner and Pickering (1994:46), encouraged the following rather unfortunate misconception:

> ". . . that $F1$ and $F2$ directly reveal the size of the cavities behind and in front of the highest region of the tongue, respectively. For example, the highest region lies towards the front of the mouth for /i/, creating a large back and a small front cavity. Supposedly, $F1$ is related to the back cavity and therefore has a low frequency, while $F2$, affiliated to the front cavity, is high in frequency . . . However, this approach is oversimplified. Formant-to-cavity affiliations only hold to a limited extended".

With respect to the representation of acoustic vowel quality the general trend is to represent F_1 on an inverted y-axis and F_2 on a similarity inverted x-axis. This captures the fact that vowel height is inversely proportional to F_1 (i.e. the higher the vowel the lower the F_1) while, with respect to F_2, the backer the vowel the lower the frequency. A typical acoustic formant plot would thus look something like Figure 2.3 on the next page, which, to pre-empt a little, plots the mean values for the first two formants of the short-vowels of the GenSAE speakers who were the subjects of the current research. This figure clearly captures the high and front quality of SAE /e/, the central quality of certain allophones of SAE /ɪ/ ('kit5'), the back(ish) quality of /ʊ/ and the open quality of the /æ, ʌ, ɒ/ vowels.

Finally, the independence of the source and filter components of this model needs to be stressed again. In practice, this means that while the F_0 of a speaker might vacilate,

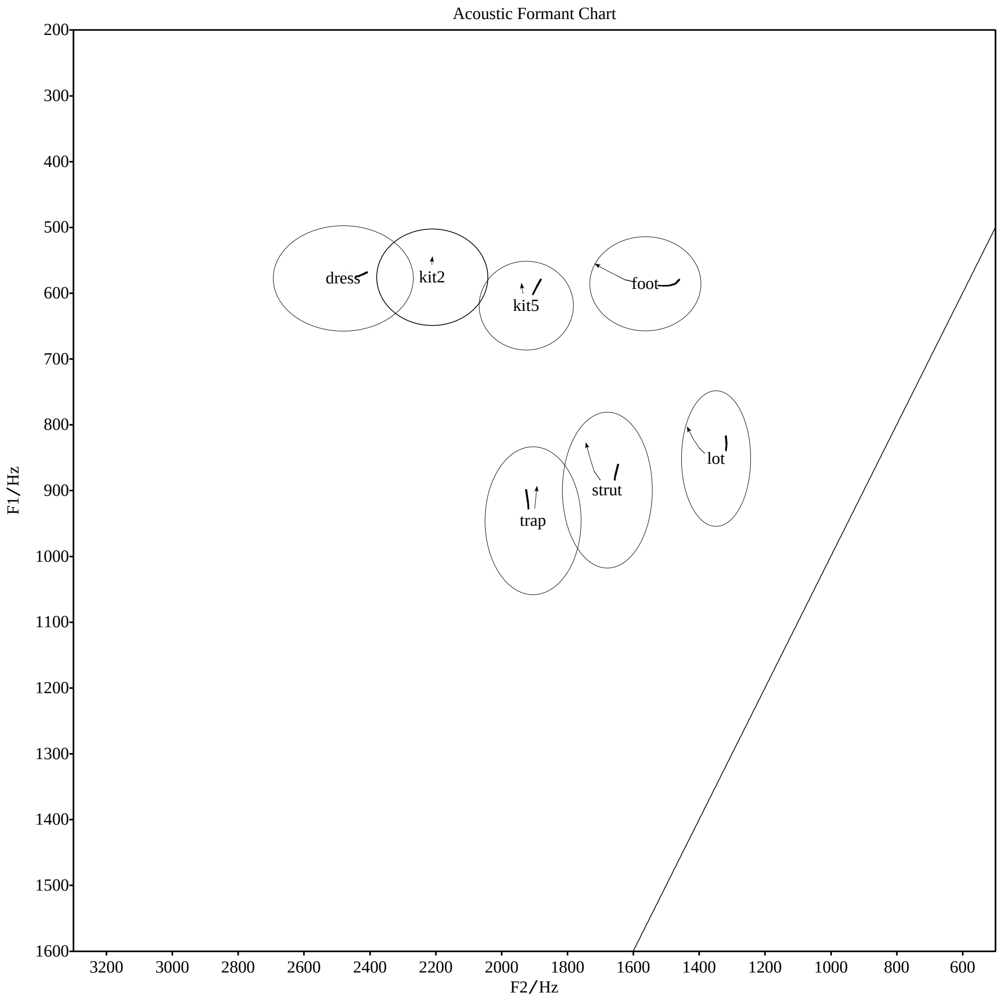


Figure 2.3: Acoustic formant chart

for intonational purposes for example, the filter can, for all intents and purposes, be held constant. This accounts for the simple fact that the same vowel quality can be produced at different frequencies.

2.2.2.2 Vowel Height

The two tense, [+high] vowels of English, /iː/ and /uː/, are characterized by a high tongue position. As a result of this high tongue position, the cross-sectional area of the vocal tract anterior to the velopharyngeal port is decreased (in comparison with the articulatory configuration for a neutral schwa-like vowel) while the section of the vocal tract posterior to this port is increased. The main effect of this configuration is a decrease in the frequency of F_1. This high tongue position is accompanied by a fronted tongue position in the case of /iː/, but with a relatively retracted tongue position in the case of /uː/[2].

Non-high vowels, i.e. [-high], are, of course, distinguished from high vowels by a higher F_1. In terms of the distinctive-feature scheme provided for by, for example, Stevens (1998) or Odden (2005), [-high] vowels are themselves divisible into two sub-groups: low vowels and non-low vowels. Low vowels, [-high], [+low], such as /ɑː/ and /æ/, are characterised by a decrease in cross-sectional area of the posterior pharyngeal region and an increase in cross-sectional area in the anterior region. Both these effects are the result of tongue lowering and retraction. The overall result (which is in effect the opposite as that for high vowels) is a dramatic increase in the frequency of the F_1.

Non-low (non-high) vowels, i.e. [-high], [-low], such as /e/ and /oː/, have articulatory configurations which are intermediate between the two extremes represented by the high and low vowels discussed above. So, for example, the cross-sectional area of the posterior region will not be as constricted as it is for low vowels, nor as large as it is for high vowels. As a consequence the F_1 is also intermediate between the values common for high and low vowels.

Another factor that needs to be taken into account is the fact that in the case of front vowels the cross-sectional area of the anterior portion of the vocal tract decreases with vowel height i.e. as the tongue is lowered to produce lower vowels the area increases. Thus the F_2 maximum of [+high] front vowels is higher than the F_2 maximum of [-high] front

[2] This state-of-affairs, however, needs some qualification with regard to many accents of English, including SAE, since they are often characterized by a substantially fronted /uː/ vowel i.e. [ʉː] to [yː] – see §8.2. In extreme cases of fronting it is conceivable, therefore, that the distinction between /iː/ and /uː/ might be more a matter of absence or presence of lip-rounding rather than the positioning of the tongue body on the horizontal plane – see §2.2.2.4 for more on the acoustic effects of lip-rounding. The possibility for arguing, in extreme cases of GOOSE-fronting, for a *phonological* reanalysis of the difference between FLEECE and GOOSE as involving [$\pm$ round] rather than [$\pm$ back] is an interesting one, but beyond the scope of this thesis.

vowels and in general the lower the vowel the lower the F_2 maximum. This of course explains the characteristic 'slope' of the front part of the traditional vowel chart. On an instrumental level, Labov (1994:160), for example, defines vowel height in terms of the following formula, which captures the fact that "the great majority of raisings are not simply alterations of F1, but rather combine changes in both F1 and F2 along this dimension":

$$\sqrt{F2^2 - (2 * F1)^2}$$

As can be seen from this formula, the raising of a vowel along the front diagonal usually, in fact, involves a greater shift in F_2 than F_1.

2.2.2.3 The Front-Back Distinction

In very general terms, the acoustic effect of the tongue on the front-back dimension is quite simply that the more fronted the position of the tongue the higher the F_2, while a backed tongue position leads to a lower F_2. As a general principle, this applies across all vowel heights. On a phonological level often only one feature, e.g. [$\pm$ back], is used to capture this dimension.

In the previous section, the front diagonal of the traditional vowel chart was related to instrumental correlates of vowel height along this diagonal. This front diagonal is, of course, in contrast with the more perpendicular 'back line' of the traditional IPA chart – see Appendix A. There seems to be some consensus that the straightness of this line is more indicative of the *difference* between F_2 and F_1 (F_2 - F_1) than of the value of F_2 itself and, in fact, some representations of acoustic vowel quality employ an F_2-F_1 transform in order to bring the acoustic representation closer to that of the traditional vowel chart.

In charts employing F_2 alone there is, on the contrary, a characteristic slope from a low central position to a high back position (in the case of monophthongs usually ending with /oː/ in high-back position given the fronted articulation of /uː/ in many accents of English). This can be clearly seen in Figure 2.4 on the following page, which provides data on the /uː/, /əʊ/ and /aʊ/ vowels of GenSAE. As will be seen in later chapters, a final, tautosyllabic 'dark' /l/ in SAE (i.e. [ɫ]) has a substantial retracting influence on preceding vowel quality. In Figure 2.4, a subscript 'l' has been added to the end of the pre-/l/ allophones of each phoneme. As can be seen these allophones follow a path which corresponds roughly with the F_2=F_1 line that slopes from the bottom of Figure 2.4 to the top right. A chart employing a F_2-F_1 transform for the x-axis would thus effectively 'straighten up' these three retracted allophones. While an F_2-F_1 representation of the front-back dimension has the advantage of bringing the acoustic vowel chart more in line with the traditional one, the use of F_2

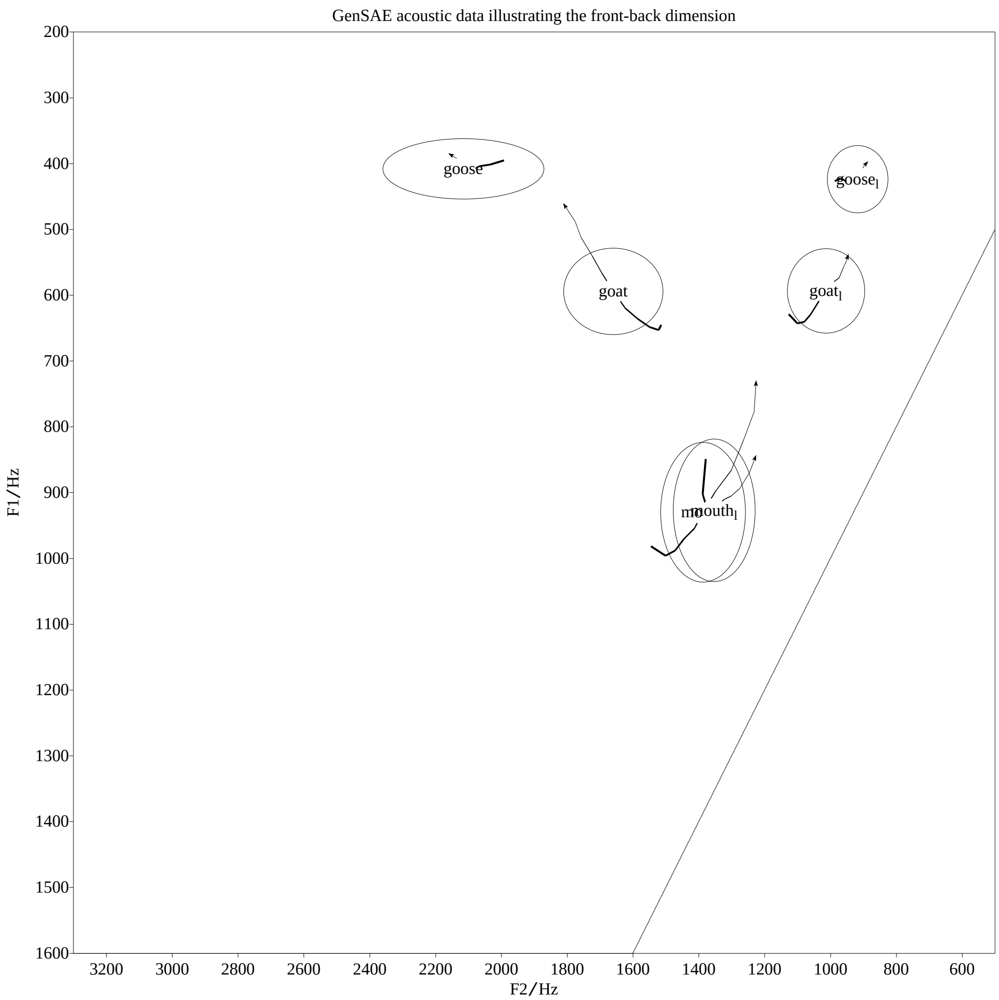


Figure 2.4: GenSAE acoustic data illustrating the front-back dimension

values alone is more in-line with the procedures employed by Labov (1994), the F_2=F_1 line being used in conjunction with the formula provided in §2.2.2.2 to define the outer limits of a triangular phonetic space.

As will become clearer in §2.2.2.4, the position and articulation of the lips often has a dramatic effect on the F_2 value of a vowel, and to a lesser extent on the F_1 as well.

2.2.2.4 The Effects of Lip-Rounding

According to Docherty and Foulkes (1999:53), "the effect of lip-rounding is ... difficult to represent unambiguously". Furthermore, while for most purposes the division of vowels into rounded and unrounded suffices, it should be noted that the classic representation of rounding on the IPA vowel chart, e.g. [i] vs. [y], is highly idealized given that lip-rounding is, on phonetic grounds at least, a continuum. Jones (1934:39), for one, provides a more subtle framework and distinguishes between a spread, neutral and rounded lip position and divides lip-rounding into close and open varieties. According to Ladefoged (2007:16), even on a phonemic level "we probably need ... three possible values for Rounding, [neutral], [rounded], and [over-rounded]. ... Further investigation may show that we need a fourth value [spread]". For phonological purposes, this thesis, while recognizing the possible over-simplification it implies, will retain the more traditional [$\pm$ round].

As has been suggested above, the fundamental effect of lip-rounding is a decrease in the F_2 of the vowel concerned i.e. given an equivalent tongue configuration the F_2 will be lower than it would have otherwise been. Figure 2.5 on the next page, which provides the spectra of two non-low back vowels and which is taken from Fig 6.19 in Stevens (1998:291), illustrates this effect nicely. It is clear from this figure that the F_2 of the rounded-vowel spectrum is closer to the F_1 than in the case of the non-rounded-vowel spectrum.

The effect of lip-rounding on front vowels is somewhat similar. Again this is best illustrated with examples of spectra. Figure 2.6 on page 37, taken from Fig 6.21 in Stevens (1998:293), shows the difference between the unrounded high front vowel [iː] and the rounded high front vowel [yː]. This difference in F_2 between these two vowels is important to emphasise since it implies that the same tongue configuration (i.e. degree of *articulatory* frontness) produces a higher F_2 when the lips are not rounded but a relatively lower F_2 when lip-rounding is added. This is of particular relevance to accents such as SAE which display substantial fronting of /uː/.

The lip-rounding of vowels also usually causes a lowering of F_1. While this is generally true for most vowels, the effect is most clearly visible with respect to low vowels. Thus the rounding of a low back vowel (e.g. /ɒ/) would cause a lowering of F_2 plus, in addition, a relatively substantial lowering of F_1 (i.e. greater vowel height), both a result of the increase in size of the anterior cavity.

The overall effect of lip-rounding is captured beautifully in Figure 2.7 on the next page, taken from Fig 6.2.2. in (Stevens 1998:292). In this diagram the solid-dot vowels are the four unrounded vowels [iː,eː,æ,ɑː] and the two rounded back vowels [oː,uː]. Arrows are indicative of the effect of rounding. As can be seen the main effects of rounding are a decrease in F_2 and (particularly in the case of the low vowels) a decrease in F_1.

The opposite of lip-rounding, i.e. lip-spreading, would presumably create the opposite effect of lip-rounding i.e. an increase in both F_2 and F_1. Thus, by way of example, for low front vowels such as /æ/, Stevens (1998:276) mentions that "the lips can be spread to decrease the effective length of the vocal tract and thus to maintain as high an F1 and F2 as possible".

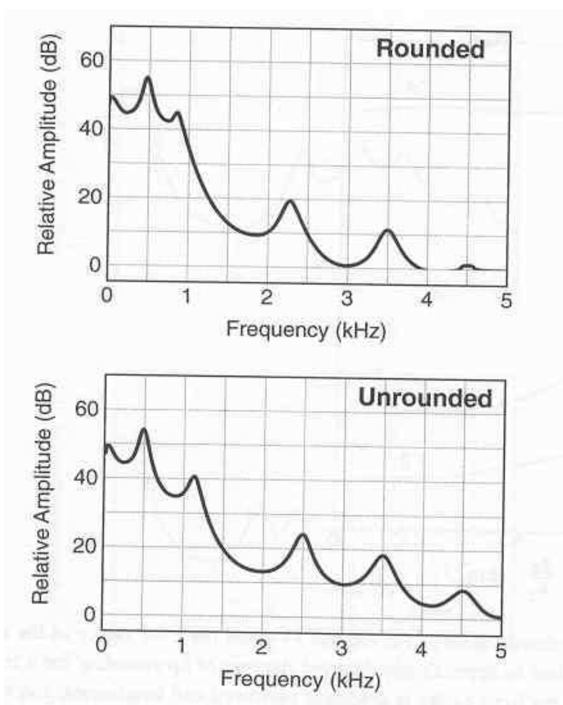


Figure 2.5: Spectra of two non-low back vowels, one rounded the other unrounded

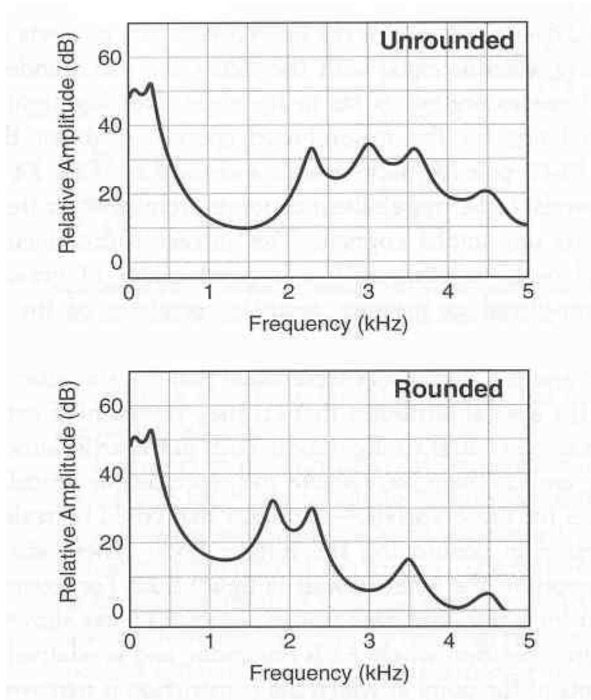


Figure 2.6: Spectra of unrounded [iː] and rounded [yː]

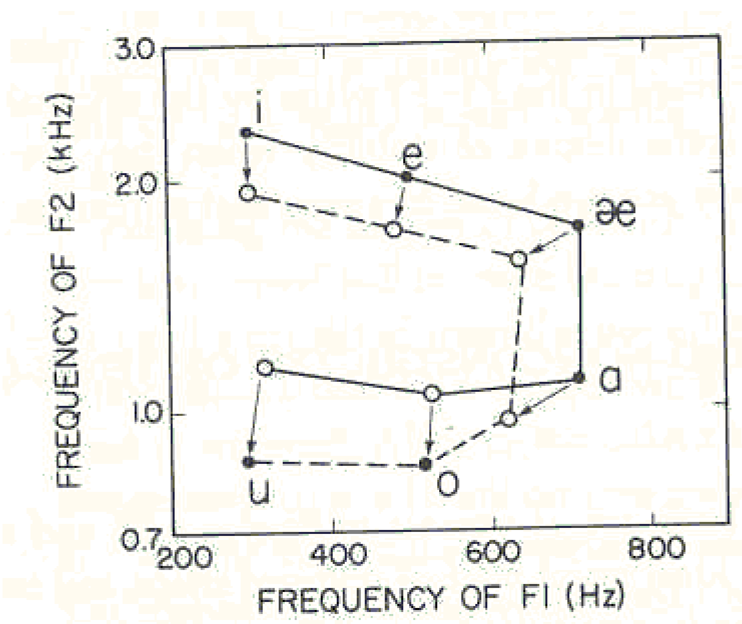


Figure 2.7: The effects of lip-rounding

2.2.2.5 Vowel Length

It is well-known that vowel duration in English is affected by the following consonant. In many accents of English, vowels are lengthened prior to voiced consonants and shortened before voiceless consonants. Thus, in SAE, the vowel in *bad* is substantially longer than the vowel in *bat*. Above and beyond this allophonic patterning, however, there is also what Peterson and Lehiste (1960:701) refer to as "instrinsic durations of syllable nuclei in English". In essence certain vowels are instrically longer than others. Thus, in SAE, the vowel in *keen* is longer than the vowel in *kin* although the environments are identical. This interplay between intrinsic and allophonically-determined length often leads to "durational overlap between the durations of long nucleus plus voiceless consonant and short nucleus plus voiced consonant, for example in such sets as *bead-beat-bit-bid*" (Peterson and Lehiste 1960:702).

Perhaps more centrally, length is not generally viewed as phonemically important in English since difference in length is usually accompanied by a difference in vowel quality. Thus while /uː/ is long and /ʊ/ is short, these phonemes are also generally realized differently in terms of their position in F_1/F_2 space. There appears to be some evidence, however, that the SAE vowel system has a number of cases where this general principle does not apply i.e. there are pairs of vowel phonemes distinguished principally by length i.e. by the [$\pm$ long] feature. These will be dealt with on a case-by-case basis in later chapters.

2.2.2.6 The Tense-Lax Distinction

Vowels positioned on the periphery of the vowel quadrilateral can be thought of, in articulatory terms, as representative of those tongue configurations that are furthermost away from the configuration which produces schwa, [ə], where schwa is represented by the mid-point of the vowel quadrilateral and characterised by a "uniform vocal tract cross-sectional area" (Stevens 1998:284). Such peripheral vowels are also known as tense vowels and are contrasted with so-called lax vowels which do not show such extreme movement away from the neutral position. In the case of non-low vowels, the tense /uː/ and /iː/ have, for example, their lax counterparts /ʊ/ and /ɪ/, the latter vowel being produced, in comparison to /iː/, with a less wide pharyngeal region and an increase in size of the anterior oral region.

More importantly for the purposes of the current discussion, Labov (1994:174–5) appropriates the notion of a [$\pm$ tense] category, relating it, in turn, to his notions of the peripheral and nonperipheral tracks of acoustic vowel space. The tense vs. lax distinction and its relationship to peripherality play an integral part in Labov's (1994) general principles of vowel shifting, as further elaborated on in §2.3.2. In his conception, the feature [$\pm$ tense] is related to a range of different phonetic phenomena, and applies, in the case of diphthongs, not to the whole vowel but to the nucleus and the glide as separate entities. As far as the phonetic manifestation of [$\pm$ tense] is concerned, the most important connection for now is the link between tense nuclei and a position "closer to the periphery of the two– or three–formant vowel space" (Labov 1994:175). Since Labov (1994) generally uses F_2, and not F_2-F_1, this acoustic vowel space is representable as a triangle with two main tracks, as illustrated in Figure 2.8. In this figure, the two outer areas represent the peripheral and non-peripheral tracks respectively.

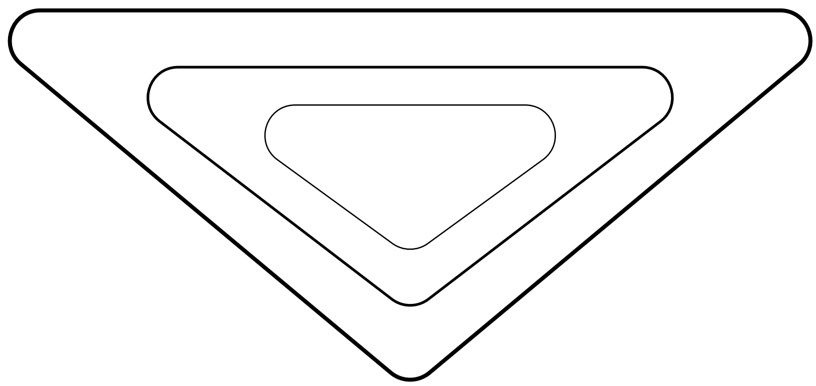


Figure 2.8: Labov's (1994) acoustic vowel triangle

2.2.3 A Brief Note on Tonotopic Scales

As emphasized by Johnson (1997:55), "the auditory system's frequency response is not linear". In other words, frequency is not the same as pitch and in practical terms this means, for example, that while the difference between 100Hz and 50Hz is equivalent on an *acoustic* level to the difference between 1100Hz and 1050Hz, the two differences are certainly not equivalent on a perceptual level. The non-linear, quasi-logarithmic relationship between frequency and perceived pitch is captured by, for example, the mel scale, which is constructed "from a series of ... judgments, ... such that frequencies that are spaced at equal distances along this scale are judged to be related by the same pitch ratio" (Stevens 1998:227).

A number of different scaling methods, or auditory transforms, have been utilized over the years to model the change from frequency to its perceptual derivative i.e. pitch (Rosner and Pickering 1994:16–19). These include the logarithmic (log) transform, the mel scale (Stevens and Volkmann 1940) as described above, the Koenig transform (Koenig 1949), the critical-band-rate (CB-rate) scale (Zwicker 1961) and the equivalent-rectangular-bandwidth-rate (ERB-rate) scale (Moore and Glasberg 1983).

Joos (1948:52), for example, uses the logarithmic scale for the design of his acoustic vowel charts: "the scales are not linear, but logarithmic like the musical scale". A log-based auditory transformation also forms the basis of a number of normalization techniques – see §2.2.4.1, likewise with the mel and Koenig scales. Of particular relevance are, however, the last two scales, the CB-rate and ERB-rate scales, since both of these have become increasingly common in sociophonetic research, perhaps as a result of comments such as that of Watt and Tillotson (2001:275–6; footnote):

> "The use of psychoacoustic frequency transforms – the Bark, Koenig or mel scales, for instance – is well established in experimental phonetic research ... but is practically absent from current work on sociophonetics. While a number of researchers in the latter field recognise the importance of treating speech sounds more as perceptual objects than as purely acoustic events ... psychoacoustic scales are conspicuous by their absence in Labov's work and the paradigm it has generated".

Thus, by way of example, the relatively recent Hawkins and Midgley (2005) use the ERB-scale in their acoustic analysis of RP monophthongs.

What Watt and Tillotson (2001) fail to discern, however, is that the logorithmic contraction, particularly in the high F_2 regions, that the use of such scales imply, while perceptually valid, often obscure important distinctions of *production* (Labov 1994:165). Simply put, "though diagrams with logarithmic second formant displays correspond more closely to even perceptual spacing, the more expanded view of the second formant is helpful in exploring the dimension of peripherality" (Labov 1994:xi).

2.2.4 The Problem of Acoustic Non-Invariance

In their article on the acoustic properties of English fricatives, Jongman, Wayland and Wong (2000:1252) point out that one of the main problems in phonetic research and, in particular, in acoustic phonetic research, is the identification of invariant acoustic cues (if any) that characterize the various classes of speech sound. While on the level of the linguistic system itself, units such as phonemes are for all intents and purposes perceptually invariant and discrete, on a phonetic level there is a contrasting high degree of variance. It is this variance that accounts, for example, for the fact that, while we attempt to *imitate* nonlinguistic and foreign sounds, we *repeat* utterances in our own languages i.e. speakers "reproduce the utterance by substituting, for the sounds they heard, the particular corresponding variants that they habitually pronounce" (Harris 1964:36).

The high degree of variance across sound signals has it roots in "a variety of sources, including speaker size, phonetic context and speaking rate" (Jongman et al. 2000:1252). The influence of phonetic context, as derived from the coarticulation of succeeding phonemes, has in particular troubled researchers. As summarized by Sussman, McCaffrey and Matthews (1991:1309; my parenthesis):

> "The pervasive coarticulatory nature of speech has contributed to . . . [the] . . . theoretical puzzle known as the "noninvariance problem" – i.e., perceptual constancy despite physical variation in the signal".

While the non-invariance problem is often discussed in the light of attempts to understand how humans 'extract' invariant features from the acoustic substrate, focussing thus on the perception of phonemic units, the problem of non-invariance has equal (though indirect) relevance to intraphonemic distinctions, particularly those of a sociolinguistic nature. When it comes to vowel perception in particular we note in the next section that the challenge is to not only neutralize differences between perceptually identical vowel tokens, but at the same time to preserve relevant distinctions not only on an allophonic, but also on a sociolinguistic level.

2.2.4.1 The Vowel Normalization Problem

The problem of vowel normalization (Pisoni and Luce 1986:10–1) relates to two sources of variation mentioned above i.e. speaker size and speaking rate. The former kind of normalization will be the focus of what follows, given that the data collection procedures employed in this research impose a degree of control on the latter[3].

Different speakers have different vocal-tract shapes and sizes. Their articulation of the same speech sound and the corresponding acoustic output will therefore differ in absolute terms. Any invariance encoded in the speech signal must, as a result, be of a relative nature.

The problem of normalization, although discussed by Joos (1948), is generally traced back to Peterson and Barney (1952), who provided the first clear picture of the effect of the vocal tract sizes of men, women and children on the formant values of vowels. Interestingly enough, there is some evidence for the possibility of the vocal-tract changing its shape and size over the course of one individual's lifespan, and not just on a physiological level. Thus, Story (2008), for example, provides compelling evidence to show how one individual's vocal-tract setting has changed over the course of a few years.

Since Peterson and Barney (1952), there have been a plethora of vowel normalization techniques, each with its different emphasis and theoretical assumptions. Following Adank (2003), we can adopt a four-way classification of such techniques. In what follows, a technique is vowel-intrinsic if it uses information contained in only one vowel to normalize that

[3] See Chapter 4 for more on this and other methodological issues.

vowel and vowel-extrinsic if it uses information across vowels. A formant-intrinsic method of normalization derives dimensions (in practical terms, axes of the relevant plot) based on information derived from one formant, while formant-extrinsic methods derive dimensions from information across formants or from the relationship between formants. Note that Adank (2003:14) uses the term 'formant' somewhat loosely to include F_0; the same practice will be employed in what follows (but not elsewhere).

Vowel-intrinsic / formant-intrinsic techniques: These are, essentially, the various auditory transforms discussed in §2.2.3, the point being that they do, to some degree, 'smooth out' differences between speakers as a result of the fact that there is greater contraction as frequency increases. As pointed out by Watt and Fabricus (2003:161; my parenthesis), however, these scales do "not in fact fully permit direct comparison of one speaker's vowel sample with another speaker's vowel sample in the way that we would wish. This is because the influence of VTL [Vocal Tract Length] is not actually wholly eliminated". While the previous remarks are aimed specifically at the bark-scale they are meant to stand as a critique of all auditory transforms. The main criticism is that in the frequency domain in which the F_1 of a vowel finds itself, "between c. 200Hz and 1kHz – the mapping between Hz and Barks is effectively linear" (Watt and Fabricus 2003:161). It should be pointed out in this regard that the last comment is only partly true, if one accepts that this is meant as a general critique against auditory transforms. The log-scale is, by definition, not linear at all, and the same would appear to apply to the ERB-scale (Moore and Glasberg 1983).

Vowel-intrinsic / formant-extrinsic techniques: The example provided by Adank (2003:19–20) is Syrdal and Gopal's (1986) normalization technique based on bark-differences, mainly between bark-transformed F_1 and F_0 (representing vowel height) and between bark-transformed F_3 and F_2 (representing the front-back dimension). The method is vowel-intrinsic because the resultant dimensions are based on ratios derived from formant measurements (and the accompanying auditory transformation) from the vowel in question i.e. each vowel is normalized independently of the other. The technique is formant-extrinsic because each dimension is based on a *relationship* between formants.

Vowel-extrinsic / formant-intrinsic techniques: Here the examples mentioned by Adank (2003:21–3) are those of Gerstman (1968), Lobanov (1971) and the individual log-mean model in Nearey (1977). The latter two normalization techniques are particularly common in sociophonetic research of the Labovian mould. They all rely on information *across* vowels, but each normalized dimension relates to only one formant. Thus, for example, Lobanov (1971:606) provides the following formula:

$$F_i^N = (F_i - Mi)/\sigma_i$$

which basically says that the normalized value of any formant of any vowel-token (F_i^N) is equal to the mean frequency value for that formant for all the vowel-tokens of the relevant speaker (Mi) subtracted from the frequency value of the vowel-token formant in question (F_i) and then subsequently divided by the standard deviation of the formant frequency values of all the vowel-tokens for the same speaker (σ_i).

Vowel-extrinsic / formant-extrinsic techniques: The normalization methods included under this rubric by Adank (2003:23–5) are Nearey's (1977) shared log-mean model and those contained in Norström and Lindblom (1975) and Miller (1989). Each of these methods relies on information from more than one token. So, for example, Nearey's (1977) CLIH$_{s4}$ model uses a mean derived from the log-transformed frequencies of the F$_0$, F$_1$, F$_2$ and F$_3$, Norström and Lindblom's (1975) vocal-tract transformation model derives a mean F$_3$ across vowel tokens in order to estimate the relevant speaker's vocal-tract length, while Miller's (1989) formant ratio transformation includes a Sensory Reference (SR) point based on the "geometric mean of the current speaker's voice pitch" (Miller 1989:2121).

According to Thomas and Kendall (2007), another useful distinction can also be made between so-called speaker-intrinsic and speaker-extrinsic normalization methods. The first kind normalizes each individual's vowels separately, while the second kind normalizes a number of speakers together. While all of the above normalization techniques are, in most guises, speaker intrinsic, many of them can be 'turned' into speaker-extrinsic methods, as discussed in Thomas and Kendall (2007). The most well-known speaker-extrinsic method is that used in Labov, Ash and Boberg (2006). It is accepted, however, that this method is generally effective only with a very large number of subjects.

As stressed by Docherty and Foulkes (1999:53), "there is no inscrutable algorithm for transforming the mathematical differences between speakers". Still, it would appear that some normalization techniques are better than others. Adank (2003) is specifically focussed on analyzing the utility of the various normalization techniques mentioned above for sociolinguistic purposes. At the end of the day the technique needs to meet the following criteria, as expressed in Hindle (1978:167):

"…it is essential to have a transformation that will minimize formant differences between individuals due to inherent physiological factors, but will preserve distinctions that correspond to perceptibly different vowels. This need suggests a third criterion in addition to clustering and separation ... namely, that a successful normalization will reveal socially significant differences. For example, in a speech community where there are perceptible differences between the vowels of young and old speakers ... the normalization procedure must preserve this relation".

While the exact details of Adank's (2003) analysis is beyond the scope of this discussion, her conclusions are pertinent. In essence, the vowel-extrinsic forms of normalization are unambiguously seen to perform better than the vowel-intrinsic methods at reducing the influence of vocal-tract length while preserving relevant sociolinguistic differences. It was, in particular, the vowel-extrinsic, formant-intrinsic methods that stood out i.e. those of Lobanov (1971), Gerstman (1968) and the CLIH$_{i4}$ model of Nearey (1977). This result is reflected in more recent work such as Escudero and Bion (2007:1414) who choose "not to include any vowel-intrinsic normalization procedure ... because previous studies ... have shown that they perform much worse than the extrinsic procedures".

As such it was decided, in line with the above-mentioned review of the literature on this topic, to employ a vowel-extrinsic, formant intrinsic method of vowel normalization, in particular the Lobanov method, the formula for which has already been provided above[4].

2.2.4.2 Coarticulation and Vowel Targets

Vowel quality is in almost all cases a dynamic phenomenon, the basis for this dynamism being CV or VC transitions as well as vowel-inherent spectral change (VISC) (Nearey and Assmann 1986). At least for monophthongs, and for short monophthongs in particular, and

[4] As an aside, the irony of this situation is that many of the vowel-extrinsic methods "were designed for improving vowel classification, often for automatic speech recognition" and are not necessarily focussed on "modeling human vowel perception" (Adank 2003:14). The most successful methods of normalization therefore appear to be those which have only a tangential relationship with the normalization procedures employed during *human* vowel-normalization. As emphasized by Harrington and Cassidy (1999:77), "the evidence suggests that extrinsic cues can *influence* vowel quality". Nearey (1989), for example, includes "extrinsic (transsegmental) relational properties" as one of four types of information impacting on the perception of (English) vowels. "What is less clear is whether extrinsic information is *necessary* to judge vowel quality accurately" (Harrington and Cassidy 1999:77). The experimental evidence seems to point to a negative answer. Without entering into the experimental details, Verbrugge, Strange, Shankweiler and Edman (1976:208) conclude that "talker-dependent acoustic variation does not pose a major perceptual problem within a common dialect group. Listeners can identify a high proportion of vowels spoken in citation-form syllables by talkers with whom they have little or no previous experience"; while Assmann, Nearey and Hogan (1982) found that the difference in terms of vowel recognition was marginal between contexts where the same speaker produced the vowels or where the speakers differed. According to these authors, "the extreme version of the relative formant normalization hypothesis is not supported by the data ..., since speaker-randomized isolated vowels are well recognized by listeners" (Assmann et al. 1982:979).

especially in sociophonetic studies in the Labovian mould, it is common practice to determine vowel quality on the basis of formant measurements (F_1 and F_2) at the vowel-target i.e. that segment or point of the vowel that is deemed to constitute the target of the articulatory process. Thus, Harrington and Cassidy (1999:59) claim that the "the acoustic vowel target, which in monophthongs typically occurs near the vowel's temporal midpoint . . . is presumed both to be the section of the vowel that is influenced least by context effects and to be relatively steady-state (that is, unchanging)". Various methods in the sociophonetic literature have been used to establish the vowel targets of monophthongal vowels. As confirmed in Labov (1994:165), the most common method involves establishing points of inflection in the formant trajectories, "and the most common such point of inflection is the F1 maximum reached when the consonantal influence of the onset or coda is minimal". In a more recent paper the term 'turning point' has been used as an equivalent for point of inflection (Langstrof 2006:143). Boberg (2005:7–8) employs the following set of criteria in his measurements of the short monophthongs of Canadian English:

> "The measurement was made at the point that best represented the central tendency of the vowel. This was usually the maximum value of F1, or the middle of a steady state of F1. In a few cases, the trajectory of F2 was used as a guide to establishing a more precise point of measurement within an F1 steady state."

The use of the F_1-max as a sampling point is particularly useful, given that it is not dependent on "the determination of the onset and offset of the vowel, and consequently on measurements of its duration" (Di Benedetto 1989:58), as, for example, the use of the temporal midpoint would be.

The term 'vowel target' has an unfortunate ambiguity in the literature though. On the one hand, and from one perspective, each vowel has "an ideal articulatory configuration represented acoustically by an ideal formant pattern" (Di Benedetto 1989:59). Underlying this perspective is the search for potential invariant cues for vowel identity and, as mentioned, the target values in question are often "ideal" ones e.g. those found in hVD, #Vd (Di Benedetto 1989:56) or isolated citation-form contexts (Harrington and Cassidy 1999:69). According to Andruski and Nearey (1992:390), it is often the case that in consonantal contexts such a vowel target is not present; more particularly, the formant values fail to reach the target values due to coarticulatory pressures, a phenomenon known as formant undershoot (Harrington and Cassidy 1999:69), a form of target undershoot which is defined by, for example, Oh (2008:363), as "deviation of a sound from its target value toward those of other segments in the context". Similarly, Halle and Stevens (1964:605) claim that for speech in general, "when utterances are produced at any but the slowest rates, a given articulatory configuration may not be reached before motion toward the next must be initiated". Formant undershoot is thus just one instance of a more general phenomenon and target undershoot is a natural and common reflex of coarticulation. As will be touched on later, target undershoot also appears to provide a fertile ground for language change. Bauer (2008), for example, establishes a clear link between such processes and the general diachronic process of lenition. Interestingly enough, Di Benedetto (1989) also provides examples of ideal vowel targets being overshot, thus a possible source of historical changes involving fortition.

The fact that vowel targets, as conceived of in the ideal sense, are often not realised, has resulted in some dissatisfaction with this ideal notion and has led more recently to a search for "other cues to vowel identity beyond the traditional, steady-state formant targets" (Andruski and Nearey 1992:390). Di Benedetto (1989:55), for one, shows that F_1 maximum is not a sufficient mechanism for disambiguating vowels spoken by the same speaker with regard to the height feature:

> "Examination of the $F1$ trajectories of the vowels for which confusion occurs shows variations in the way $F1$ reaches its maximum among different vowels. In particular, if two different vowels such as [ɪ] and [ɛ] have the same maximum $F1$, then $F1$ for the lower vowel reaches its maximum value earlier".

From a sociophonetic perspective, on the other hand, where interest lies in intrinsic variation of different sorts rather than the extraction of invariant features or 'targets', the allophonic variety assumed by the various consonantal contexts is important to capture. From this perspective, the vowel target is far from ideal and is definable as that point of the formant trajectory *least* influenced by consonantal effects. It is thus, perhaps, best to distinguish between the ideal, abstract *phonemic* vowel target and the actual *phonetic* target reached in any particular instance. In an acoustic analysis such as the current one, the focus would fall (mainly) on the latter.

Diphthongs, on the other hand, have generally been viewed as containing two vowel targets, an onset target and an offset target. There has, however, been some debate about the exact relationship between these two targets. Thus, according to Harrington and Cassidy (1999:66), there are essentially three hypotheses relating to this relationship. Firstly, there is the *dual target* hypothesis which claims that the actual values of the two targets forms the primary basis for diphthong identification. Secondly, the *onset plus slope* hypothesis according to which "diphthong quality is presumed to depend on the first target and rate of spectral change towards the second" (Harrington and Cassidy 1999:66) and, lastly, the *onset plus direction* hypothesis, the nature of which should be self-evident. Andruski and Nearey (1992:390) use the term "compound target approach" as a cover-term for all three hypotheses and, where necessary, the same device will be employed here. The general consensus, however, is that both targets are crucial for the identification of diphthong vowel quality. More interestingly, a recent attempt to provide an improvement on the ideal notion of a vowel target, as discussed above, has involved the extension of this model to monophthongal vowels (Nearey and Assmann 1986, Andruski and Nearey 1992). Andruski and Nearey (1992:390) summarize this approach in the following manner:

> "... the "vowel-inherent dynamic" or compound target ... approach ... states that properties of both *nucleus* and vocalic *offglide* sections must be considered. This theory suggests that not only phonemic diphthongs, but the majority of English vowels, exhibit characteristic kinds of formant movement and that vowel-inherent spectral change (VISC) ... is important for the perception of English vowels".

While the relevant research is, again, focussed on the identification of acoustic cues for vowel identification on a phonemic level, many of the arguments apply equally well to sub-phonemic vowel quality. In fact the inclusion of information above-and-beyond the vowel-target formant values (however defined) seems particularly warranted in descriptive accounts which, ideally, deal with qualities that are more finely-grained than the purely phonemic. Thus, to provide a hypothetical example, while the systematic difference in *degree* of monophthongization of the PRICE vowel between GenSAE and Cultivated SAE subjects might have no effect on the identification of the various tokens as members of /aɪ/ per se, it might constitute an important indicator or marker of the social identity of these two groups.

Two other 'improvements' to the notion of an ideal target are also reviewed by Andruski and Nearey (1992:391–2). The first, known as the *undershoot compensation* theory basically claims that in cases of formant undershoot, the perceptual system, in effect, 'reconstructs' the relevant vowel target, a process which most often relies on knowledge of the coarticulatory effect that the relevant consonantal context has on the vowel in question. The third theory is the *dynamic cospecification* one, which claims, perhaps surprisingly, that the main cues for vowel identification lie in the VC and CV transition periods – see below for more on this.

Another *prima facie* reason for extending the analysis of all, particularly phonemically monophthongal, vowels in order to include VISC, is that vowels do change in terms of their status as monophthongal or short. In the case of long monophthongs, there is ample evidence to suggest that the boundary between such vowels and so-called diphthongs is a tenuous one at best, and even in the case of short 'monophthongal' vowels things are not intrinsically stable. Short vowels become tense, or in Labov's (1994) terms, move onto the peripheral track, and often (although not necessarily) gain length and VISC in the process. A famous example of this is short **a** in American English, which has, in some dialects, risen on a wholesale basis along the peripheral track, often gaining an inglide in the process e.g. [mɪən] for *man*; notice that a sub-category of this phoneme (/æh/) are the traditionally short vowels found in words such as *cap*, *mat* etc. On the whole, therefore, there seems to be sufficient grounds for not assuming the monophthongal or diphthongal status of a particular vowel and, in addition, for indeed assuming that a descriptively adequate (acoustic or impressionistic) account of a particular accent, as is one of the foci of this research, requires the tracking of vowel quality over time.

An acceptance of the position argued for in the previous paragraph, however, raises two additional related issues, that of coarticulation and segmentation. It follows logically that if one wishes to track vowel quality over time one needs to make a decision about when the vowel 'begins' and when the vowel 'ends'; and perhaps even about whether these are valid expressions given the coarticulatory nature of speech i.e. the fact that "a given phonetic parameter cannot change instantaneously from one value to another; the transitions from one target configuration to the next must be gradual, or smooth . . . Thus, the configuration at any given time may be the result of instructions from more than one phoneme" (Halle and Stevens 1964:605). The question, ultimately, is about whether or not (and *how*, *if possible*) to exclude such coarticulatory effects from one's description of vowel quality.

As mentioned above, in a sociophonetic study it seems trivially obvious that one would at least attempt to *identify* the effects of coarticulation as far as possible i.e. the relevant vowel's allophony. This relates to certain well-known facts concerning the 'reach' of coarticulatory effects. As emphasized by Lehiste and Peterson (1961:271), in their study of, among other things, CV and VC transitions, "there seems to be no evidence that in the interaction between two sounds in sequence one will remain constant", the essential point here being that coarticulation very often effects not only the traditional transitional 'margins' of the syllable, but the nature of the vocalic nucleus itself. However an important distinction should be drawn between what Wells (1982:41–4) calls accommodatory (instrinsic) and non-accommodatory (extrinsic) allophones of a phoneme, the former being "natural" reflexes of coarticulation while the later are part of the (abstract) linguistic system i.e. the outcomes of phonological rules. In one's analysis of vowel quality one would, presumably, want to distinguish the former from the later. As already touched on in §1.4, it appears that certain coarticulatory effects are, in a sense, 'embraced' by the system, a case in point being the influence of a final dark-/l/ on vowel quality in SAEP. The velarized nature of this /l/ has a 'natural' accommodatory tendency to lower formants of the preceding vowel, particularly the F_2, but the *degree and extent* of this lowering is, ultimately, an optional effect. Thus, for example, while a dark-/l/ certainly has a retracting influence on preceding vowels in RP (Cruttenden 2001:203) the influence is, in most cases, far less than in SAEP – the vowel in *pill*, for example, is far more retracted in SAEP than in RP. One might wish to argue that the degree of retraction of the pre-/l/ vowel is determined by the degree of velarization of the lateral i.e. the high degree of vowel retraction in SAE is the result of a higher degree of /l/-velarization in SAE. This, however, does not seem to be borne out by the literature. Cruttenden (2001:203) reports a high degree of variation for RP "in the quality of the back vowel resonance associated with [ɫ] . . . with a range extending from [ö], [ʊ], or [ɤ] to [ɔ̈] or [ʌ̈]" (i.e. there are substantially retracted variants), and while the (recent) general consensus is that final-/l/ in SAEP is dark, there is at least one (prominent) commentator who claims that "although South African /l/ *is not really dark in any environment*, it nevertheless exerts a lowering and retracting effect on certain vowels" (Wells 1982:617; my emphasis). While the facts are, thus, a little unclear, it seem likely that the degree and extent of vowel retraction before /l/ in SAEP is not simply an accommodatory effect, but is part of the abstract linguistic system. The reader will recall from §1.4, that this procedure of phonologization has been evoked to explain phenomenon of just such a nature i.e. the 'jump' from accommodatory to non-accommodatory being initiated by a reanalysis of the phonetic data by the new language learner. In essence the child 'creates' a new target for the relevant allophone, instead of leaving the articulation of the vowel to 'natural' articulatory processes.

In the previous paragraph, however, I have placed the term 'natural' in scare quotes in deference to theorists such as Lass (1980:143) who claim that in linguistics "'naturalness' is not an explanatory notion". Important in this regard are certain so-called weak theories of speech production which claim that "coarticulatory influences may be activity managed to accommodate perceptual mechanisms of rather limited and specific computational capacity. This suggests that coarticulation is *stylized* to produce stylized covariance patterns in speech output"(Nearey 1997:3244). The coarticulatory processes being mentioned here are of the traditionally accommodatory kind, and are viewed as constrained (by both production and perceptual demands) but certainly not determined. As such, "coarticulation and other phonetic details that are language-specific should be specified in the grammar of each language" (Oh 2008:362). From this perspective, of course, the distinction between Wells' (1982) accommodatory vs. non-accommodatory distinction becomes somewhat 'fudged'. If it is, in addition, possible that if even in the case of accommodatory allophony the coarticulatory transition periods between consonant and vowel are, for example, stylized and not fully determined by relevant articulatory pressures, there is room not only for information assisting in the phonemic identification of the vowel, but also for sub-phonemic, sociolinguistic information. Different groups might, hypothetically at least, be differentiable on the basis on how transition and coarticulation is ‘managed’. In support of this, there is a substantial body of literature, collectively subsumed under the rubric ‘dynamic cospecification theory’ by Andruski and Nearey (1992:390), reporting on research focussed mainly on the use of silent-center syllables, which claims to show that it is in fact the areas of transition that contain the most relevant cues to vowel identification (Strange, Jenkins and Johnson 1983, Verbrugge and Rakerd 1986). If these theories are correct there seems little reason, again, to exclude, on *a priori* grounds, the possibility of sociophonetic information being encoded in the transition periods.

The existing theory, therefore, seems to ask of a rigorous acoustic analysis of the sort proposed here that it remain sensitive to both coarticulatory effects and the possibility of VISC, especially with respect to phonemic monophthongs. Still, while it is theoretically possible that sociophonetic value is contained in the CV and VC transitions, one would prefer to exclude these components of the acoustic signal as part of an initial analysis such as the current one i.e. to focus on the *phonetic* vowel target as defined above and as is the norm in sociophonetic research; but with certain provisos that allow for the representation of VISC. In particular, and as far as the later requirement is concerned, one would want to allow for the measurement of spectral change in quasi-stationary periods that lie outside of transition. Thus, with regard to coarticulation, the point is to capture coarticulation effects in terms of their effect on the phonetic vowel target (and thus capturing both acommodatory and non-accommodatory forms of allophony), but at the same time allowing for the capture of some VISC. In an initial analysis such as the current one, capturing sociolinguistically-relevant information within the CV and VC transitions strikes one as overly ambitious. With respect to diphthongs the tracking of VISC is of course an integral part of their analysis and description. Section 4.5.2 is specifically tasked with describing the methods used for putting these general principles into practice.

2.3 Vowel Systems

This section is focussed on reviewing two frameworks which sub-divide the vowels of English into various sub-systems. The first subsection provides a brief overview of Wells’ (1982) various Part-Systems; a descriptive framework which will be utilized in further chapters. The second subsection reviews Labov’s (1994) division of vowel-shifting into a number of relatively discrete patterns. These various patterns are described and exemplified and linked to broader principles of vowel-shifting. This section also includes a recent critique of some aspects of Labov’s (1994) model i.e. that provided by Torgersen and Kerswill (2004).

2.3.1 Wells' (1982) Framework

Since part of the task of this research will be to compare SAE with other accents of English, a notational system for referring to vowels across different accents of English is required. While a purely phonemic form of notation is one possibility, it is cumbersome for a number of reasons, not least because accents differ systemically. Thus, while most accents of English have a phonemic distinction between what would be traditionally transcribed as /ʌ/ and /ʊ/, accents in the northern areas of England generally do not; the vowel in *strut* is identical to the vowel in *foot*. Referring to this 'merged' vowel as /ʊ/ creates an obvious source of confusion, while any other option seems equally deficient. Wells (1982:122; my parenthesis), on the other hand, provides a very useful "framework of standard lexical sets ... for describing the lexical incidence of vowels in all the many accents [of English]". Wells (1982:xviii) explains as follows:

> These enable one to refer concisely to large groups of words which tend to share the same vowel, and to the vowel which they share. They are based on the vowel correspondences which apply between British Received Pronunciation and (a variety of) General American, and make use of keywords intended to be unmistakable no matter what accent one says them in. Thus 'the KIT words' refers to 'ship, bridge, milk ...'; 'the KIT vowel' refers to the vowel these words have (in most accents, /ɪ/); both may just be referred to as KIT".

The various phonemes (and related lexical sets) have, in addition, been divided into four "part-systems" (Wells 1982:168–78). The four part-systems are as follows:

Part-system A: The short-vowels of English. In stressed monosyllables these vowels cannot be word-final, so, for example, [sɪ] is an impossible word in English. Current RP has a six-vowel part-system A: /ɪ,ɛ,æ,ɒ,ʌ,ʊ/ or, in terms of Wells' (1982) keywords, KIT, DRESS, TRAP, LOT-CLOTH, STRUT and FOOT. In Labov (1994), these vowels are transcribed as /i/, /e/, /æ/, /o/, /ʌ/ and /u/. In the American tradition, /o/ is often transcribed as /a/ reflecting the fact that it has undergone a process of unrounding, lowering and centralization. According to Labov (1994:161; footnote), "in many dialects that do not preserve distinctive length or inglide in the low central vowels, /a/ also includes the vowels of *father*, *calm*, *pajama*, *rajah*, *pa*, *ma*, etc." i.e. the so-called PALM lexical set – see the description of Part-System D below. Some American dialects, such as New York, do however maintain a tense-lax distinction between /o/ (e.g. *not* and *hog*) and /ah/ (e.g. *God*, *father* and *bomb*). LOT and CLOTH are separate to reflect the fact that while in so-called flat-CLOTH accents (like General RP) LOT and CLOTH have merged, in broad-CLOTH accents (e.g. General American) CLOTH has the same vowel as THOUGHT. Where the distinction between LOT and CLOTH is irrelevant to the discussion, LOT will be used as a cover term for both LOT and CLOTH; following the same logic, where necessary the distinction will be made explicit.

Part-system B: A group of long vowels having a front mid to close quality, either for the whole vowel in the case of monophthongs or for the up-glide in the case of diphthongs. In RP, as well as SAE, there are four members of this part-system: /iː, eɪ, aɪ, ɔɪ/, or, in terms of Wells' (1982) lexical sets, FLEECE, FACE, PRICE and CHOICE. Such vowels can occur word-finally in stressed monosyllables, e.g. *see*, *say*, *tie*, *boy*. In Labov (1994), the vowels are represented as /iy, ey, ay, oy/, the first three also sometimes referred to as *long E*, *long A* and *long I*.

Part-system C: A group of long vowels having a back mid to close quality, either for the whole vowel in the case of monophthongs or for the up-glide in the case of diphthongs. SAEP shares with contemporary RP a three-member part-system with the relevant phonemes being /uː, əʊ, aʊ/, or in Wells' (1982) terms GOOSE, GOAT and MOUTH. Such vowels can occur word-finally in stressed syllables e.g. *moo, so, cow*. In Labov (1994), these vowels are transcribed as /uw, ow, aw/, the first two also referred to as *long U* and *long O*. The American phonetic tradition often also includes the /iw/-vowel which is used contrastively in lexical items belonging to GOOSE. The example provided by Wells (1982:173) is *threw* [θrɪu] vs. *through* [θruː]. While found in certain conservative varieties of North America, England and Wales it is absent from RP and the accents of the Southern Hemisphere. According to Wells (1982:539), it *is* found in the southern USA.

Part-system D: A group of long vowels having "a relatively open quality or (if diphthongal) endpoint, including under 'relatively open' the mid central quality [ə]". The six relevant phonemes for contemporary RP (and SAEP) are /ɪə, ɛə, ɜː, ɑː, ɔː, ʊə/ or NEAR, SQUARE, NURSE, BATH-START-PALM, THOUGHT-NORTH-FORCE and CURE. Such vowels can occur word-finally in stressed syllables e.g. *fear*, *fair*, *sir*, *father*, *thaw*, *pure*. In Labov (1994), the relevant symbols (for a non-rhotic USA accent) would be /ih/, /eh/, /ɜh/, /ah/, /oh/, and /uh/. For rhotic American accents many of these vowels are equivalent to a sequence of vowel plus /r/; thus, in the case of CURE, /ʊr/, or, in the case of NEAR, /ɪr/[5]. Unless the distinctions are necessary

[5] An exception to this pattern is found in New York English where the historical post-vocalic /r/ is making a

for the discussion, in what follows the convention will be to employ THOUGHT as a cover term for THOUGHT, NORTH and FORCE and BATH as a cover term for BATH, PALM and START. This will be the case particularly when SAE is discussed given that, by all accounts, the relevant distinctions have been lost in this accent. Where THOUGHT or BATH etc. need to be understood in their strict sense this will be clear from the context or will be made clear as part of the discussion.

In later chapters, each of the SAE vowel phonemes has been described in terms of these keywords. It should be stressed that in some cases there is not a one-to-one correspondence between keyword and phoneme. Thus the /ɑː/ phoneme in SAE corresponds to the BATH, PALM and START lexical sets, while in other accents these lexical sets differ systematically.

2.3.2 Labov's (1994) Principles and Patterns of Vowel-Shifting

Labov (1994) provides a number of basic principles of vowel-shifting, with a particular focus on chain-shifting. These principles are, essentially, part of a broader effort to determine the constraints operative on phonetic and phonological change. The principles are as follows:

Principle I: In chain-shifts, [+tense], peripheral vowels rise along the peripheral track. These mostly include long monophthongs and long-ingliding vowels (Labov 1994:137). While the development of in-glides is common among vowels moving along the peripheral track, upgliding vowels can also rise along the same 'path'.

Principle II: In chain-shifts, [-tense], non-peripheral vowels fall along the non-peripheral track; this includes the nuclei of upgliding diphthongs, i.e. those diphthongs which glide to [i] or [u]. The application of Principle II to upgliding diphthongs is often referred to as Principle IIa.

Principle III: In chain-shifts, back vowels move to the front.

The above principles apply, as should be evident, to chain-shifts and not to simple vowel movements into 'empty spaces'. Thus Labov (1994:30) stresses that such movements show very few constraints. There is nothing unprincipled therefore about a short vowel (e.g. STRUT) moving 'backwards' or a short (lax) vowel (e.g. LOT) rising if these movements

comeback. As a result, New York often has an inglide plus /r/. Labov (1994) uses transcriptions such as /ohr/ to capture this phenomenon.

do not constitute a link in a larger chain-shift. Equally important is to distinguish the principles underlying chain-shifts from those underlying "parallel shifts, or shifts that restore parallelism" (Labov 1994:30).

The above principles are, in fact, conflated in Labov (1994:262; 601) into one overarching Vowel Shift Principle:

(2.1) **In chain shifts peripheral vowels become less open and nonperipheral vowels become more open.**

While this rule clearly captures the basic 'spirit' of Principles I and II, the movement of back vowels to the front, as *per* Principle III, appears to be missing. Labov's (1994) reduction of the three Principles into one involves, however, a reanalysis of peripherality in *articulatory* as opposed to acoustic terms. The technical details are not of particular relevance to the current research, but the basic idea can be explained with relative ease.

If the reader returns to Figure 2.1 on page 25, he or she will note that the most open vowel is [ɑ], while the closest vowel is [iː]. Other (acoustically) high, back vowels, such as [u] are in fact more open (from an articulatory perspective) than [i], and while it is not perfectly clear from Figure 2.1, Labov (1994:256–64) shows convincingly that articulatory vowel space constitutes an ellipses with the 'highest' point at [i] and the 'lowest' point at [a]. The one 'track' of this elliptical space runs therefore from [a] through [ɔ] and [u] to [i] (or *vica versa*) and the other runs from [a] through [æ] and [e] to [i] (or *vica versa*). The manner is which these articulatory notions are reflected in acoustic vowel-triangle terms is reflected in Figure 2.9.

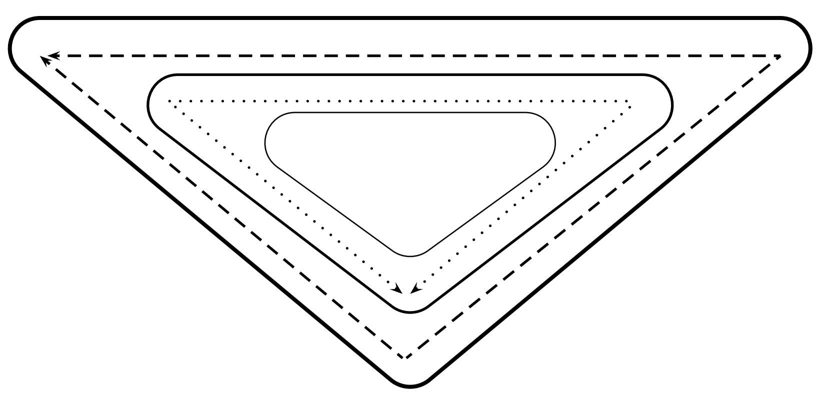


Figure 2.9: Labov's (1994) acoustic vowel triangle showing the Vowel Shift Principle

Two other principles that require brief recognition, deal with the movement of vowels from one sub-system to another, in particular the reanalysis of peripheral, tense vowels as non-peripheral, and short, lax vowels as peripheral. According to Labov (1994:280–1), these are as follows:

(2.2) **In chain shifting, low nonperipheral vowels become peripheral.**

(2.3) **In chain shifting, one of two high peripheral morae becomes nonperipheral.**

Example 2.2 is known as the Lower Exit Principle while 2.3 is referred to as the Upper Exit Principle and both capture common patterns of reanalysis, examples of which will be provided in §2.3.2.1.

It is perhaps important to stress that the various principles, in either their separate forms or dissolved into the more General Vowel Shift Principle, do allow for exception. Labov's (1994) work is, in general, characterized by an inductive approach which "creates generalizations slowly as the data base grows, moving step by step to statements of increasing generality" (Labov 1994:13). The inductive approach does not, by its very nature, assume the universality and exceptionlessness of the observations and principles that it generates. Thus the reader will, no doubt, in what follows, find cases where the data does not fit the principles in a completely 'snug' fashion; or even cases where the data seems to reveal inconsistencies between the principles conceived of as separate and the principles conceived of as integrated under the General Vowel Shift Principle. This is partly to do with the fact that there *are*, in all likelihood, exceptions to Labov's (1994) model; but it is also a result of the model's complexity, which cannot be presented here in all its subtlety, but can only be gleaned from a close reading of Labov (1994). The current section and that which follows does not, therefore, attempt to provide a defense of what is, by all accounts, a *working* model. Its aim is to rather introduce the reader to the basic principles and patterns that Labov (1994) has introduced in order to capture undeniable regularities and similarities across different languages and accents.

2.3.2.1 The Four Basic Patterns

For all languages, although the emphasis in what follows will be on English, the principles discussed in the previous section are seen to be operative in a number of combinations or basic patterns, which are, in turn, as follows:

Pattern 1: This essentially involves the raising, in accordance with Principle I, of back and front long (tense) monophthongs, "with the high vowels leaving the system of long monophthongs to become upgliding dipthongs" (Labov 1994:123) i.e. in accordance with the Upper Exit Principle. Following Labov (1994:123), the classic example of this highly symmetrical pattern is the English Great Vowel Shift (GVS), which involved the following two broad movements:

1. /ǣ/ > /ē/ > /ī > iy/ and /ɔ̄/ > /ō/ > /ū > uw/ which shows the raising of both front and back long monophthongs and their eventual 'breaking' (i.e. diphthongisation); and
2. /iy > ey > ay / and /uw > ow > aw/ which shows how the nuclei of the two diphthongs fall.

The end result is thus a fall of ME **ī** and **ū** to a relatively low position. Note that this "fall" (unlike the initial raising) is not a chain-shift but a simple lowering of the phonetic quality of the vowel from a high to low position and does not therefore constitute, in Labov's (1994) terms, an example of Principle IIa in action[6]. According to Labov (1994:167), "many conservative English dialects now rest at the end point of this process, displaying a symmetrical pair of upgliding diphthongs /ay/ and /aw/ with the identical low central nucleus [a]".

While the symmetrical GVS-pattern is not evident as a change-in-progress in any current dialects of English, mainly due to the influence of Principle III which fronts the back vowels instead of allowing their nuclei to fall along the back non-peripheral track, a current *extension* of this pattern leads to further backing of PRICE (/ay/) and fronting of MOUTH (/aw/), often explained in terms of a more general principle of increased "nucleus-glide differentiation" (Labov 1994:167). Also known as the PRICE-MOUTH Crossover, this often leads to further *raising* of the nuclei of both these vowels along the peripheral track, as is common, for example, in vernacular New York English. An intermediate step relates to the nonperipheral vs. peripheral status of the nuclei of these two diphthongs i.e. in their most conservative states the nuclei of both PRICE and MOUTH are nonperipheral (lax). Before raising occurs these nuclei are tensed in accordance with the Lower Exit Principle and become "the most open vowels of the system" (Labov 1994:177). A similar process is found in many American accents with respect to the tensing of TRAP i.e. the first step is a *lowering* process.

An important distinguishing feature of the Pattern 1 extension appears to be that the /iy, ey, uw, ow/ vowels remain in a conservative, peripheral position. This is in

[6]The movement of ME **ī** and **ū** in the GVS would appear to be based on the application of Labov's (1994) Expansion Convention, which refers to "the mechanism by which [a] sound crosses an unoccupied phonological space" (Labov 1994:266; my parenthesis).

contrast to those patterns constitutive of the Southern Shift (i.e. Patterns 3 and 4), as dealt with below.

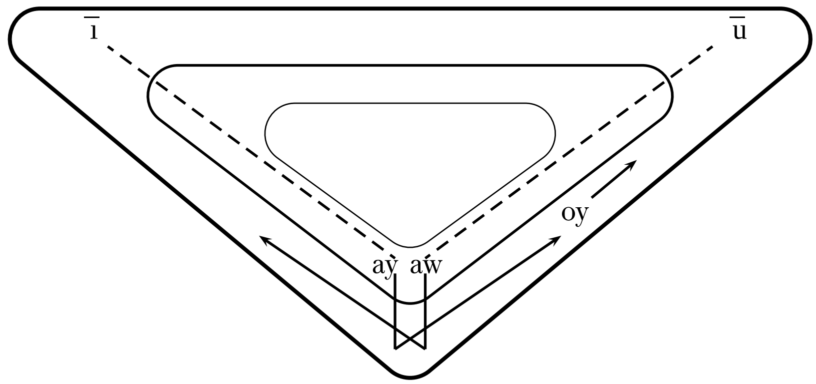


Figure 2.10: Labov's (1994) Pattern 1 Extensions

The Pattern 1 extension is illustrated in Figure 2.10. In this figure, Labov's (1994) /ay/, /aw/ and /oy/ can be viewed as roughly equivalent to Wells' (1982) PRICE, MOUTH and CHOICE respectively. The dashed lines indicate the initial GVS movement, while the sold lines indicate subsequent extensions. The movements indicated in this figure relate solely to the nuclei of the various vowels.

It is interesting to note that in his description of two other non-English examples of Pattern 1, Czech and Old Prussian, Labov (1994:124) indicates that in both cases two vowel systems co-exist i.e. one which displays the pattern in its full or most advanced form (Common Czech and Catechism II respectively) and one which displays a more conservative pattern (Catechism I) or no Pattern 1 at all (Standard Czech). Two other examples of this pattern are to be found in Middle High German and Western Yiddish.

Pattern 2: these chain-shifts, as described by Labov (1994:125–8; 177–201), archetypally involve the fronting of a long low back vowel (such as /ɑː/), in accordance with Principle III, which in turn displaces a long low front vowel, i.e. /æː/, onto the peripheral track. This vowel then rises along the front periphery and most often develops an in-glide to distinguish it from /iy/. It can also merge with /iy/. At the same time, the front short vowels fall along the non-peripheral track (in accordance with Principle II). Languages displaying Pattern 2 movements include North Frisian, Old Korean, Greek and Vegliote, but the classic (although not archetypal) English example of Pattern 2 is the so-called Northern Cities Shift in the USA which involves the tensing and raising of short **a** and the anti-clockwise movement of some of the other short vowels. In the Northern-Cities Shift, like in Pattern 1, but unlike in Pattern 4 below, the high and mid upgliding diphthongs (/iy, ey, uw, ow/) are usually stable, with /iy/ and /uw/ often remaining monophthongal.

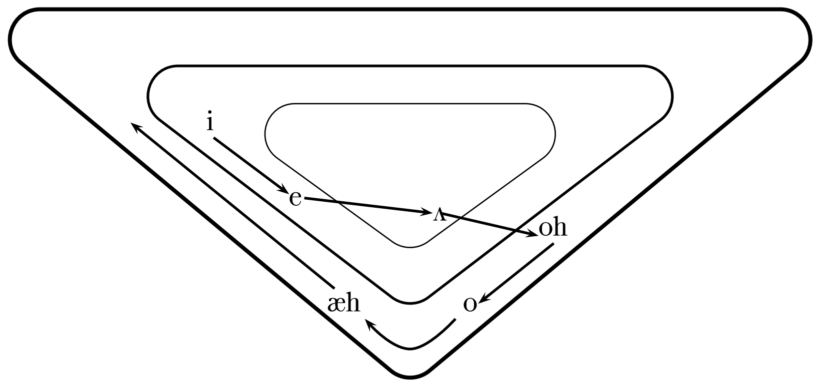


Figure 2.11: Labov's (1994) Northern Cities Shift

Figure 2.11 provides the basic architecture of this shift which is initiated by the tensing and raising along the peripheral track of the entire short **a** class (*bath*, *hat*, *pack*, *last*, *man* i.e. TRAP and BATH), unlike in the Southern U.S. and Middle Atlantic States (New York, Pennsylvania) where there is a lexical split between lax /æ/ and tense /æh/. In all dialects, including those participating in the Northern Cities Shift, /æh/ appears to be most susceptible to raising in the context of a following nasal (*man*, *hand*) while "the least advanced ... are those with initial liquid clusters (*flat*) or following velars (*bag*, *back*)" (Labov 1994:180; footnote). The raising of /æh/ along the peripheral track in turn leads to the fronting of /o/ (e.g. *locks* becomes [læks]), followed, in turn, by the centralization and fronting of /oh/, in a classic drag-chain. As mentioned above, the raising of /æh/ is accompanied by the lowering of KIT, "to mid (or even lower mid) position" (Labov 1994:188) as well as the lowering and backing of DRESS (such that *Debbie* is pronounced as [dʌbi]). DRESS in turn displaces STRUT so that it occupies the position originally reserved for /oh/.

Pattern 3: According to Labov (1994:129–33; 201–8), this pattern involves the raising of tense back vowels and their subsequent fronting, in accordance with Principles I and III. In some cases diphthongization takes place before the nucleus is fronted. There are numerous examples of this pattern in other languages, both Indo-European and non-Indo-European. In English, it translates into the raising and subsequent fronting of long back monophthongs, in its simplest form /a/ > /ɔ/ > /o/ > /uw > üw/. Labov (1994:202–6) provides examples of this pattern in action using data from London, Philadelphia, Central Texas, New York City and Norwich. Pattern 3 has been schematized in Figure 2.12.

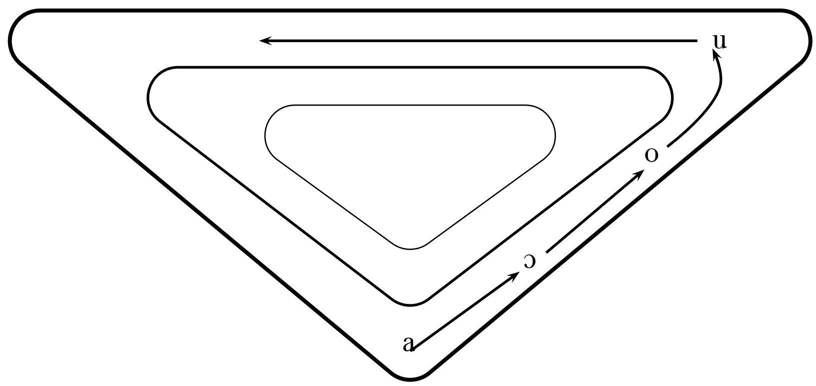


Figure 2.12: Labov's (1994) Pattern 3

In many dialects which show Pattern 3 movements, there is also a fronting (or lowering) of the /ow/ vowel, GOAT in Wells' (1982) terms. According to Labov (1994:208), GOAT fronting is a generalization of the fronting of GOOSE, "when /ow/ is fronted, it is always in parallel with /uw/ and considerably behind it" i.e. an example of a parallel shift as mentioned briefly in §2.3.2. The other alternative is for the nucleus of /ow/ to fall, a phenomenon which, for example, occurs in London and Texas. While this resembles a Pattern 1 type movement, Labov (1994:208) stresses that "it is not associated with any chain shift. Instead it appears to represent a generalization of the lowering of the nucleus of /ey/, part of the chain shift . . . Pattern 4".

Pattern 4: According to Labov (1994:208–18), this pattern deals, firstly, with the front upgliding diphthongs i.e. /iy/, /ey/, /ay/ and /oy/, and "is the dominant mode of chain shifting of upgliding vowels that is found in present-day English dialects in the Southern Shift" (Labov 1994:209). It needs to be carefully distinguished from the extension of Pattern 1 discussed above, which involves the *independent* backing of PRICE and fronting of MOUTH as a result of the position these vowels find themselves at the completion of the GVS. Thus, in those accents which display the Pattern 1 extension, PRICE and MOUTH appear to participate in an independent tensing, fronting/backing and raising. In a Pattern 4 movement, however, the initiating change is rather the laxing of the /iy/ and /ey/ vowels: "The nucleus of /ey/ falls, usually to the most open position, and the nucleus of /iy/ follows" (Labov 1994:209). There are two possible effects on /ay/:

1. It can front and monophthongise to [aː]; this is the pattern found in "the Gulf

States, Texas, and many areas of the Upper and Lower South" (Labov 1994:209). This basically terminates Pattern 4, since there is no pressure on /oy/ to raise as a result of a backed and raised /ay/.

2. /ay/ can, however, back and raise, pushing /oy/ further up the vowel chart as well. This is the pattern common in New Zealand, Australia, southern England and "the coastal areas of the Eastern United States" (Labov 1994:209).

It should be mentioned at this point, that while Labov (1994) separates out the backing of PRICE in terms of Pattern 1 and the backing (and fronting) of PRICE in terms of Pattern 4, these two separate movements are potentially reconcilable. An alternative interpretation is that the GVS and Pattern 4 (as well as Pattern 3) are intimately connected, with the backing or fronting of PRICE being the initiating movement of the subsequent lowering of the FACE and FLEECE nuclei in a pull-chain dynamic, while the fronting of MOUTH is conceptualized as a pull on both GOAT and GOOSE often leading, in turn, to the diphthongization of these vowels. The notion of Pattern 3 and Pattern 4 movements simply both being an extension of the GVS is implicit in Wells' (1982) notion of Diphthong Shift and the suggestion that PRICE is one of the initiating loci of these extensions is, in fact, elaborated upon in Trudgill (2004:50), who shows that, in relation to his NZE data, that "Diphthong Shift occurred in the following order: /au/, /ai/, /ou/, /ei/, /uː/ and, finally, /iː/. That is, there is an implicational scale such that, for example, speakers who have shifted /ou/ will necessarily have shifted /ai/ but not necessarily /ei/. Britain (forthcoming) also argues that shifting of /au/ came first and of /iː/ last". Under this conception, the Southern Shift is simply a logical continuation of the GVS. This, in fact, has the added benefit of accounting in a more principled fashion for the lowered nucleus of GOAT so often found as part of Pattern 3 i.e. the lowering of MOUTH 'pulls' GOAT down in the same way that PRICE pulls FACE down. It also possibly accounts for the fronting of GOOSE in SAE vs. the monophthongal status of FLEECE, given that "it ... seems that Diphthong Shift of GOOSE occurred before that of FLEECE" (Trudgill 2004:59)[7].

Returning to Pattern 4, of equal importance to the lowering of FACE and FLEECE, is the ripple-effect that this lowering has on the traditionally short front vowels. Basically, according to Labov (1994:212), "the highest F2 values are those of the original

[7]It is unclear, however, whether the fronting of GOOSE (and GOAT), as opposed to the lowering of their nuclei, is linked to the Diphthong Shift. Trudgill (2004:55), for one, considers "these two possible types of GOOSE Shift – diphthongisation and centralisation – as separate but not necessarily mutually exclusive processes."

short vowels, /i, e, æ/ . . . we have yet to find a dialect where there are not some nuclei with a tense quality that impressionistically fall into the range of cardinal [i, e, ɛ, æ]". In the southern states of the USA, the short-vowels are analyzed as tense and are phonetically long with a tendency to develop inglides. On the other hand, "in London, Australia, and New Zealand, they remain phonetically short, so that they are clearly short, tense [+peripheral] vowels" (Labov 1994:212). In its basic form, Pattern 4 involves a 'swopping' of positions of the upgliding diphthongs with the front short vowels of English. Since the nuclei of the former become lax and fall along the non-peripheral path, the short vowels are free to occupy the peripheral position. The overall architecture of Pattern 4 is provided in Figure 2.13.

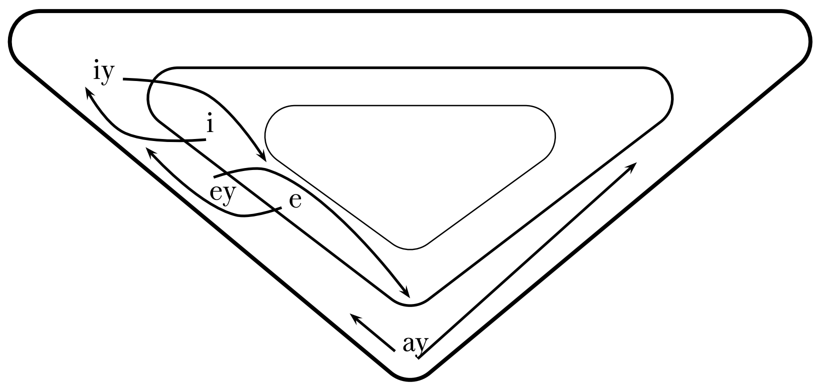


Figure 2.13: Labov's (1994) Pattern 4

It should noted that while /æ/ is missing from Figure 2.13, it is clear from Labov (1994) that TRAP is meant to acquire [+tense] status and overall membership in a Pattern 4 shift. Unlike KIT and DRESS, however, there appears to be no overt mechanism leading to the tensing of this vowel[8]. What is clear, however, is that the displacement of these originally lax vowels onto the peripheral track place them in a position for further raising, a phenomenon well-attested to in, for example, NZE, for which there is clear evidence of post-settlement raising of TRAP and DRESS in particular (Langstrof 2006). In the case of AusE and SAE it is, however, unclear whether the raised status in Broad varieties of these accents are simply a relic of the original colonial input or whether these vowels were raised even further post-

[8]This is not completely true in the sense that for such phenomena Labov (1994:184) makes recourse to a more 'basic' chain-shifting principle, which applies to 'entering' elements i.e. those elements which move to fill the gap left by a 'leaving' element. According to Labov (1994), "the direction of the leaving element follows Principles I-III, while the direction of the entering element may be governed by the basic chain-shifting principle". Even in Labov (1994), however, this strikes the reader as a somewhat *ad hoc* account.

settlement. A particularly tricky issue in this regard relates to the centralization of KIT in both NZE and SAE, a phenomenon which fits uncomfortably into the overall picture of the Southern Shift and Pattern 4 in particular. For NZE, KIT centralization has generally been viewed an "an endemic phenomenon" (Langstrof 2006:142) i.e. post-settlement, and related to secondary (i.e. endemic) raising of TRAP and DRESS. A further problem relates to the clear monophthongal status of FLEECE in SAE.

Regardless of the above-mentioned difficulties, however, SAE has been named as one of the dialects participating in the so-called 'Southern Shift' (Labov 1991:22–8), a series of sound changes that purportedly links English dialects in southern England, Australia, New Zealand and South Africa as well as the southern states of the United States. With respect to the latter, Labov (1994:202) mentions "the southern Middle Atlantic States, the Upper and Lower South, the South Midland, the Gulf states, and Texas". The Southern Shift is, essentially, a blending of Patterns 3 and 4, with, on occasion, some Pattern 1-*type* features, in particular GOAT-lowering[9]. South African English, as will be seen in latter chapters, displays a large number of these features. It should be emphasized that the existence of Pattern 4 features in a dialect in no way implies that the dialect in question is currently participating in a shift. As stressed by Labov (1994:210;footnote), Pattern 4 features in the speech of Londoners have "been well established for over 100 years". Different accents naturally show different stages of the Pattern 4 shift. Thus while Labov's (1994) Cockney examples show diphthongization of both FACE and FLEECE as well as relatively high values for both KIT and DRESS, Labov's (1994) Norwich data, on the other hand, has an /iy/ "which does not show any clear movement away from the high front area" (Labov 1994:210).

Labov (1994:137–40) also reviews a number of apparent counter-examples to the basic principles outlined in §2.3.2. Of particular interest is his description of The Reverse Parisian Chain-Shift (Labov 1994:139-40). The details are unimportant for the present purposes, but in general terms this chain-shift reverses a previous chain-shift (intuitively enough, The Parisian Chain-Shift) which involved the raising and fronting of back vowels i.e. Pattern 3 as a composite of Principles I and III. In the reverse chain-shift these same vowels are backed and lowered. It is not so much the reversal itself that is of interest here, but Labov's (1994) analysis of its motivation:

> "The change is led by the most conservative group in Parisian society: upper middle class females . . . it demonstrates that none of the principles of chain shifting is either

[9] Labov's (1994) Southern Shift is, for all intents and purposes, thus the equivalent of Wells' (1982) Diphthong Shift plus the front short vowels.

absolute or isolated from social factors. If social pressures are strong enough, phonetic processes that are deeply rooted in the history of the language and the functional economy of the system can be reversed ... In Paris, the internal factors that lead to change from below are [weak] ... and the changes that led to the local vernacular all showed the properties of completed changes, at high levels of social awareness. It is under these conditions that the conservative factors in the speech community can operate in a systematic manner to reverse chain shifts as a whole" (Labov 1994:139–40; my parenthesis).

With respect to the Southern Shift (and Pattern 4 in particular) there appear to be cases of similar reversals. Thus, for example, while Philadelphia speech had a lowered FACE vowel, and has (to a degree) a backed and raised PRICE as well as a raised CHOICE, "recent developments show a retreat from this alignment ... with the retrograde raising of /ey/ in checked position" (Labov 1994:211). It should also be mentioned that Philadelphia does not appear to display FLEECE diphthongization.

2.3.2.2 Torgersen and Kerswill's (2004) Critique

In an important article, Torgersen and Kerswill (2004) provide convincing evidence that, for London (and surrounding areas), Pattern 4, an important part of Labov's (1994) Southern Shift, does not apply. Instead, these authors identify an anticlockwise chain shift of the short front vowels which "appears to be common to much of south-east England' (Torgersen and Kerswill 2004:45). In other words, instead of being raised, as would be predicted by Pattern 4, the short vowels are being lowered. What these authors do conclude, however, is that the front short vowels *used* to be raised in the London-area; thus traditional Cockney values include a raised KIT, DRESS and TRAP (and fronted STRUT). Torgersen and Kerswill (2004:31) continue in the following manner:

> "In hypothesizing that south-eastern British English had closer realizations of KIT, DRESS and TRAP than are current now, we do in fact find some support in southern-hemisphere English vowel systems ... southern-hemisphere English has retained a conservative short vowel system with quite close front vowel realizations. Typically this is noticed in DRESS and TRAP".

As we will see later in §3.4.3, Trudgill (2004) provides some support for the notion that Southern Hemisphere Englishes are simply behind the south-east of England in the lowering of these vowels; what Trudgill (2004) calls 'colonial lag'.

Important too, is the fact that the impetus for this chain-shift, which, following Torgersen and Kerswill (2004:46), one might wish to call 'the south-east English short vowel chain shift' (henceforth SECS), appears to be a lowering of the TRAP vowel. Such lowering firstly leads to the retraction of a fronted STRUT vowel (an old Cockney value) to a backer position, and then to the lowering of DRESS (and KIT).

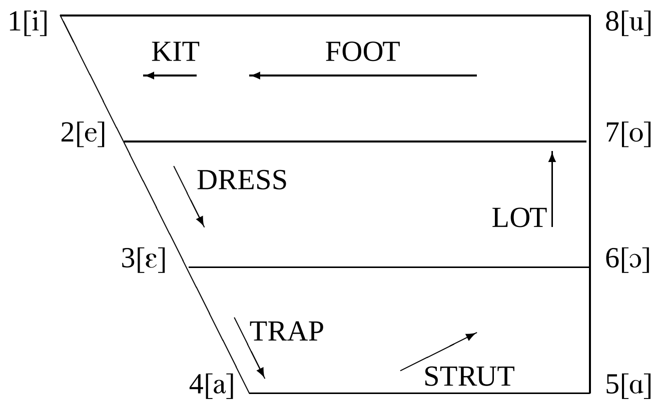


Figure 2.14: A Vowel Chart showing Torgersen and Kerswill's (2004) SECS

A reasonably advanced SECS-shift, as conceived by Torgersen and Kerswill (2004), is illustrated in Figure 2.14. What is most important about this evidence is that it implies, in essence, a reversal of the Southern Shift as provided for by Labov (1994), at least with respect to the short-vowel system. Coupled with well-attested phenomenon such as the centralization of KIT in SAE and NZE, along with clear monophthongization of FLEECE in SAE (as well as Norwich and Philadelphia) leads to the possibility that the Southern Shift is possibly more of a relic than an operational system in most (relevant) accents of English. The investigation of the Southern Hemisphere Englishes is particularly important in this regard given that they, in all likelihood, contain evidence for 'structurings' that the English of the south-east of England presumably went through before the full adoption of the SECS shift. This is particularly true if we can find evidence for at least an incipient SECS shift in the Southern Hemisphere Englises. Particularly relevant in this regard is the TRAP vowel, which, if we accept its tensing as a result of the Southern Shift, must have somehow become reanalyzed as [-tense], thus allowing it to fall along the non-peripheral track and thus initiate the SECS shift. Important too, in this regard, is a comparison of SAE with those 'southern' accents which display the Southern Shift in most, if not all, of its particulars i.e. numerous accents of the southern USA and AusE.

2.4 Conclusion

This chapter has reviewed a number of theoretical and descriptive frameworks, each providing some insight into the nature of vowel quality. Firstly, the acoustic 'reflexes' of vowel quality in general, and its various manifestations, have been elaborated upon, as well as the link between impressionistic, articulatory and acoustic representations of vowel 'space', including Labov's (1994) notion of peripheral vs. non-peripheral tracks, an essential element in his model for explaining the mechanics of vowel chain-shifts. Moreover, two aspects of the greater problem of noninvariance have been tackled in order to provide theoretical support for methodological decisions taken as part and parcel of the acoustic analysis which forms a core element of this thesis i.e. normalization and issues surrounding co-articulation and vowel targets. In particular, support for the adoption of a vowel-extrinsic, formant intrinsic normalization technique, such as that provided by Lobanov (1971), has been provided, as well as for a focus on capturing VISC during an acoustic analysis, as opposed to the sole use of vowel targets. In the process, the link between the phonetics of coarticulation and the diachronic process of phonologization has been highlighted. The use of 'raw' formant data, as opposed to the application of any number of available tonotopic scales, has also been justified.

In the second main section of this chapter, Wells' (1982) descriptive framework for English vowels has been touched on, while the various principles and patterns which form such an essential element of Labov's (1994) theory have received a fair degree of attention. Lastly, the critique provided by Torgersen and Kerswill (2004) has been summarized. In the process, some of the anomalies of SAE *vis-à-vis* Labov's (1994) Southern Shift have been dealt with (e.g. KIT centralization and a monophthongal FLEECE vowel), while, with respect to Torgersen and Kerswill's (2004) SECS shift, SAE has been targeted as a potential source of confirming evidence i.e. if TRAP lowering and related phenomena are identifiable in the SAE data, this provides added support for Torgersen and Kerswill's (2004) implicit critique of Labov's (1994) Southern Shift and, in a number of important respects, for its status as a relic rather than an 'operational' chain-shift. Labov's (1994) notion of a reverse chain-shift has been alluded to as one tantalizingly possible 'route' for Southern Shift varieties such as SAE.

CHAPTER 3

The Vowels of SAE in Context

> "…the evolution of the current SAE complex consists of a multiplex and as yet poorly understood network of contacts, within which endogenous innovation continues, and there is considerable archaic survival" (Lass 2004:384).

3.1 Introduction

This chapter, as well as Chapters 5 through 9 provide, amongst other things, a review of the literature relevant to a thorough description of the SAE accent. While this chapter begins with a brief excursus into matters terminological (§3.2), its main function is to place SAE within its broader sociohistorical and linguistic context. The above-mentioned five chapters, on the other hand, provide (again, amongst other things) an in-detail review of the literature on SAE, as it pertains specifically to the vowel system of this accent.

Section 3.3 of this chapter thus looks at the broader sociohistorical factors that led to the development and nature of SAE. This is the task mainly of §3.3.1. Section §3.3.2 and §3.3.3 then deal, respectively, with the prestige values of SAE, as reported on in the extant literature, and with SAE's 'place' in the Dynamic Model as provided by Schneider (2003; 2007).

Section 3.4 is then dedicated to placing SAE within the broader *linguistic* context of English accents in general. It begins, in §3.4.1, with the identification of SAE as a particular kind of Extraterratorial English (ETE) i.e. as Southern Hemisphere and as 'Southern' more broadly. This, in turn, dovetails with the task of the following section (§3.4.2), which is to review the few attempts that have been made to contextualise SAE within the broader framework of chain-shift patterns provided in Labov (1994), particularly *vis-à-vis* his notion of the 'Southern Shift', as touched on in Chapter 2. Lastly, putative links between SAE and the other Southern Hemisphere Englishes are drawn, drawing heavily on the recent work of Trudgill (2004) in this regard (§3.4.3).

3.2 Terminology

In this section a number of terminological issues relating to SAE are dealt with.

Firstly, a working definition for the term 'South African English' needs to be provided, given that it is used interchangeably in the literature to refer to all varieties of English spoken in South Africa, all *mother–tongue* English varieties in South Africa or, more often, to the English spoken most commonly by 'white' mother–tongue speakers. This thesis accepts the observation that the use of ethnic labels cannot imply strictly homogenous groups (De Klerk 1996) and follows Da Silva (2007:2) in taking note of the fact that the traditional terminology used to segment SAE[1] is no longer valid given that "one is no longer able to determine a variety on the basis of ethnicity alone . . . with the dissolution of the apartheid system, particularly the changing structures within the educational system of the country, the current varities of English spoken in South Africa require a new system and recategorisation". The focus of this research is, however, not on establishing a more sensitive and accurate nomenclature; as such, the default practice here will be to use the term 'SAE' to refer to the dialect (mainly) used in the apartheid past by 'white' speakers, while acknowledging, of course, that more recently many non-white South Africans have acquired native or near-native proficiency in this dialect. Occasionally, though, the term 'SAE' will be used to refer to all varieties, both L1 and L2, of English in South Africa. It will be made clear when this non-default meaning is being used.

With respect to SAE itself, Lanham, in his 1978 publication, and elsewhere, uses the terms 'Extreme', 'Respectable' and 'Conservative' to refer to the three main sociolects of this dialect. In line with more recent work in this field, e.g. Bowerman (2004), and following a suggestion made originally (although implicitly) by Lass (1995:94), the more neutral terms 'Broad', 'General' and 'Cultivated' are preferred in this thesis. These terms are derived from Mitchell and Delbridge (1965)'s work on AusE. The term 'General SAE' is thus equivalent to 'Respectable SAE'. Where necessary a 'P' will be added to the relevant

[1] i.e. 'white' South African English, Afrikaans English, Coloured English, South African Indian English as well as Black South African English.

acronym in order to distinguish accent from dialect e.g. General South African English Pronunciation (General SAEP or, more often, GenSAEP).

Lanham and Macdonald (1979) also use the term ‘General’, but with reference to variables i.e. General SAE (GE) variables which, unlike Natal English (NE) and Cape English (CE) variables, “are . . . more pervasive, more generally shared in the SAE community and have weaker differentiating properties in correlation with social variables, including region” (Lanham and Macdonald 1979:36). NE and CE variables, as their names suggest, have historical geographic loci in the KwaZulu–Natal and Eastern Cape provinces respectively[2]. Within the framework of Lanham and Macdonald (1979), Respectable SAEP (GenSAEP), is characterized, for example, by a high prevalence of NE variables or both NE and GE variables, while Extreme SAEP (Broad SAEP) is characterised by the high prevalence of CE features coupled with GE variables. Generally, it appears that CE variables are receding, while NE variables are advancing through the SAE speech community. Briefly, and by way of example, CE vocalic variables include a fronted and glide-weakened MOUTH and a backed, raised and rounded BATH. NE variables include significant vowel retraction before final, tautosyllabic /l/ and fronted, monophthongized PRICE.

3.3 Sociohistorical Considerations

Relevant sociohistorical considerations are presented in this section, with particular attention to three concerns. The first, §3.3.1, is the social history of SAE and the related development and indexicality of the three SAE sociolects that have played such a prominent role in the literature on SAE. The second, §3.3.2, relates to the prestige variables within GenSAE, as generally reported on in the existing literature. This focus is motivated by the fact that the subjects recorded for this thesis are from the upper-end of the socioeconomic scale and, secondly, by the fact that the style of elicitation was a formal one i.e. word-list style – see Chapter 4 for more on these and related issues. Thirdly, in §3.3.3, an attempt is made to integrate SAE within the broader sociohistorical model provided by Schneider (2003; 2007). In the light of the evidence provided in the preceding two sections, Schneider’s (2007) description of SAE in terms of its development along the various stages provided in the so-called Dynamic Model is critically reviewed. In the process the issue of regional variation within SAE is briefly considered.

[2]See Figure 4.3 on page 121 for a map of South Africa.

3.3.1 General and Historical Background

SAE, as defined in §3.2, arguably constitutes the main input to all other mother-tongue varieties of SAE (e.g. Coloured SAE and SAIE[3]). The relative influence of the L1 on non-mother-tongue varieties of SAE, e.g. so-called 'Black South African English' (BSAE), is, no doubt, variable from individual to individual[4].

SAE has its roots in successive waves of British immigration, mainly during the 19th-century. According to Mesthrie (1993:27), English was first brought to South Africa in 1795, but it was only in 1806, the second British occupation, that "a sizeable number of English speakers – comprising several thousand officials and soldiers, and some traders and farmers – was established".

The first permanent settlement was, however, in 1820, with the arrival in the Eastern Cape (see Figure 4.3 in Chapter 4) of approximately 5000 immigrants, generally of working- and lower-middle class origin and mainly from "London and the Home counties, with accents characteristic of this segment of English society" (Mesthrie 1993:27)[5]. Lass (2004:371) adds that there were also "sizeable contingents from the West Country and Yorkshire, and some from Scotland and Wales". While the input into Cape English (CE) thus consisted mainly of the dialects of London and broad environs, it appears, however, that the "overtly stigmatized pronunciations such as h-dropping, g-dropping (in *-ing* endings) which must have been prevalent in Settler speech, were apparently eliminated at an early stage" (Lanham and Macdonald 1979:73). According to at least Lanham and Macdonald (1979:73), contact between the early Eastern Cape settlers and the Dutch/Afrikaans-speaking community of the area, led to the adoption of certain Afrikaans features which characterize SAE to this day, i.e. Lanham and Macdonald's (1979) CE variables found mostly in Broad SAE (BrSAE), e.g. obstruent /r/; as well as features found more generally in SAE e.g. the prevalence of schwa in positions (both stressed and unstressed) more commonly carrying an [ɪ]-like quality in other accents. The prior presence of the Afrikaans population does complicate the picture somewhat, but it is still clear that in terms of English per se, the Eastern Cape settlement constituted an example of Trudgill's (2004) so-called 'tabula rasa context', a concept dealt with briefly at the end of §1.3.3.

[3]South African Indian English.

[4]While I acknowledge that there is some debate on this issue, particularly with respect to the institutional fossilization of this variety and the implications of this for its L1 or L2 status, this is hardly the place, given the focus of this thesis, to provide arguments for the position taken here.

[5]This standard picture of the class status of the Eastern Cape settlers perhaps requires some revision, especially given statements such as that of Welsh (1998:127), who claims that the selection criteria for choosing individuals for emigration "were rather too strict, in that whilst they produced a high proportion of educated and responsible citizens, there were too few of the labourers and artisans needed for the pioneering work".

The next main settlement occurred in Natal (now KwaZulu–Natal), between 1848 and 1862, with the proportion of middle– to upper–class individuals being far higher than during the Eastern Cape settlement. Thus Hattersley (1950) confirms that "except for the East Riding of Yorkshire, the majority of those who emigrated to Natal, even from predominantly rural counties, did not make a living directly from the land". According to Mesthrie (1993:27–8), there were also more individuals from "the Midlands, Yorkshire and Lancashire. Natal English speakers also had fewer contacts with Afrikaans than their Cape counterparts".

While "a small settlement of traders and hunters established itself at Port Natal in 1824" (Davenport and Saunders 2000:113), and while there was, for example, some influx from the Eastern Cape and elsewhere after the British Occupation of Port Natal in 1842 and throughout the forties, the small and transient nature of the English-speaking population of this time is confirmed by Hattersley (1950:92), who claims that "at the close of the 'forties it seemed doubtful whether Natal would retain a permanent population of white colonists". The first major settlement began, therefore, with "the Byrne scheme of 1849–50 [which] brought some five thousand English and Scots" (Davenport and Saunders 2000:115; my parenthesis). Hattersley (1950) confirms the north-of-England bias with respect to the Natal immigration and the areas from which many of the settlers originated include Lancashire[6] and the West, East and North Riding of Yorkshire. That a Scottish presence was also felt is reflected in the fact that this author dedicates a chapter to these immigrants, claiming that most came from Glasgow and the Highlands while "the eastern side of the country showed less interest in immigration" (Hattersley 1950:178). There was some influence from the southwest (Cornwell and Devon) while "East Anglia showed little interest in emigration to Natal" and, contrary to Mesthrie's (1993) claims above, "not many of the Natal settlers came from Midland towns" (Hattersley 1950:164; 131). In conclusion, therefore, while not as clear-cut as the Eastern Cape scenario, there is much evidence to suggest that the so-called 'Founder Effect', as originally conceived by Mufwene (1991)[7], was minimal in Natal, and that the outcome of any dialect mixture would have been a distinctly north-of-England sounding variety, although tempered no doubt to some degree by a 'Standard English'-influence (given the overall class-status of the immigrants)[8] and, perhaps, in deference to

[6] Hattersley (1950) says that "Lancashire men were predominant among the emigrants of 1849–50".

[7] As explained by Trudgill (2004:163), "the founder effect implies that the linguistic founding population of an area has a built-in advantage when it comes to the continuing influence and survival of their speech forms, as opposed to those of later arrivals".

[8] The possibility that many of the Natal immigrants were, on the basis of their social class, Standard English speakers should not lead one to assume, as perhaps some commentators have done, that their accents were unaffected by their regional provenance. Only a few years after the main Natal migration, Ellis (1869: 1215), cited in Jones (2006:291), observes that "in the main the most educated pronunciation in English is local, with

the possibility of some residual Founder Effect, by the effects of an early-CE variety, given that there is some evidence of immigration from the Cape before the main influx occurred.

The last major influx of British subjects occurred during the latter half of the 19th-century, as part of a more general influx of immigrants to South Africa. Thus Lanham (1982:327) mentions "400,000 immigrants who arrived between 1875 and 1904 . . . mainly from Britain and eastern and western Europe," the major influx beginning after the discovery of diamonds in Kimberley and the 1886 discovery of gold on the Witwatersrand. Many of the mainland European immigrants were of Jewish extraction, thus constituting the first un-ancestral population to be integrated into the SAE speech community. As intimated above, this influx was linked mainly to the discovery of diamonds in Kimberley and the development of the gold industry in the Witwatersrand area (roughly-speaking, Johannesburg and environs). This influx was accompanied by the more local migration of CE, NE (and Afrikaans) speakers to these same areas. As far as proportions are concerned, Welsh (1998:302) provides 1896 census figures which show "the white population as 50,907, of whom 6,205 were Transvaalers, well outnumbered by the 16,265 from England and the 15,162 from the Cape Colony (the next-largest population was that of 'Russian' – more properly Baltic – Jews)". This quotation requires comment. Firstly, the omission of Natal as a source of immigrants is undoubtedly an error and these immigrants should probably be counted under 'the Cape Colony' although this is, for obvious reasons, unclear. Secondly, and according to Kaplan and Robertson (1991), "the majority of South Africa's East European immigrants came from Lithuania and Latvia, particularly the provinces Kovno, Vilna and Grodno". The vast majority were Yiddish-speakers. These authors add that the "Johannesburg Jewry is unique among Jewish communities throughout the world because its people joined with others to become the founders of a city". Furthermore, the quote obscures the fact that "into this cauldron of capitalist development poured men, women and children drawn from all over the world" (Van Onselen 1982:xv) and perhaps, more importantly, that by 'English' one should assume the transplantation of a wide array of dialects, among which were, in all likelihood, non-English but British dialects such as Scottish English and Irish English.

It seems, furthermore, reasonably clear that, for all intents and purposes, Johannesburg constituted another tabula rasa context for dialect-mixing. Johannesburg was, after all, created as a direct result of the discovery of gold[9]. Important in this regard is the observation by Kaplan and Robertson (1991:16) that "at the diamond diggings, among all sections of the

its corners more rubbed off than it was fifty or a hundred years ago, but still essentially local".

[9]Thus Leyds (1964:3) explains that "the original village of Johannesburg was laid out on the farm Randjeslaagte but there is no record of there having been a house on the land, as it was government ground".

white population, many of the concepts of class fell away and the same pattern was perceivable on the Transvaal gold-fields". Yet while this state-of-affairs perhaps characterized the very early years of the new settlement, Johannesburg very quickly adopted a class-structure typical of an industrial city. Van Onselen (1982:2; my parenthesis) explains as follows:

> "Given the company that the parents [the immigrants] kept, it is scarcely surprising that the child [Johannesburg] lost its innocence at an early age. With white workers ranged against black, skilled miners against the mine owners and the Randlords against the state, Johannesburg was racked by class conflict during much of this period [1886–1914]."

The same author also mentions "clusters of workers ... [which] ... produced, reproduced and accentuated several elements of late-nineteenth century working-class culture – and, in the case of white miners, elements of British male working-class culture in particular" (Van Onselen 1982:5), mentioning, importantly, Cornwall, Cumberland and Lancashire as prominent areas from which these workers came[10]. Also relevant in this regard is Van Onselen's (1982) observation that during this early period there were very few women and children, meaning that "early Johannesburg was largely devoid of working-class family life" (Van Onselen 1982:5). As such, it is highly likely, in this particular tabula rasa setting, and contrary to Trudgill's (2004) model, that identity (social class and other) played a central role in the mixing process that must have occurred.

The first two 'waves' mentioned above (the Eastern Cape and Natal settlements), or 'prongs' as Jeffery (1982:252) puts it, have generally been viewed as constituting the main input into the development of a distinctly South African accent, while the late-19th-century influx is seen as the primary determinant in the development of social-class dialects. "By the turn of the 19th century ... while RP-like English commanded the highest prestige in the mining territories, Natal English was more readily available as a reference dialect, and gained prestige and spread at the expense of Cape English" (Mesthrie 1993:28). This gradual adoption of a local standard (GenSAE), based mainly on Natal English variables, appears, according to Lanham and Macdonald (1979:84–5), to have been particularly prevalent among Cape colonial and Eastern European descent-group females bent on social and economic advancement[11]. It was particularly "the European Jewish community [on the Witwatersrand], prominent users of Respectable South African English, [that were] primary agents in promoting this local standard in the mining-industrial society" (Lanham

[10] Naturally, there were also working-class immigrants from the southeast of England, including London.

[11] Relatedly, and according to Kaplan and Robertson (1991:14), "it is evident from various interviews with pioneers that the Russian Jewish immigrants found the company of Afrikaaners more attractive than that of the British-born". The sociolinguistic history of this group certainly demands further research.

1982:328–9; my parenthesis), with women in the forefront of the advancement of the relevant Natal English variables (Lanham 1982:336). This is summarized beautifully, by Jeffery (1982:253), in his review of Lanham and Macdonald (1979):

> "So later on there were plenty of people from Natal who could associate effortlessly with the MC and UC fortune-seekers from Victorian England drawn by the gold and diamonds: and when it was the Natalians who were getting the large slices of the new cake it was natural that their way of speaking English should be chosen as a model by the many non-British immigrants, mostly East European Jews. Anybody could see that a NE-accent opened wider doors than a CE-accent. So developed another variety, Respectable South African English . . . which as the "local" standard . . . co-exists with the "authentic" standard . . . that is RP, or the RP-like Conservative South African English".

Two aspects of this quote require emphasis. Firstly, the role played by NE in the development of GenSAE and, in particular, the influence, as mentioned above, of the more Northern varieties of England on the NE dialect and accent[12]. Secondly, the emphasis placed by Jeffery (1982), on the development of another (i.e. *new*) variety, is one which, I believe, has not received its due. As Trudgill (2004:84) emphasizes, "certain sorts of sociolinguistic situation involving contact between mutually intelligible dialects – colonial situations, new towns, rapid urbanisation – can lead to the development of new dialects". Given the large number of immigrants to Johannesburg etc, the rapid urbanisation that took place, and, more importantly, given that Cape and Natal colonials were also, in large measure, 'immigrating' from the original Cape and Natal colonies to the gold and diamond fields (and to, effectively, another country[13]), it is reasonably clear that early Johannesburg constituted just such a sociolinguistic situation. There was, therefore, ample opportunity for new forms of 'mixing' to occur and thus for the emergence of SAE varieties different in important respects to the original colonial varieties. If we assume, furthermore, and as will be detailed later, the ever-increasing social and geographic growth of GenSAE at the expense of both BrSAE and Cultivated SAE (CulSAE), then it is not of little consequence that this form of SAE could, from the above-mentioned perspective, and along with the other sociolects of SAE, be regarded as a *late*-19th-century as opposed to an early or even

[12] It should be stressed here, that the NE 'dialect' brought to Johannesburg was, in all likelihood, a somewhat unfocussed one. The 37 years spanning 1849 (the beginning of permanent settlement in Natal) to 1886 (the discovery of gold on the Witwatersrand) was unlikely to have produced a high degree of levelling in NE. This 'dialect' was, therefore, probably still characterised by a high degree of variability.

[13] i.e. to the Zuid-Afrikaanse Republiek, one of two 'Boer'-states established in the 19th-century and eventually overthrown by the British as a result of the Second Anglo-Boer War.

mid-19th-century 'new' English. While this 'model' is still rather speculative, we will see later that it has some merit as an explanatory framework[14]. This model which, for lack of a better expression, I term the 'three-stage koinéization model' of the formation of SAE is represented schematically in Figure 3.1 on the next page.

In Figure 3.1, the triangles represent the three separate koinéization stages (CE, NE and Johannesburg) and the labeled arrows represent the inputs, with the size of the arrows representing roughly estimated strength of influence. It should be emphasized that this is a working model and is eminently open to revision, particularly in terms of the implied strength of any one input. Its utility lies in allowing one to distil, as represented in Figure 3.1, the various stages and (exogenous) influences that were part and parcel of the development of SAE.

Jeffery (1982), as quoted above, also mentions the role played by RP or, at least, near-RP varieties in the history of SAE. From at least post-WWII times, there has, however, been a trend away from the use of a British or near-British reference point:

> "The main effect of the second wave of English immigration[15] to Southern Africa was: firstly to establish more firmly than before the dimension of social dialect in South African society, and secondly to assign the highest social value to standard British English, or, in more recent times, forms of South African English fairly close to this status dialect of Great Britain . . . At the present time the process of change is working the other way. South Africanisms are encroaching rapidly on near-British-English in this country and the difference between child and parent in the professional, better educated group has never been more marked that it is today" (Lanham 1964?:21–2; my footnote).

Generally, therefore, and as confirmed, for example, in Schneider (2007:180), the story of SAE involves, *very roughly*, the absorbtion and reanalysis of regional dialects (Cape and Natal English) as sociolects (BrSAE and GenSAE) as a result of the development of the modern industrial city, particularly Johannesburg, and the diffusion of this new class-based form of linguistic stratification throughout South Africa as a whole; this picture, essentially what Trudgill (2004:87–8) would term 'reallocation', is, however, somewhat complicated

[14]For one, and as a brief aside, one of the merits of this model is that it allows one to take seriously the possibility that the learner-English of the Eastern European Jewry contributed significantly to the feature pool out of which GenSAE arose. Bekker (2007), which provides evidence for the existence of a fronted, dental or denti-alveolar /s/ in some subvarieties of GenSAE, ends by speculating on the possible source of this feature being the Yiddish brought over by these immigrants. A similar hypothesis regarding the nasalization of vowels, particularly in certain areas of Johannesburg, is provided in §6.2. Naturally, these and other possibilities require further research.

[15]The author is referring here to immigration to Johannesburg and not to Natal.

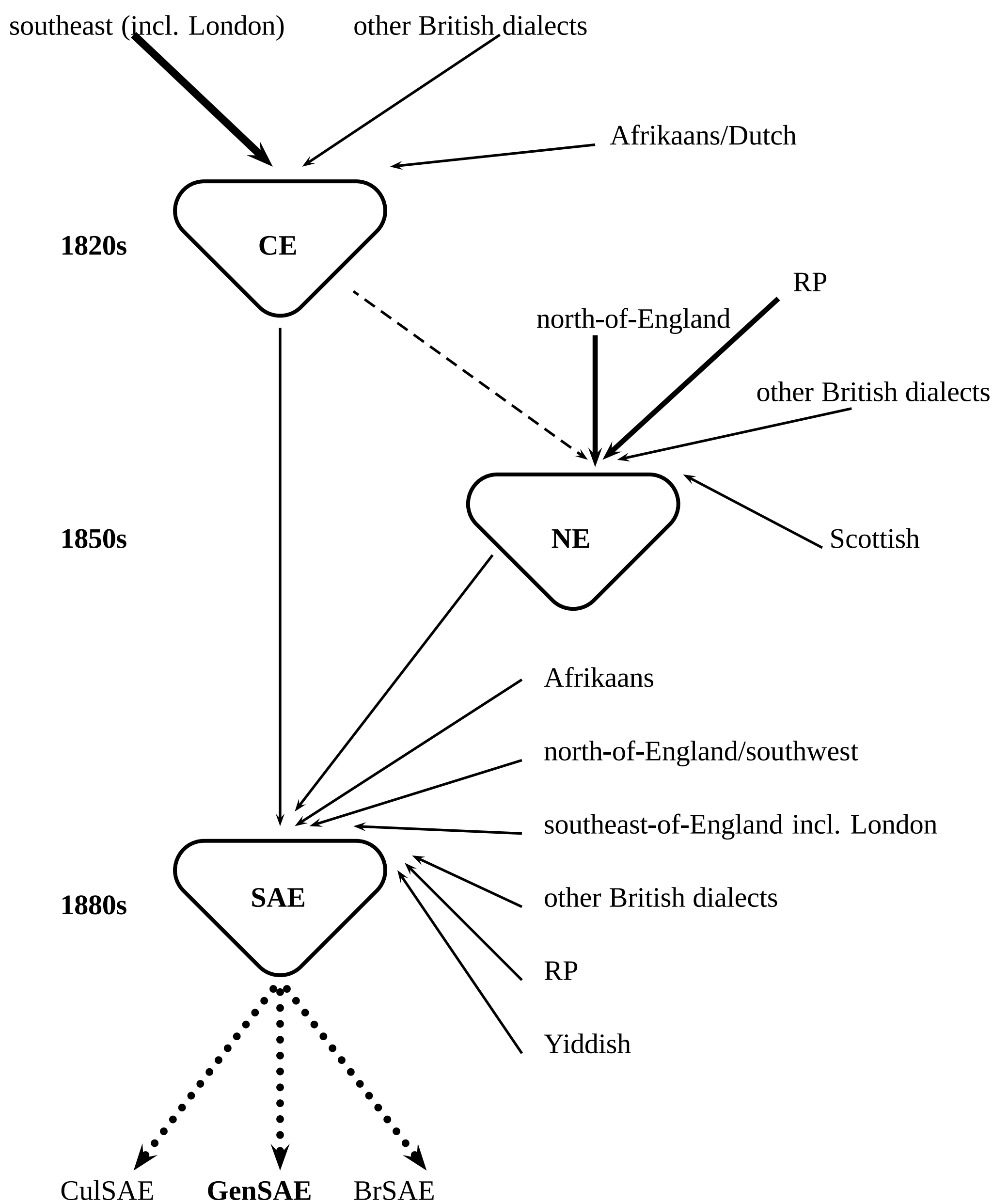


Figure 3.1: The Three-Stage Koinéization Model of SAE Development

by the existence, for some time, of two competing standards, one local (GenSAE) and one with an external focus (CulSAE). Substantially based on Lanham and Macdonald (1979), is is also, to some degree, speculative in the absence of "strong historical evidence for the period before 1870" (Jeffery 1982:257) which might serve to confirm the hypothesis that SAE is essentially the 'merger' of what were previously two regional varieties. It is also, I believe, complicated by the fact that the koinéisation process involved in this reanalysis, was, due to the 'weakness' of the Founder Effect, open to the inclusion of other features from other dialects, such as late 19th-century RP, a variety of non-prestigious British dialects as well as the L2 varieties used by Afrikaans speakers and the Eastern European Jewry. With respect to the potential influence of RP and NE, Trudgill (2004:87; my parenthesis and emphasis), in his description of new-dialect formation and interdialect forms, mentions hyperadaptation, "in which [*adult*] speakers attempt to use forms from higher status accents, but employ an incorrect analysis and extend changes to items where they are inappropriate". As a whole, therefore, a simplistic equation of GenSAE with NE, or BrSAE with CE, needs to be avoided.

With respect to the social value of the two competing standards, Lanham (1982:331) makes the following claim:

> "...sociolinguistic change in English-speaking society does not represent a major change in values, a disregarding of a British tradition in social values and a British identity in favour of those obviously South African ... Standard British English retains its social meaning as "standard" – equatable with high social prestige and correctness – with even greater consensus in English-speaking South African society today than in Britain. But in present-day South Africa, it is inaccessible as a model, particularly for those whose links with Britain are remote or nonexistent – a majority in the higher socioeconomic strata. Originally taken as a local substitute for authentic standard English, Natal English is, in the form of Respectable South African English, accepted as an informal, local standard expressing high social status if not correctness in English speech. Standard British English is today only acquired in childhood as a peculiarly individual social experience. Respectable South African English is the viable standard, and it expresses symbolically the antithesis of "typical local man" with his obviously local speech and his values inherited from the frontier".

While the above *might* have validly encapsulated the state-of-affairs of the time, although I suspect it overstates the degree of allegiance to all things British, it certainly seems incongruous as a description for the (particularly young) English-speaking community of South Africa today. With respect to its pronunciation norms at least, Standard British English, or its local equivalent, CulSAE, hardly seems to provide a model for correctness and, while this remains a purely impressionistic assessment, it is probable that Respectable (General) SAE has become virtually the *only* standard, expressing *both* social prestige and correctness. It should also be mentioned that for Lanham and Macdonald (1979) at least, GenSAE was, at the time, advancing at the expense of BrSAE, or, as Jeffery (1982:256) puts it, "CE is receding and NE advancing". If valid, this is an important observation, given the standard Labovian distinction between changes-from-above and changes-from-below. The former are often characterisable in the following terms:

> "*Changes from above* are introduced by the dominant social class, often with full public awareness. Normally, they represent borrowings from other speech communities that have higher prestige in the view of the dominant class. Such borrowings do not immediately affect the vernacular patterns of the dominant class or other social classes, but appear primarily in careful speech, reflecting a superposed dialect learned after the vernacular is acquired" (Labov 1994:78).

That Lanham and Macdonald (1979), for one, clearly viewed the 'importation' of NE variables into, for example, the speech of early Johannesburgers, as a case of change-from-above is clear from the following quote:

> "The 'lower class', in the part they play in advancing NE variables, must be seen in the peculiarly SA form of class structure; their nearest equivalent in Labov's analysis of New York City is the lower middle class" (Lanham and Macdonald 1979:86).

As we will see more clearly in Chapter 4, it is particularly the hypercorrective behavior of the interior classes (including the lower-middle) that is instrumental, from the classical Labovian perspective, for changes-from-above.

Yet, while CulSAE might be said to clearly constitute a superposed dialect, Labov's (1994) description, as provided above, does not apply in all its detail to GenSAE. Firstly, and as will be seen below, the link between the BrSAE–GenSAE continuum and social class is a loose one at best. Secondly, the NE variables that are isolated by Lanham and Macdonald (1979) as differentiating GenSAE, were not, in many cases, *borrowings* first adopted by the dominant class of Johannesburg, but (in all likelihood and in part) constituted the vernacular used by many Natalians who migrated to the Johannesburg-area. This is even more likely given Lanham and Macdonald's (1979) observations regarding the negative evaluation that PRICE-fronting and monophthongisation has been subject to *in Natal*. What occurred in Johannesburg was thus, most probably, a case of dialect contact and dialect-mixture rather than the imposition of a 'superposed dialect' unrelated to the vernacular. While it is clear that one result of this contact and mixture was the development of a sociolectal continuum of sorts and not a complete 'leveling' and integration of the various dialects, it is also reasonably certain that both retained (and retain) vernacular status for a sizeable number of speakers. It should thus not be surprising if *some* GenSAE features have undergone shifts more typical of so-called changes-from-below i.e. changes that appear first in the vernacular, are driven by 'internal', endogenous factors, and which speakers (and even phoneticians) only become aware of once the changes in question have almost reached completion (Labov 1994:78).

More generally, what SAE shares with most Southern Hemisphere colonial Englishes (as opposed to Northern Hemisphere colonial Englishes like the US and Canadian dialects) is, according to Lass (1995:93), the development of three distinct sociolects; what this author refers to as the "the great trichotomy". These are, in the case of SAE:

1. A standard with an external British reference: Received Pronunciation (RP) or near-RP with respect to pronunciation. Cultivated SAE in South Africa. This variety is hardly found amongst the young any longer (Lass 2002:110). Thus, for example, as early as 1967, Lanham (1967:61) confirms that "a feature of English in South Africa is the rapidity with which more extreme SAE is displacing the conservative, near-British-English forms". Still, one needs to be cautious in this regard. Lanham and Macdonald (1979:37), for example, claim that "Cons SAE remains a functioning variety in SAE; it is still transmitted to children ... and is a formal-style accent of a small proportion of adolescents and young adults, particularly women". The possibility of CulSAE appearing in the acoustic data collected for this research has thus not been completely discounted, even though close to 30 years have transpired since these comments. On a phonetic level, Lanham and Macdonald (1979:36) define CulSAE as that variety of SAE characterized by a complete lack of NE or CE variables – see §3.2 above;

2. A local or 'provincial' standard; GenSAE in South Africa, characterizable mainly by NE variables according to Lanham and Macdonald (1979); During Lanham and Macdonald's (1979) time, although co-existing with CulSAE, it did not evoke any social stereotypes *vis-à-vis* CulSAE[16]. The reverse would, in fact, be true in South Africa today i.e. CulSAE evokes stereotypes. For Lanham and Macdonald (1979) both CulSAE and GenSAE were, however, associated with "rejection of South Africanism

[16]Although admittedly the use of certain marked variants, such as PRICE fronting and monopthongisation to [aː], are often stereotyped as typical of the so-called 'Kugel' of Eastern European descent – more on this in §3.3.2.

in favour of links with the wider Anglo-Saxon world, a low level of patriotism, and hostility towards Afrikaners" (Jeffery 1982:254). Jeffery (1982) augments this analysis of the indexicality of variation in SAE by positing a more complex link between linguistic variation and politics. With respect to GenSAE, he explains as follows:

> "Resp SAE speakers, on the other hand, are not so stereotyped: if you uphold the British liberal tradition, which is to say if you are progessive, left-wingish, anti-Afrikaner-Nationalist and pro-African, then you will very likely be U/MC and speak with a Resp SAE accent; but nobody would be led by your accent alone to assume that that is how you think, for everybody knows people with respectable accents and illiberal opinions" (Jeffery 1982).

3. A local vernacular; Broad SAE in South Africa, characterised, according to Lanham and Macdonald (1979) by the prevalence of CE variables. "The more extreme a variety is, the harder it becomes to distinguish it from second-language Afrikaans English" (Lass 2004:373). According to the interpretation of Jeffery (1982:252; my parenthesis) of Lanham and Macdonald (1979), essentially an "uprooted [Cape English] with some influence from Afrikaans, and which has an automatic [lower-class]-association in present-day white SA society". Its indexicality is, however, somewhat more complex than simply that of lower-class. Thus Jeffery (1982:253) mentions Lanham and Macdonald's (1979) association of BrSAE with a stereotype involving, for men at least, the attributes of being "tough, manly, sport-mad, sociable, patriotic and other things beside". On a more political level, "Ext SAE is loaded with political-ideological meaning as well as social: the South African tradition is to be not only tough etc. but also conservative, right-wing, authoritarian, unsympathetic to African aspirations, and so on; and Ext SAE speech reliably predicts such views, which are anything but covert (whether in the Labovian or lay sense) in white SA society, and which are a significant part of the stereotype of the "typical local man". And indeed you do not have to be LC to conform to the stereotype" (Jeffery 1982:255).

Interestingly enough, Jeffery (1982:255) adds that the political-ideological indexicality of varieties of SAE has "important implications for orthodox sociolinguistic theory, which at present appears to know nothing of political-ideological values associating with varieties rather than languages". This is echoed in, for example, Deumert (1997:159) who argues that one of the weaknesses of sociolinguistics in the Labovian tradition is its *a priori* division of the population under investigation into various socio-economic categories "and it is taken for granted that these social groups share a common linguistic behaviour". As hinted at in

§1.3.3, however, this is certainly no longer the case in much sociolinguistic work.

One important putative difference between SAE and other Southern Hemisphere Englishes is highlighted by Lanham and Macdonald (1979:21; footnote):

> "The peculiar properties of English in SA society are highlighted by a comparison with Australian English as another variety of 'Commonwealth English' of similar vintage and origin as SAE. A first point of difference is the fact that an accent close to Standard Southern British English ... is a functioning variety in the SAE community, but apparently not in Australia".

The point here is not that AusE does not have a Cultivated sociolect, but that the later is viewed as indigenous: "whatever accent type has the status of standard in Australian society, such a form is typically Australian speech" (Lanham and Macdonald 1979:91). In the South African context this difference also explains, in Lanham and Macdonald's (1979) time at least, the existence of two *competing* standards in the SAE community and the rapid disappearance of CulSAE in modern SAE society. Jeffery (1982:254), however, downplays the interest-value of the existence of two co-existing standards, mentioning Ireland and Scotland, "where Scottish Standard English and RP co-exist in similar fashion ... nothing very interesting seems to emerge from it". Furthermore, Lanham and Macdonald's (1979) claims in this regard are cast in doubt by, for example, Harrington, Cox and Evans (1997:156) who describe Cultivated AusE as that sociolect which "most closely approximates Received Pronunciation of British English".

It should also be noted that these three categories are not internally homogenous. So, for example, Lass (1995:94–5) talks about 'Extreme Conservative' (read 'Extreme Cultivated') to refer to a sub-type of CulSAE which is closest to an old variety of RP; and, in relation to BrSAE, about "the *range* of accent-types associated with relatively low socioeconomic status" (Lass 1995:95; my emphasis). The fact that these three sociolects in fact imply a continuum, with considerable overlap, should also be emphasized; that the same consideration applies to, for example, AusE is just as clear (Harrington et al. 1997:156).

Since the subjects in the acoustic component of this research are young (thus in all likelihood excluding CulSAE) and from the higher end of the socioeconomic scale, they are likely to use the *General* variety of SAE, the focus of the next section.

3.3.2 GenSAE and its Prestige Values

In the previous section, the non-homogeneity of CulSAE and BrSAE was emphasized. The disappearance of CulSAE from the speech of the young was also mentioned, with GenSAE, in the process, replacing CulSAE as the new standard in South Africa. As with the other sociolects of SAE, it is, however, not perfectly homogenous and, as is the case elsewhere in the English-speaking world, it is likely to vary with the socioeconomic status (or identity) of its speakers. The subjects recorded for the acoustic analysis component of this research are from the highest end of the socioeconomic scale and it is therefore pertinent to briefly review some of the literature relating to prestige values in GenSAE. A higher level of technical detail is indulged in in later chapters.

Lanham (1978:153) uses the expression "a feature of "refined" female speech" to characterize the fronted, monophthongized variant [aː] of the PRICE vowel in GenSAE. It is characterized by this author as "the defining variable of Resp SAE" (our GenSAE). Lanham (1978:148) confirms the prevalence and origin of this feature in Natal, other examples being (substantial) vowel retraction before /l/, a raised, fronted and rounded NURSE vowel and, possibly, centralised to fronted GOOSE (Lanham and Macdonald 1979:43–5). It is interesting to note, with respect to the latter, that Limbrick (1989) provides (tentative) evidence to suggest that the affrication of /j/ (which often occurs before SAE GOOSE in words such as *tune* or *duke*) might also have a Natal bias[17].

The role of NE variables, as opposed to CE variables, as loci of prestige in GenSAE is emphasized in Lanham and Macdonald (1979:35) in the following passage:

> "CE variables are the most salient characterizing properties of stigmatized, non-standard Ext SAE . . . style-shifts in formal speech reverse rather than advance the trends. NE variables on the other hand are associated with Resp SAE, they receive social approbation . . . and characterize the 'provincial standard'; there is evidence that they are advancing in the speech community and formal style shifting is towards them rather than away from them".

In short, "the SAE community treats fronted ai – and probably other NE variables in their common co-occurance with ai – as expressing social prestige and, in this sense, a standard variety" (Lanham and Macdonald 1979:62). It should be emphasized, however, that the role of Natal variables or input more generally is not accepted by all commentators. Lass (1995:93), for one, states that while "Lanham has claimed that the Natal input is particularly important as the source of certain variables like 'glide-weakening' (i.e. monophthongisation) of certain diphthongs . . . these processes are so widespread in English that they cannot be argued to have a specific regional input".

[17] Any attempt to link GOOSE-fronting and the affrication of /j/ ('yod') as in [ʧʉːn] needs to, however, take account of the fact that American English, which commonly drops its 'yod', [tʉːn], has also undergone GOOSE-fronting – see §8.2.

Lanham and Macdonald (1979) stress that the role of PRICE and other variables as prestige values in GenSAE is most prominent among young (Jewish) females in Johannesburg. In Natal, at least during Lanham's time, they are subject to negative attitudes. So, for example, Lanham and Macdonald (1979:63) state that "old females of the Natal sample show negative attitudes and the suppression of NE variables in formal speech, although young Natalians apparently holding similar attitudes, are nevertheless in step with the wider SAE community in maintaining these variables in formal speech". The two authors elaborate further:

> "Results ... support a hypothesis that, except in Natal, the SAE community treats fronted ai – and probably other NE variables in their common occurrence with ai – as expressing social prestige and, in this sense, a standard variety. Fronted ai is highly valued by, in ascending order of degree of approbation: old Cape females; young non–Jewish Witwatersrand females; young Jewish females on the Witwatersrand" (Lanham and Macdonald 1979:62).

They also point out in a table the "advance through age–grades of the Fronted ai in Witwatersrand female speech" as well as its prevalence in "formal speech behaviour" (Lanham and Macdonald 1979:62–3); and refer to the "spectacular advance of Fronted ai in the speech of Witwatersrand females" (Lanham and Macdonald 1979:57). More generally, "a significant group of exponents of Resp SAE [and thus NE variables in general] comprises young and middle–class females of the Eastern European descent group, particularly in Johannesburg" (Lanham and Macdonald 1979:149; my parenthesis). As should be clear from the following quotation, however, even during Lanham and Macdonald's (1979) time, the fronted PRICE vowel (and thus perhaps advanced values for the other NE variables mentioned above) was not only a feature of the Eastern European (Jewish) descent group in the Johannesburg area:

> "The Eastern European descent group has the closest association with the Kugel stereotype analysed in caricature in the Johannesburg Sunday Times, September 19, 1976, and defined as: 'single, pretty, often rich girls; university graduates ... but only there to hunt for a husband amongst medical and law students; mainly Jewish - *although anyone can join*"' (Lanham and Macdonald 1979:57; my emphasis).

More recently, Bekker and Eley (2007) conducted an acoustic analysis of the monophthongs of GenSAE, using data elicited from two sets of subjects: young females from private schools in Johannesburg and young females from public schools in East London.

Results suggest that a lowered TRAP vowel is a new prestige value, particularly for the Johannesburg area, and more specifically for one of the more wealthy areas of Johannesburg; the so-called 'Northern Suburbs' – see §4.3.2 for a definition of this term.

How do we conceptualize these 'prestige' variables in terms of Labov's (1966) distinction between changes-from-below and changes-from-above? As is common with the latter, these variables have been used (if not introduced) by the dominant social classes and, by all accounts, appear in careful speech. On the other hand, it is unclear that, apart perhaps from fronted /ai/, they have reached any real level of public awareness. More importantly, they do not, in any real sense, "represent borrowings from other speech communities that have higher prestige in the view of the dominant class" (Labov 1994:78). While it is clear that CulSAE is easily conceivable, for many SAE speakers, as a "superposed dialect learned after the vernacular is acquired" (Labov 1994:78), the same does not appear to apply to GenSAE as already hinted at in §3.3.1.

3.3.3 SAE and the Dynamic Model of Schneider (2003; 2007)

Schneider (2003:235) provides "a new, coherent framework to explain the emergence of, and thus the relationships among, New Englishes," among which he includes mother-tongue colonial varieties such as AusE, NZE and SAE. Unlike Labov (1991) and Trudgill (2004), the emphasis in Schneider (2003; 2007) is on broader sociolinguistic development through a number of stages, beginning with an initial stage of koinéization which he defines as "the emergence of a relatively homogenous 'middle-of-the-road' variety" (Schneider 2003:224) from a diverse linguistic input, such as that which prevailed in most settler communities in colonies such as New Zealand and Australia. The end-point of said sociolinguistic development, on the other hand, is a process of regional and ethnically-based diversification. The so-called 'Dynamic model of the evolution of Postcolonial Englishes' posits the following five phases. The rough descriptions of these phases, as provided below, are based on Schneider (2007:33–55):

Phase 1, Foundation: The relevant settlers regard themselves as British and completely separate from the indigenous population. There is dialect contact between the various English varieties transplanted to the relevant colony, and thus the related processes of koinéization, leveling and the construction of interdialectal (compromise) forms. Minimal language-contact *vis-à-vis* the indigenous languages.

Phase 2, Exonormative stabilization: Settler-identity is described as ""British plus": genuinely British no doubt but seasoned with the additional flavor of the colonial experience" (Schneider 2007:37). A small but important segment of the indigenous population becomes bilingual, with the use of English being associated with high status. Overall, norm-orientation is towards an external source i.e. RP and 'Standard' English, while "English in its spoken form begins to move toward a local language variety" (Schneider 2007:39). For the indigenous populations, structural nativizations occur, linked largely to transfer phenomena.

Phase 3, Nativization: A stage of growing political independence for the colony concerned. "The gap between immigrants and indigenous population groups is significantly reduced at this stage" (Schneider 2007:41), although on a linguistic level most of the 'work' of accommodation is conducted from the indigenous side, thus "leading to widespread second-language acquisition of English, and sometimes an almost complete language shift or even language death" (Schneider 2007:42). The growing separateness of the relevant forms of English leads to a 'complaint tradition'. Linguistically, the gap between the English of the settler-strand and that of the indigenous population is reduced, often to a sociolinguistic (i.e. class) distinction i.e. "lower-status members of the [settler] community are quite likely to become active agents in the transfer of features from the [indigenous] strand, diffusing local, transfer-based innovations into native-language dialects" (Schneider 2007:45; my parenthesis). This stage is also characterized by the appearance of mixed-codes/languages[18] often associated with specific identities. According to Schneider (2007:48; my parenthesis), "mixed codes apparently originate when the native language of the [indigenous] strand is still strongly rooted in the community (and possibly receives some official support) and English also enjoys high prestige (but access to it is limited)".

Phase 4, Endonormative stabalization: This phase follows from the previous one as a result of a "new identity construction that follows political separation" (Schneider 2007:48). This necessary but not sufficient condition of political independence is highlighted by Schneider (2007) by pointing out that both Australia and New Zealand remained essentially British for decades after *de facto* political independence. The transition from phase 3 to 4 thus requires an "exceptional, quasi-catastrophic political event which ultimately causes the identity alignment of [settler] strand speakers to switch from a self-association with the former mother country . . . to a truly independent identity" (Schneider 2007:48; my parenthesis), termed 'Event X' by the author. Identity construction includes members of the indigenous population and a

[18] To be distinguished, according to Schneider (2007:47), from code-switching i.e. a mixed-code is one variety-as-system, while code-switching is the simultaneous use of two systems. As Schneider (2007) puts it, "code-switching represents a performance phenomenon".

new, local norm is used to express that identity. On a linguistic level "the new indigenous language variety is perceived as remarkably homogenous" (Schneider 2007:51), although this perception might reflect a political need rather than a reality.

Phase 5, Differentiation: As opposed to the previous stage, in which the focus of identity-construction was in opposition to the 'mother' country, here the focus moves on to the various social divisions existing *within* the new nation: "the focus of an individual's identity construction narrows down, from the national to the immediate community scale" (Schneider 2007:53). This leads to the development of new regional and social dialects meant to reflect these internal divisions.

It should be added that Schneider (2007) emphasizes the fact that the model needs to be regarded as a 'fluid' one, allowing overlap between the various phases.

The question is, of course, where SAE (broadly-conceived[19]) belongs in this model. Schneider (2007:173–88), in fact, dedicates a section to South Africa, but the account is not convincing in all respects. In essence, Schneider (2007:188) considers SAE to have "made deep inroads into phase 4" with the 1994 move to full democracy constituting the relevant 'Event X'. While it is certainly and uncontroversially true that SAE has undergone Phases 1 and 2, it is unclear to me that SAE has come close to completing Phase 3 and, as a consequence, I believe that the 'inroads' into Stage 4 are not as 'deep' as proposed. Essentially, it seems clear, to me, that the nativization process (Phase 3) has only begun with respect to the 'black' community, largely as a result of the *de facto* or legislative 'Apartheid' that has characterized South African history until very recently; it is thus only recently that the 'gap' between the various population groups has *begun* to be bridged. By Schneider's (2007) own admission, SAE, as broadly conceived, is hardly characterizable as homogenous and "no local variant of English is the carrier of this newly emerging national identity"; furthermore, the outcome of "a focussing process toward a pan-South African norm ... is still a few generations down the road" (Schneider 2007:187).

While the above hardly constitutes a substantive critique of the model *per se*, it does constitute a critique of its application to a particular variety of English. I would like to suggest that a less rigid adherence to the temporal *sequencing* of the various phases as well as the adoption of an approach that allows for various sub-varieties of SAE to run their own 'course'[20], 'fits' the SAE data better than the one taken by Schneider (2007) himself. On this application of the model, the development of SAE as a '*white*' variety is characterized

[19] i.e. as including all forms of English in South Africa, not simply that spoken by the white minority. The non-default reading is being used here.

[20] At least until Phase 4, which is, by definition, a phase characterised by at least perceived homogeneity.

by a Phase 1 koinéization process that arguably culminated in the mixing and reallocation that took place in Johannesburg and the other 'boom' towns of the late-19th-century[21]. An important element of this koinéization process was, undoubtedly, the influence of extensive English-Afrikaans bilingualism, at least in the CE and, to a degree, Johannesburg settlements. There was, in addition, a clear phase of exonormative orientation, at least until WWII. The nativization phase would appear to have run, for some time at least, concurrently with the 'earlier' two phases i.e. the reanalysis of CE and NE into BrSAE and GenSAE to various 'points' on the social-class or at least broad-indexical scale was, in all likelihood, the direct outcome of the mixing and reallocation that took place in Johannesburg and was followed by the *spread* of GenSAE, as the new standard, at the expense of CulSAE and BrSAE, a phenomenon which appears have gained momentum particularly in the second half of the 20th-century. It is this spread that has, presumably, also led to the gradual, although perhaps not complete, recession of the settlement-based regional markers which defined CE and NE and thus, consequently, also to growing homogeneity coupled with an endonormative orientation, both characteristics of Phase 4 of Schneider's (2007) model[22].

As far as possible developments on a phase 5 level for 'white' SAE are concerned, the pronouncements are scattered and difficult to interpret. We return to the possible development of *modern* regional variety in SAE below. At this point, I would like to suggest, when applied to SAE as a more broadly conceived concept, and, in particular, in terms of the role played by the '*black*' majority, that SAE has not completed Phase 3 and that, in fact, the process of nativization has only begun, an interpretation that gains support from some of Schneider's (2007) own observations, as reported on above. While Black South African English[23] has gained some support, it seems unlikely, as far as I can tell, to develop into

[21]Thus, as already hinted at above, this conceptualization of the development of 'white' SAE, views the process of koinéization as a three-stage one in the South African context. Two original koinés, i.e. CE and NE, are transplanted, from their original base and mixed, in Johannesburg etc., with other varieties, both L1 (e.g. British dialects) and L2 (e.g. Afrikaans English and the Yiddish-influenced English spoken by the Eastern European Jewry). The new varieties i.e. CulSAE, GenSAE and BrSAE, are, as a result, similar to but not identical with their original sources (i.e. RP, NE and CE) and together form a sociolinguistic continuum.

[22]For now, I leave as an open question whether either the BrSAE or GenSAE varieties have been in any way affected by Afrikaans on a phonetic or phonological level. That both have been affected by Afrikaans on a lexical and lexico-grammatical level is certain; thus the adoption of loans such as *lekker* and *ag* (which are certainly no longer restricted to Broad variants). Note too the use of *busy* as part of the progressive construction, as in *I was busy sleeping*, which has also been mooted as a possible transfer from Afrikaans.

[23]Defined here as those varieties of South African English which show evidence of L1-transfer from the Bantu languages, regardless of whether these features are the result of individual language-learning or of a broader process of institutionalization. Such varieties should be carefully distinguished from other emerging varieties that seem to be based on "the "Respectable" accent, with some additional influence from "American sitcom English"" (Schneider 2007:187), as explored, for example, in Da Silva (2007).

a pan-South African norm[24]. On the other hand, and as reported on in Da Silva (2007), a potentially new and non-racial norm, derived mainly from GenSAE but 'spiced' with what appear to be a number of Americanisms, has *begun* to develop. While its use still appears to be largely ethnically-based, impressionistic evidence as well as some of the data collected as part of this research project, seems to indicate that this new variety is spreading into the 'white' community. Some of these features will be reported on in later chapters.

Returning to the development of new variation in 'white' SAE, *a là* Schneider's (2007) Phase 5, we find a lack of consistency in the available literature. Thus, on the one hand, the early Lanham (1964?:33) states bluntly that "English in South Africa has no clearly marked regional dialects" and in terms of a hierarchy proposed by Lanham and Traill (1962:1), region is regarded as the least important determining variable after social class, age and gender. Wells (1982:622) mentions a "trend towards greater homogeneity in South African English pronunciation". On the other hand, the more recent Bowerman (2004:935) makes the following claim:

> "Regional variation in WSAfE is naturally associated with the strongest concentrations of White English speaking communities. These can broadly be divided into (Western) Cape, Natal and Transvaal (Gauteng) English, and recognisable Namibian and Zimbabwean varieties".

This difference of opinion could, of course, be linked to possible (Phase 5-like) developments over the past few decades. Thus the relatively early Lanham and Traill (1962:4) claim that "there are many indications that SAE is changing fairly rapidly;" and while these authors are referring here to change in SAE *as a whole*, as opposed to its separate development in the various urban centers, it seems unlikely that SAE[25] has been completely immune to forces that are fully operative in other areas of the English-speaking world. It is thus at last conceivable that over the last few decades SAE has 'come into its own' in the sense of developing not only an endonormative orientation (Phase 4) but, in addition, the potential for entering "a vigorous phase of new or increased, sociolinguistically meaningful internal diversifaction" (Schneider 2003:254).

As should be clear from the review of Schneider (2003; 2007) and the applicability of the Dynamic Model to SAE as provided above, it is important to keep the original, settler-based regionalisms separate from 'putative' Phase-5 developments. Thus while Lanham and Macdonald (1979) and Branford (1994:472) refer to certain features that correlate with regional provenance (particularly the Eastern Cape and Natal), these differences

[24]I do, however, confess that this perspective is not shared by all commentators.
[25]I return here to the default reading of 'SAE' i.e. 'white' SAE.

have been largely obscured by the reallocation of regional markers for social-class purposes and, through the spread of GenSAE and the establishment of what Schneider (2003) terms "endonormative stablilization" i.e. Phase 4. By way of example, the influence of final /l/ on preceding vowel quality, was, according to Lanham (1967:66-7) more prevalent among Natal-speakers over 50 years of age than similar speakers from other areas of South Africa. "/l/ influence ... however ... is at present discernable in the speech of the majority of younger SAE speakers, being particulary noticeable in more extreme forms of SAE in Natal and the Transvaal" (Lanham 1967:66-7). Similarly, Lanham and Macdonald (1979:22) make reference, with respect to regional differences linked to the Eastern Cape and Natal, to "older (over 45), long-resident subjects".

Still, subsequent post-Phase 4 developments that might serve to distinguish SAE speakers from the different urban centers have certainly been alluded to in the literature. Thus Lass (1990:272; my emphasis), for one, implies (modern) regional variation in his description of "Respectable (=(Upper)Middle Class) White *Cape Town* English". He goes further in a footnote:

> "Local folk linguistics suggests distinct 'Atlantic' and 'False Bay' (western and eastern) varieties, as well as northern vs. southern suburbs, not to mention finer subdivisions, like 'Atlantic Jewish' (Sea Point) vs 'False Bay Jewish' (Muizenberg). Intuitively my guess is that these terms correspond to genuine natural kinds" (Lass 1990:283).

On a more informal level, many people claim an ability to identify whether a South African comes from, say, Johannesburg or Cape Town, although interesting in this regard is Lanham's (1964?) claim that "those who claim that they can identify English speakers as coming from Cape Town, Johannesburg or Durban usually fail when put to the test" (Lanham 1964?:34). Generally, however, there has been no attempt, apart from the research by Lass (1990) mentioned above, to isolate regional dialects based in the various urban centres of South Africa and even in this particular case it is not always clear which features are specifically 'Cape Townian' or characteristic of GenSAE more broadly. This lack of clarity is in stark contrast to, for example, the situation in the USA today where, in a recent interview (Gordon 2006:348), William Labov is able to confidently and uncontroversially pronounce that "regional dialects are getting stronger and more diverse as language change is continuing"[26].

[26]As an aside, while, according to Lanham (1982:350;footnote1), there is a case for differences existing between SAE and the English spoken in Zimbabwe (formerly Rhodesia), Wells' (1982) claim that Rhodesian English "has not been systematically studied" (Wells 1982:611) is still valid and it is thus a dialect worthy

Thus while it is, in all likelihood, true that, “it is the entrance of new ethnic and racial groups into the community that provides the motivating forces behind renewed diversification” (Labov 1980:263) and that, more broadly, “a well-accepted sociolinguistic principle is that the fluctuating course of linguistic change is correlated with and indirectly caused by social changes that alter the structure of the speech community” (Labov 1990:206), it is still unclear whether this has translated into a true Phase-5 development. Rather, it seems clear that a process of nativization (i.e. the linguistic reflection of the social integration that has taken place since 1994) is being referred to when Mair (2006:201) concludes with reference to GenSAE that “its future as the national pronunciation norm for a new multiracial state remains uncertain”, and Da Silva (2007:2) is moved to conclude that the traditional ethnic-based terminology commonly used to separate out the various sub-varieties of SAE is no longer valid given the entry, over the last two decades or so, of black individuals particularly into the previously whites-only educational system.

The methodological implications of all of the above is briefly dealt with in 4.3.2.

3.4 SAE and other English Accents

As emphasized by Labov (1994:9), “geographic separation naturally and inevitably leads to linguistic separation”. Thus, while an understanding of the inner-diversity of SAE, as provided in the previous sections, is important, it is also essential (and central to the focus of this thesis) to place SAE within the larger context of English accents in general and, in particular, in the context of those accents of English from which SAE has been, in a sense, separated. This section provides just such a context, with a specific focus on the vowels of English. Section 3.4.1 begins by showing how SAE qualifies as a so-called ‘Extraterritorial English (ETE)’ and what features it shares with other ETEs. Section 3.4.2 then reviews already existing attempts to place SAE within the context of Labov’s (1994) various patterns, as provided in §2.3.2.1. This is followed by §3.4.3, which concentrates on *Southern Hemisphere* ETEs, showing in what ways SAE is similar to (and differs from) other Southern Hemisphere ETEs like AusE and NZE. The section gives much space to the work contained in Trudgill (2004).

It should be stressed that the aim of this section is to provide a relatively ‘rough’ account of the similarities and differences between the various accents dealt with. In Chapters 6 thought 9, which deal with each of the SAE vowels in a thorough and step-by-step fashion,

of future attention. Some preliminary investigations in this direction have recently been made by Marshall (n.d.), but nothing thorough or substantive is yet on offer. The term ‘Rhodesian English’ is used here with no intended political overtones but rather to distinguish the L1 variety from L2 varieties which could, collectively, be labelled ‘Zimbabwean English’.

a more fine-grained analysis of said similarities and differences is provided, particularly through the comparative use of acoustic data.

3.4.1 SAE as an Extraterritorial English

Although unique in a number of respects, SAE is, according to (Lass 2004:369), like all ETEs except "some varieties of Irish English", characterisable as a dialect "of Southern British English (SBE)" (Lass 1995:89–90), and all 'Southern' dialects of English share the following phonetic/phonemic features:

1. According to Lass (2004:368), diphthongisation of MOUTH i.e. [aʊ] or something similar. This includes later diphthong-shifting and/or glide weakening, such as that found in SAE and Southern dialects of the USA. Dialects of the north of England often have a [uː]-like quality in MOUTH i.e. they have not undergone the final stages of the GVS, as explained under Pattern 1 in §2.3.2.1.

2. "A reflex of Middle English (ME) /ɔː/ in GOAT, rather than ME /aː/: the North has /eː/ or /e/ in words of this class (Scots *hame* = *home*)" (Lass 2004:368). Southern varieties range from [oː] to [ou] to [œu], including later developments.

3. [æ∼ ɛ] in so–called TRAP words (Wells 1982:129-30). The vowel used in words such as *trap*, *tap*, *back* etc. ranges from [æ] to as high as [ɛ], a range which is higher than the vowel [a] used in more 'northern' varieties such as those found in the north of England;

4. A STRUT-FOOT Split. Thus words such as *cud* and *could* are not homophonous as they are in the local accents of the north of England (Wells 1982:198). The FOOT vowel is in the region of [u] – RP, for example, has [ʊ] –, while the STRUT vowel is in a broad area ranging from [ʌ] to "central [a̶]{[ɐ̞]} to centralized front [ä] or even raised [ɛ̈]" (Lass 1995:90)[27].

5. What Lass (2004:369) calls Lengthening I: the BATH vowel is longer than the TRAP vowel and is usually different in quality. Thus in SAE the TRAP vowel varies from [æ] to [ɛ] (see above) while BATH is longer (different in length) and different in quality i.e. "centralized back [ɑ̈ː] or back [ɑː]" (Lass 1995:91). This can be fruitfully compared with Australasian English (combined AusE and NZE), which has [æ∼ ɛ]

[27] Note that The symbol [a̶] is not found in either the modern IPA (International Phonetic Association (1999)) or Pullum and Ladusaw (1996). It is, however, a logical analog of, for example, [ʉ], and indicates a fully–open central vowel. The best 'modern' transcription would, therefore, I believe, be [ɐ̞].

and the more front [aː], and American English, which in many cases only has a length distinction, i.e. [æ] in TRAP and [æː] in BATH. Note that with respect to the latter accent, in START, where there is a postvocalic /r/, and PALM, [ɑ] is used; and

6. What Lass (2004:369) calls Lengthening II: the TRAP vowel is lengthened before voiced stops and nasals (except /ŋ/); so, [kæp] *cap*, but [kæːb] *cab* (Lass 1995:91).

7. Lass and Wright (1985:137) also mention a contrast between the FOOT class, typically with [ʊ], and the GOOSE class, typically with [uː], as an additional defining characteristic.

Note that more recently Lass's (2004) position has been challenged, for example in Trudgill (2008:242), who argues for a polygenetic (as opposed to monogenetic) origin for SAE as well as other dialects which have been transplanted to other regions through colonization. From this perspective, therefore, SAE was "in origin a mixed dialect" (Trudgill 2008:250), resulting in a "new feature pool [which set] in competition variants which did not compete directly or regularly with each other in the motherland . . . new competition dynamics usually result in new varieties" (Mufwene 2008:255; my parenthesis). From this perspective, although many ETEs will predominantly show features with a southeast-of-England bias, circumstances do arise in which variants from other sources (such as the north of England or Scotland) persist, sometimes even to the exclusion of any others. Trudgill (2004:116–120) attempts to show that this is, in fact, the case for many Southern Hemisphere Englishes, showing how features such as H-retention (*heart* is [hɑːt] and not [ɑːt]) and a Weak Vowel Merger (*Lennon* and *Lenin* are homophonous) all provide evidence for the influence of non-southeast-of-England varieties on these ETEs[28]. It is perhaps worth pointing out at this early stage that SAE displays both of the above features and evidence will be provided at various points below for the influence of such non-southeastern varieties on the development of GenSAE in particular. It will be recalled above that GenSAE is seen to have been heavily influenced by NE, a (relatively unfocussed) proto-koiné which was, in its turn, characterised by a north-of-England (and middle-to-upper class) bias. It could also be hypothesized, given an acceptance of the notion that the creation of Johannesburg and other boom-towns of the late-19th-century provided the conditions for further mixing and

[28] Differences between the various Southern Hemisphere Englishes themselves are, presumably, also (partly) attributable to differential input from non-southeast-of-England dialects. Still, while Trudgill (2004:117) points to the absence of Glide Cluster Reduction (*which* and *witch* begin, respectively, with [ʍ] and [w]) in NZE, and attributes this fact to "the Scottish, Irish and Northern English form", he provides little explanation for its absence from AusE and NZE. Any answer would, no doubt, lie in a more in-detailed comparison of the immigration patterns to the respective areas. From an admittedly superficial reading of the matter it appears that one possible difference lies in the greater influence of Scottish English on NZE, while "the north of England . . . was underrepresented" (Trudgill 2004:16).

koinéization, that the other British dialects (RP and well as non-standard ones) as well as L2 varieties of the time had a role to play as well. Important in this respect is Trudgill's (2004) notion of hyperadaptation, an adult characteristic which has already been alluded to above; a process which often involves "an incorrect analysis" (Trudgill 2004:87) and thus, potentially, a reanalysis of features of the 'source' variety. In this regard, we have seen above that early Johannesburg was characterisable as heavily adult-male dominated. Naturally, the possibility that endogenous development has had a hand in the development of SAE has to be kept firmly in mind as well, and one of the challenges for the reconstruction of the linguistic history of an ETE such as SAE is to, indeed, separate out endogenous developments from those based on language and dialect contact.

3.4.2 SAEP and Labov's (1994) Patterns

There have, surprisingly, been very few attempts to describe the SAE vowel system in terms of Labov's various patterns of chain-shifting; and perhaps even more intriguingly, the first of these attempts makes recourse to Pattern 2 i.e. the pattern characterizing the so-called Northern Cities Shift in the USA. Thus Jeffery (1982:258) highlights the possibility that the raised DRESS and TRAP vowels, and the raised and monophthongised SQUARE vowel of SAE are related to each other as well as to fronted and monophthongised PRICE and MOUTH vowels and the characteristic polarization of the SAE KIT vowel. His analysis is as follows:

> "glide-weakening of /ai/ and /au/ made it expedient to move them out of the phonological space of /aː/{/ɑː/} so they came to be realized further forward ... causing the front series including /ɪ/ to move up ahead of them ... in this chain-shift realizations of /ɪ/ were conditioned by an adjacent velar consonant, with the effect that allophones not so conditioned reached [i] first and then moved on down the non-peripheral path towards [ə] as the velar conditioned allophones kept on coming up towards [i] ... All this is to propose that a chain-shift of the Pattern 2 type ... is taking or has taken place in SAE" (Jeffery 1982:258–9; my parenthesis).

While Jeffery's (1982) proposed explanation certainly does not classify as a prototypical example of a Pattern 2 shift, the proposed initiating power of a glideless fronted PRICE and MOUTH makes sense and it also has the advantage of 'taking care' of the fact that SAE displays a 'pure' FLEECE vowel as well as a degree of KIT centralization, both features of, for example, the Northern Cities Shift. It is clear, too, that SAE has, at least in its Broader varieties, a raised TRAP vowel. On the other hand, the general trend of the Northern Cities

Shift is for KIT and DRESS to remain on the non-peripheral path and for the later to become more open too. This certainly doesn't 'square', although for different reasons, with the facts about SAE (which if anything has a substantially raised DRESS vowel) or with Jeffery's (1982) notion of "the front series" moving up.

Indirectly, the slightly later Bailey (1984:11–3), also brings up the possibility of a Pattern 2 chain-shift in SAE. He focusses on /æ/-lengthening in SAE and at least one example of /æ/-breaking, about which he makes the following claims:

> "The breaking of this vowel is a widespread feature of SAE and might be diagnostic of SAE. It is almost certainly part of the same phenomenon that occurs more widely in American dialects and just one of several instances of the influence of velars on vowels. In SAE this breaking occurs in one environment only, i.e. before /ŋ/ . . . as in . . . bang, bank, tanker, languish" (Bailey 1984:13).

The context is in fact even more restricted than this, only occurring in /æ/ before historical 'eng' i.e. in *bang* but not *pancake*. One thus assumes, that by 'widespread' (see above quote), Bailey (1984) means socially and geographically as opposed to systematically. In any event, he gives [e$^{>j}$]{[ë^{i}]} as its realization, as in [be>jŋ]{[bë^{i}ŋ]} *bang*, emphasizing the shortness of this 'diphthong' in line with his previous comments about [ŋ] not being an environment that conditions long allophones of the TRAP vowel. As far as I can tell, such a quality is unattested in any of the other sources on SAE (see §6.3); it also, appears, impressionistically to be a stereotype-feature of a (very) Broad SAE accent, if anything. A certain degree of skepticism is also generated in this regard as a result of Bailey's (1984) confession that phonetics and/or SAE "is not an area of specialization of the writer" (Bailey 1984:1), as well as a number of other uncertain analyses relating to TRAP-lengthening more generally. Thus he claims, for example, that *ham*, *lad*, *pad*, and *wag* "may be long but are most often short" (Bailey 1984:11) and that *badge* is short, whereas in my idiolect these are all invariantly long. Regardless of the particulars, however, Bailey's (1984) focus on TRAP in SAE and its lengthening and possible breaking, highlights the potential merits of investigating Pattern-2 like shifts in SAE, as originally suggested by Jeffery (1982).

Webb (1983), on the other hand, focusses on the 'Southern Shift' nature of SAE (i.e. Patterns 3 and 4). In terms of Pattern 3, he points out the fronted GOOSE and GOAT vowels and the raised and peripheral nature of the THOUGHT and BATH vowels. In terms of Pattern 4, he emphasizes the lowering of FACE, but points out that SAE "does not exhibit the typical positional exchange between /iy/ and /ɪ/ which is typical of Southern U.S.A. dialects. The /iy/ is still higher and further forward". Importantly, this author stresses that the features mentioned should not, necessarily, be viewed as changes-in-progress but, in all likelihood, as relics of the British input into SAE.

3.4.3 SAE as a Southern Hemisphere English

According to Lass (2004:369–40), ETEs can generally be subdivided into Northern Hemisphere (US and Canada) and Southern Hemisphere types (Australasian English and SAE). The different phonetic features of these two main types are traceable back to their periods of colonization; thus, by way of example, and referring back to §3.4.1, Lengthening I went through a number of stages:

1. Prior to the mid-eighteenth century, only lengthening with no change in quality; thus [pæt] *pat* but [pæːθ] *path* in US English;

2. Late eighteenth-century BATH-lowering to [aː]; thus [pæt] *pat* but [paːθ] *path* in Australasian English;

3. Nineteenth century retraction of BATH to [ɑː]; thus [pæt] *pat* but [pɑːθ] *path* in SAE.

Thus SAE is characterisable, as it is by Lass (1990:273), as "a relatively 'evolved' southern variety," (whether Southern British or Southern Hemisphere); a fact attributable to the relatively late settlement of South Africa by the British. This is reflected in its relatively close TRAP (§6.3) and DRESS (§6.4) vowels (like most Southern Hemisphere Englishes), but, more particularly, by its rather back value for BATH (see §8.2), unlike the other Southern Hemisphere Englishes.

Trudgill (2004:63–64) provides a particularly intriguing argument for the presence of his so-called 'START-backing' in SAE. The puzzle is essentially that the absence of such backing in AusE, NZE and Falklands Island English as opposed to its presence in SAE and Tristan da Cunha, does not provide a neat chronological pattern, particularly given that Falklands Island English was in all likelihood established after SAE, at least in the latter's 1820-settlers form. His explanation for its presence in SAE is one that has interesting implications for other features of SAE and deserves to be quoted in full:

> "The absence of START Backing from Falklands Island English, which probably developed latter than South African English, might suggest that it *is* a later and independent innovation in South Africa. However, Lanham (1978:153) indicates that backing may be a feature of South African English of long standing. If he is correct, this may be due to the settlements of the Natal and Transvaal anglophone communities

at dates later than that of New Zealand – the additional decade or three allowing for START Backing to have become more common in urban southern England. It could also be due, if the backed variant had reached the upper-middle class by the appropriate date . . . to the greater role played by this class in the formation of South African English: according to De Klerk (1996:10), the largely vernacular-speaking settlements of the 1820s were followed by the 'largely standard-speaking Natal settlements of the 1840s"' (Trudgill 2004:63).

SAE differs, furthermore, from the two Australasian varieties in that it lacks diphthongisation of FLEECE (see §6.4) or "vocalisation of /l/ in syllable codas" (Lass 1990:273). In other words, in SAE [l] does not become [ʊ] in words such as *fill*: /fɪl/ is not [fɪʊ]. According to Lass (1990), it does, however, in the case of BrSAE, share with Australasian English an open onset for both FACE (§6.4) and GOAT (§8.3) as well as the so-called PRICE / MOUTH crossover (§6.2)[29].

Like NZE and AusE, SAE has, accordingly to Lass (1990:273), been the subject of a "chain–shift of the short front vowels". This has lead to both the raised DRESS and TRAP vowels mentioned above, as well as to the centralisation of KIT (Chapter 7) in certain contexts. As summarised by Mesthrie (1993:30), the front vowels in SAE "are raised in comparison with RP vowels (with the highest vowel being centralised)". In order to illustrate this chain-shift, Mesthrie (1993:30) provides a diagram – see Figure 3.2 – in which "the two sets of reflexes are geometrically related . . . (with SAE in caps, RP in lower case)".

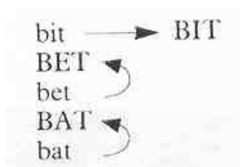


Figure 3.2: The SAE Front Vowel Chain-shift, from Mesthrie (1993:30)

The effect of this chain–shift on SAEP is explained humorously by Mesthrie (1993:30–31) in the following manner:

"Loosely speaking, the SAE pronunciation of *bat* sounds closer to RP *bet*, the SAE pronunciation of *bet* sounds intermediate between RP *bet* and *bit*; and the pronunciation of *bit* sounds intermediate between RP *bit* and the sequence bVt, where V stands

[29] SAE also differs from many other English accents in lacking T-glottaling i.e. the replacement of /t/ in final position with [ʔ] (Lass 2004:378).

> for the vowel in *hook*. There is a grain of truth in the old joke that for some South Africans *sex* is what you carry coals in, while *six* is necessary for procreation".

With respect to the origins of this difference between SAEP and RP, Lass and Wright (1986:202) argue for "a reallocation of features spread out in a 'mixed' input, and recodified into an emergent ETE system". What this means more simply is that close values for both TRAP and DRESS as well as centralized (and close) values for KIT are to be found in the original settler input i.e. they were found on mainland England at the time of settlement and were likely to be well represented in the speech of the settlers. With reference to certain specific vowels, e.g. KIT, Lass and Wright (1985) and Lass and Wright (1986) offer this as an alternative to the hypothesis contained in much of Lanham's work that language-contact between Afrikaans and English (and resultant extensive bilingualism) was the source of many typical SAE values. The approach is primarily, therefore, a exclusively endogenous one with a focus on explaining SAE values in terms of settler input and the structural dynamics of the vowel system in particular. As pointed out by Branford (1994:487), however, "in the case of a variety or variety cluster with as complex a history as that of SAEP, we are unlikely to be able to construct a single monolithic explanation that will fit the observable facts". While accepting the role played by purely endogenous, systemic factors Branford (1994) contends that language contact and "standardisation encouraged by formal education" are two additional factors that need to be taken into account, particularly the former and the related "input from many competent bilinguals".

One last feature that SAE shares with all other Southern Hemisphere Englishes is non–rhoticity. Like RP, AusE and (mostly) NZE, SAE is (except in some BrSAE accents and Afrikaans English) non-rhotic i.e. /r/ "is subject to the severe phonotactic constraint that it can occur only before a vowel" (Wells 1982:125). Note, however, that rhoticity is found in NZE in "the Southland/Otago area, centered in the lay mind at least around the town of Gore" (Bauer and Bauer 2002:170); see also Lass and Wright (1986:204–7) for a discussion of the possible origins of rhoticity in SAE. BrSAE accents are better viewed as *semi*-rhotic since "*variable* rhoticity increases with descent down the socioeconomic scale"(Lass and Wright 1986:205; my emphasis) and there appears to be some evidence to suggest that at least 1820-settler Cape English was rhotic (Trudgill 2004:68)[30]. This has implications for the pronunciation of vowels before a so-called 'historical /r/' and, in particular, for lexical sets such as SQUARE, NURSE, START, NEAR, CURE etc.

[30]The fact that SAE often displays a retroflex [ɻ] in pre-vocalic position rather than the more RP-like postalveolar approximant [ɹ], as well as the use of a tapped [ɾ] in Broader varieties, also speaks to the retention of conservative features in SAE (Trudgill 2004:71–2).

Of particular relevance to the present discussion is Trudgill (2004), which places SAE within the context of Southern Hemisphere Englishes as a whole, as well as within the broader debate about the nature of language change. Many commentators have pointed out the similarities between the various Southern Hemisphere Englishes, and Trudgill (2004) attempts to explain these similarities by making recourse to a number of more general principles.

One of these principles is 'colonial lag' which he defines as "a lag or delay, which lasts for about one generation, in the normal progression and development of linguistic change" (Trudgill 2004:34). Of particular relevance in this regard is this author's discussion of the KIT, DRESS and TRAP vowels and their status in Southern Hemisphere Englishes. Like Lass and Wright (1986), he believes that, in general, the quality of these vowels can be explained by reference to the input brought to the various colonies by the settlers, and that the values generally found in Southern Hemisphere Englishes are actually of a conservative nature. This is, for example, reflective of the fact that on the 'mainland' KIT, TRAP and DRESS have, in many cases, been subject to progressive *lowering* – as, for example, already pointed out in the discussion of Torgersen and Kerswill's (2004) SECS-shift in §2.14. Southern Hemisphere Englishes are thus behind in this process. Mair (2006:165) implies a similar analysis in his description of the progressively lowered values for RP TRAP and DRESS:

> "Phonetically, words such as *bet* and *bat* would now best be rendered as [bɛt] and [bat], while the conservative early twentieth-century pronunciations – now current only in Southern Hemisphere varieties of Australia, New Zealand and South Africa – would have been [bet] (for *bet*) and [bɛt]/[bæt] (for *bat*)".

The notion of colonial lag is, in Trudgill (2004), also applied to other features, such as START-backing (as explained above) as well as the relatively old-fashioned retention of THOUGHT in some CLOTH items, i.e. *off* as [oːf] and not [ɒf]. In general, the Southern Hemisphere Englishes are all seen to retain rather conservative variants of these features, that is in comparison with independent developments in London and the south-east of England. Some of these have already been alluded to above, but all will receive separate and more in-detail treatment in later chapters.

Another principle that Trudgill (2004) draws upon is 'language drift', a process based on language varieties deriving from their common source shared tendencies or propensities for change; and thus tending to develop in parallel even after geographical separation. Naturally, there is an inherent sympathy between this notion and theories such as that of Labov

(1994) which posit an important role for endogenous, structurally-driven, or ‘natural’ language change. Importantly, Trudgill (2004:132–3) posits two forms of drift:

1. a first type (type-1) in which “linguistic changes that are already in progress in the common source continue after separation” (Trudgill 2004:132). Examples relevant to the Southern Hemisphere varieties are Diphthong Shift and loss of rhoticity. With respect to the former it will be shown that SAE has been unusual in its resilience with respect to this development.

2. a second type (type-2) in which “no linguistic change was underway in the common source, but varieties derived from it shared tendencies or propensities which . . . lead to the development of similar but new changes” (Trudgill 2004:132).

Beginning with the first kind of shift, Trudgill (2004) adds, with particular reference to the Southern Hemisphere Englishes, that in many cases these already-in-progress changes advanced at a greater rate in the colonial varieties as a result of less pressure from prestige varieties such as RP. This was particularly the case, Trudgill (2004) claims, with Diphthong Shift. As pointed out, however, SAE (at least in its non-BrSAE variants) constitutes an anomaly in this regard, the lack of advancement of the Diphthong Shift in this variety, compared to AusE and NZE, attributable perhaps to the influence of prestige variants, in the form of NE as well as, more arguably, RP.

The other example of type-1 drift provided by Trudgill (2004) is that of fronted and lowered STRUT. Of all modern-day English short vowels the STRUT vowel “is the result of qualitative changes greater than those which have affected any other” (Cruttenden 2001:114). As confirmed by Trudgill (2004:133), “the STRUT vowel is a recent arrival in the phonological inventory of English, and many local varieties in England and south-western Wales do not have it yet”. Its historical development, beginning with the late 16th-century unrounding of [ʊ] into [ɤ] in the south-east of England, is outlined in Figure 3.3 on the following page, which is taken from Figure 4 in Trudgill (2004:134).

As can be seen, after unrounding the STRUT vowel lowered to [ʌ], which, according to Trudgill (2004), was its value in the 18th-century. By the beginning of the 20th-century, the RP norm was slightly more front, [ʌ̟], and has subsequently fronted even further to [ɐ], “which is the pronunciation in RP today” (Trudgill 2004:133).

Modern-day Cockney approaches [a] (Cruttenden 2001:113). Likewise, Southern Hemisphere Englishes also show open and front realisations:

> “Modern New Zealand English typically has open and front realisations of this vowel. This is true of Australian English also. STRUT in Tristanian is also generally fronter

than RP ...*STRUT fronting is not a prominent feature of South African English* or Falklands Islands English ...” (Trudgill 2004:135–6; my emphasis).

STRUT-fronting in the Southern Hemisphere Englishes is therefore seen as a continuation of a process that began in the 1500s and while Trudgill (2004) is in the main correct in his assessment of this development across the various relevant accents, his assessment of STRUT in SAE appears tenuous. While the principle value of STRUT in SAE is [ɜ] (which while not fully open, is at least in a front-central position), Lass (1990), for example, mentions [ä] as one possible value, while Lass (1995) gives even [ɛ̈] as being used by especially young, female, GenSAE speakers i.e. the subjects of the acoustic analysis component of this research.

Moving on to type-2 instances of drift, Trudgill (2004:136–47) divides these into 19th- and 20th-century innovations. Beginning with the first group we have happY Tensing, Glide Weakening and changes to the NURSE vowel:

happY Tensing: This feature relates to the final syllable of words such as *happy* or *money*, and is dealt with in more detail in §9.3. According to Trudgill (2004), happY Tensing refers to the use of [i $\sim$ iː] instead of [ɪ] in the above-mentioned position and is meant to occur in AusE, NZE, Tristanian English, Falklands Islands English and SAE. With regard to its use on the ‘mainland’, Trudgill (2004:137) claims that it was “very rare in mid–nineteenth century Britain but is currently very rapidly becoming the norm”. This is confirmed in a recent study by Fabricius (2002), the results of which are also discussed further in §9.3.

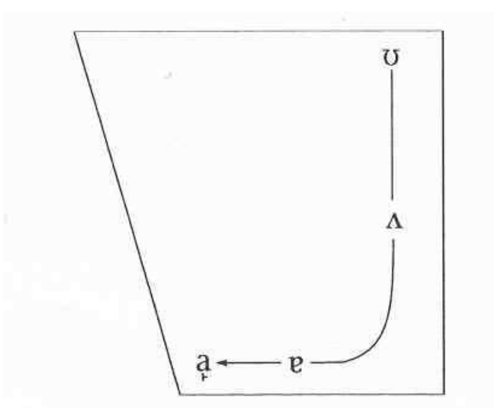


Figure 3.3: Change in the STRUT Vowel, taken from Figure 4 in Trudgill (2004:134))

Glide Weakening: Glide weakening is a secondary effect of a more profound shift of vowel quality i.e. the Diphthong Shift, as a consequence of which the "first elements ... are increasingly removed from their second elements" (Trudgill 2004:139). So, for example, in AusE, the first element of FACE has moved to a more open position than the traditional [e], thus encroaching on the space originally 'inhabited' by PRICE. Initially, the distance between the onset and offset is thus further than was originally the case. Glide-weakening is a reaction against this latter phenomenon i.e. the distance between the two targets is reduced by 'weakening' the glide to the second. This implies either (or both) less distance between the two diphthongal vowel targets ([æe] instead of [æɪ] for AusE), or an increase in the length and prominence of the first vowel target at the expense of the second, leading possibly to monophthongization. Trudgill (2004) claims Glide Weakening as being characteristic of all Southern Hemisphere Englishes, including SAE. The presence of glide-weakening in SAE is, for example, confirmed in Lanham (1967) who provides the following analysis:

> "The average length of the glides ... in SAE diphthongs is noticeably shorter than in RP ..., with the possible exception of /oi/. In more extreme forms ... barely discernable glides usually fluctuate with a simple length distinction" (Lanham 1967:62-3).

This process of diphthong-glide weakening and monophthongization will be investigated in latter chapters on a vowel-by-vowel basis, but it should perhaps be mentioned at this point that *fronted*, glide-weakened variants of PRICE are, on Trudgill's (2004) analysis, unrelated to Glide-Weakening (as a more general phenomenon) for the simple reason that fronted variants of this vowel have, by definition, not undergone Diphthong Shift.

The NURSE vowel: According to Trudgill (2004), the Southern Hemisphere Englishes share a fronted, raised and rounded NURSE vowel, different to the better-known, although less prevalent, RP [ɜː]-value. A feature unique to (some) non-rhotic accents, Trudgill (2004:144) explains its development in terms of the compensatory value lip-rounding plays *vis-à-vis* the loss of rhoticity that has characterized RP and the various Southern-Hemisphere Englishes; in particular, lip-rounding has a similar (although less dramatic) effect to rhoticity with respect to the lowering of formants. The idea is, presumably, that the more conservative RP speakers have resisted these processes of heightening, fronting and rounding, while the Southern Hemisphere Englishes have been given freer reign to develop in this direction[31].

Turning finally to 20th-century innovations, we note the following, in accordance with Trudgill (2004:145–7):

- The so-called Second FORCE Merger which involves the increasing use of /ɔː/ (THOUGHT) for words in the CURE lexical set. According to Trudgill (2004), while prevelant in NZE, AusE, Tristanian English and Falklands Island English, "it appears ... to be absent, so far, from South African English"; see, however, §8.2 for more on this.

- The NEAR-SQUARE merger, which while particularly characteristic of NZE, is also found in Tristanian. There is some indication that NEAR and SQUARE have undergone some degree of merger in SAE ; see §9.2 for more.

While Trudgill (2004) rightly emphasizes the similarities between the various Southern Hemisphere Englishes, the peculiar status of SAE needs to be brought briefly into focus again. Thus, by way of example, an interesting perspective on the relationship between SAE and AusE in particular in provided by Lanham and Macdonald (1979:91; my parenthesis) who claim that "present trends are therefore in the process of reducing the similarities between two speech communities [SAE and AusE] speaking varieties of Commonwealth English originally closer than any other two major varieties". These authors claim that in SAE English, the various Cape English variables (which define BrSAE) are receding. It is these CE variables that generally provide the overlap between SAE more generally and AusE e.g. fronted MOUTH, backed and raised PRICE, lowered FACE and lowered GOAT. NE variables have, however, been advancing through the SAE community, bringing values that are unlike those found in AusE, particularly the fronted, glide-weakened PRICE vowel. We have already noted above the possible influence of northern England and more middle-to-upper class varieties on the development of SAE and GenSAE in particular.

3.5 Conclusion

After dealing with a few terminological issues, this chapter charts the sociohistorical development of SAE as a colonial variety and sets up a schema in which SAE as we know it

[31] Still, this hardly explains, as will be shown in §9.2 and as illustrated in Figure 9.4 on page 381, the generally less fronted (i.e. *lower* F_2) values for RP NURSE; in other words, if the loss of rhoticity implies increased values of F_3 and F_2 one would expect a relatively front vowel for RP after R-Dropping given that it does not employ a compensatory lip-rounding mechanism. This is, of course, not the case, with RP instead displaying a less-fronted NURSE vowel *despite* a lack of lip-rounding.

today is the result of three stages of dialect mixing and koinéization, culminating, in the crucible that was Johannesburg, in the development of three sociolects: BrSAE, CulSAE and GenSAE, with the third of these expanding, more recently, at the expense of the first two. A new perspective on the development of SAE is provided by proposing that Johannesburg constituted a tabula rasa environment (i.e. with a minimal Founder Effect) from which new dialects emerged. From this perspective SAE should be viewed as a *late* and not early 19th-century ETE. There are also broad hints contained in this chapter that, indexicality, in the form of the association of particular linguistic variants with certain identities and prestige (or its lack) is, contra Trudgill (2004), in all likelihood to have been a force in the shaping of the three new sociolects.

The chapter then moves on to provide an account of the Dynamic Model as set out in Schneider (2003; 2007), followed by a critique of Schneider's (2007) application of this model to SAE (broadly conceived). In particular it is suggested that 'white' SAE is on a different course to various other SAE varieties (particularly those associated with the black majority). 'White' SAE is, furthermore, shown to have reached a stage of endonormative orientation, but to be only at the very beginning of a move towards further differentiation.

In the second half of the chapter, SAE as a *linguistic* system is placed in the context of other varieties of English. It is shown to constitute a 'Southern' variety of English (as are most ETEs) and one perhaps characterized by a range of influences, including mid-19th-century north-of-England varieties (via NE) as well as late-19th-century prestige dialects (RP), various other non-standard British dialects from the same period as well as L2-variants such as that employed by L1-Yiddish speakers.

Attention is then placed on the few attempts that have been made to situate SAE within the broader framework of the chain-shift patterns provided by Labov (1994). It is shown that while SAE shows many of the expected Southern Shift characteristics, there is at least tentative evidence for Pattern 2-like shifts in this dialect, as advanced by Jeffery (1982) and Bailey (1984).

SAE is then contextualized as a relatively evolved Southern Hemisphere English as can be seen, for example, from its rather backed BATH vowel, a fact which provides additional support to the notion of modern SAE as a late-19th-century development. Other similarities and differences between SAE and the two Australasian varieties are then briefly dealt with, with a particular focus on the front short-vowel 'chain-shift', and then focus is shifted onto Trudgill's (2004) account of the Southern Hemisphere Englishes. His notions of colonial lag and drift are explained and exemplified, in the case of colonial lag with reference to the above-mentioned short-vowel shift and the nature of BATH, and, in the case of drift, with reference to the Diphthong Shift, STRUT lowering and fronting, happY Tensing, Glide Weakening and the Second FORCE Merger. The chapter ends by re-emphasizing the anomalous nature of *GenSAE* (as opposed to BrSAE) *vis-à-vis* other Southern Hemisphere varieties.

CHAPTER 4

Methodology

4.1 Introduction

This project involved, among other things, the collection and analysis of acoustic data from 27 subjects: all white females between the ages of 18 and 19 from a variety of urban centers in South Africa and from the higher end of the socioeconomic scale. This chapter deals mainly with a range of methodological issues relating to the choice of subjects as well as the collection and analysis of this acoustic data.

As implied above, the selection of subjects for this research was guided by a number of relatively strict controls, most of them relating to a number of well-known sociolinguistic variables. Such control is related directly to objective no.1 of this research, as provided in §1.6, i.e. to isolate and describe, using instrumental techniques, one sub-variety of SAE i.e. GenSAE. Section 4.2 provides a brief defence of the particular value chosen for each of the variables concerned e.g. females over males with respect to gender. In the case of social class (see §4.2.3), a defense of the method employed for controlling this variable is also provided. The next section, §4.3, provides more in-detail information on the research subjects, including information, in §4.3.2, on the regional distribution of the subjects as well as a brief defence of why this variable was not subject to control.

Section 4.4 then turns to describing and defending the actual methods of data collection, while §4.5 focusses on dealing with various issues relating to the analysis of the acoustic data.

Lastly, the acoustic analysis as described in the aforegoing sections, is, in §4.6, placed in the context of the broader methodology of the thesis. In particular, emphasis is placed on the fact that the acoustic analysis constitutes only one component of a broader analysis which involves a comparison of the acoustic data with a variety of other data sources, including other SAE acoustic data from Webb (1983), acoustic and impressionistic data from other accents of English as well as with data derived from a critical analysis of the available impressionistic literature on the SAE vowel system.

4.2 The Variables

It is a well-known fact that *age*, *sex/gender*, *region*, *race/ethnicity* and especially *class* are all sociolinguistic variables; see, for example, Chambers (2003), as well as Labov (1994:2) who also adds *community size* to this list. What needs to be defended, however, is the choice of young, white women from the upper end of the socioeconomic scale i.e. the values of the variables concerned. In the following few sections just such a defense will be provided.

4.2.1 Race/Ethnicity

The focus of this (acoustic) component of the project is on a particular accent of English i.e. GenSAE, which was (and is perhaps still) spoken predominantly by ‘white’ (however defined) individuals in Apartheid-era South Africa. While the growing use of English as a mother tongue by other population groups in South Africa is recognized, the simplest way to guarantee that the accent used was GenSAE, was by restricting the subjects to those who qualified unproblematically, and in the view of the researcher, as ‘white’. This was necessary since while on “the whole middle-class Black English varieties are fairly close to the accents of their [white] schoolmates” (Mesthrie 2008:15; my parenthesis) there is growing evidence for the emergence of a new black (middle-class) ethnolect that constitutes “a young persons’ educated English that goes beyond apartheid restrictions, but avoids many of the variables characteristic of White South African English” (Mesthrie 2008:17). While research into this lect is still in its infancy it appears to be characterised by the adoption of Americanisms such as rhoticity and ‘t-tapping’. There is, of course, the possibility that the direction of influence is in the opposite direction (i.e. white speakers adopting new ‘prestige’ variables from black speakers) and we will touch on some interesting possibilities later on in this thesis.

More generally, and as pointed out by Da Silva (2007:57), “the idea of a *shared norm* as a motivation for variable use becomes problematic when studying South African societies,” especially with respect to English, used as it is by a range of different ethnic groups each, at least potentially, with its own set of indexical norms; and, within each ethnic group, used in different ways and at differing levels of proficiency depending on factors such as educational background, social class etc. One way of ensuring a *relatively* high degree of homogeneity in terms of shared linguistic (and other) norms was thus to control the study for ethnic group. Even in post-1994 South Africa it is an unfortunate truism that much of society is still structured in terms of the ethnic divisions that formed such an integral part of the Apartheid-past. While studies such as Da Silva's (2007) show that these divisions are slowly but surely being eroded, it was felt that the specific requirements of the current study required a form of control in this regard.

Naturally, many L1-Afrikaans-speakers (as well as the few fluent English-Afrikaans bilinguals) are also traditionally included in the 'white' population group. As will be seen in §4.4, one of the aims of the interviews that formed the basis for the data collection was to ensure that the subjects were L1-English. An additional form of control focussed on the exclusion of students with a Jewish background. This was of particular importance with respect to subjects from the Johannesburg area given that "'Johannesburg Jewish private school' is a type that would be recognized by any native" (Lass 2004:373).

4.2.2 Gender

It is perhaps because "women, deprived of access to real power, must claim status through the use of symbols of social membership" (Eckert 1989:265) that they are, all things being equal, more sensitive to prestige norms in language than men, and tend to deviate less from these prestige norms. Thus Lanham (1982:336) claims that "women are more assiduous in acquiring standard speech" and Dubois and Horvath (2000:288) claim that "women are more sensitive than men to the social evaluation of speech and use more of the positively evaluated variants and less of the negatively evaluated variants than do men"; and this, with a few minor exceptions such as provided by Dubois and Horvath (2000:289) and a few other cases reviewed in Labov (1990:212–3), appears to be the rule. This is particularly true of stable linguistic variables i.e. those not undergoing change and "those which are well established as indicators in a community" (Chambers 2003:121). Although women are, thus, from a certain perspective, certainly more conservative in their use of language, they are also, in a seemingly paradoxical manner, equally adept at being at the forefront of linguistic change. Labov (1990) summarizes the sociolinguistic role played by women in the following two principles:

1. "In stable sociolinguistic stratification, men use a higher frequency of nonstan-

dard forms than women" (Labov 1990:205).

2. "In the majority of linguistic changes, women use a higher frequency of the incoming forms than men" (Labov 1990:206).

These two tendencies (i.e. the conservative role of women vs. their role in language change) are less contradictory than they might at first appear; but, in order to be understood, need to be contextualised in terms of the sociolinguistically crucial intersection between gender and social-class as well as in terms of the two different kinds of linguistic change most commonly reported on in the literature.

Beginning with the latter, a distinction is often made in the literature between 'changes-from-above' and 'changes-from-below'. These terms refer to both the level of awareness of the participants in the change (changes-from-below are, for example, beneath the level of consciousness) as well as the social origin of the change. According to Mesthrie et al. (2000:118; my parenthesis) changes-from-below "involve sounds that are originally part of the vernacular ... are given full play in working-class dialects ... [and] ... often arrested or suppressed by the middle class". Changes-from-above, on the contrary, are introduced at the upper end of the social scale.

The issue of class, and its intersection with gender, will be taken up further in §4.2.3. More important for the present focus of discussion, however, are Mesthrie et al.'s (2000) assertions to the effect that changes-from-above are generally driven by females whereas men are more prominently involved in changes-from-below. While it is no doubt true that females are more conservative with regard to stable linguistic variables and are, in addition, with respect to changes-from-above, the main originators and participants in linguistic change, the perspective provided above by Mesthrie et al. (2000:118) on the relationship between gender and change-from-below is in contrast with much of the available literature. So, for example, while there are a few examples of men at the forefront of such changes[1], Labov (1990:215) stresses that "in change from below, women are most often the innovators". Similarly, Cameron (2003:190; my emphasis) claims that "the leaders *of most changes* are women rather than men", i.e. changes-from-above *and* changes-from-below. With respect to the few exceptions to this general principle, Labov (1990:219) adds the following important observation: "the male-dominated changes are all relatively isolated changes ... they do not include chain shifts such as the Southern Shift or the Northern Cities Shift that rotate the sound system as a whole. All those cases of chain shifting that we have been able to examine with quantitative means are dominated by women".

[1] Labov's (1963) famous study of the Martha's Vineyard community being one of them.

For a full understanding of the role of females in initiating and participating in language change, however, the relationship between gender and other sociological variables needs to be fully understood. The first of these (and perhaps the most prominent in the variationist literature) is that of social class.

4.2.3 Social Class

This section begins, in §4.2.3.1, with a (selective) review of the literature dealing with the intersection of class, gender and linguistic variability and change, while §4.2.3.2 considers class in the South African context and its control in the acoustic-analysis component of the thesis.

4.2.3.1 Social Class in the Sociolinguistic Literature

As emphasised by Kerswill (2006:51), different "discussions of class place different emphases on economic factors and more broadly cultural factors". A definition or review of 'social class' is, however, not the object of this section. The more specific aim is to summarise the variationist perspective on social class, to outline its links with gender and other relevant variables, and to situate the current study within this broader sociolinguistic context. Since, on a sociolinguistic level, social class and gender interact in important ways there will be some overlap between this section and the previous one.

With respect to both main forms of change-in-progress (those from above and those from below) and stable sociolinguistic variables, the differential between men and women at the upper end of the socioeconomic scale is small compared to that found in the so-called 'interior classes' i.e. the lower-middle and upper-working classes. Beginning with changes-from-above, we find that it is particularly the hypercorrect behaviour of lower-middle class females that initiates such change (Labov 1972:Ch. 5). This, very broadly-speaking, refers to a phenomenon whereby, in formal contexts, the speech of lower-middle class female speakers 'crosses-over' that of female speakers at the higher end of the socioeconomic spectrum. This is well-illustrated by Figure 4.1, taken from Figure 5 in Labov (1990:224).

As explained by Labov (1990:224–4; my parenthesis), while "systematic style shifting is aligned towards a particular target, usually a level set by a higher status group [here referring to behaviour aligned to a *stable* linguistic variable] . . . "hypercorrect behavior" is . . . a shift of a linguistic variable that passes beyond this target". As can be seen in Figure 4.1, in the most formal context (i.e. word-lists), lower-middle class females show the highest index of use of the (oh) prestige variable. It should be pointed out, however, that upper-middle class females still show a relatively high index of use in this regard.

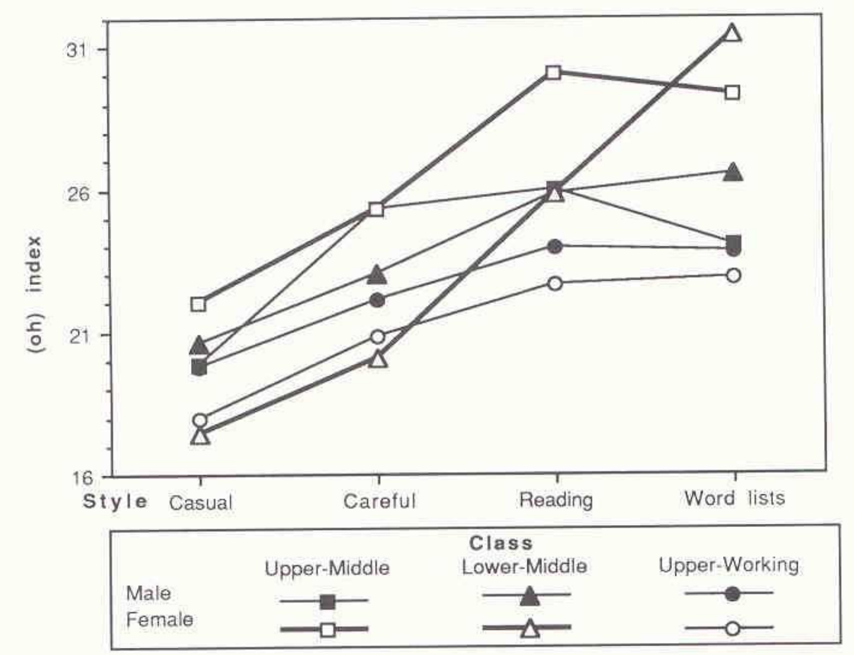


FIGURE 5. Style shifting of (oh) by three socioeconomic groups in New York City.

Figure 4.1: Change from above: the ‘crossover’ pattern

The usual view of change-from-above is that the relevant feature is borrowed from another speech community, introduced by the dominant classes, and often superposed on top of the vernacular. Thus, Labov's (1966) well-known study of rhoticization in New York City in conjunction with the 1986 replica study by Fowler[2] shows how the effect of the use of the prestige final and preconsonantal /r/ "was not to modify the vernacular of most New Yorkers, but rather to produce variable behavior in formal situations" (Labov 1994:90). Yet, as pointed out by Labov (1994:90), the new prestige norm in New York City is effectively being superimposed upon a previous prestige norm that has become part of vernacular speech i.e. lack of rhoticization was clearly a London-influenced prestige norm of the 19th- and early 20th-centuries. Thus changes-from-above *can* become part of the vernacular and there is, perhaps, a link between this process of vernacularization and the hypercorrective behaviour already mentioned above:

> "One suggestion that was advanced in 1966 looked to the hypercorrective behaviour of the lower middle class for this mechanism. Since lower middle class speakers use even more [r] in their most formal style than the upper middle class, they may impress this form upon their children in instructional styles, and so give them the early practice that leads to consistent motor patterns" (Labov 1994:90).

While this hypercorrective behaviour has not yet led to an acceleration of [r]-usage in New York City, it is likely that such a process of change-from-above did lead to the earlier adoption of non-rhoticity as in integral part of the New York vernacular. If we generalize from this one example, it seems clear that changes-from-above can, therefore, become integral parts of the system, a fact emphasized by recent research on dialect leveling e.g. Torgersen and Kerswill (2004). In addition, and with respect to vowels in particular, there seems no reason why, once a change-from-above has become 'settled' into the vernacular that it cannot initiate (or participate in) system-internal vowel movements such as chain-shifts.

With respect to change-from-below, the general consensus, with respect to social-class, is that they "generally arise in centrally located social groups" (Labov 1980:260) or the so-called interior classes, generally the upper-working class and the lower-middle class: It should also be noted that "whenever age distributions and earlier reports indicate that there may be sound change in progress, the highest social class lags behind" (Labov 1980:254). The so-called 'curvilinear pattern' so characteristic of changes-from-below is captured in Fig 4.2 , taken from Figure 7 in Labov (1990:233). As can be seen in Fig 4.2, the 'Skilled

[2] cited in Labov (1994:86–94).

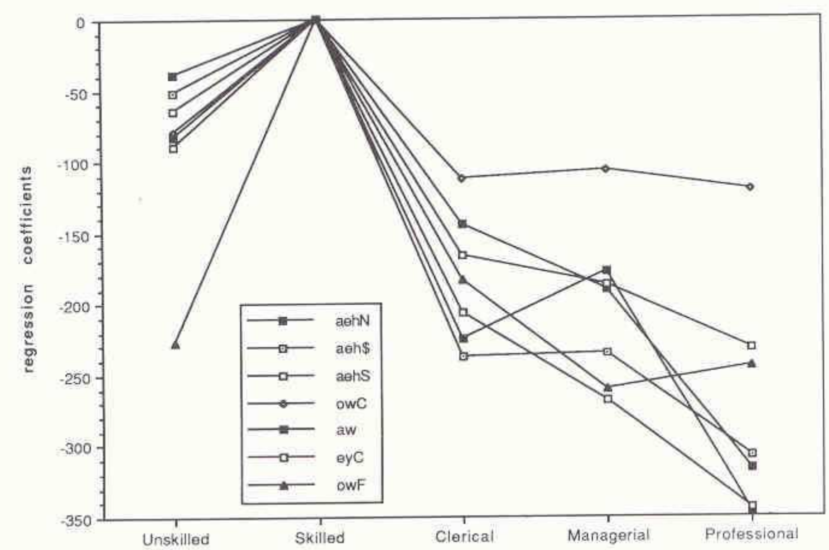


FIGURE 7. The curvilinear pattern in the Philadelphia Neighborhood Study as reflected in occupational coefficients of multiple regression analysis, weighted for the number of tokens for each speaker. *Skilled* occupation is fixed at 0.00 as the reference group.

Figure 4.2: Change from below: the curvilinear pattern

Labour' group (roughly, upper-working class) is ahead in the use of the various changes-in-progress reported. The higher the social class the less the variant is used.

More generally, much recent work in the Labovian mould argues for a distinction between internal vs. external change, the former linked to general (endogenous, structural) principles of language change, the later more a matter of language-contact. Generally, the implication has been that internal change (i.e. change-from-below) is of greater theoretical interest as well as being more pervasive and vigorous. Furthermore, and as discussed above, the general consensus is that such change is usually driven by the so-called "interior social classes" and that the upper classes are generally the last to feel the effects of such change. The question does arise, therefore, whether research focussed on a prestige variety of SAE (i.e. on the speech of the middle- to upper-classes) is at all likely to capture the putative vigorous language changes taking place in South African urban centers – particularly change that might have taken place since the fall of Apartheid. While this perspective is accepted as generally valid, it should be stressed that, firstly, the capturing of recent trends in SAE more generally is to be viewed as a possible positive 'spin-off' of the the more focussed goal of capturing the current status of GenSAE in comparison to other accents of English. Secondly, the role of external change should not, perhaps, be so summarily dismissed. So, for example, dialect leveling (an undoubtedly contact-based phenomenon) is currently playing a major role in the fate of many British (particularly English) dialects. Of relevance in this regard is Labov's (2007) distinction between transmission and diffusion, the first linked to internal change from below and the transmission of language from caregiver to child, the second being linked to external change based on contact mainly between adults. While transmission is seen as the primary mechanism for the development of complex sound change phenomena (e.g. *short a* raising in New York City which involves a complex set of phonological and morphological conditions), diffusion can lead to change on a more superficial (e.g. purely phonetic) level. Thus changes in the phonetic realization of a particular phoneme are just as likely to be the result of change-from-above than change-from-below.

Still, and as will become evident in following chapters, the acoustic component of the current research has been most successful in capturing what are in all likelihood stable sociolinguistic variables which are "aligned with ... class hierarchies in a monotonic fashion. For a prestige marker, the higher a speaker's socioeconomic status, the higher the frequency of use", a famous example of such a stable sociolinguistic variable being that of (ing) in Norwich, as reported on in Trudgill (1974:91–5). The 'capturing' of such stable variation has a clear value, best expressed in the words of Labov (1972:243; my emphasis):

> "*Without a base line of stable sociolinguistic markers, there is no basis for investigating more abstract questions*: the contrast between change and stability; between phonological and grammatical features; between fine and sharp stratification, or between abstract higher-level rules and low-level phonetic patterns; the role of referential function vs. expressive sociolinguistic information; the interaction of sex with social class and ethnic status; the hypercorrect pattern of the second-highest status group; and *many* other *questions all presuppose that we have laid out the basic sociolinguistic orientation of the community with stable sociolinguistic markers*".

The results of the acoustic component of the current research are envisaged as just such a "base line" i.e. as the basis on which further research of the above-mentioned nature can be conducted.

The choice of the 'standard' also makes sense in the context of the existing literature on SAE. There are a number of publications in the literature with an explicit focus on the more standard varieties of SAE. Two examples, both reviewed in later chapters, are Lanham and Traill (1962) and Lass (1990). With perhaps the exception of the very early Hopwood (1928), there are none which focus explicitly on BrSAE. A comparison of the acoustic results of this study with the mainly impressionistic results of earlier studies would also be a likely basis for tracing some changes that have taken place over the last couple of decades. Although the relevant changes would, no doubt, be limited to those which have reached virtual completion (given the focus of this study on the upper end of the socioeconomic scale), and perhaps unlikely to capture current changes in progress in SAE, the results of such a comparison would, I believe, still be of merit in terms of broader debates on the direction of language change in Southern Hemisphere Englishes and the Southern Shift more generally.

Lastly, it should be mentioned that the speech of the upper classes has received scant attention in the sociophonetic literature. So, for example, Fabricius (2002:212), in the introduction to her study on weak vowels in modern RP, is moved to state that "sociolinguistic studies in general remain remarkably silent on the subject of the speech of the upper middle and upper class". Mair (2006:5) echoes this sentiment by claiming that standard varieties of English have, in general, been neglected in the sociolinguistic literature; no doubt mainly as a result of the emphasis Labov has placed, in his various works, on the so-called 'vernacular'.

4.2.3.2 Controlling Social Class

The control of social class for the current research project has been conducted in a relatively intuitive manner. All the subjects come from a private-school background, have begun their tertiary-level education at a relatively prestigious English university and, in the case of the Johannesburg subjects, had, at the time of the recordings, families residing in an area of Johannesburg commonly associated with privilege and wealth. It is highly unlikely, therefore, that the subjects were anything less than middle-class and highly *likely* that on any number of socioeconomic scales, find themselves at the top end.

This relative "casualness" towards and imprecision with respect to social class or socioeconomic status is also based on the overall nature of the English-speaking 'white' community in South Africa and on developments in SAE itself.

According to Lanham (1982:336), and with respect to the SAE community, "three social classes are distinguished, but categorizing is difficult. There are no South African English-speaking manual laborers". Wells (1982:622) echoes this perspective when he claims that "the social-class composition of English-speaking South Africa is quite different from that of other English-speaking societies in that it contains no real manual labouring class"[3].

It is, furthermore, probably still valid, that "taken as a whole ... English-speaking South Africans as a group come from the most culturally advantaged home background of any ethnic or subcultural group in South Africa" (Watts 1976:87). While there has been a shift in political power in South Africa it is undoubtedly true that, on the whole, the white population group retains its prominence on the socioeconomic scale. Thus the following comment from Watts (1976:60–3), which refers to the position of White English-speaking South Africans during the rule of the Afrikaans-centered Nationalist regime, seems, ironically perhaps, an apposite description for the position of these individuals in the New South Africa as well:

> "... whereas until fairly recently the English-speaking whites held a virtual monopoly of economic power, they are now having to share it, and the day is not far distant when not only will they have lost political power ... but they will also find themselves in danger of losing control of important sections of economic power in the country".

It is thus instructive to note, for example, that in 2001 the white population group was still far ahead of the other population groups in terms of individual monthly income.

[3] This author, in fact, bases the absence of an *-ing* variable and *h*-dropping on this fact. In other words, SAE speakers do not omit the /h/ in words such as *here* or *house* while all SAE speakers use the velar nasal [ŋ] and not the alveolar equivalent [n] (*-in*) so prevalent across English accents in words such as *talking* and *walking*; indeed Wells (1982:622) claims that SAE might be unique in having all its speakers consistently using the velar variant.

Thus data gleaned from Statistics South Africa (2006) shows that the white population group, which constituted only 9.6% of the total population in 2001 (Statistics South Africa 2006:2.9), was still ahead *in absolute terms* with respect to the six highest salary brackets[4]. It is unlikely that much has changed over the intervening few years. The same source shows a similar picture with respect to occupation and education.

It should also be noted that CulSAE is a rapidly disappearing variety and unlikely to be found among the young (Lass 2002:110). In an earlier publication, Lanham (1964?:35) implies that as a result of the disappearance of CulSAE, "dialect as a social differential is weakening in South African society". Lanham (1965:90; my parenthesis) provides a similar perspective when claiming with regard to social-class in South Africa that "it is even difficult to correlate points on the scale [i.e. CulSAE to BrSAE] with social differences . . . South African English, even in extreme forms, can be heard at any level in the scale of social class". Thus while GenSAE has, as pointed out in §3.3.1, been seen as advancing at the expense of BrSAE, in order to ensure the capturing of *Gen*SAE it is perhaps therefore still important to control a number of variables relating to gender, social-class and style, as has indeed been done for this research.

As already mentioned at several points above, the notion of social class is to some degree a construct that varies from society to society (and from research design to research design); in this research social class has been controlled in terms of the educational (and regional) factors already mentioned. In fact, Wells (1982:622) echoes this choice when he claims that "a relatively lower social class can nevertheless be distinguished from a higher one within South African English-speaking society on grounds of restricted education and also by place of residence". The more recent Da Silva (2007:49) does "not include class as a social variable but rather educational background in investigating linguistic behaviour" among South African English speakers[5].

A further rationale for a relatively lax approach to the measurement of social class lies in the growing relaxation of the link between standard varieties and the upper social-classes. Mair (2006:201) summarizes this beautifully:

> "Processes of linguistic standardization have a functional and an ideological dimension. Linguistic standardization is functional, inevitable and necessary in large and technologically advanced communities of speakers because it ensures easy commu-

[4]i.e. the difference here is not simply a proportional one; there were more 'white' individuals in the top brackets than 'black' individuals.

[5]Admittedly, given that Da Silva (2007) focusses on the English of both 'black' and 'white' subjects, her range of educational backgrounds is broader than in the current study. I would submit, however, in agreement with Wells (1982), that the general utility of educational background as a rough index of social class still applies equally well to the more traditional 'white' group of SAEP speakers as well.

> nication across large geographical distances, across social classes and different ethnic groups, and across national boundaries. Beyond what is necessary in these purely functional terms, standardization is also ideologically driven . . . Certain pronunciation variants are enforced or stigmatized by dominant elites as social markers, symbolic correlates of membership (or lack thereof) in a dominant group, so that a proper pronunciation becomes an element of proper social conduct and, ultimately, even of the speaker's moral integrity. At the risk of oversimplifying, one could summarize twentieth-century developments by saying that the functional pressures for standardization have strengthened further as a result of the rise of the audiovisual media, whereas the ideological pressures have weakened as a result of the egalitarian, democratic, and to an extent anti-authoritarian, ethos that has come to characterize public discourse in the industrialized Western world in the second half of the twentieth-century".

This relaxation of class-differences is perhaps reflected in the growing interest in the (particularly UK) literature in so-called 'leveling', a process which should be clearly distinguished from standardization given that (modern) speakers "do not automatically abandon their local forms in preference for the standard. Rather, there appears to be a tension between speakers' desire to continue signalling loyalty to the local community by using local speech norms, and a concurrent urge to appear outward-looking or more cosmopolitan" (Foulkes and Docherty 1999:13). The pertinent difference, in the UK at least, is, therefore, between local and non-local, as opposed to non-standard vs. standard. Thus, for example, Estuary English, a leveled variety in its own right, could easily be (and probably is) viewed as a indicator of urbanity and cosmopolitanism. Watt and Milroy (1999), for example, show how in the Tyneside area of north-east England, local accent features are being replaced by a non-local, yet regional variety "localised to the north of England" (Watt and Milroy 1999:43). Thus, by way of example, particularly young, female and middle-class Newcastle speakers tend to adopt unmarked monophthongal variants of the FACE and GOAT vowels i.e. [eː] and [oː] (Watt and Milroy 1999:40) above the local [ɪə] and RP-related [eɪ] in the case of FACE and the local [ʊə ∼ ɵː] and RP-related [əʊ] in the case of GOAT:

> "Although localised variants in Newcastle are vanishing, the dialect as a whole is not moving towards a nationally prestigious homogenised variety. And it is difficult to know in what sense the monophthongs might be described as 'prestige' variants" (Watt and Milroy 1999:42).

I would like to suggest that the SAE-'reflex' of such processes has, in fact, been the social and geographic advancement of GenSAE at the expense of both CulSAE and BrSAE, although certainly more the former than the latter.

An interesting perspective on social-class and its relationship with linguistic variation is highlighted by Da Silva (2007:46–50), who emphasizes the distinction between the functionalist and Marxist approaches to class, the former being based on the notion that "all groups within a society work on a basis of *shared norms*" (Da Silva 2007:46), and providing the basic framework for sociolinguistic classics such as Labov (1966), while the latter focusses rather on class division and linguistic divergence, allowing for a potentially more elegant explanation of the maintenance and even spread of stigmatized linguistic features. The potential relevance of the latter approach to SAE sociolinguistics is an intriguing one, but needs to take into account the above-mentioned claims that class-consciousness has played a comparatively small role in 'white' *English* South African society as well as observations such as Jeffery's (1982), as outlined in §3.3.1, regarding the political-ideological associations of the two main sociolects of SAE (BrSAE and GenSAE) rather than their strict correspondence with class, whether viewed from a functionalist or Marxist perspective. This in turn relates to the fact that "English-speaking South Africans apparently do not perceive themselves as a clearly defined social group *per se*" (Da Silva 2007:73). These and related issues await further elaboration and research. For the purposes of this study is was decided to take an agnostic position with respect to conceptions of class, while operationalizing it in terms of educational background, as already explained above.

4.2.4 Age

The literature seems, as a whole, quite unclear with respect to a link between adolescence and the use of prestigious or non-prestigious variants. On the one hand, Lanham and Macdonald (1979:51) provide the following explanation of the link between the acquisition and use of standard variables:

> "In the age of social perception i.e. middle adolescence to young adulthood, standard variables are acquired, or quantities of advanced variants increase. This constitutes a measurable shift in speech habits most marked in middle class . . . and most advanced in formal speech behaviour".

On the other hand, "many well-established sociolinguistic variables exhibit . . . age-grading . . . adolescents and young adults use stigmatized variants more freely than middle-aged speakers, especially when they are being observed" (Labov 1994:73).

The apparent contradiction here is perhaps resolvable in terms of Labov's (1994) already-mentioned distinction between changes-from-above and changes-from-below. Quite simply, in both instances young speakers, particularly female ones, are the most assiduous in acquiring ‘incoming’ variants. More generally, the following can safely be claimed about the link between adolescence and speech production:

> “The process of attuning phonological behavior to individual social identity is perhaps most transparent during adolescence. It is unlikely to be a coincidence that adolescence, the time at which an individual’s self-definition vis-à-vis the peer group is under intense development, is also the time when individual stylistic variability is most strongly apparent in speech production . . . Just as other aspects of behavior are geared by adolescents to projecting variable identities to particular audiences, evidence suggests that phonological patterning can be attuned to the same ends” (Foulkes and Docherty 2006:431).

It is thus imperative when for example using word-list data, such as is the case in the acoustic-analysis component of this research, to not make generalisations too freely regarding the full stylistic range of the subjects concerned, particularly if these subjects are adolescents or young adults (as is again the case in the current research). This brings us to a discussion of the last sociolinguistic variable that requires attention i.e. style.

4.2.5 Style

Since the perspective of this research is mainly on a *comparative* analysis of SAE, particularly *vis-à-vis* other English accents, and since much of the *available* acoustic data in the various accents of concern is in word-list (or carrier-sentence) style, the elicitation of data from the formal end of the stylistic spectrum was focussed on. In light of Labov and other’s work on change-in-progress, the current research is thus relatively ill-equipped in terms of capturing vigorous changes, if any, current in the various urban centers of South Africa. As will become clear in later chapters, however, the citation-form data collected as part of the current research project is embedded within a critical review of the extant impressionistic literature, and carefully contrasted with Webb’s (1983) earlier *spontaneous-speech* data as well as (mainly) citation-form data from other accents of English, primarily RP, AusE, NZE as well as Torgersen and Kerswill’s (2004) Ashford data, as briefly mentioned in §2.14. As will be seen, this ‘confluence’ of different forms of data and evidence provides a useful perspective on both the synchronic and diachronic status of SAE, particularly with respect to the various ‘shifts’ and patterns described in Chapters 2 and 3. Furthermore, and as argued for briefly in §1.3.2, the results will be seen to provide a useful point-of-departure for further Labovian-type studies of change-in-progress, an issue further embellished upon in Chapter 10.2 where, among other things, suggestions for future research are provided. On a more philosophical note there are also substantially defensible theoretical reasons for at least beginning a thorough analysis of a dialect with a focus on the more ‘controlled’ end of the stylistic scale. As already mentioned in §1.3.2, Labov (1994:63), in fact, advocates the use of ‘conservative’ reference points for the study of language change. More forcefully, and, perhaps, more controversially, Lass and Wright (1985:158; footnote 8) have the following to say on this issue in relation to their use of citation-form data in their study of the SAE short front vowel chain-shift:

> “We also do not apologize for using words in isolation as primary evidence: in terms of ‘psychological reality’, ‘canonicalness’ or whatever, these are surely the primary input to *any* phonological study, synchronic or diachronic . . . ” (Lass and Wright 1985:158; footnote 8).

4.3 The Subjects

As should already be clear from the reading of previous sections of this chapter, all the subjects recorded for this interview were young female ‘white’ adolescents (18–19 years of age) with a private-school background. Subjects were also limited to those who claimed L1-English status, were not of Jewish extraction, and have for all intents and purposes spent their whole lives in the relevant region (i.e. excluding brief holidays overseas). Section 4.3.1 and 4.3.2 provide further relevant information on these subjects.

4.3.1 General Details

All the students were first-year Linguistics students at Rhodes University at the time of recording and the latter took place either during the registration period at the beginning of the academic year or directly after one of the extended breaks in order to avoid any form of accommodation that might take place among students at Rhodes University. That a form of accommodation does take place even where there is no “indigenous” community (such as at Rhodes) should be obvious from the history of SAE itself, where the original input was also diverse and where there was no indigenous community to accommodate to; see Trudgill (1999*b*) for an interesting discussion of this and related phenomena. Thus one possible outcome of a situation such as at Rhodes is the creation of an entirely new accent; it was important, therefore, to control for such a possibility.

Two different groups of subjects were recorded. One set was recorded during the course of 2005 and 2006, mostly after the extended July vacations. Some of this material has already been used in a number of research projects e.g. Bekker and Eley (2007) and Bekker

(2007). The other group of subjects was recorded during the 2007 Registration period. Henceforth the first group of subjects will be referred to as the Old-Group subjects (O-G subjects) and the second group as the New-Group subjects (N-G subjects). Overall, 27 subjects were recorded.

4.3.2 The Regional Provenance of the Subjects

Figure 4.3 provides a map of South Africa and its main urban centers as well as provinces.

Figure 4.3: Map of South Africa copied from www.SA-Venues.com – accessed on 26 July 2007

The regional provenance of the various subjects is provided in Table 4.1 on the next page. As can be seen, the vast majority of subjects come either from Johannesburg (or more specifically the 'Northern Suburbs' of Johannesburg) or KwaZulu-Natal, both traditionally

Region	Number
Johannesburg ‘Northern Suburbs’	12
KwaZulu-Natal	10
Eastern Cape	2
Johannesburg ‘Other’	1
Pretoria	1
Cape Town	1
Total	**27**

Table 4.1: Regional Provenance of the Subjects

English areas in South Africa. Figure 4.4 on the facing page provides a overview of Johannesburg and its main regions. In the popular consciousness, the so-called “Northern Suburbs” includes much of region 3 in Figure 4.4 and parts of 4, 2 and 7. It extends from just north of the inner city (the upper part of region 4) to the border with Midrand (region 2). The various areas and suburbs are generally associated with wealth and prestige.

It should be added here, that no attempt has been made to control region in this research. As stressed in §3.3.3, the evidence for Phase 5 regional etc. diversification, as per the model provided by Schneider (2007), is scant at best and the focus on young, privileged speakers largely precludes the possibility of settlement-based regionalisms occurring in their speech, particularly in word-list style. The use of the aforementioned style also mitigates against the occurance of new regionalisms that might, in fact, differentiate the different urban centers from each other. Since, apart from a few isolated features[6], these have not been noticed impressionistically they are, if in existence at all, probably highly variable and particularly recent. The issue of (modern) regional variation in SAE as a whole requires further research and lies outside the ambit of the current research, which focusses on the ‘broader picture’ i.e. how SAE relates to other accents of English. In comparison to the difference between SAE and, by way of example, NZE, the potential difference between a putative ‘Cape Town English’ and a putative ‘Johannesburg English’ would appear to be insignificant.

[6] Such as the prevalence of a central value for KIT in pre-palatoalveolar contexts in Natal English i.e. fish [fəʃ], the association of fronted, monophthongised PRICE with Johannesburg and a few others that are mentioned in passing in further chapters. Still, it is unclear to what degree even these features have remained solely in their respective regions and not spread more broadly.

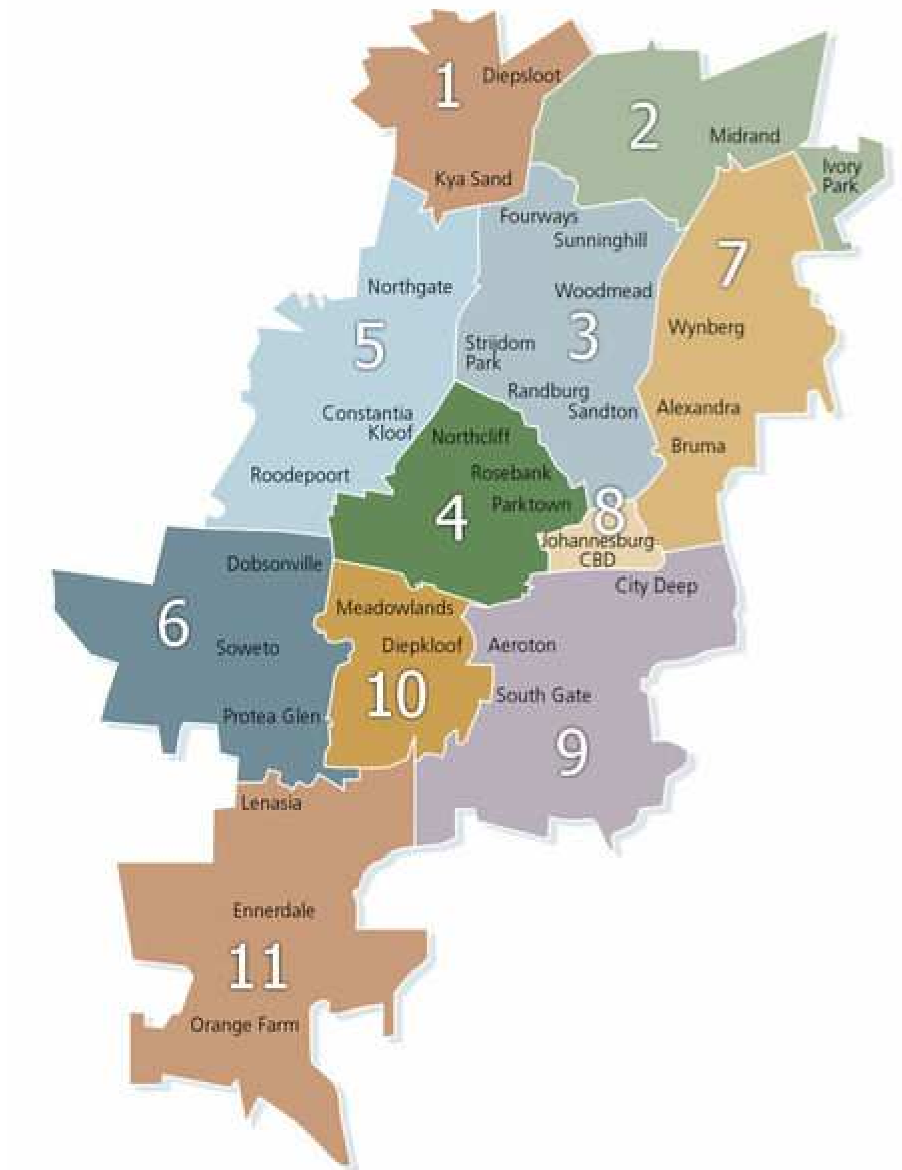


Figure 4.4: Map of Johannesburg copied from the Wikipedia (www.wikipedia.org) entry "Regions of Johannesburg" - accessed on 5 April 2007

4.4 Data Collection: The Recordings

The relevant word-list was read by the subjects in a sound-attentuated room in the English Language and Linguistics Department at Rhodes University, using standard recording equipment and specifications. The word-list used differed across O-G and N-G subjects. In the case of O-G subjects, either the whole or part of the word-list provided by Lass (1990:285) was utilised, a copy of which is included as Appendix B. In the case of N-G subjects a new word-list was constructed, a copy of which is included as Appendix C. A perusal of these appendices will show that different emphasis is placed on different vowels; motivated mostly by the results obtained from research already conducted on the O-G data. So, for example, Bekker and Eley (2007) highlight TRAP and STRUT as particular vowels of interest. Words carrying these vowels are, accordingly, given prominence in the N-G word-list. The word-lists were presented randomly and through a glass panel using cue-cards. This was to avoid so-called list intonation (Watson, Harrington and Evans 1998:189).

4.5 Data Measurement and Analysis

This section focusses on the measurement and analysis of the recorded acoustic data. Section 4.5.1 deals with the first steps of the acoustic analysis, while §4.5.2 focusses on the problem of segmentation given that an attempt was made to capture vowel-internal spectral change (VISC), as defined and defended in §2.2.4.2. Lastly, §4.5.3 deals with issues relating to formant measurement, plotting and normalization.

4.5.1 Preliminaries

The main tool used for measurement and analysis was the open-source computer program `Praat` (Boersma and Weeknink 2007), useful for the analysis, synthesis and manipulation of digital speech signals and, in what follows, a certain degree of familiarity with this computer program on the part of the reader is assumed. One of the main advantages of this program is that it runs so-called scripts, which are essentially 'add-on's' tailor-made for specific purposes. A number of such scripts were written for this project in order to automate the segmentation, analysis and plotting functions as far as possible. In the process, however, certain existing scripts were 'tinkered with' in order to meet the particular requirements of the current research.

Thus, the first script began by calling up another (slightly amended) script called `mark_pauses.praat` (Lennes 2006*a*) in order to automatically generate a `TextGrid` and the necessary interval boundaries for all the words in the recorded word-list. The mentioned program runs "a series of intensity analyses on a LongSound object and mark[s] boundaries at pauses into a new TextGrid object" (Lennes 2006*b*:my parentheses). It thus, indirectly, creates intervals which, in a rough manner, isolate the speech sounds in a sound file. After allowing the user to make any necessary minor adjustments after running `mark_pauses.praat`, the most important of which was to create extra interval boundaries for the few disyllabic words which were included in the word-list, the next step was for the main script to call another (slightly amended) script called `label_from_text_file.praat` (Lennes 2002) which uses an existing text (`.txt`) file (i.e. the word-list) to label the intervals created by `mark_pauses.praat`. Once the various tier intervals are labeled, the main script then prompts the user to refine the segmentation process. This latter process was, in fact, the one that required an almost exclusively 'hands-on' approach. Since the approach adopted to the segmentation of the vowels was a considered one, it has been dealt with separately in the next section. Section 4.5.3 then continues with the exposition of the procedures involved in the measurement and analysis of the data.

4.5.2 Determining Vowel Boundaries

The determination of the boundaries of a vowel is, of course, linked to a determination of its duration. While length in English is commonly viewed as a redundant aspect of the system (i.e. any difference in length between two vowels is accompanied by a difference in quality), there are exceptions, a recent example being that reported on by Labov and Baranowski (2006). It is interesting to note in this regard that one common view of the SAE vowel system is that it often employs length more often for contrastive purposes than other accents. Thus, for example, Lanham and Traill (1962:18) claim that, unlike in RP, "in SAE . . . however, the contrast is partly 'long : short' and partly 'complex : simple"'.

Apart from its possible contrastive value, vowel length can have obvious sociolinguistic value as well. With respect to SAEP, Hopwood (1928), for example, claims a high degree of general vowel shortening for Part-Systems B to D i.e. the 'long' vowels. This vowel shortening thus, presumably, constitutes a sub-phonemic (sociolinguistic) feature of SAE. As such vowel duration will, on occasion, be one of the acoustic features included as part of the more general analysis contained in later chapters. The guidelines provided by Peterson and Lehiste (1960) will be used in this regard to determine the beginning and endpoints of the relevant vowels, with the tacit acknowledgement that the following comments still apply today and are thus relevant to the current research:

> "An essential problem in the measurement of the duration of syllable nuclei is that

of segmentation. Segmentation has long been and continues to be a major problem in speech analysis ... There are many instances in which the cues signalling the beginning and end of a syllable nucleus are relatively unambiguous, but there are many other instances where it is very difficult to specify the point of segmentation ... It should be emphasized that the procedures employed ... sometimes involved a great deal of human judgement"(Peterson and Lehiste 1960:694).

However, given that the focus of the research falls mainly on vowel *quality* (as opposed to duration) it was felt necessary, as explained in §2.2.4.2, to ensure that quality measurements excluded consonant-to-vowel formant transitions, such as those reported on in Cooper, Delattre, Liberman, Borst and Gerstman (1952) and Delattre, Liberman and Cooper (1955), as much as possible; and while it is recognized that formant transitions (particularly that of the F_2), and related concepts such as formant loci (Delattre et al. 1955), are possibly not, in a direct fashion, the main acoustic cues for place of articulation with respect to initial stop consonants (Kewley-Port 1983:324)[7], there is no doubt that F_2 formant transitions constitute the most basic acoustic 'reflex' of the articulatory transition from consonant to vowel position and thus constitute a useful basis for segmentation. As stressed by Sussman, Fruchter and Cable (1995:3112), "the F2 transition is the best known acoustic correlate of the articulatory kinematics underlying obstruent-vowel production". As such, the heuristics provided by Peterson and Lehiste (1960), for example, focussed as they are on determining vowel *duration*, provide too much information from the surrounding consonants. So, by way of example, in the case of initial voiced plosives the measurements were taken from the centre of the release 'spike', "so that the frication period was included in the duration of the vowel" (Peterson and Lehiste 1960).

Segmentation in this research was thus conducted in a flexible manner, guided at all times by the requirement that a balance should be sought between determining the length of the vowel, ensuring that as much of the nucleus be included in the analysis as possible (in order to capture VISC) while at the same time ensuring that the eventual formant measurements were focussed solely on the vowel nucleus proper.

As confirmed by Harrington and Cassidy (1994:360), it is, however, "difficult to separate transitions from vowel targets on a principled basis". As such, it was decided, in the case of Part-System A vowels, to employ arbitrary cut-off points of 25% and 75% of the total vowel duration, the latter determined by implementing Peterson and Lehiste's (1960) heuristics mentioned above: i.e. the central 50% of the vowel was used as a basis for

[7] See also Stevens and Blumstein (1981), Kewley-Port (1983), Lahiri, Gewirth and Blumstein (1984) and Sussman et al. (1991) for some more recent research into the acoustic cues for stop consonant place of articulation perception.

determining vowel quality. This would, in most cases, capture the traditional vowel target (whether defined as a single time point or a period of time) as well as any VISC in the monophthongal vowels. There are certainly disadvantages to this approach. Thus, for example, and according to Lehiste and Peterson (1961:272), labials have a much shorter transition period than linguals. Essentially the tongue is able to pre-empt the vowel sooner in the case of labials. Similarly, "shorter transitions are, in general, more closely associated with voiced initial consonants than with voiceless initial consonants" (Lehiste and Peterson 1961:273). Thus, depending on the relevant consonantal context, the application of a middle-50% cuttoff-mechanism will differ in how much or how little of the nucleus or transition it includes. Still, the advantages of the proposed method outweigh the disadvantages not least because while it, like most other methods, captures the vowel target it also at least provides for the possibility of tracking VISC. The current approach is defended by, for example, Andruski and Nearey (1992:394) in the following manner:

> "While the arbitary … cutoff times may result in less than the best possible approximations to "target frequencies," we believe that a simple criterion involving as little observer judgement as practical is preferable to more complex and highly stimulus-dependent criteria".

In the case of Andruski and Nearey (1992), the cutoff point is an absolute one and not a percentage of total vowel duration i.e. 40ms after release and 40ms prior to closure. These authors are, however, focused on one particular context, bVb, and are therefore in a position to employ such a rigid approach to segmentation. In the present case, however, there are vast differences in the overall duration of the vocalic segments and thus a percentage-based approach seemed more appropriate.

In the case of Part-Systems B to D (i.e. the long monophthongs and diphthongs) it was felt that the middle-50% cuttoff-mechanism would not do. In the case of diphthongs this seemed obvious enough given that VISC is often found to occur outside the middle-half of the vocalic segment, given the dual-target nature of most diphthongs. In terms of their status as tense vowels (see §2.2.2.6), the mentioned mechanism would not do for long monophthongs either given that one of the primary differences between lax (short) and tense (long) monophthongs is the behaviour of the off-glide transition i.e. in the former they are longer (and thus constitute a greater percentage of the total vowel) while in the latter they are relatively brief. In such cases therefore, the initial cuttoff point was placed at 15% of the total vowel duration and the final cuttoff point at 85% of the total vowel duration.

Returning now to the establishment of vowel boundaries, the following framework is based on the guidelines provided by Peterson and Lehiste (1960):

Initial and Final Plosives: i.e. initial and final /p,t,k/ as well as the syllable-initial glottal stop, [ʔ], a characteristic feature of SAEP (Lass 2002:122).

In the case of initial voiceless plosives the point of segmentation was the onset of voicing which "could be determined relatively accurately ... by observing the first formant ... usually clearly distinguishable as the moment in time at which periodic striations started" (Peterson and Lehiste 1960:694). Peterson and Lehiste (1960) in fact provide two measurement points, one from the mentioned onset of voicing, the other from the center of the release spike of the relevant stop. The former was chosen since it was felt that the inclusion of both the periods of post-release frication and aspiration would mean that, particularly in the case of the middle-70% cuttoff-mechanism used for tense vowels, there was the potential of capturing too much information outside of the vowel nucleus i.e. the consonant-to-vowel transitions as well as periods of aspiration.

In the few cases where /p,t,k/ followed an /s/ (e.g. *speck* and *stuck*), and were consequently unaspirated, the guideline used was the same as that for initial voiced plosives – see below.

In the case of initial glottal stops, Cruttenden (2001:169) claims that "there is no acoustic manifestation of the glottal plosive other than the abrupt cessation or onset of the adjacent sound". In the case of pre-vocalic word-initial glottal stops, therefore, the vowel segment boundary was placed at the onset of phonation.

For final /p,t,k/, the end of a vowel prior to a voiceless plosive was mostly based on the cessation of formants, as suggested by Peterson and Lehiste (1960:695). It was soon noted, however, that the automatic formant tracker provided by `Praat` often tracked formants well into what appeared to be the period of articulatory closure. Fig 4.5 on the next page provides an example of this phenomenon.

As can be seen, formant tracking continues after closure (i.e the F_1 minimum), no doubt due to continued voicing during the closure period. As pointed out by Stevens (1998:344), "once closure is achieved ... energy in the vicinity of the first formant can be radiated from the surfaces of the neck and face, but radiation of energy for higher formants is much weaker". This seems to explain perfectly the phenomenon provided for in Figure 4.5. As can be seen from the `TextGrid` underneath the spectrogram, in such cases segmentation was based on what was assumed to be the moment of occlusion, rather than the cessation of formants. As confirmed again by Stevens (1998:334), "immediately prior to closure ... there is a fall in $F1$".

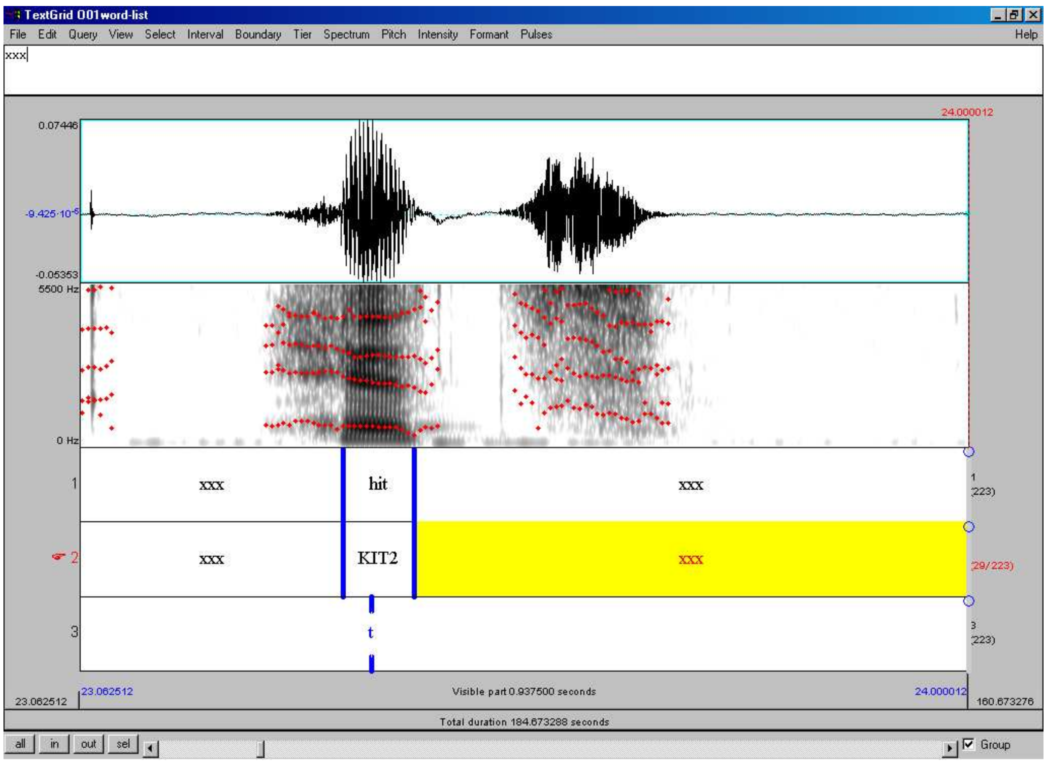


Figure 4.5: Praat edit window of the word *hit*

The segmentation problem was further complicated by the fact that in many cases there were periods of creak towards the end of the vowel as well as the presence prior to the occlusion of what might be considered pre-aspiration. Figure 4.6 on the facing page, a spectrogram of the word *but* as spoken by an O-G subject, gives an example of this phenomenon. In was felt that especially the use of the middle-50% or even the 70% cuttoff-mechanisms mentioned in §4.5.2 would exclude at least some if not most of the aspiration, thus at least partially avoiding any problems that the accompanying 'noise' might cause for formant measurement[8].

In the case of initial *voiced* plosives, i.e. initial /b,d,g/, "the measurements were taken from the center of the spike, so that the frication period was included in the duration of the vowel" (Peterson and Lehiste 1960:695). For final voiced plosives, i.e. final /b,d,g/, the boundary is considered by (Peterson and Lehiste 1960:694–695) to be that point where there was a decrease in energy in the higher harmonics (as determined by using a narrow-band spectrogram). The presence of a F_1 minimum, as in the case of final voiceless stops, was an additional cue drawn upon.

Initial and Final Nasals: i.e. initial /m,n/ and final /m,n,ŋ/.

In the case of initial nasals, the change in formant frequencies was the cue to the velar closure which would, articulatorily-speaking, define the beginning of the vowel i.e. "the abrupt change from steady formant pattern to rapid onglide movement" (Peterson and Lehiste 1960:695).

For final nasals the beginning point of a following nasal was determined by either

[8]As an aside, Foulkes and Docherty (2006:417–9), as well as Docherty and Foulkes (1999:62–6), provide examples of two phenomena, both of which are labelled preaspiration: firstly, what they "term 'extended frication', a period of fricative energy preceding the stop gap, presumably created by a relatively slow tongue tip closure" (Docherty and Foulkes 1999), and showing a discernable energy concentration in the higher frequency range (as is indeed the case with the frication *post*-release in Figure 4.6 on the next page). The second form of preaspiration has energy which is more spread, indicating aspiration rather than frication. According to Harrington and Cassidy (1999:101), "[h] shows spectral characteristics that are typical for vowels: peaks (corresponding to averaged formants) in the 0-4000Hz region and a spectrum that falls off with increasing frequency". The period of 'noise' in question seems to fit this description: the spectrogram in Figure 4.6 shows clear evidence of formants and a spectral slice taken at the center of this period confirmed the decrease in energy from low to high frequency. Both the 'creak'-effect as well as the aspiration is, presumably, explained by "the speaker . . . opening [her] vocal cords at the end of phonation" (Lieberman and Blumstein 1988:59; my parenthesis). Nearey (1997:3243) confirms such preaspiration as a common acoustic cue for final voiceless (as opposed to voiced) stops. It would appear, however, that such pre-aspiration is well within the control of the speaker and is thus potentially employable for indexical use. Thus Foulkes and Docherty (2006) find that in Newcastle speech preaspiration (of both kinds) occurs most often in the speech of females. The fact that preaspiration is potentially under the control of a speaker is further confirmed by comparative data from Derby, in which there is no evidence for preaspiration at all. The possibility that preaspiration might have a similar indexical role to play in SAE is an intriguing one and well worth investigating.

using the above guideline in reverse or using a narrow-band spectrogam, the segmentation point in the latter case being defined as the point during which the harmonics not contained within the formant regions began to loose their intensity; practically-speaking, became lighter on the spectrogram (Peterson and Lehiste 1960:695–6). This guideline was used by Peterson and Lehiste (1960) to specifically overcome the problem of nasalized vowels. Particular care was exercised in 'looking out' for such cases, particularly since the accent of those from the Johannesburg area is often characterized, in the general linguistic 'folk-lore' are least, as having a particularly nasal quality. According to Stevens (1997:484), the most salient acoustic cues for the nazalization of a vowel include an increase in the F_1 bandwidth and the appearance of extra peaks in the spectrum, usually one in the 800 – 1,100 Hz frequency range and occasionally one below the first formant. The values for F_1 are also shifted up or down depending on the vowel in question (Stevens 1998:312). Since Peterson and Lehiste (1960:695) also confess that the relative ease with which they were able to segment the initial nasals might have had much to do with the low number of syllable-initial nasals requiring analysis, the use of the above-mentioned guideline in reverse

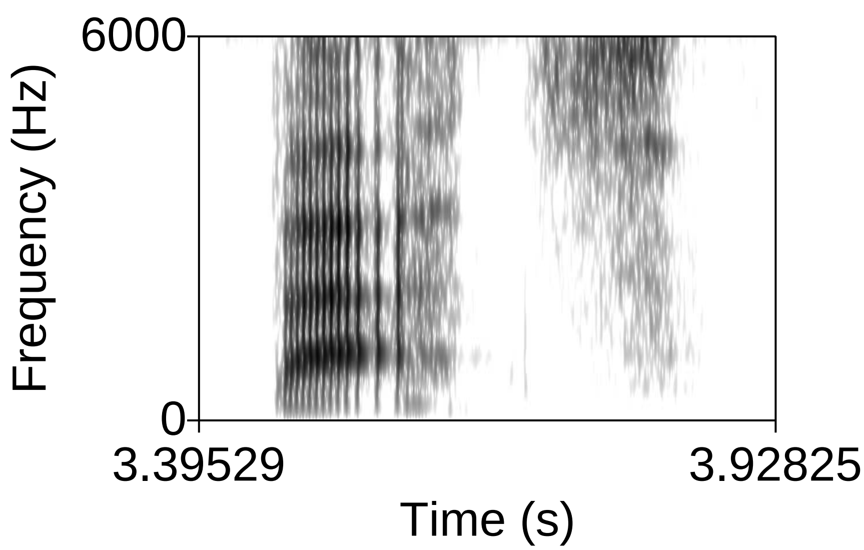


Figure 4.6: Spectrogram of the word *but* showing creak and pre-aspiration

(i.e. the point at which the non-formant harmonics *increase* in energy) was used as a 'back-up' for the segmentation of syllable-initial nasals as well.

Initial and Final Fricatives: i.e. initial and final /θ,f,s,ʃ,ð,v,z,ʒ/. There is, of course, no final-/h/ in English.

The beginning of a vowel after an initial voiceless fricative was identified by Peterson and Lehiste (1960:696) as "the onset of voicing in the region of the first formant" with a minumum in the intensity curve often providing an additional cue. The same guidelines applied to /h/. With respect to initial voiced fricatives "the superimposed noise usually ended abruptly"(Peterson and Lehiste 1960:696–7).

The beginning of final voiceless fricatives were determined by the onset of noise, "even though voicing in a few low harmonics continued for a few centiseconds in most cases" (Peterson and Lehiste 1960:696–7). According to Peterson and Lehiste (1960:697) final voiced fricatives created particular problems. They explain as follows:

> "In broad-band and narrow-band spectrograms, the transition between vowel and consonant appeared rather gradual, but the onset of high-frequency energy in the case of /z/ and /ʒ/ provided a clear boundary on the intensity curves. The boundaries preceding final /v/ and /ð/ were recognized chiefly by the rapid decrease of energy that could usually be detected on the intensity curves".

The intensity contour across the vowel and following voiced fricative was thus used as the main basis on which segmentation took place.

Initial and final affricates: i.e. [ʧ] and [ʤ]. Affricates were given, for segmentation purposes, either fricative or plosive status depending on whether they were initial or final. Thus, by way of example, [ʧ] was considered to be a fricative if initial and a plosive if final.

Initial /w/ and /j/: In the case of an initial /w/ three cues were employed. The first cue related to a sudden increase in the frequency of F_2: "the region in which the slope of the second formant acquired a positive value was considered the boundary" (Peterson and Lehiste 1960:697). The second cue, which often accompanied this change in F_2, but was relied on particularly when the following vowel had a low F_2, was the increase in energy in the harmonics not in the formant regions. An increase in the intensity contour was an additional cue drawn upon when necessary.

For initial /j/, a similar heuristic was used, with the change in harmonic intensity being focussed on when the main cue failed to provide a clear boundary. This main cue relates to the fact that the end of a /j/ ordinarily involves a sharp 'dip' of the F_3 "before rising back to the third-formant position of the vowel" (Peterson and Lehiste 1960:697). The minimum of this 'dip' was considered to be the boundary between the /j/ and the following vowel. In the case of a /j/ following a plosive such as /k/, as in *cute*, the /j/ was generally fricated and thus dealt with in the same manner as fricatives – see above[9].

Initial and final /l/ and initial /r/: Initial /l/ and /r/ provided few problems, the first being identified by the increase in the intensity of non-formant harmonics from consonant to vowel. Initial /r/ was often segmentable from the following vowel on a similar basis, but the "the position at which the third formant begins to rise rapidly in frequency in the broad-band spectrogram" (Peterson and Lehiste 1960:698) was the most obvious cue. In the case of /r/ in post-/t/ or post-/d/ position, the latter was also characterized as containing a degree of frication; the presence of superimposed noise (or more particularly its cessation) was thus an additional cue that could be drawn upon.

In the case of final /l/ things were far trickier. In some cases the intensity curve across both vowel and following consonantal resonant assisted, given that there was often a discernable difference in this respect between the two components. With respect to final-/l/ there were cases where "a frequency minimum or a relatively rapid rise in the frequency of the third formant was sometimes present and was used as the basis of the segmentation" (Peterson and Lehiste 1960:698). In most cases, however, it was necessary to resort to information relating to the pitch of the vowel and following /l/; in particular the point at which the pitch reached its minima and (usually) continued at a relatively steady-state in the /l/ portion of the word was considered to be the boundary between the vowel and the consonant. A spectrogram of such an example is provided in Figure 4.7 on the following page.

As can be seen neither the F_3, nor the intensity contour provide much of a segmentation cue. There is a pitch minimum, however, as pointed out by the arrow.

[9] In this regard, it should also be noted that in SAE "/j/ often strengthens to [ɣ] before a high front vowel: *yield* [ɣɪːɫd]" (Bowerman 2004:940). As pointed out by an examiner of this thesis, however, Bowerman (2004) is probably incorrect here, with a palatal fricative being more likely i.e. [j̊].

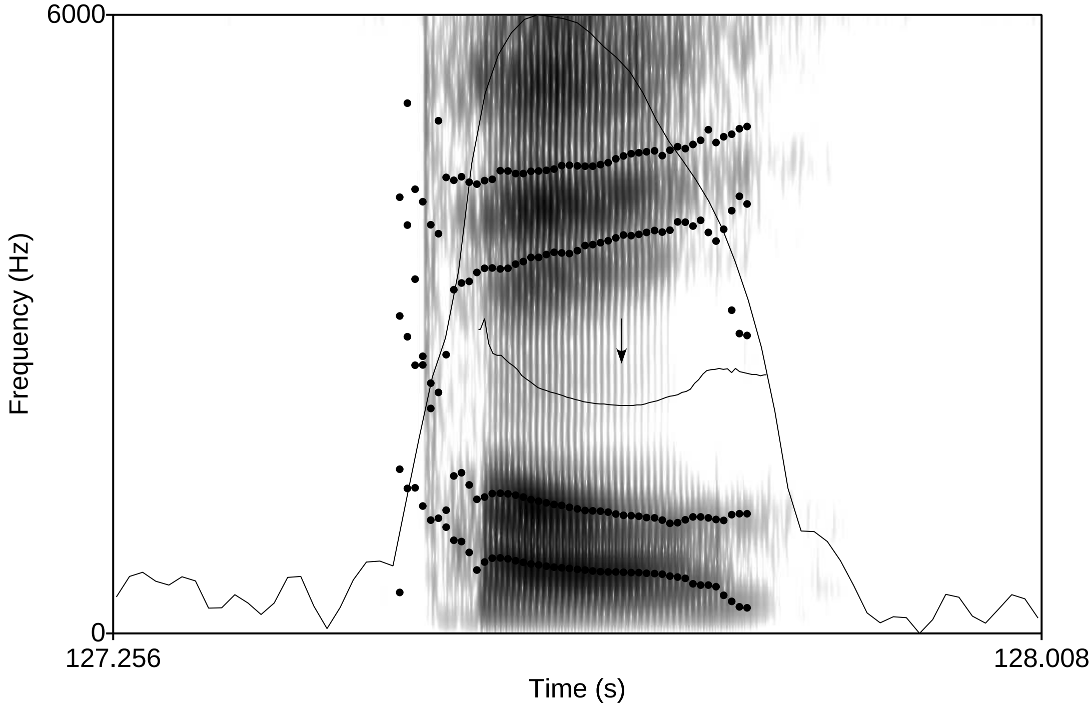


Figure 4.7: Spectrogram of the word *pull*, with formant tracking, as well as amplitude and pitch contours

4.5.3 Measurement and Analysis Continued

Once the segmentation process had been completed, the next step was to tag those items which required 'manual' intervention with respect to formant measurement. These were mostly cases where it was obvious from visual inspection that the automatic formant tracker provided in `Praat` was mistaken. Once the relevant items had been manually tagged the main script was resumed. The first step after the segmentation hiatus was the creation of a duplicate `TextGrid`, on a second tier, the interval boundaries of which were directly matched to those established during segmentation for the original `TextGrid` on the first tier. The script `label_from_text_file.praat` was then called again, this time using a `.txt` file listing the lexical sets (KIT, GOOSE etc.) corresponding to the vowels of the various words in the original word-list. The end-result, therefore, was a properly segmented sound file with both word and lexical-set labels.

The next step was the determination of the duration of each vowel, which the script does by simply subtracting, for each interval, the time of the first boundary from that of the second. This duration is then used, more importantly, to establish the various 'points' of formant measurement i.e. a number of formant measurements were taken for each vowel, regardless of whether it has traditionally been considered monophthongal or diphthongal. It was decided, furthermore, to institute a form of duration normalization by using percentages of duration rather than absolute time-steps (e.g. every 5 milliseconds). This was done primarily in order to ensure that the number of data for each vowel was equivalent. Non-equivalence would have created unnecessary problems for latter analysis and plotting.

The time point of initial measurement (at the 15% cutoff in the case of long monophthongs and diphthongs and at the 25% cutoff in the case of Part-System A vowels) as well as each 10^{th} percentile after this point up to the final point of measurement (again 15% or 25% 'in' from the point of segmentation) was established, providing for 11 measurement points per vowel. The method employed here simply 'circumnavigates' the problem of establishing the vowel target by establishing and plotting vowel quality across a proportion of the duration of the nucleus. There were, of course, no doubt cases where the vowel target, defined, for example, in terms of a F_1 or F_2 minimum or maximum for example, lay between any two of the percentiles used. The degree of error caused by this eventuality, though, is, I believe, small enough to be discounted given the degree of 'resolution' offered by using every 10^{th} percentile; as opposed to, for example, taking quality measurements at the 25^{th}, 50^{th} and 75^{th} percentiles only, in which case valuable information might indeed have been lost.

Once the measurement points had been established, the script then 'asked' `Praat` to conduct a `formant(burg)` analysis on the whole sound file. The relevant algorithm first resamples the sound file to twice the `Maximum formant` setting, which in this case was set at 5500Hz, suitable for adult females. The default values for the rest of the settings were also used i.e. the number of formants was 5, the (Gaussian) window length was 0.025 and pre-emphasis took place from 50Hz i.e. there was no pre-emphasis at 50Hz, but from then on 6dB were added for every octave. The algorithm then conducted an LPC-analysis of the sound signal, deriving the relevant coefficients on the basis of the algorithm originally provided by Burg (1978). As a result, formant values (F_1 to F_5) are extractable at any given time value. However, the script written for this research extracted only the F_1 and F_2 data at each measurement point and along with the time of each measurement point, the duration of the vowel as a whole, the word-name and lexical category of the vowel, stored this information in a `.txt` file that could be opened by `Excel`. One `Excel` spreadsheet per subject was created. Previously it was mentioned that during the process of segmentation, problematic cases had been tagged for later manual checking. It is at this point that these problematic cases were dealt with, the relevant Hertz values in the `Excel` cells being replaced with the manually-derived values.

Once the data for each subject had been extracted and, where necessary, manually revised, a further `Praat` script was run in order to provide the normalized values for each token for the subject in question. The normalization technique utilized in this research was the vowel-extrinsic, formant intrinsic method proposed by Lobanov (1971), a commonly-used normalization technique in sociophonetic research. The formula has already been provided in §2.2.4.1. Following Labov (1994), and as discussed and defended in §2.2.3, the normalization procedure was conducted on raw Hz values.

As will become clearer later, the normalized values were used exclusively for the comparison between the current data (derived from females) and that of Webb (1983) (derived from a male). In the case of the cross-accent comparisons (discussed in §4.6), raw unnormalized Hz values were used however, but only 'female' data from the other accents was employed as a basis of comparison. One of the reasons for employing unnormalized data in the fashion was that according to Thomas and Kendall (2007), for example, one of the weaknesses of the Lobanov method, and other vowel extrinsic techniques, "is that it may be impaired when different dialects or languages that show different vowel systems are compared". The use of raw Hz values from subjects *from the same gender* is exemplified (and defended) in, for example, Torgersen and Kerswill (2004).

Once the data had been subject to normalization it was ready for plotting. Various plotting scripts were written, each with a different purpose. The details of the programming are unimportant but it should be mentioned that all the vowel plots provided in Chapters 6 to

9 are linear, following Labov (1994) again. As is the norm in sociophonetic research, the F_1 values are plotted on the y-axis (with the lower values beginning at the 'top') while the F_2 values are plotted on the x-axis with the lower values beginning on the right. This renders the plots *roughly* interpretable in terms of the traditional IPA vowel chart. On the plots using raw Hz values a F_1=F_2 function is, following Labov (1991; 1994) plotted to provide a sense of the extremities of the vowel space, at least in terms of the back periphery. In the case of all the vowels, both traditionally monophthongal or diphthongal, the script used *all* the F_1 and F_2 values to plot the relevant vowel i.e. a line connects each of the eleven time points, with an arrow being added onto the end-point in order to indicate the direction of the glide (the term 'glide' being used here in the broad sense of VISC).

4.6 The Broader Methodology of the Thesis

The acoustic analysis as described in the sections above, needs to be considered as only one component of the overall research reported on in this thesis. One of the methodological strengths of this research is that it brings together different data sources for comparison and it is this broader comparative analysis that I regard as the 'broader' methodology and the ultimate goal of the research, given that it is only through the analysis of all these data sources together that one can draw reasonable conclusions about the synchronic and diachronic status of the SAE vowel system. Apart from the obvious impressionistic research, therefore, which is, particularly in its SAE-form, subjected to a critical and thorough analysis in Chapters 6 to 9, the research draws on two other sources of *acoustic* data as a basis for comparison: Webb's (1983) spontaneous-speech SAE data (§4.6.1 below) and data from a variety of studies on other accents of English (§4.6.2 below).

4.6.1 Webb's (1983) Acoustic Data

As far as I can ascertain, this thesis is the first to reference Webb (1983). The work's status, or rather lack thereof, is perhaps attributable to the fact that it is based on one interview between the author and another scholar; the author's own idiolect is thus the basis for analysis and description. It, however, forms a useful basis for comparison with the current data not least because, unlike the current data, it is based on spontaneous speech from a male over twenty years ago. Interestingly enough the research was conducted under the direction of William Labov and much emphasis is placed on potential links between the data and Labov's principles and patterns of chain-shifting as outlined in §2.3.2 and §2.3.2.1. More detail regarding this particular study is provided in §5.1.3.

4.6.2 Acoustic Data from Other Accents of English

What follows is a brief description of the acoustic data drawn from research on other relevant accents of English. In all cases female-subject data and, except in two cases, only citation-form data have been used.

Deterding's (1990) RP data: i.e. as reported on in Deterding (1997) and Cruttenden (2001). Citation-form and connected-speech data for all the short-vowels is provided. Both sets of data have been utilized, but, of course, only formant values relating to female subjects. According to Cruttenden (2001:100), the citation-form data was taken from 8 females, while the connected-speech data was taken from 10 females. According to this author, with respect to the citation-form data "one female showed slight London influence [while in the case of the connected-speech data] one speaker showed similar influence from the north of England" (Cruttenden 2001:101; my parenthesis).

Cruttenden's (2001) RP data: Relating solely to (some of the) English diphthongs, in citation form. The subject pool here was limited, with only three females having been recorded, one of which showed "influence from the London region" (Cruttenden 2001:101)

Torgersen and Kerswill's (2004) Ashford data: i.e. focussing on the SECS-Shift (§2.14). Note that these values appear to constitute a combination of word-list and "continuous interview speech" data. Two data sets have been utilized for comparative purposes: the data from older female speakers (in the 70s and 80s) and data from female speakers 14-15 years of age. If, as hypothesized, South Africa shows a 'conservative' implementation of the SECS-Shift due to colonial lag, then the SAE data for the short vowels should be closer to the older speakers and behind the younger speakers.

Butcher's (n.d.) AusE data: More particularly from South Australia (Adelaide). The data was derived from 92 young female students, and was produced in /hVd/ word-list form. "Three tokens of each vowel were read from randomized lists" (Butcher n.d.:450). All the short and long monophthongs were included in the study as well as the upgliding diphthongs (both front and back).

Cox's (2006) AusE data: Focussed on the speech of adolescents from Sydney. 60 males and 60 females were recorded as part of the study, with all words being recorded in /hVd/ context. Only the female data has been 'mined' for present purposes. "No examples of Broad or Cultivated speakers could be found in the subject pool" (Cox 2006:152); thus all were speakers of General 'standard' AusE.

Watson, Harrington and Evans's (1998) AusE data: the formant details are actually provided in Easton and Bauer (2000); the data was taken "from ... isolated word materials ... we used speakers from the 18–30 years range ... there were 21 female speakers ... their accent types were evenly distributed across the AE accent continuum" (Watson, Harrington and Evans 1998:188). These authors only used tokens in /hVd/ or /hV/ contexts (Watson, Harrington and Evans 1998:189).

Watson, Harrington and Evans's (1998) NZE data: formant values are provided in Easton and Bauer (2000). The data relates to 10 female NZE speakers, between 16 and 33 years old. Most (8) subjects had had a tertiary-level education. Although, "most of the speakers were students of the University of Otago" (Watson, Harrington and Evans 1998:188) they originated from all over New Zealand. The data used in Watson, Harrington and Evans (1998) was citation-form.

Easton and Bauer's (2000) NZE data: Based on word-list data. Although a far greater number of speakers were recorded, in this thesis only the data collected from five (5) "middle-class Pakeha women"[10] (Easton and Bauer 2000:94) between the ages of 15 and 19 was utilized, for obvious reasons.

Palethorpe and Cox's (2003) AusE data: focussed on the effect of final /l/ on preceding vowel quality. The data was collected in word-list style, in /hVd/, /hV/ and /hVl/ contexts, from 15 young female speakers. Data for a single target was established in the case of monophthongs and, with a number of exceptions, for two targets in the case of diphthongs. The formant trajectories across both the vowel and the following /l/ are also illustrated.

Flemming and Johnson's (2007) weak-vowel data: focussed on weak-vowel data from non-southern American English accents. The data is based on nine student female subjects.

4.7 Conclusion

This chapter has focussed, in §4.2 through to §4.5, on methodological issues related to the acoustic-analysis component of the research. In §4.2 it provides a defense of the various controls that guided the selection of subjects for the acoustic analysis, focussing respectively, on race/ethnicity, gender, social class, age and style. Particularly in the case of

[10] The term 'Pakeha' refers to New Zealand individuals of European descent (Easton and Bauer 2000:94).

gender (§4.2.2) and social-class (§4.2.3), the variationist literature is drawn upon as part of this defence and in the case of the relationship between variation in SAE and social-class, various sources are drawn upon to create a particularly South African focus on this issue.

The chapter then moves on, in §4.3, to provide additional demographic data concerning the subjects of the acoustic analysis, as well as a brief defence of drawing on subjects from more than one region.

Section 4.5 is, on the other hand, focussed on the measurement and analysis procedures, providing much attention to the issue of segmentation (§4.5.2), but also providing detail on the formant measurement, plotting and normalization procedures utilized.

Lastly, in §4.6, emphasis is placed on the 'broader' methodology of the thesis and, in particular, on the fact that the acoustic analysis forms only one component of a broader comparative analysis which begins, effectively, in the following chapter.

CHAPTER 5

Introducing the SAEP Literature

> "The first step for a well-trained investigator is to search for any previous study that bears on the subject under investigation" (Labov 1994:74).

5.1 Introduction

As mentioned in the conclusion to the previous chapter, the acoustic-analysis component of this research is to be viewed as only one aspect of a broader comparative analysis of SAE vowel quality. As stressed at various points in this thesis, another component is a critical analysis of the existing SAE literature. While much of this analysis is contained in Chapters 6 to 9, on a vowel-by-vowel basis, the various approaches, assumptions and data sources subsumed or drawn upon in said literature requires a relatively brief overview. The various differences between the studies covered in this chapter need to be kept in mind when the vowel-by-vowel analysis is provided in the following four chapters, and where particularly relevant to the discussion these differences will be alluded to again.

The literature on SAEP can be subdivided into a number of reasonably-clear stages. Naturally, these are not completely clear-cut categories and have simply been employed for ease of exposition. They are as follows:

1. An early stage, represented by the work of Hopwood, Hooper and Breckwoldt, and as reviewed in §5.1.1.
2. The work of Lanham and his co-workers; dealt with in §5.1.2.

3. The acoustic analysis provided by Webb (1983), which although limited in various ways, will provide an important basis for comparison in later chapters. This is dealt with in §5.1.3.

4. The work of Lass and colleagues, as reviewed in §5.1.4.

5. Other, generally recent work, focussed on SAE, including Wells (1982), Jeffery (1982), Bailey (1984), Taylor (1991), Mesthrie (1993), Branford (1994), Bowerman (2004) and Da Silva (2007). This is covered in §5.1.5.

5.1.1 Hopwood, Hooper and Breckwoldt

Hopwood (1928; 1952; 1961): the earliest, seriously linguistic attempt to describe SAEP is Hopwood (1928). This document requires careful reading, since the author, in places, conflates both L1 and L2 (specifically Afrikaans English) accents under the rubric 'South African English Pronunciation'. Thus, for example, he claims that SAEP [θɛ̃ŋks] for *thanks* "exhibits a substitution of the nearest sounding Afrikaans speech position ɛ, for an unpronounceable English position æ" (Hopwood 1928:vi). Later, however, with reference to MOUTH, he distinguishes between Afrikaans English (AE) speakers who retract the first element of this diphthong to [ɑ] and other, presumably L1, speakers of SAE who front this element: "definitely a Cockney E. characteristic" (Hopwood 1928:5). In general, the values given for SAEP pronunciation are those which were "common [at the time] to the speech of any five boys in their 'teens, chosen at random from any Primary School in the Union . . . It will make no difference whether their mother tongue is English or Afrikaans, or what their parentage is, provided they learned to speak English in South Africa" (Hopwood 1928:7; my parenthesis). The values provided by this author for SAEP are, consequently, fairly Broad ones. The same, in fact, applies to a much latter work by the same author: Hopwood (1961) mentions, for SAEP, the devoicing of word-final consonants (*cup* for *cub*), epenthetic schwa in *film* ([fələm]) and 'ees' for *is*, all of which are representative of extremely Broad if not AE, while the values he provides for 'Standard English' strike one as rather conservative. This somewhat conservative (and prescriptivist) approach is also evident in Hopwood (1952), in which the author produces such gems as "the vowel-sounds represented by the dictionary spellings 'meelk' and 'mulk' are *un-English* pronunciations of the word 'milk"' (Hopwood 1952:8; my emphasis). From this article it is apparent that Professor Hopwood received his M.A. from Oxford, a fact which *perhaps* explains some of his transcriptions for 'Standard

English'. It should also be noted that Hopwood's (1928) symbols differ in certain important respects from those of the modern IPA[1]:

1. The [ɑ] symbol is described as low and middle (Hopwood 1928:89), whereas [ɑ] in modern terms would be a fully-back vowel. Hopwood's (1928) **ɑ** would thus be best translated into the modern [ɐ̝].
2. 'Standard English' FLEECE and GOOSE are narrowly transcribed as ˫**ìˑj** and ˧**ùˑw** respectively. **j** and **w** relate to the "second element [of] the vowel" (Hopwood 1928:89; my parenthesis). The grave accent is used to "denote slackness in a vowel" (Hopwood 1928:84). The more modern, although somewhat broader, transcriptions [iː] and [uː] have generally been settled on. Where necessary, the distinction between Hopwood's (1928) **ì** and **i** will be maintained by contrasting [ɪ̆] with [i] with the former indicating a slightly less than cardinal position.
3. Hopwood (1928) uses a turned-*i* symbol (i.e. upside down *i*) to refer to the sound occurring in Afrikaans words such as *wit*. As he admits, "it is not found among the symbols of the Phonetic Association. It is really ˔˧**ë́**" (Hopwood 1928:86; footnote 3). While this has some relevance to the discussion of the SAE KIT vowel in Chapter 7, no attempt will be made to provide a modern equivalent. Needless to say it is a quality that is somewhat raised and fronted from a centralised [e] position.

Hooper (1945): this is the first study to attempt to "establish more accurately, by statistical methods, the actual spread of variant pronunciations of English in South Africa". The author provides correspondences between a number of linguistic and sociolinguistic variables. The former include MOUTH, CHOICE, FACE, BATH and KIT, while the latter include gender, age and (broadly defined) class. The study was limited to Transvaal[2] English-medium high-schools.

Hooper (1951; 1952): Hooper (1951) is focussed on the broader socio-political role of English in South Africa, particularly *vis-à-vis* Afrikaans. The author advocates a local standard and emphasizes "the strong and increasing influence of Afrikaans on the English of those born and brought up in South Africa, not only in vocabulary, but in

[1]See Figure 5.1 on the next page for a copy of Hopwood's (1928) vowel chart.

[2]The term 'Transvaal' refers to a pre-1994 province and area which would include the current Northern Province, North-West Province, Mpumalanga as well as Gauteng – see Figure 4.3 on page 121 for a map of South Africa.

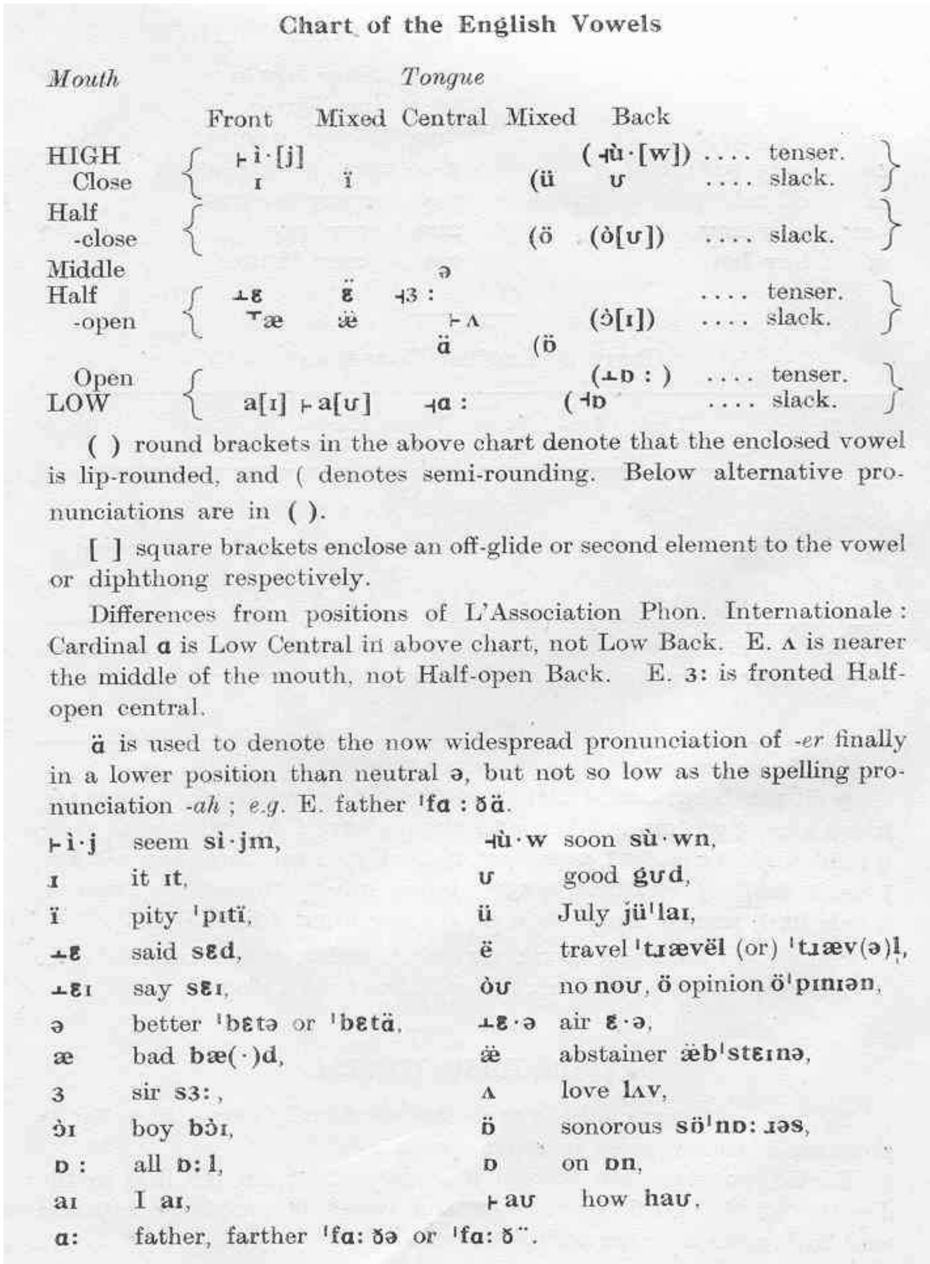


Chart of the English Vowels

Mouth			*Tongue*				
	Front	Mixed	Central	Mixed	Back		
HIGH Close	⊦i·[j] ɪ	ï		(ü	(⊣ù·[w]) ʊ	.... tenser. .... slack.	
Half -close				(ö	(ò[ʊ])	.... slack.	
Middle Half -open	⊥ɛ ⊤æ	ɛ̈ æ̈	ə ⊣3 : ⊦ʌ ä	(ŏ	(ɔ̀[ɪ])	.... tenser. .... slack.	
Open LOW	a[ɪ] ⊦a[ʊ]		⊣ɑ :		(⊥ɒ :) (⊣ɒ	.... tenser. .... slack.	

( ) round brackets in the above chart denote that the enclosed vowel is lip-rounded, and (denotes semi-rounding. Below alternative pronunciations are in **( )**.

[ ] square brackets enclose an off-glide or second element to the vowel or diphthong respectively.

Differences from positions of L'Association Phon. Internationale : Cardinal **ɑ** is Low Central in above chart, not Low Back. E. **ʌ** is nearer the middle of the mouth, not Half-open Back. E. **3:** is fronted Half-open central.

ä is used to denote the now widespread pronunciation of *-er* finally in a lower position than neutral **ə**, but not so low as the spelling pronunciation *-ah* ; *e.g.* E. father **ˈfɑ : ðä**.

⊦i·j	seem **si·jm**,	⊣ù·w	soon **sù·wn**,
ɪ	it **ɪt**,	ʊ	good **gʊd**,
ï	pity **ˈpɪtï**,	ü	July **jüˈlaɪ**,
⊥ɛ	said **sɛd**,	ë	travel **ˈtɹævël** (or) **ˈtɹæv(ə)l̩**,
⊥ɛɪ	say **sɛɪ**,	òʊ	no **noʊ**, ö opinion **öˈpɪnɪən**,
ə	better **ˈbɛtə** or **ˈbɛtä**,	⊥ɛ·ə	air **ɛ·ə**,
æ	bad **bæ(·)d**,	æ̈	abstainer **æ̈bˈstɛɪnə**,
3	sir **s3:**,	ʌ	love **lʌv**,
ɔ̀ɪ	boy **bɔ̀ɪ**,	ŏ	sonorous **sŏˈnɒ: ɹəs**,
ɒ :	all **ɒ: l**,	ɒ	on **ɒn**,
aɪ	I **aɪ**,	⊦aʊ	how **haʊ**,
ɑ:	father, farther **ˈfɑ: ðə** or **ˈfɑ: ð¨**.		

Figure 5.1: Hopwood's 'Chart of the English Vowels'

idiom and pronunciation" (Hooper 1951:82). A similar tolerance for SAEP *vis-à-vis* an English standard is displayed in Hooper (1952).

Breckwoldt (1961): this author is completely sold on the Afrikaans influence on SAE. He summarizes by claiming that "South African English vowels . . . show the following characteristics:

1. Diversification (through contact with the Afrikaans language);
2. "Positional raising" (Breckwoldt 1961:12)

Some of the values he provides are extremely surprising (and unlikely); even in the most extreme SAE. So, for example, "ʌmbrɛlə > œmbərɛlə" (Breckwoldt 1961:7) for *umbrella*. The exact source of Breckwoldt's (1961) data is unclear.

5.1.2 Lanham, Traill and MacDonald

Lanham and Traill (1962): these authors are focussed on providing a phonetic description of what they term 'South African Received Pronunciation (SARP)' which they define as "'educated English' – the 'prestige dialect' as far as the English community in South Africa is concerned". The authors' description of their subjects is also pertinent:

> "The SARP group was identified according to occupation (professional men and women, managers etc.) and *education ('private' schools and/or university)*. All were South African born mother–tongue speakers of English from unilingual homes. Children of SARP parents down to the age of 18 were included" (Lanham and Traill 1962:6; my emphasis).

Interestingly, the authors find it necessary, based on the patterns of pronunciation of their subjects, to divide SARP into two systems: SARP 'A' and SARP 'B' (Lanham and Traill 1962:7), the first system closely approximating the RP of the time, with the second system showing more distinct features of SAEP. Important is their observation that "the great majority of younger people are distributed over the SARP 'B' end of the spectrum . . . [and] many young people investigated [were] clearly outside the SARP group" (Lanham and Traill 1962:7; my parenthesis). Generally, however, few women fell outside the SARP group. Lanham and Traill (1962:37) also claim that "the more exclusive private schools whose pupils a generation ago were identified by their near-British-English pronunciations, to-day present overwhelmingly SAE patterns of pronunciation”.

Loosely–speaking, Lanham and Traill’s (1962) SARP ‘A’ variety is equivalent to CulSAE, while SARP ‘B’ is GenSAE (GenSAE of the time, of course).

Lanham and Traill (1962:8) provide their own vowel “grid”, a copy of which is provided in Figure 5.2. Some of the phonetic values utilized in Lanham and Traill (1962) are of such a precise nature that the modern IPA symbol-set (International Phonetic Association 1999) is not quite suited for their representation. In the following chapters, the closet IPA equivalents have in all cases been provided, as well as the original symbols offered by Lanham and Traill (1962).

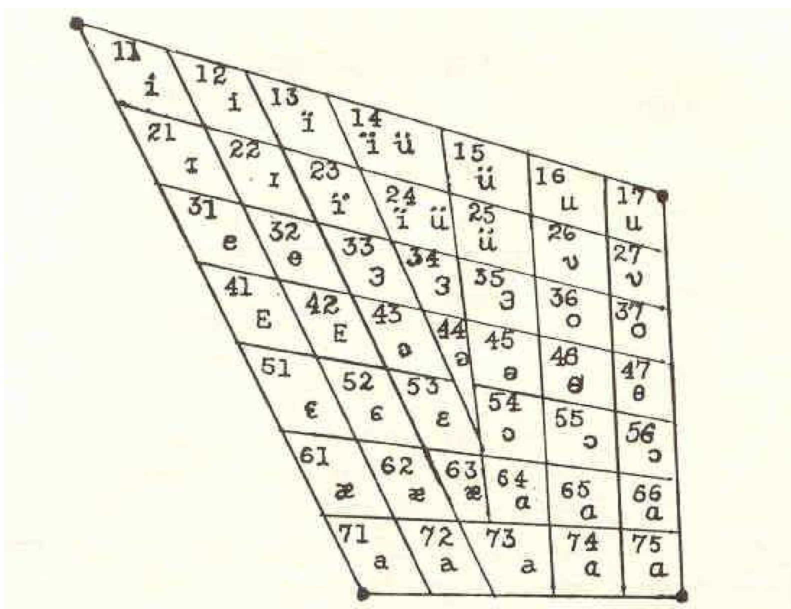


Figure 5.2: Lanham and Traill’s (1962) Vowel Grid

The following explanation from Lanham and Traill (1962:8–9) is relevant:

> “Seven tongue heights and seven points along the front–to–back scale are recognized. These intersect to give 45 points of vowel articulation . . . An even finer grid is necessary at times, however, and [ɛ]$^{51/61}$ and [ɛ]$^{51/62}$, for example, are points of articulation respectively somewhat lower, and lower and more central than ‘Cardinal’ [ɛ]. In such cases the highest, nearest phonetic symbol and the appropriate number are used for identification. The dimensions ‘tense : lax’ and ‘closely rounded : unrounded‘ are unrepresented on the chart . . . .

Typographical limitations prevent a distinction being made between central [ɪ] and central [i] and both are represented by [ï]. Similarly central [ʊ] and central [u] are both represented by [ü], and [*a*] represents vowels at two different levels in the low back position".

Lanham (1964?): this work is a relatively light excursus into the history and social role of SAE. On a phonetic level it deals with some of the most obvious hallmarks of SAEP e.g. KIT centralization, glide-weakening, a raised DRESS vowel and vowel-retraction before /l/.

Lanham (1965): this work is focussed on providing a prediction of change in SAEP i.e. how SAEP will 'look like' in 2065 A.D. Thus Lanham (1965:91) provides the following general predictions:

1. "Loss of diphthongs and the rise of long vowels . . . length . . . appears to be extending its function at the expense of diphthongs" (Lanham 1965:91). Thus SAE is meant to be undergoing an increasing monophthongization of its (historical) diphthongs and an increasing reliance on length alone. The vowels focussed on by Lanham (1965) are the PRICE, GOAT, NEAR, CURE, SQUARE, THOUGHT, BATH and (to a lesser extent) the MOUTH vowel. The only diphthongs that appear to be retaining their status are FACE and CHOICE.
2. The continued influence of 'final'-/l/ on preceding vowels in the sense of lowering and retracting these vowels as well as neutralizing them in certain cases.
3. The reduction of "the 7 short simple stressed vowels of S.A.E. to 6" (Lanham 1965:93). In other words the values of KIT will be in perfect complementation – see Chapter 7 for more on this matter.
4. The continued raising of the KIN vowel (or, alternatively put, the fronted, raised allophone of the KIT vowel), DRESS and TRAP vowels: "present pressures are certainly in the direction of raising all 3 front vowels".

Lanham (1967): this is very much an extension of Lanham and Traill (1962). i.e. while Lanham and Traill (1962) focus on SARP, Lanham (1967) extends the description to SAE more generally. Again, the author provides his own vowel grid, provided in Figure 5.3 on the following page.

The above grid is identical to that of Lanham and Traill (1962) except with respect to blocks 23, 24 and 25 in which the symbols [ï] and [ü] have been replaced with the more commonsensical [ɪ̈] and [ʊ̈]. Degree or lack of rounding is not indicated on the above chart and, unless otherwise indicated, front vowels are assumed to be unrounded and back vowels rounded. The only extra symbol used in Lanham (1967) is [ʌ], the unrounded version of [ɔ] (Lanham 1967:4). Lanham (1967:8) also provides the following values for Daniel Jones' primary cardinal vowels:

1. [i]^{11}
2. [e]^{31}
3. [ɛ]^{51}
4. [a]^{71}
5. [ɑ]^{66}
6. [ɔ]^{56}
7. [o]^{37}
8. [u]^{17}

All of the above is unsurprising bar cardinal no. 5 which has a less-than-fully peripheral value. According to Lanham (1967:113), "a lower position than that given on

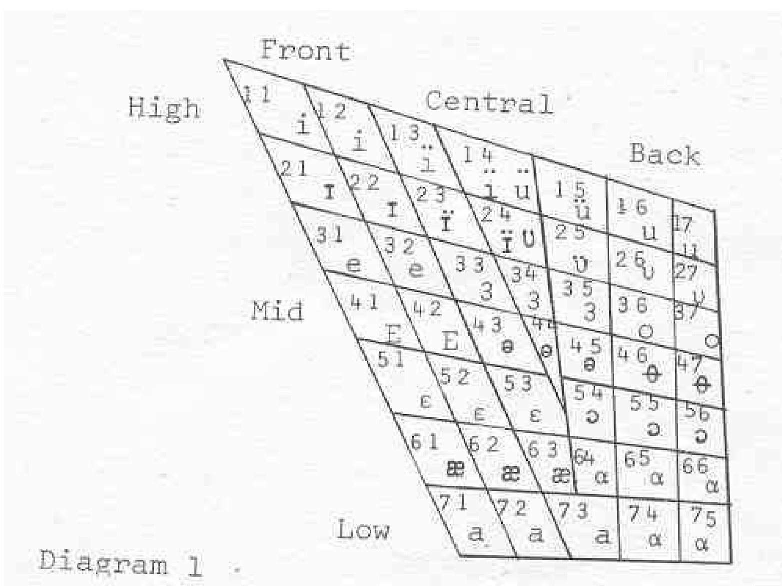


Figure 5.3: Lanham's (1967) Vowel Grid

the record is possible i.e. [ɑ][75] is a more open vocoid than that produced by Daniel Jones".

An important feature of this text is the inclusion of a phonetic description of the RP of the time which "differs at a limited number of points from standard descriptions such as those of Jones and Ward because (a) it refers to present-day pronunciation, and (b) the possibility that certain allophones as described by them were not in fact the prevailing norms at the time" (Lanham 1967:114). An important component of this description of RP is that of the KIT vowel, from which it becomes clear that the so-called KIN-PIN split (as a *phonemic* split) is seen, by this author, as having its basis in RP, and that, in fact, SAEP is characterized by the devolution of what is phonemic in RP into what it purely allophonic in SAEP. It is also important to note, that in Lanham (1967), the author generally divides SAEP into extreme and conservative varieties. The later would cover what we would now call CulSAE while the former would relate to both GenSAE and BrSAE.

Lanham (1978), Lanham and Macdonald (1979): Lanham (1978) constitutes, in essence, a precursor to (and, barring a few details, is 'contained in') Lanham and Macdonald (1979). The descriptions of the latter work are of particular relevance to this research given that their analysis is focussed particularly on formal speech (Lanham and Macdonald 1979:30) and thus standard varieties of SAE. One important difference, however, is that the authors exclude "the reading of word-lists from the analysis of style-shifting" (Lanham and Macdonald 1979:33). More generally, these authors are focussed on making connections between variation in SAEP and the broader social history of South Africa. As highlighted, however, in Jeffery (1982:251; my parenthesis), while Lanham and Macdonald (1979) are often focussed on speculating about the British or Afrikaans origins of certain SAE features, a weakness of the study is a lack of consideration of sound-change, particularly those of an endogenous systemic nature: "it is only [Labov's] synchronic work in New York City that L&M draw on, in spite of their historical stance. What they have done in fact is to blend social history with sociolinguistics in the manner of Labov's work on Martha's Vineyard, though without his particular interest in sound-change".

Lanham (1982) constitutes a general overview of the sound-system of SAE (conceived broadly[3]).

[3] i.e. as including varieties other than 'white' SAE.

5.1.3 Webb

Webb (1983) is the first and only other attempt to apply acoustic techniques to the analysis of the SAE vowel system. The methodology employed by Webb (1983) has one advantage in that it focusses on relatively spontaneous speech. On the downside it analyzes the speech of only one subject (the author) and only provides data for 249 tokens, "generally about 10 tokens per class" (Webb 1983:136). Where relevant, vowel allophony is taken into account. The author, in a footnote, admits to being Afrikaans-speaking (Webb 1983:161; footnote 8), although it is unclear what effect on the English spoken by the author this is meant to imply. Earlier on, though, the author mentions that "Natal English . . . is the variety of English acquired by the author" (Webb 1983:154)[4].

In Figure 5.4 on the next page, Webb's (1983) short vowels have been plotted, while in Figure 5.5 on page 152 the long monophthongs and diphthongs have been provided[5]. In order to aid later comparison with the data collected as part of this research, the values have been subjected to Lobanov normalization (see §2.2.4.1). Normalization was particularly necessary given that, as a male, Webb's vocal-tract would have, in all likelihood, been larger on average than those of the present subjects. Another adaptation of the data involves Webb's (1983) procedure for determining allophony in the data i.e. in Webb (1983) tokens are grouped in terms of phonetic similarity based on formant values rather than in terms of the expected allophony as derived from the existing literature. Thus, unlike the case in the present research, tokens were not assigned on an *a priori* basis to allophonic groups, but were essentially allowed to 'assign themselves'. Thus, by way of example, the *give* token was 'assigned' to the centralized KIT allophone based on its formant values (F_1 = 453Hz; F_2 = 1584Hz), not on the basis of it having a velar context. Both procedures appear equally valid. The procedure followed in the current research is focussed on testing whether the allophony proposed by the existing literature is confirmed by the acoustic data, whereas the procedure employed by Webb (1983) seeks to determine allophony inductively i.e. from the 'bottom-up'. The latter method does, however, run the risk of incorrectly assigning a token to an allophonic grouping on the basis of incorrect formant data. Formant measurement is notoriously prone to error in measurement. The advantage of *a priori* assignment is that it at least alerts the researcher to possible sources of error e.g. if only one token fails to fit the overall picture then the possibility of measurement error with respect to this token needs

[4]Webb (2008) provides the following description: "my linguistic biography is: am Afrikaans-speaking, however, went to English-schools from day 1 (in Pietermaritzburg) and to university (also in Pmburg). Taught linguistics in English from 1975 to the present. So I am Natalian, thus, I suppose, my accent."

[5]FLEECE before final /l/ does not appear in this graph because it is 'hidden behind' the unmarked allophone of this vowel.

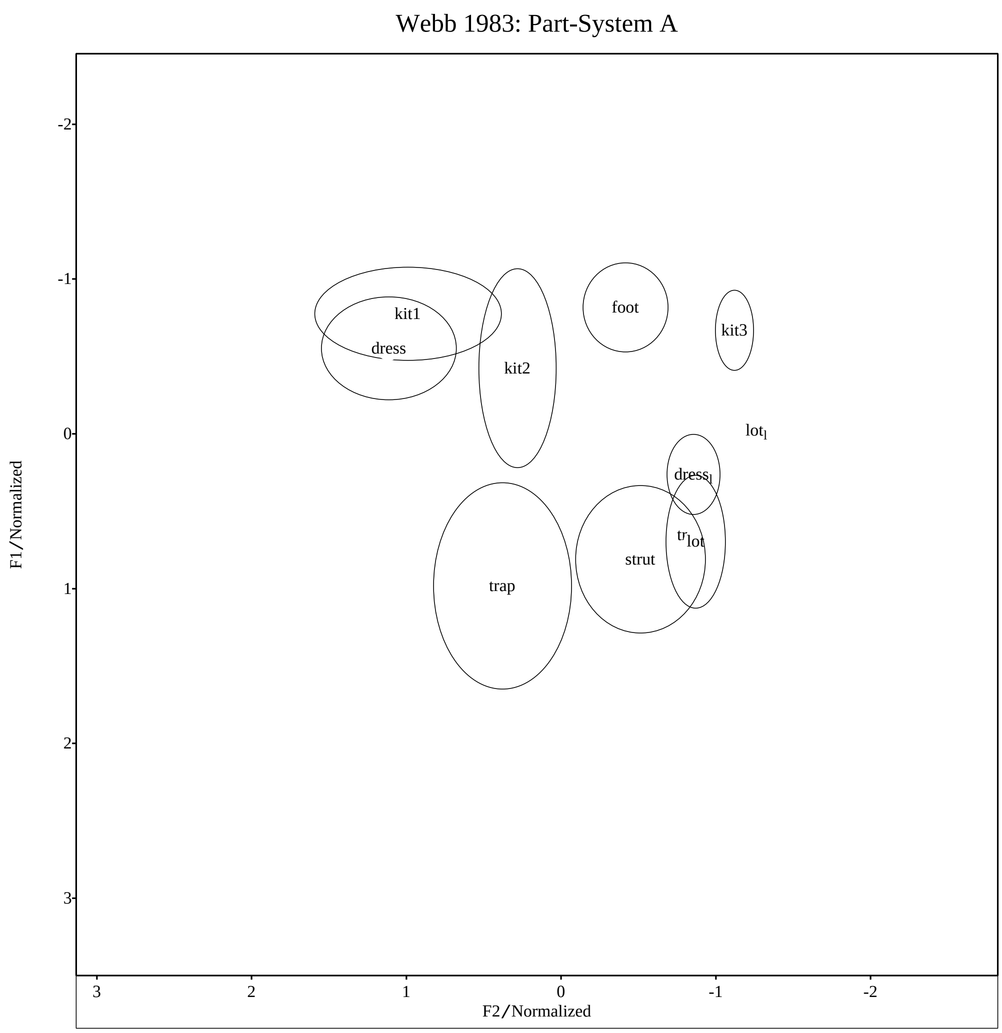


Figure 5.4: The short, lax vowels of Webb (1983); Part-System A

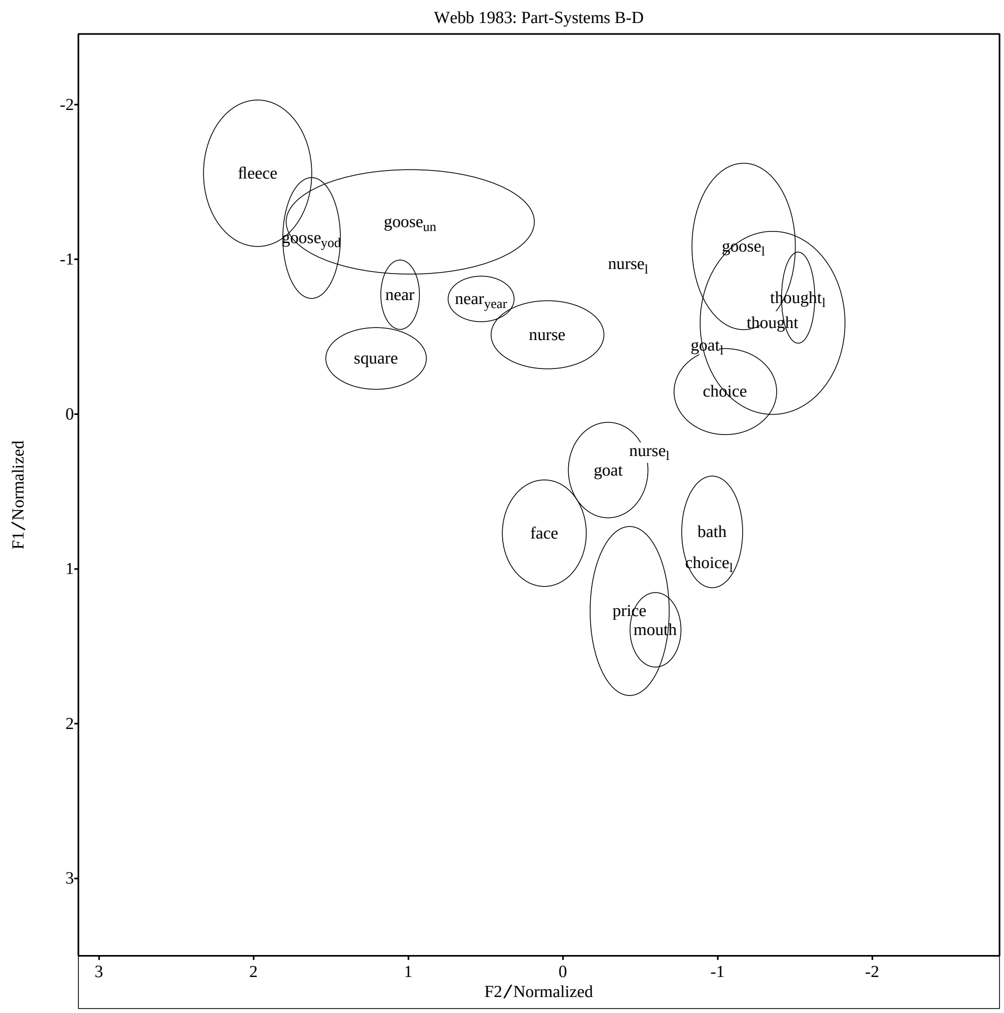


Figure 5.5: Webb's (1983) Long Vowels and Diphthongs

to at least be entertained. In conclusion, it was decided, for purposes of comparability, to recalculate Webb's (1983) means and standard deviations on the basis of the mentioned *a priori* basis of assignment.

5.1.4 Lass and Wright

Lass and Wright (1985) and Lass and Wright (1986) are mainly of relevance to the discussion of the SAE KIT vowel and will be dealt with in more detail in Chapter 7.

Lass (1990) has a general focus on "the upper socioeconomic stratum of . . . White native speakers" (Lass 1990:283; footnote 2) and a more specific regional focus on "one local but (intuitively) highly typical standard SAE variety: Respectable (=(Upper)Middle Class) White *Cape Town* English" (Lass 1990:272; my emphasis). It should also be noted that the description of this variety provided in Lass (1990) is that of male and female speakers, at the time "between the ages of 33 and 46" (Lass 1990:284; footnote 7) i.e. born roughly between 1944 and 1957. The subjects used in the present study would not have been born earlier than 1985. One 'problem' with this article, however, is that it is not always clear whether the descriptions of the various vowels are meant to be exclusively 'Cape Townian' or reflective of GenSAE as a whole.

Lass (2002) is, apart from some minor inconsequential details, the same article as Lass (1995) and will, therefore, not be referred to any longer. The reader is, in fact, encouraged to use Lass (1995) rather than Lass (2002) since the latter has a number of rather irritating errata. Both constitute general overviews of SAE, as does Lass (2004).

5.1.5 Other Works on SAE

Wells (1982), Mesthrie (1993), Branford (1994) and Bowerman (2004) all provide broad overviews of SAEP. Bailey (1984) isolates some (potentially) interesting features of SAEP some of which have already been touched on in previous chapters. The same applies to Jeffery (1982), which also provides a review of Lanham and Macdonald (1979). Since Taylor (1991) is focussed exclusively on the KIT vowel, more details regarding this study will be provided in Chapter 7.

Da Silva (2007) follows Horvath (1985) in using Principal Components Analysis to isolate two lects used by students at the University of the Witwatersrand in Johannesburg. While it is clear that ethnicity or race can no longer be employed as an absolute predictor of accent[6], it is equally clear that Da Silva's (2007) lect 1 is an overwhelmingly 'white' accent. It most respects its features resonate with the previous impressionistic literature. While, in

[6] Although, if truth be told, it never could.

the final analysis, the main interest of Da Silva's (2007) study lies in its description and delimitation of the so-called 'lect 2', used mostly by 'black' speakers, a number of interesting observations relating to lect 1 are made. Amongst others, these include observations regarding the potential influence of lect 2 on lect 1, particularly in terms of the spread of American-like variants. The main weakness of this study is the use of impressionistic as opposed to instrumentally-derived data.

5.2 Conclusion

This chapter has provided a brief sketch of the main foci, methodologies, assumptions and transcription methods used in the various sources on SAEP. As pointed out in the introduction, these differences need to be kept in mind in the following four chapters, which, collectively, provide a vowel-by-vowel comparative and critical analysis of vowel quality in SAE, with a focus both on its synchronic status and diachronic development.

CHAPTER 6

Labov's (1994) Patterns 1 and 4

> "There are two basic approaches to the problem of accumulating real-time data. The simplest and most efficient is to search the literature dealing with the community in question and to compare earlier findings with current ones" (Labov 1994:73).

6.1 Introduction

This chapter begins the vowel-by-vowel analysis of SAEP. As stressed at numerous points in previous chapters, the emphasis of this analysis is to bring together a range of different data in order to position SAEP vowel quality synchronically, diachronically and in terms of a number of descriptive and theoretical frameworks. This chapter begins by focussing on most of the vowels which are the primary 'focus' of Labov's (1994) Pattern 1 *extension* and Pattern 4 chain-shifts, as outlined in §2.3.2.1 i.e. MOUTH, PRICE, CHOICE (§6.2) with respect to the Pattern 1 extension, and TRAP (§6.3), FACE, DRESS and FLEECE (§6.4) with respect to Pattern 4. Given the extensive treatment KIT has received in the impressionistic SAE literature this vowel is dealt with in the separate Chapter 7.

6.2 MOUTH, PRICE and CHOICE (MPC)

As described in §2.3.2.1 and §3.4.1, the diphthongization and the laxing and lowering of the nuclei of ME ū and ī is an essential feature of the GVS and ETEs in general. A tensing, fronting (and raising) of MOUTH along with a tensing, backing (and raising) of PRICE along the front and the back peripheral tracks respectively, and by way of the Lower Exit Principle and Principle I, is a common extension of the GVS and Labov's (1994) Pattern 1, also known as the PRICE–MOUTH Crossover. A less dramatic effect involves the lowering (and tensing) of the nuclei of both MOUTH and PRICE onto the peripheral track so that they jointly become the most open vowels of the system, but without the PRICE–MOUTH Crossover.

However, PRICE backing and raising is also seen by Labov (1994) as central to a Pattern 4 movement i.e. one of the patterns constituting the 'Southern Shift'. In this case it is viewed as a reaction to the laxing and lowering of the nucleus of FACE, and another option for PRICE is fronting and monophthongization (see Figure 2.13 on page 61). The difference between the Pattern 1 PRICE movement and the Pattern 4 backed-PRICE movement appears, therefore, to be that in the former PRICE is viewed as moving independently while in the later it is a 'knock-on' effect of the lowering of FACE and FLEECE. In both cases a potential side-effect of the backing and raising of PRICE is the raising of CHOICE.

Given that SAE is considered to be participating in the Southern Shift it might, therefore, appear odd to include SAE PRICE under Labov's (1994) Pattern 1 extension and not Pattern 4, particularly given that there is extensive evidence to suggest that SAE does not retain conservative values of FACE and GOAT, particularly in Broader accents. Although this last comment will be seen to require a degree of qualification, more importantly, and as already mentioned in §2.3.2.1, the movement of PRICE to the back periphery in both the Pattern 1 extension and the Pattern 4 movement raises the question whether Pattern 4 should not simply be viewed as an extension of Pattern 1, as is implicitly done in, for example, Wells' (1982) notion of Diphthong Shift, which also conveniently 'links' the fronting and raising of MOUTH with Pattern 3 developments (see Chapter 8). It was also noted, in §2.3.2.1, that, under this Diphthong Shift conception, the backing of PRICE and fronting of MOUTH constitute the initiating moves of the shift (i.e. not FACE or GOAT), a view which appears to be supported by historical facts (Trudgill 2004:50).

Thus, while it is clear that Labov (1994) in a sense requires a distinction between a Pattern 1 extension and a Pattern 4 movement to distinguish those accents of English which show a PRICE–MOUTH Crossover but no movement of FACE or GOAT[1], from those accents which show a PRICE–MOUTH Crossover but *also* a lowered FACE and GOAT[2], from the viewpoint of the Diphthong Shift 'thesis', the distinction between a Pattern 1 extension

[1] Like New York English.

[2] As well as, potentially, a diphthongized FLEECE and GOOSE. These are, of course, the archetypal characteristics of Southern Shift Englishes like that of the southern USA or Australia.

and a Pattern 4 movement is purely arbitrary and is thus applied here mainly as a convenient mechanism by which to divide the various vowels up into manageable ‘chunks’.

Two *substantive* consequences of the adoption of the Diphthong Shift ‘thesis’ of these various vowel movements is, however, that, firstly, the movements of PRICE and MOUTH are seen to initiate the shift and, secondly, in line with comments by Trudgill (2004), that a fronted (monophthongised) PRICE vowel is not viewed as part of this shift, unlike in Labov’s (1994) Pattern 4, in which this movement is viewed as an alternative to PRICE-backing and as an integral part of the pattern as a whole. The implications of these assumptions should become clearer as the section proceeds.

6.2.1 MPC: Other Accents of English

Received Pronunciation: beginning with RP MOUTH, Lanham (1967:41) claims an unrounded [a]^{73}{[ɐ̝]} for the first element and a lip-rounded [o]^{36}{[ö]} for the end of the off-glide: [ɐ̝ö]. Lanham (1967:39) gives “a position further back at [a]^{74}” for what he refers to as “affected “Oxford English” often referred to as “advanced RP”” (Lanham 1967:29–30)[3], but a quick glance at Figure 5.3 on page 148 will convince the reader that there is no such possibility; Lanham (1967), in all likelihood, meant [ɑ]^{74}{[ä]} i.e. a slightly more retracted starting point than for General RP. This is, however, in contrast with Wells (1982:281; 292), who claims a relatively *front* starting point for MOUTH in his U-RP, although I suspect this is meant to reflect an older value[4]. Aside from these slight discrepancies, what does appear to be clear is that for RP as a whole the starting-point of MOUTH has undergone a process of retraction. In mainstream RP, there is a reasonably large degree of variation with values ranging from a retracted cardinal 4 to an almost fully-back cardinal 5. Furthermore, for RP in general, “the starting point of /aʊ/ is never fronter than the starting point of /aɪ/” (Wells 1982:292) i.e. there is no PRICE–MOUTH Crossover. Wells (1982) also claims a tendency toward unrounding on the glide (for both MOUTH and PRICE),

[3]As already mentioned in §5.1.2, this author notes various discrepancies between his analysis of RP and the more standard ones. This perhaps accounts for some of the differences between Lanham’s (1967) account of RP and those of other authors reviewed.

[4]Wells (1982:279–80) contrasts **U-RP** (upper-crust RP) with **Mainstream RP** (equivalent to what is also known as **General RP**). This author views **Conservative RP** and **Advanced RP** as, for all intents and purposes, two chronological variants (older and younger respectively) of U-RP. According to Wells (1982:280–1), “the term Oxford English, sometimes encountered, is best avoided because of its vagueness: although it may refer to a form of RP ‘in which certain tendencies are (sometimes affectedly) exaggerated’ (*Chambers twentieth century dictionary* 1972) – i.e. perhaps to U-RP – it has also been used to refer both more narrowly to the don stereotype and more widely to RP in general”. It is clear though that Lanham (1967) is not using the term in the latter sense.

and “there is now one variant of /aʊ/ were the second element is not only unrounded but also fronted, [ɨ]; this is heard by many as /aɪ/ . . . [leading to] . . . claims that the upper classes are ‘ite and abite dine tine”’ (Wells 1982:292).

Cruttenden (2001:137) generally concurs with the above analysis, mentioning a large degree of latitude in terms of the placement of the nucleus between C[4] and C[5] adding that “for many speakers, the first element of /aɪ/ and /aʊ/ may in fact be identical” (Cruttenden 2001:137) and, like Lanham (1967), also mentions, with respect to the glide, that “the tongue may not be raised higher than the close-mid level, i.e. [ö]”; the lips are only weakly rounded. Like Lanham (1967) and *contra* Wells (1982), Cruttenden (2001) gives his Refined RP[5] a more retracted value “sometimes reaching C[ɑ]” and often with glide-weakening and loss of lip-rounding leading to near homonymy between, for example, *loud* and *lard*. Said retraction is viewed by Cruttenden (2001) as a reaction to the fronted values so characteristic of Cockney (see below).

For RP PRICE, Lanham (1967:41) claims a value of [a]^{72} to ^{73} {[ä ∼ ɐ̝]} for the nucleus (thus slightly fronter than MOUTH – see above) and a value of [E]^{42}{[ɛ̞̈]} for the end of the glide: [äɛ̞̈ ∼ ɐ̝ɛ̞̈]. The author also claims that the glide is never lost. A more retracted variant is given for this author’s Oxford English, in contrast again with Wells (1982:281) who gives a fronter starting point for U-RP (who again is, in all likelihood, referring to an older variant). Wells (1982:292) emphasizes the great degree of variability in RP PRICE, but claims that it has, in general, undergone a process of retraction:

> “Older textbooks sometimes explained the quality of cardinal 4, [a], by identifying it with the first element of RP /aɪ/; but so front a quality is now unusual . . . mainstream RP currently embraces a range from retracted front [a] to an advanced back [ɑ]”.

Cruttenden (2001:131–2) generally confirms this variability and, for General RP, gives values as front as [æɪ] (often for those who have a close FACE vowel) and with a nucleus as back as a retracted [a] or fronted [ɑ] (often for those with a lower FACE vowel). This author claims that, as far as the glide is concerned, “the tongue is not usually raised to a level closer than [ë]” (Cruttenden 2001:132). In accordance with Lanham (1967), Cruttenden (2001:132) confirms a particularly retracted

[5]Cruttenden’s (2001) definition of Refined RP as “that type which is commonly considered to be upper-class” would appear to be close to Lanham’s (1967) Oxford English and Wells’ (1982) U-RP.

and monophthongized variant for Refined RP: [ɑ̈ː], “which may be only marginally differentiated from /ɑː/ realized as C[ɑ]”.

For RP CHOICE, Lanham (1967:41) gives [ɔ]^{55}{[ɔ̈]} for the first element, “with very little lip rounding”, and [E]^{42}{ɛ̝̈} for the end of the glide, which “is always clearly heard”; thus [ɔ̈ɛ̝̈]. For Wells (1982:293), the value of the nucleus in RP ranges “from [ɒ] to cardinal [ɔ] or somewhat closer”. For Cruttenden (2001:133), the first element has a similar value to that given by Wells (1982), with open rounding; the glide does not reach “a level closer than [ë̞]”.

London (Cockney) and Other Accents of England: the quality of MOUTH is of particular sociolinguistic relevance to the London area and a Cockney accent is commonly stereotyped along these lines. Cockney MOUTH is best represented in terms of a cline from RP to Cockney, as provided by Wells (1982:310); thus RP [ɑʊ], Popular London [æʊ] and Cockney [æː]. Trudgill (2004:52) provides convincing evidence to suggest that a fronted Diphthong-shifted MOUTH was prevalent in the broader south-east of England in the nineteenth-century. We note, along with Britain (2008:213–4), that north-of-England varieties would have had, in the main, at the time of the development of SAE, monophthongal values for MOUTH, such as [uː] or in some cases [aʊ]. Wells (1982:359) adds that “it seems … that the monophthongal types to be heard in some accents of the middle north, [aː ∼ ɑː], are synchronic stylistic variants of diphthongal types”.

For Cockney PRICE, Wells (1982:308) gives [ɑ̟ ∼ ɑ ∼ ɒ] for the nucleus, “while the second element may be reduced or absent (with compensatory lengthening of the first element), so that we have variants such as [ɑ+ə, ɑ+ː]{[ɑ̟ə ∼ ɑ̟ː]}”; thus, interestingly enough, very close to some of the the Refined RP values given above. Trudgill (2004:52) suggests that a Diphthong-shifted PRICE vowel was only really evident in the London area in nineteenth-century England. With respect to Glide Weakening as a result of Diphthong Shift, Trudgill (2004:140–1) claims that while glide-weakening and monophthongisation are found in Cockney, it “applies much more extensively in New Zealand English and is most unlikely to derive solely from lower-working-class East End of London speech”. Trudgill (2004) is of course referring here to his hypothesis, as presented in §3.4.3, that Glide Weakening was not imported to NZE, and the other Southern Hemisphere varieties, but constituted, instead, an endogenous development. Overall, “Glide Weakening has not (yet) occurred to any great extent in Britain” (Trudgill 2004:142). Trudgill (2004:140) also confirms that a fronted, monophthongised PRICE vowel is characteristic of Lancashire in the north-west of

England, while Wells (1982:358) gives “a front [a] in the middle north . . . the diphthong . . . not moving much beyond [ɛ], and often being (variably?) monophthongal . . . so that *price* appears as . . . [praɛs]”.

For Cockney CHOICE, Wells (1982:308) gives a raised [ɔ̣ɪ ∼ oɪ]{[ɔ̝ɪ ∼ oɪ]}. It is interesting to note that during the 18th-century and for most of the 19th-century (i.e. the period during which the Southern Hemisphere Englishes were established) CHOICE had the same value as PRICE in the English of London i.e. *oil* and *isle* had (virtually) the same phonetic value, a situation that by all accounts still prevails today in Essex, and which Labov (1994) identifies as a case of near-merger.

AusE and NZE: both AusE and NZE have front starting values for MOUTH with the former ranging from [a-ʊ]{[a̠ʊ]} in Cultivated AusE, [æo] in General AusE and [ɛːo ∼ ɛ̃ːɣ] in Broad AusE (Wells 1982:597). Harrington et al. (1997:172) provide acoustic evidence to show that the second target of MOUTH in AusE is, in fact, more likely to be associated with LOT than with FOOT. For NZE, and by way of example, “Conversational Wellington English . . . on average, uses [ɛ̝^{ə}]” (Britain 2008:190), where we notice the front and raised nucleus with a substantially reduced offglide. Perhaps more importantly, Britain (2008) provides relatively convincing evidence, in support of Trudgill’s (2004) comments above on Diphthong-Shifted MOUTH in 19th-century England, that a fronted and raised MOUTH in NZE (and AusE for that matter) is not the result of an endogenous development but was, effectively, brought over from England, mainly from the south.

For Australian PRICE, Wells (1982:597) gives [a-ɪ̜ ∼ ɒɪ̜ ∼ ɒːɪ̜]{[a̠ɪ̝ ∼ ɒɪ̝ ∼ ɒːɪ̝]} for the Cultivated, General and Broad sociolects respectively. Harrington et al.’s (1997) acoustic data provide evidence, though, for a much weaker glide in AusE PRICE, generally ending closer to DRESS than to KIT; they propose [ɑe] as a symbol suitable for Australian transcriptions.

Cruttenden (2001:133) gives a closer-than-RP starting point for AusE CHOICE.

USA Accents: turning to the USA, vernacular ‘white’ New York has the classic PRICE–MOUTH Crossover, with typical values for the former being [ɒi] and the latter being in the region of [æʊ]. Both PRICE and MOUTH are indicators i.e. below the level of consciousness with no style-shifting. In conservative New York speech a central [aʊ]-value for MOUTH is retained (Labov 1994:50) and [ai] for PRICE. CHOICE in New York has a particularly high nucleus: [ʊ^{i}] or [u^{i}] according to Labov (1994:101; footnote).

The PRICE–MOUTH Crossover is even more pronounced in Philadelphia. The raising and fronting of MOUTH is extreme: Labov (1994:60;81) gives [æʊ] as the most conservative form, with the AusE-like [e̞ɔ] being the most advanced, a shift which links the Philadelphia accent to the US South. PRICE, on the other hand, has backed from [ay] to [ʌy] in the particular context of a following voiceless consonant (Labov 1994:60; 82); by all accounts a relatively recent change in Philadelphia. As in New York, CHOICE in Philadelphia has "risen to high position, and beyond: the nucleus of /oy/ is so close that it is often heard as a semivowel, with a shift in syllabicity, so that *choice* and *twice* can be confused" (Labov 1994:211).

As pointed out in §2.3.2.1 and in §6.2 above, the PRICE–MOUTH Crossover is viewed in Labov's (1994) terms as an extension of Pattern 1. In New York City and Philadelphia this appears, according to Labov (1994:209), to be an isolated chain-shift and a further extension of the GVS; in particular, the other upgliding front vowels remain unaffected.

Turning to Labov's (1994) Pattern 4, this author mentions that "many Southern U.S. . . . dialects show /aw/ as [eü] with a fronted nucleus and /ay/ as a fully backed and raised [ɔɪ]" (Labov 1994:169). Thus, and as pointed out by Zhang (2008:204), "a raised and backed /ay/" is characteristic of "the 'hoi toider' dialect of Ocracoke"; also see Labov (1994:211) in this regard. Dialects with PRICE-backing are generally rural and/or conservative and characteristic of specific areas of the eastern seaboard, however, and a fronted, monophthongised [aː], or glide-reduced [aæ] PRICE vowel is more characteristic of the southern states of the USA.

With respect to fronted PRICE, and according to Wolfram, Carter and Moriello (2004:345), "there are two patterns of /ai/ glide reduction," one involving its reduction categorically, the other in non-prevoiceless position i.e. in final position and before voiced consonants. In general, according to Labov (1994:214), it is often invariantly monophthongal for lower-class individuals, while for other speakers PRICE monophthongisation is restricted to non-prevoiceless position. Thus the use of [aː] in words such as *mice* or *price* is clearly indexical of lower-class (and rural) status in many parts of the Southern USA. It is worth mentioning that dialect-leveling to a full diphthongal quality in all contexts seems to be a growing feature of many Southern urban areas e.g. in Raleigh (Wolfram et al. 2004:349). A fronted and monophthongised PRICE vowel is also common among 'black' speakers in the USA more generally e.g. in New York City (Labov 1994:50–1).

Importantly, Trudgill (2004:140) claims that the monophthongisation of fronted PRICE in the US South (as well as in Lancashire – see above) is unrelated to post-Diphthong-Shift Glide Weakening, as dealt with in §3.4.3; simply put "no Diphthong Shift has occurred in these cases – that is, the first element of the diphthong is (was) an unshifted [a]".

A monophthongal [aː] for MOUTH is a well-known indicator of Pittsburghese (Johnstone and Kiesling 2008).

6.2.2 MPC in SAE: The Impressionistic Data

Hopwood (1928): "**aʊ**{[a̠ʊ]} > **æw**{[æu]}, *˧**aˑ(w)**{[aˑ(u)]}, or **ɑˑù**{[ɐ̞ˑŭ]}" is given by Hopwood (1928:23) for SAEP MOUTH, with the value before the 'wedge' representing the RP value, the values to the right representing the SAEP values and the most common value being marked by a star (Hopwood 1928:7)[6]. The relatively fronted [a̠] for Hopwood's (1928) RP value makes sense in terms of Wells' (1982) comments, as paraphrased in §6.2.1, concerning the process of retraction MOUTH has undergone in RP i.e. the older the speaker the more fronted the nucleus. Hopwood (1928:23) also claims that "occasionally before a consonant there is a tendency for the off-glide . . . to disappear; e.g. out **aʊt**{[a̠ʊt]} > SAE. **'aˑ(w)t**{[ʔaˑ(u)t]} . . . Cp. "Punch's" spelling "aht" for Cockney E. of "out""[7].

Hopwood (1928) claims a Cockney origin for the first two alternatives given at the beginning of the previous paragraph, i.e. **æw** and ˧**aˑ(w)**, while **ɑˑù** is the Afrikaans English pronunciation.

For PRICE in SAE, Hopwood (1928:16) gives "˫**ɑˑj**{[ɑˑi]} (or **ɒˑj**{[ɒˑi]})," noting again that Hopwood's (1928) **ɑ** "is Low Central . . . not Low Back" (Hopwood 1928:89) and that ˫**ɑ** is thus (approximately) the equivalent of IPA [ɑ]. We also note a more tense off-glide in SAEP than the value Hopwood (1928) gives for RP i.e. [aɪ][8]; this is somewhat surprising given that the general stereotype of a BrSAE PRICE vowel is that it is monophthongised i.e. [ɑː]. As is usual for this author, Hopwood (1928) claims both Afrikaans and Cockney influence in the development of the quality of this vowel. The PRICE–MOUTH Crossover is clearly indicated in Hopwood's (1928)

[6]Note that as far as the translations, provided in curly brackets, are concerned, Hopwood (1928:89) provides a narrow transcription of RP **a** as ˫**a** i.e. modern IPA [a̠]. A fronted ˧**a** thus, presumably, implies modern [a]. Lastly, and as mentioned in §5.1.1, Hopwood's (1928) **ɑ** symbol is translatable as [ɐ̞].

[7]Note that Hopwood (1928) is somewhat inconsistent here in applying his transcriptions: SAEP **'aˑ(w)t** should strictly-speaking be **'˧aˑ(w)t** otherwise his original comments make little sense.

[8]Note that Hopwood's (1928) value for PRICE supports the notion of an older, fronted (although diphthongised) version of RP PRICE.

analysis.

For CHOICE, Hopwood (1928:17) claims “**ɔɪ > oì**{[oi̽]} . . . a raising and slight tensifying of the first element (sometimes accompanied by lengthening), from the half-open position to the half-close”. The author does not advance any possible origin for this value although it, of course, clearly relates, in chain-shift terms, to a raised and backed PRICE.

Hooper (1945): This author highlights the variable nature of MOUTH in SAE, showing that relatively lower social-class status as well as being male tends to correlate with fronting and raising of the MOUTH vowel, while being from a relatively higher social-class or female correlates with the ‘standard pronunciation’. The range of pronunciation includes ‘standard’ [aʊt][9], “though the first element of the diphthong may vary permissibly between cardinal vowels 4 and 5 . . . [in contrast to] a high [a], to [æ] or even to cardinal vowel 3” (Hooper 1945:478; my parenthesis) as well as a backed [ɑʊ] as used by the majority of the female higher social-class subjects.

Hooper (1945) also shows that a raised CHOICE vowel is correlated with relatively lower social class and (more marginally) with being male. The ‘standard’ [ɔɪ] pronunciation is, of course, correlated with the opposite values. What this seems to indicate is, that at the time at least, CHOICE was a linguistic indicator[10]. There seems to be no evidence from earlier or later authors that this variable had or has reached a level of awareness among SAE-speakers.

Lanham and Traill (1962), Lanham (1965): Lanham and Traill (1962:29) provide the following qualities for the nucleus of SARP[11] MOUTH: [æ]$^{63/64}${[ɐ]} or [a]^{64}{[ɐ̠]}, seen as between the two extreme poles of “‘affected’ British English,” which has [*a*]^{75}{[ɑ]} (see §6.2.1), and non-SARP SAE which in “its extreme forms has a vowel of [æ] or [a] ‘Eliza Doolittle’ quality” (Lanham and Traill 1962:29). As far as the up–glide is concerned, Lanham and Traill (1962) give the RP-like [o]$^{35/46}${[ö̞]} or [o]^{36}{[ö]} for SARP, while claiming “a much weaker, shorter glide [as] a feature of non–SARP SAE” (Lanham and Traill 1962:29; my parenthesis)[12].

For PRICE, Lanham and Traill (1962:29) give [*a*]$^{63/64}${[ɐ]} for the onset for both SARP ‘A’ and SARP ‘B’, with [ï]$^{24/34}${[ə̞]} for the up–glide. For non-SARP SAE,

[9]Note the relatively fronted nucleus, in support of Wells’ (1982) notion, mentioned above, that U-RP MOUTH is characterised by a front starting-point.

[10]Or perhaps even a marker; unfortunately Hooper (1945) does not deal with stylistic variation.

[11]The reader is referred back to §5.1.2 for a definition of this concept.

[12]With respect to the value [o]$^{35/46}$, I suspect a typographic error; this should rather be [o]$^{36/46}$, which explains my translation into [ö̞].

Lanham and Traill (1962:29) claim a degree of monophthongisation as well as a degree of lip–rounding. Lanham (1965:91) confirms this general trend towards glide-weakening:

> "For many young South Africans ... this diphthong is now replaced by the long vowel /aˑ/ with the post-vowel glide completely lost. The quality of this vowel may be either [ɑ] or [a˃]{[ä]}; thus [naˑs] or [nɑˑs], [raˑd] or [rɑˑd]".

We note here the two values that have become stereotypical of respectively BrSAE, [ɑː], and certain prestige variants of GenSAE i.e. [aː], the latter already mentioned in §3.3.2.

According to Lanham and Traill (1962:29), SARP 'A' CHOICE is characterised by "slit–rounded [ɵ]$^{46/55}$" while SARP 'B' has a "slightly higher [ɵ]^{46}", although the authors stress that there is substantial variability in the quality of this vowel. The glide ends at [ɜ]^{33}. In modern terms, therefore, the values in question are approximately [ɔ̞̈ɘ̟] for SARP 'A' and [ö̝ɘ̟] for SARP 'B'. We note that the accent furthest from RP has the slightly more raised value.

Lanham (1967): Lanham (1967:93) echoes the earlier work summarized above, giving, for conservative SAE MOUTH, [a]$^{63/73}${[ɐ]} as the value for the onset, and [o]^{36}{[ö]} as the value for the terminus of the glide in CV syllables, with the glide being shorter in other contexts. For non-conservative SAE, an archetypal value is the fronted [æ]^{62}{[æ̈]}, "with a glide almost devoid of lip-rounding ending at approximately [ɜ]^{35}{[ɘ̠]}"(Lanham 1967:93).

For PRICE, Lanham (1967:63) confirms the option of glide-weakening in non-conservative varieties of SAE. Completely glide-weakened values include [aː]^{73}{[ɐ̝ː]} to [ɑː]^{65}{[ɑ̞̈ː]}, the latter of which threatens the distinction between PRICE and BATH (§8.2). However, "most speakers ... still seem able to retain the contrast by introducing a weak glide" (Lanham 1967:64); see also (Lanham 1982:338–9) in this regard. Lanham (1967:86) thus emphasizes two possible outcomes in non-conservative SAE that arise as a result of PRICE glide-weakening:

1. PRICE and BATH are distinguishable in terms of the front-back dimension, with the former having a "slightly raised [a]^{72}{[ä]} or ^{73}{[ɐ̝]}" (Lanham 1967:86) and the latter having a [ɒ]^{65}{[ɒ̞̈]} quality.
2. An even more extreme version in which PRICE is glide-weakened and retracted to the [ɒ]^{65}{[ɒ̞̈]} position, thus threatening the contrast between PRICE and

BATH. In such cases the difference between, for example, *arm* and *I'm*, is lost.

This is, of course, an example of Labov's (1994) analysis of the two possible effects of the Southern Shift on PRICE, as outlined in §2.3.2.1.

In contrast, for conservative SAE PRICE, Lanham (1967:92) claims that "the vocoid is usually slightly raised [a]^{73}{[ɐ̹]} and the glide reaches approximately [ɜ]^{33}{[ɘ̹]}". Lanham (1967:105) adds that "the tendency to lose the glide in the diphthong /ai/ . . . may be heard in Yorkshire today in much the same way as it occurs in SAE"[13].

For conservative SAE CHOICE, Lanham (1967:92–3) gives [ɔ]^{55}{[ɔ̈]} (or a little higher) as the onset and [ï]^{24}{[ɨ̹]} as the terminus of the glide in CV syllables, with the glide being somewhat shorter in other contexts. For non-conservative SAE, "the trend is to a higher, more rounded vocoid, possibly as high as [ɵ]$^{36/46}${[ö̹]} and the glide remains prominent" (Lanham 1967:93).

Lanham (1978), Lanham and Macdonald (1979): For Lanham (1978:154), a front-ed "[æ˃]{[æ̈]} followed by unrounded glide to [ə˃]{[ə̠]}" is the most advanced variant of MOUTH and is characteristic of Broader lects. It is definitely above the level of social consciousness and is hypercorrected to [ɑʊ], thus showing second and probably third-order indexicality[14].

Similarly, Lanham and Macdonald (1979:40–1) identify fronted and glide-weakened MOUTH as a CE[15] variable and thus a defining feature of BrSAE. The most advanced values are [æ̞] for the onset, with a weak (unrounded) glide to [ə̠]. The vowel is corrected to a [äʊ]-like quality or even hypercorrected "towards backed [ɑʊ] or glide-weakened [ɑˠ] with hypercorrections most prominent among females in the Cape generally" (Lanham and Macdonald 1979:40). Importantly, and according to these authors, the trend towards a fronted MOUTH is being *reversed* in South African society, with "a retreat from æ through age grades" (Lanham and Macdonald 1979:40). Furthermore, "zero index scores, or hypercorrect values, occur *with little stylistic variation* particularly in the speech of middle-class and upper-class subjects" (Lanham and Macdonald 1979:40; my emphasis). Thus a non-fronted MOUTH vowel appears to be categorical in GenSAE i.e. a marker.

[13]And we have, of course, noted in §3.3 that the north-of-England is well represented in NE (possibly one of the primary inputs into GenSAE) as well as, more speculatively, among the lower-class British subjects who came to mine on the Witwatersrand.

[14]See §1.3.3 i.e. a marker or even a stereotype.

[15]Cape English – the reader is referred back to §3.2 in this regard.

For PRICE, Lanham (1978:148–9) identifies a *fronted*, glide-weakened PRICE vowel, as an (if not *the*) identifying feature of GenSAE. It constitutes a NE[16] variable and would appear to "express social prestige in some degree" (Lanham 1978:148), a fact reflected in terms of its correlation with both social class and style i.e. the higher the social class and the more formal the style the greater the glide-weakening and fronting.

As already mentioned in §3.3.2, Lanham (1978:153) uses the expression "a feature of 'refined' female speech in the SAE community" to describe this feature for which the most advanced value is a "hypercorrect version approaching the quality of a front, tense Australian [aː]" (Lanham 1978:148), referring here, no doubt, to the AusE quality in BATH. Favourable conditioning environments include a following [l,m,n,v,z,s]. At the time there appeared to be a regional effect with respect to this value: fronting and glide-weakening seemed to have the highest degree of prestige among middle-class women in the Johannesburg area (particularly those with an Eastern-European background) and, was in Lanham's (1978) time at least, increasing in its use across generations. Natal, on the other hand, was characterised by a stable pattern, i.e. the use of a fronted monophthongised PRICE vowel was roughly equivalent across different age groups. In general, though, this value of the PRICE vowel is seen as having overt prestige in the GenSAE speech community: "we have anecdotal evidence of the adoption of the ai variable coinciding with finding employment in a high-class women's store" (Lanham 1978:149). The fronted monophthongised variant is opposed in prestige to its retracted, raised and glide-weakened BrSAE counterpart, in its most advanced version "weakly rounded [ɒ]" (Lanham 1978:153). Contexts that favour glide-weakening for the BrSAE version are following [l, m, n].

CulSAE, on the other hand, has a central and slightly more close [a] with a glide[17]. Lanham (1978:153) mentions Lancashire as a possible source of fronted, monophthongised PRICE.

Similarly, Lanham and Macdonald (1979:41) identify a "backed, raised, glide-weakened" PRICE vowel as a CE variable and thus characteristic of BrSAE. These authors imply that, as in the case of fronted MOUTH, this variant is in retreat and subject to extensive correction. A *fronted*, glide-weakened PRICE vowel constitutes a NE variable

[16] Natal English – see §3.2.

[17] This "slightly more close" value raises the suspicion that the nucleus of CulSAE PRICE has not yet fallen, by way of the Lower Exit Principle, onto the peripheral track. According to Wells (1982:149), "a starting point that is not fully open . . . is typical of the rural south of England, of Barbados, and of parts of the north-eastern United States" a confluence of accents not particularly helpful in trying to seek a possible explanation for this phenomenon. This, perhaps, requires further research.

characteristic of GenSAE:

> "The phonetic trend is towards a fronter, tenser [aː] without trace of an off-glide in the most favouring environments of following /l,m,n/. The highest index scores have a glideless Fronted ai in any environment except before V without intervening C (e.g. *buyer, buying*); the trend is auditorily most prominent word-finally (e.g. *buy*). (Spectrographic evidence confirms complete glide loss)" (Lanham and Macdonald 1979:41)[18].

Lanham and Macdonald (1979) confirm the special status of fronted PRICE in Natal, where it is subject to correction, while "generally outside Natal highest values of fronted ai are found in most formal speech behaviour" (Lanham and Macdonald 1979:42), particularly in female speech. Interestingly enough, Lanham and Macdonald (1979:42) claim that, at the time, a fronted, glide-weakened PRICE vowel was below the level of social consciousness and thus a marker[19]. The indexicality of the fronted, monophthongised PRICE vowel in SAE, as analyzed by Lanham (1978), Lanham and Macdonald (1979), is, of course, the opposite of its value in the Southern USA (§6.2.1).

Lass (1990; 1995; 2004): According to Lass (1990:273), an extreme variant for MOUTH would be in the region of [æʊ]. For his *Cape Town* GenSAE speakers, Lass (1990:280) gives [ɑ̈ɤ] as the most common value, with the onset varying between [ɑ̣̈]{[ɑ̝̈]} and [ɐ]. Both in diphthongal and monophthongal variants the value is most often more fronted (and in the case of monophthongal variants closer) than the corresponding BATH vowel. Some speakers, however, have a weak form of lip-rounding on the off-glide, giving values such as [ɑ̝̈ʊ̈] and [ɑ̝̈ʉ̞]²⁰.

Lass (1995:99-100) confirms the social importance of MOUTH (along with PRICE) and gives the basic norm for CulSAE and GenSAE as "something in the vicinity of [ɑʊ]" and thus not displaying the PRICE–MOUTH Crossover. Unrounding of the second element, i.e. [ɑɤ], is common among younger speakers and in GenSAE, unlike in CulSAE, there is a tendency to monophthongise this vowel i.e. [ɑː]. There is usually no merger with BATH, the later generally maintaining a position that is slightly lower

[18] As an aside, it would be extremely interesting to compare current acoustic data with this "spectrographic evidence", if it still exists. I know of no other reference to it.

[19] More recently, however, it has undoubtedly been subject to some degree of stereotyping, mainly due to its association with the broader 'Kugel' stereotype.

[20] Note the fronted glide-target given for the second option.

and retracted than glide-weakened MOUTH. As in the case with PRICE monophthongisation, MOUTH monophthongisation is considered by Lass (1995:99), and thus seemingly in contrast to Lanham and Macdonald (1979), to be non-categorial and related to speech style (the faster and more casual the more monophthongisation) and more common among younger and female speakers. Lass (1995:99) claims, however, that it less common than PRICE monophthongisation to [aː]. Some male GenSAE speakers and BrSAE speakers as a whole display the PRICE–MOUTH Crossover. For the former, MOUTH onset values such as [a] are found (Lass 1995:105), while [æʊ] would be a 'fully'-Broad value (Lass 1995:99-100). BrSAE MOUTH shows little evidence of monophthongisation.

Lass (2004:377–8) adds that, with respect to Broader accents, an onset even in the region of [e] or, alteratively, a [jæʊ] triphthong, are possibilities although the latter is "particularly associated with coloured L1 and Afrikaans L2 speakers"[21].

For PRICE, Lass (1990:273) gives [ɑɪ] as the general region for the diphthongal option in BrSAE. Formal word-list style in GenSAE often produces *diphthongal* variants such as [äe ∼ äɪ]. All SAE variants (except CulSAE) have the tendency, though, to monophthongize. Thus further monophthongisation of the PRICE vowel in BrSAE into [ɑː], or more accurately [ɑ̟ː] (Lass 1990:279), constitutes a common stereotype of this variety. Likewise, according to Lass (1990:279), the most common value for his GenSAE speakers is [äː][22].

Lass (1995:99), on the other hand, gives [aɪ] as the basic value for both CulSAE and GenSAE and notes that, unlike in CulSAE, there is a tendency in GenSAE to monophthongize:

> "This monophthongisation is rarely categorical, is commoner for PRICE than MOUTH, and appears to relate to speech tempo and register: the more casual or faster, the more likely monophthongal realizations. My (corrigible) impression is that monophthongisation is greater in younger than older speakers, and in females than males".

We note in this regard that for Lanham (1978), Lanham and Macdonald (1979) the *opposite* applied i.e. the more formal the style the greater the degree of monophthongization. Lass (1995) confirms a backed onset for BrSAE PRICE: [ɑɪ]. This

[21] It should be mentioned that, in Lass (2004:377–8), [aʊ] is given as the most common value for GenSAE MOUTH and the author claims that CulSAE and GenSAE speakers "have a slightly advanced or centralised [a], which may monophthongise". This is, in all likelihood, an error, with [a] having replaced [ɑ].

[22] The text actually gives [äɪː] as the value, but this is surely an error.

Broad value is also, however, prone to monophthongisation: [ɑː]. Backed and glide-weakened varieties are commonly subject to stereotyping. As mentioned above, some male *GenSAE* speakers evince a degree of PRICE–MOUTH Crossover, "with the first element of . . . PRICE a centralised back vowel, but fronter than BATH" (Lass 1995:105; footnote 14).

Lass (2004:372) explicitly denies the existence of a rounded onset to PRICE in SAE; the most common value for GenSAE speakers is given as [aɪ].

According to Lass (1990:279), the onset of CHOICE is higher than in RP, and ranges from [ɔ] to [o̜]{[o̞]}, thus [ɔɪ ∼ o̞ɪ]. There is no tendency to monophthongize. According to this author, in CulSAE [ɒ̣̈]{[ɒ̝̈]} would appear to be a relatively typical value for the onset. Lass (1995:100) adds that the onset of the diphthong is a little more open than THOUGHT with the terminus being "the higher version of KIT". This diphthong is, for all intents and purposes, invariant. According to Lass (2004:377), the onset is never as low as LOT and is always rounded. Curiously, the most common value for GenSAE speakers is implied as being [ɔɪ] and not [oɪ].

The remaining sources: Wells (1982:614; 621) gives the RP-like [a-ʊ]{[a̠ʊ]} for CulSAE MOUTH, and [æʊ], [æɤ] or the glide-weakened [æə] for Broader accents. A corrected, and advancing variant, has a starting point close to Cardinal [ɑ], giving the glided [ɑʊ] or the glide-weakened [ɑɤ]. Important too is Wells' (1982) general comment that the PRICE-MOUTH crossover has not become as established in SAE as it has in other Southern Hemisphere varieties. Mesthrie (1993:30) claims that glide-weakened MOUTH is likely to be the result of the influence of "lower class Home Counties speech", while Branford (1994:483) confirms the tendency to front this vowel and claims a tendency to weaken its glide in Broad accents. Glide-weakening of this vowel appears, however, to be a general tendency. Bowerman (2004:938) gives [ɑ̈ʊ] as the CulSAE value, with GenSAE often having a monophthongised [ɑː]. The fronted, Broad value, with offglide retention, is [æʊ]. Da Silva (2007:116-7) restricts the possible values of MOUTH in GenSAE to [aː], [ɑː], and [aʊ]; the first value is less-than-likely for GenSAE MOUTH, while a backed, diphthongal variant, [ɑʊ] or at least [ɑɤ], is conspicuously absent from her analysis. For her lect 1, the author found a substantial preference for the monophthongal backed variant over the fronted, diphthongal one (Da Silva 2007:159).

For PRICE, Wells (1982:614) gives [a-] for CulSAE, although this is undoubtedly a typographic error and should be [a-ɪ]{[a̠ɪ]}, thus retaining the glide. For Broader accents [ɒɪ] or glide-weakened [ɒː] is provided. This author confirms the 'growth' of the fronted PRICE vowel, giving both glided [aɪ] and glide-weakened [aː] variants, with the latter being "the most typically South African variant" (Wells 1982:614). Branford (1994:481) gives [aɪ] as the CulSAE value, with fronted and backed glide-weakened variants common in non-CulSAE. Bowerman (2004:938) begins by providing [aɪ] as the CulSAE value and, then, makes the rather odd statement that "in General and Broad, the articulation of the first element is often monophthongised to [aː]"; odd, firstly, in the sense that it is the diphthong as a whole that is subject to monophthongisation and never, by definition, the first element. The above also contradicts (in terms of the Broad value) the further claim that "in Broad, *the first element is somewhat back*, but more forward and higher than in BATH, and the offglide is often retained: [ɑ̈ɪ]" (Bowerman 2004:938; my emphasis).

Mesthrie (1993:30) claims that glide-weakened PRICE "is probably of Northern English origins (and therefore originated in Natal rather than Cape SAE)". Branford (1994:482), likewise, mentions Yorkshire, but also points out that it is "a British West Country variant, so that there is a just-possible input here from the speech of Cornish miners in the later nineteenth century", this being an obvious reference to immigrants to the early Johannesburg. Perhaps more controversially, this author claims that "monophthongisation of this vowel ... is widespread in Britain" (Branford 1994:482).

Da Silva (2007:116), in her impressionistic study of the speech of students at the University of the Witwatersrand, mentions a monophthongal [aː] as characteristic of some L1 speakers: as "highly marked and particular to certain speech communities in the Johannesburg area, namely affluent northern suburbs such as Sandton, and is associated with the social variables "female" and "Jewish""[23]. Overall though, her lect 1 speakers tended to be quite variable in terms of their choice of realization, with [aː] obtaining a variant mean score of 11%, while [ɑ(ː)] and [aɪ ∼ əɪ] obtained scores of 58% and 31% respectively[24]. In this respect it should be emphasized that Da Silva's

[23]The author also mentions the apparent nasality that often accompanies this and other vowels in the speech of the social groups mentioned; she, in addition, encourages further research in this regard. On this note an intriguing possibility involves so-called 'rhinoglottophilia' (Blevins 2006:135–6) which is based on the perceptual similarity of aspiration and nasality and leads to the incorrect perception of the one as the other. It appears from a brief excursus into the literature that the form of Yiddish brought to South Africa by the Eastern European Jewry would have contained plosives with very little aspiration. It is quite conceivable that these Yiddish-speakers would have perceived the aspiration characteristic of initial plosives in SAE (as in most other varieties of English) as nasalization, thus accounting for its presence in the so-called 'Kugel' accent. My very casual impression is that similar nasality is common in other Yiddish-based lects of English; if so, this would strengthen the hypothesis, although these impressions require, of course, proper investigation.

[24]Da Silva's (2007) latter claim that "PRICE3 [i.e. [aɪ ∼ əɪ]], also present in lect 1 but with less frequency, is also a diphthong that has not been referred to in other research" strikes one, at least with respect to the first value, as odd given the literature reviewed above. The confusion is compounded when the author claims

(2007) sample was not only limited to subjects with a private-school background.

With respect to CHOICE, and in line with his similar analysis of FLEECE before final /l/ (see §6.4), Bailey (1984:16) proposes that CHOICE in the same environment in SAE leads to the creation of separate syllables; thus /toy.əl/ *toil* and /boy.əl/ *boil*. Branford (1994:482) confirms the raised status of CHOICE in SAE and Bowerman (2004:938), echoing Lass (2004), claims the usual (surprisingly low) value as being [ɔɪ] in all varieties; "the onset can be as low as LOT in older Cultivated WSAfE speakers".

6.2.3 MPC: A Comparative Acoustic Analysis

From Figure 5.5 on page 152, it is clear that Webb's (1983) idiolect does not include a PRICE–MOUTH Crossover, although there is some degree of overlap. Webb (1983) is, however, motivated to comment on the glide of MOUTH in his data since, as in the case of FACE (§6.4), there is no evidence of glide-weakening: "the glide of /aw/ is quite strongly present, being on average 108Hz "higher" and 116Hz further "back" than the nucleus" (Webb 1983:155). It must be added, though, that an inspection of the original Figure 3 in Webb (1983:153) seems to indicate a rather *short* glide for MOUTH, even shorter than PRICE, which the author claims provides support *for* the notion of glide-weakening in SAE (Webb 1983:155). In this regard, it seems that the issue for Webb (1983:155) is the presence or absence of a clear pattern. Thus while it appears that the average "distance" traveled for PRICE is higher than that for MOUTH, it has a higher degree of variability and "there is no clear pattern in the glide movement . . . the glide's "forward" position varies between 19Hz and 441Hz" (Webb 1983:155).

For PRICE, Webb (1983:154) argues for possible allophony but given that the evidence is, by the author's own admission, "extremely scant" the data is not represented here. Note that in Figure 5.5 on page 152 the onset of CHOICE is, in line with the impressionistic literature, somewhat lower that the THOUGHT vowel and, perhaps, slightly centralized.

Turning to the acoustic data collected as part of the current research[25], Figure 6.1 on the next page places MOUTH in the context of the other Part-System C vowels[26] while

that "the one variant, [aɪ], has been mentioned by Lanham as being characteristic of Extreme South African English, but according to him, it does not occur in ESSA ("Respectable") English" (Da Silva 2007:218). This is, of course, confusing the front *prestigious* diphthong [aɪ] with a backed (stigmatized) diphthongal [ɑɪ].

[25] Appendix D provides means and standard deviations for the data on which the following graphs, as well as those in latter sections, are based. The ellipses are based on the standard deviation of the (normalised or unnormalised) F_1 and F_2 values at the 50th percentile i.e. the temporal midpoint.

[26] Note that in Figure 6.1 the unmarked MOUTH allophone is 'hidden behind' MOUTH before final /l/. The unmarked allophone has the somewhat stronger glide and a slightly more front position. The ellipses here, as

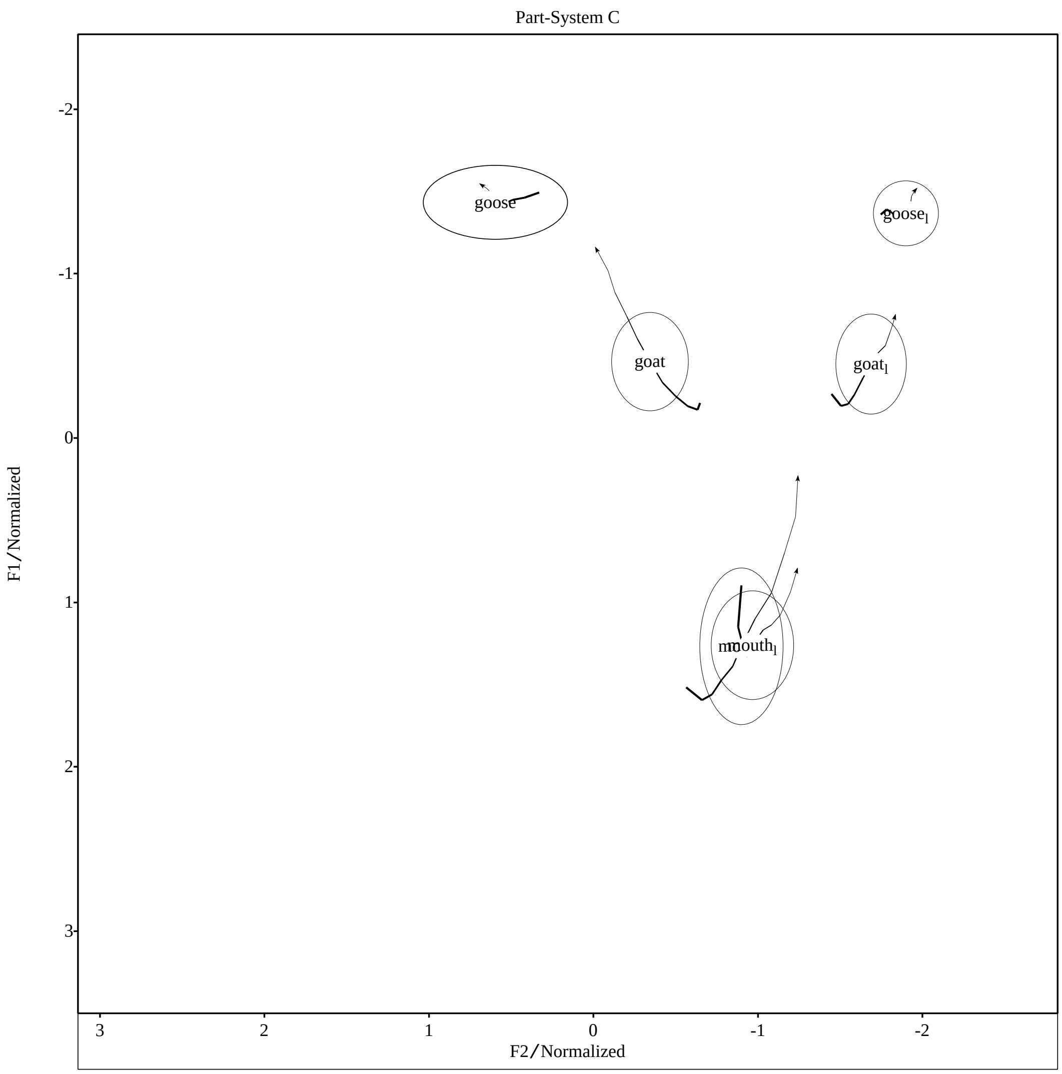


Figure 6.1: Overall Results: Part-System C (27 subjects)

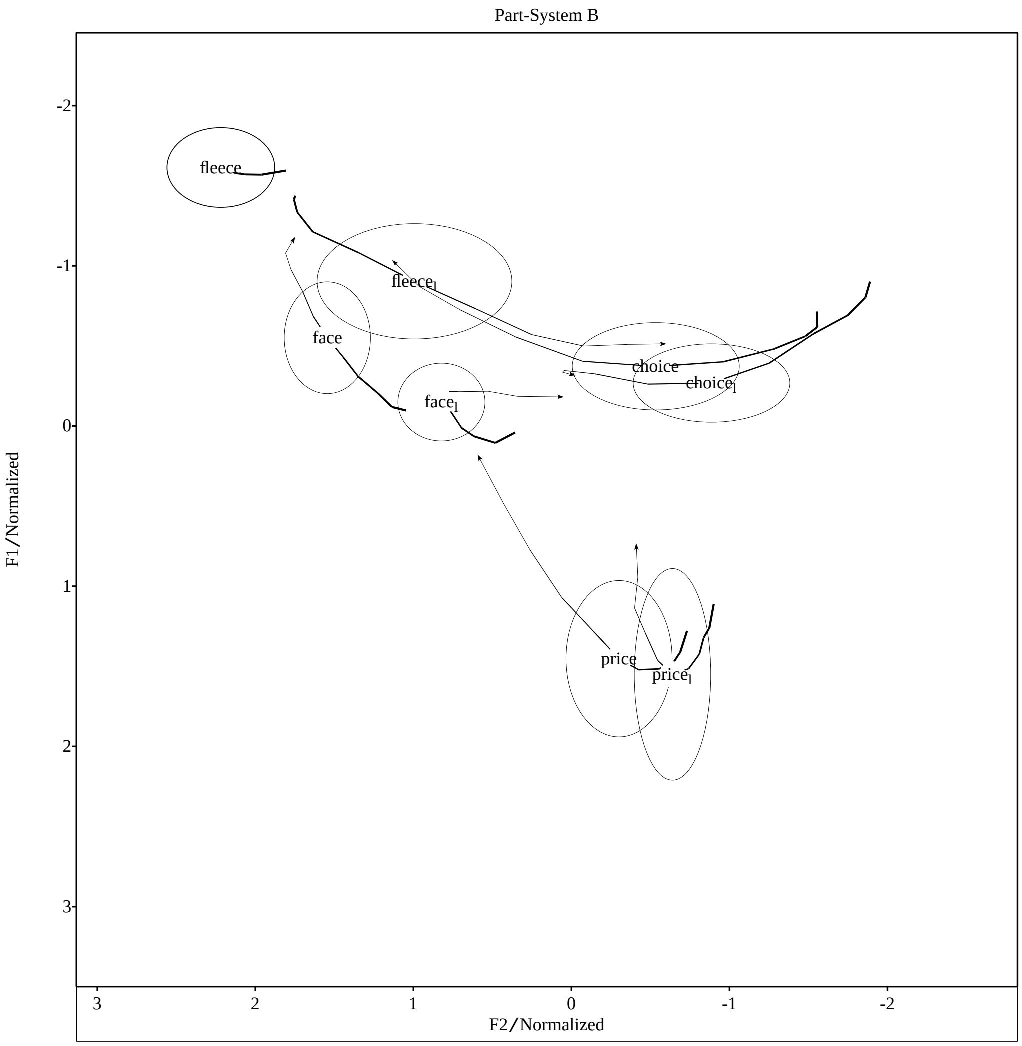


Figure 6.2: Overall Results: Part-System B (27 subjects)

Lexical Set	No. of Tokens	No. of Measurements	N-G Tokens	O-G Tokens
MOUTH	79	1738	***to**wer*, *loud*, *how*, *lout*	*out*
MOUTH before /l/	18	396	*foul*	
PRICE	177	3894	*height*, *hide*, *hype*, *hike*, *iron*, *price*, *prize*, *try*, *mine*, *mice*	*bite*
PRICE before /l/	18	396	*mile*	
CHOICE	44	946	*void*, *ex**ploit***	*boy*
CHOICE before /l/	18	396	*boil*	

Table 6.1: Summary of MOUTH, PRICE and CHOICE data

Figure 6.2 on the preceding page places PRICE and CHOICE in the context of the other Part-System B vowels. The representations of MOUTH, PRICE and CHOICE were based on the data summarized in Table 6.1. Thus the representation of non-pre-final-/l/ MOUTH in Figure 6.1 on page 172 was based on 1738 measurements across 79 tokens, with the words spoken by the N-G subjects and the O-G subjects being indicated in the two rightmost columns. Since PRICE in SAE has been the subject of some attention in the impressionistic literature, a comparatively large number of tokens per subject were collected for this vowel.

As can be seen from Figure 6.1 on page 172, while the glide of MOUTH is certainly weak it is not monophthongal, unsurprising given the citation-form style; see Lass' (1995) analysis in §6.2.2. As far as MOUTH before final /l/ is concerned, while there is not much evidence of retraction, it would appear that this allophone of MOUTH is subject to an even greater degree of glide-weakening and is, for all intents and purposes, monophthongal.

As far as PRICE is concerned, and as can be seen from Figure 6.2 on the preceding page, while there is arguably *some* glide-weakening of this vowel, it would certainly be incorrect to characterize it as monophthongal. As in the case of MOUTH, this appears to support Lass' (1990) observation that (fronted) PRICE monophthongisation (or substantial glide-weakening) is less-prevalent in citation-form than connected and spontaneous speech. This is, as mentioned in §6.2.2, in contrast to Lanham's (1978) and Lanham and Macdonald's (1979) earlier impressions that fronted, monophthongized PRICE carried overt prestige value and was thus more prone to occur in *formal* styles. There are a number of

elsewhere, indicate the standard deviation. In addition, the convention here and in most other cases is for the pre-final-/l/ allophone to be marked with a final subscript 'l'.

possible explanations for this difference, of course. Firstly, Lanham was perhaps wrong; secondly, and I suspect more believably, there has been a change in the indexical value of fronted PRICE; thirdly (and not necessarily unrelated to the second option) the indexical value of PRICE is more complex than a simple correlation with prestige (and gender), and is intertwined in a complex fashion with other variables such as region and ethnicity[27].

As expected, we note no evidence of glide-weakening in the case of CHOICE; we note too that there is no glide to FLEECE, as proposed by Hopwood (1928).

As is clear from Figure 6.3 on the next page, and as was expected, the current subjects do not show any evidence of the PRICE–MOUTH Crossover: MOUTH is discernably retracted in comparison with PRICE, and there is no overlap (at least in terms of the standard deviation) of the two vowels. This is in contrast to Webb's (1983) accent, as represented in Figure 5.5 on page 152, where, although there is no crossover as such, the two vowels share, essentially, the same nucleus. There is, as expected therefore, *some* evidence of sociolinguistic differentiation with respect to this vowel. This is confirmed in Figure 6.4 on page 177 and Figure 6.5 on page 178, where the two acoustic data sets are compared. In Figure 6.4 we note that the MOUTH value for the current data is more retracted than that of Webb (1983), while, in Figure 6.5, the current PRICE values are (slightly) more fronted than those provided in Webb (1983). Of the two vowels, however, the greatest difference is in relation to MOUTH. As was mentioned in §6.2.2, a retracted MOUTH is characteristic of GenSAE, with monophthongisation potentially threatening the distinction between MOUTH and BATH. However, for the current data at least, and as can be seen from Figure 6.3 on the next page, the nucleus of MOUTH and the value for BATH remain separate with the latter being more retracted than the former. Still, it would appear that the weak glide of MOUTH adopts a BATH quality, all of which suggests something [ɐɑ]-*like* as an appropriate transcription.

The relationship between PRICE and MOUTH, the relatively glide-weakened nature of both of these vowels, as well as their peripheral status in comparison with the short-vowel system are all illustrated nicely in Figure 6.6 on page 179. In particular, while the data shows no evidence of the PRICE–MOUTH Crossover, the nuclei of these vowels are almost certainly tense and thus on the peripheral track; they are the most open vowels of the system. This data, along with Webb's (1983) spontaneous-speech data, seems to indicate, that for GenSAE at least, PRICE and MOUTH are in a relatively conservative position

[27]I would argue, on a purely impressionistic basis, that this change in indexical value is mainly due to the stereotyping of the 'Kugel' accent more generally, an accent often associated with the *nouveau riche*, the wealthier 'Northern' suburbs of Johannesburg and the Jewish ethnic group. From this perspective I would imagine that the retention of a glide (however weakened) is meant to index dissociation *vis-à-vis* this group. This hypothesis and the indexical value of this feature requires further research.

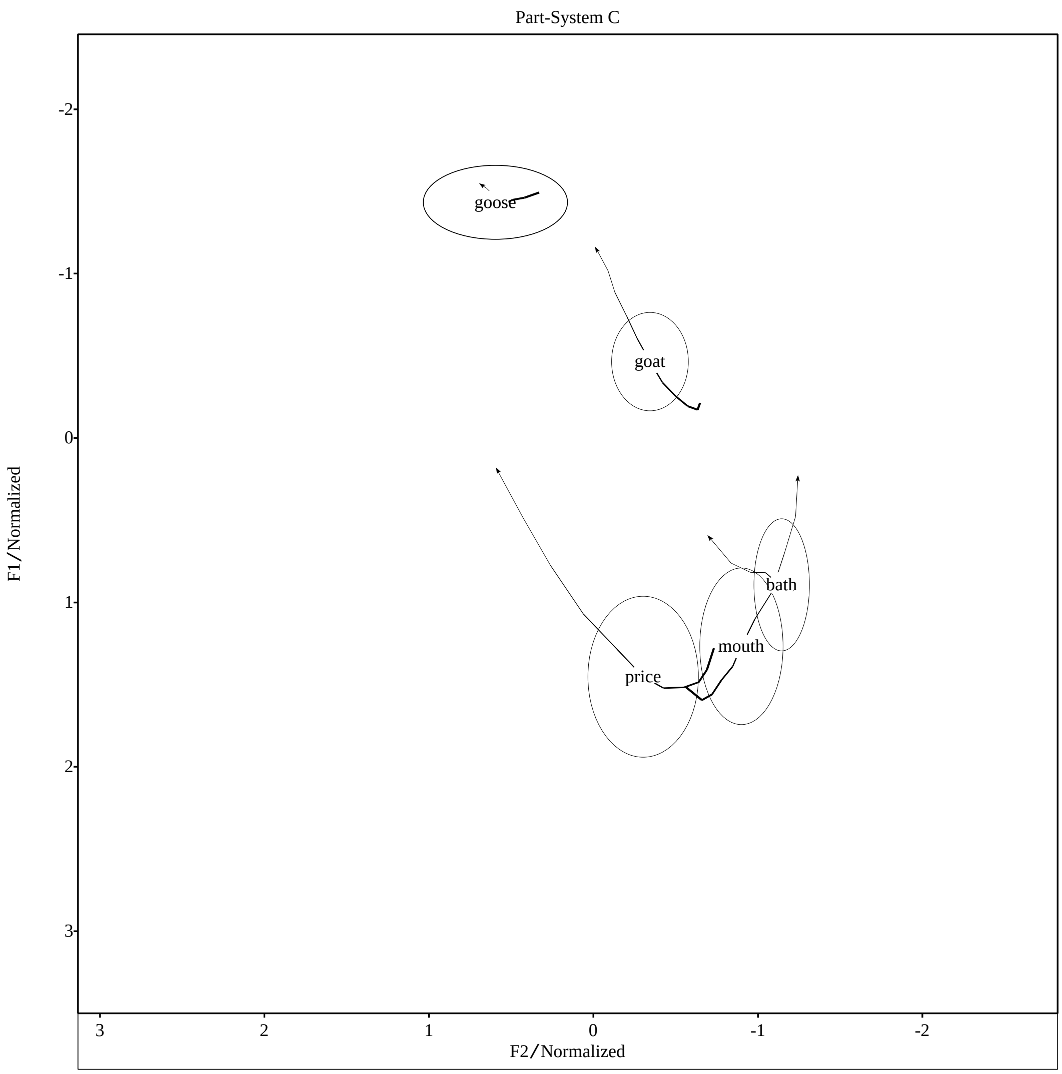


Figure 6.3: MOUTH: Comparison with PRICE and BATH

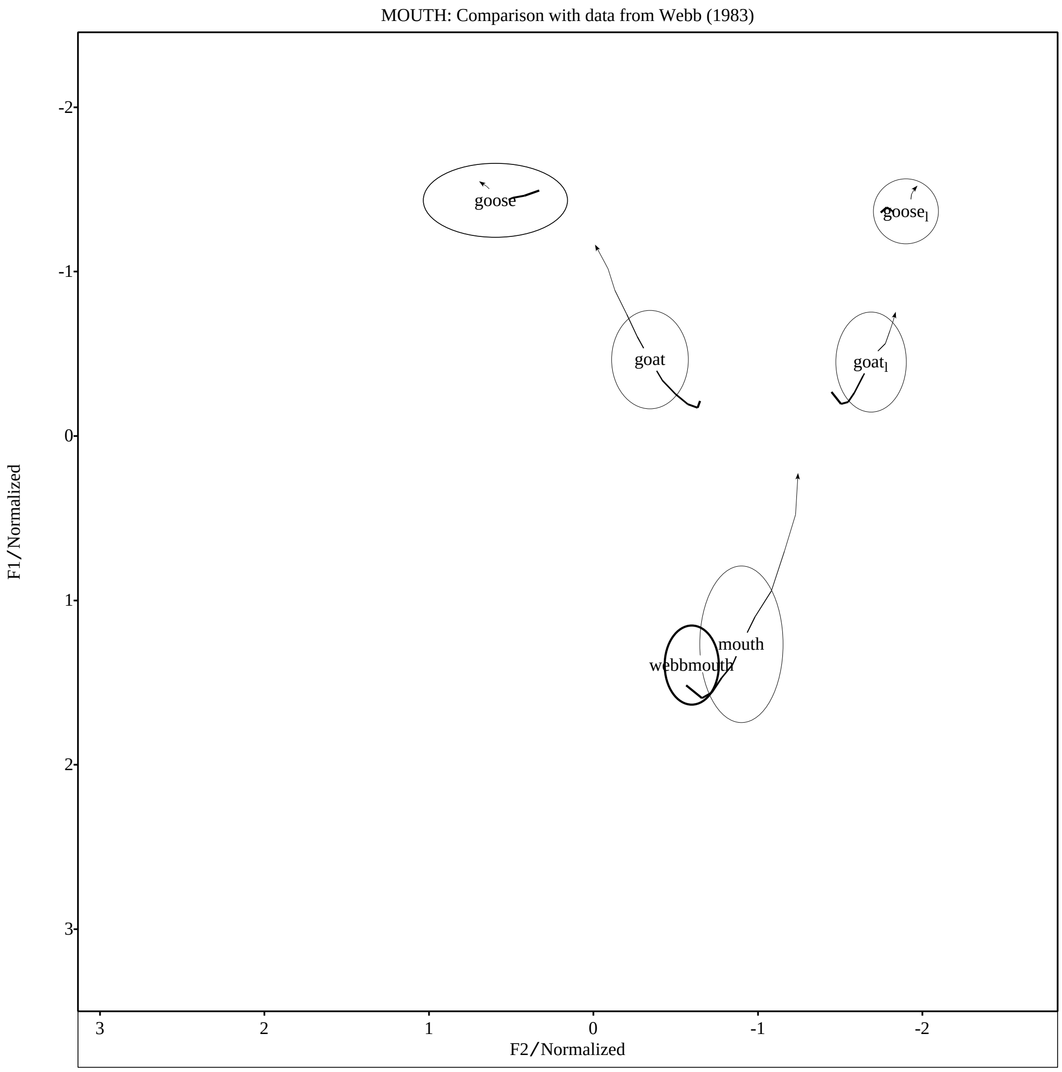


Figure 6.4: MOUTH: Comparison with data from Webb (1983)

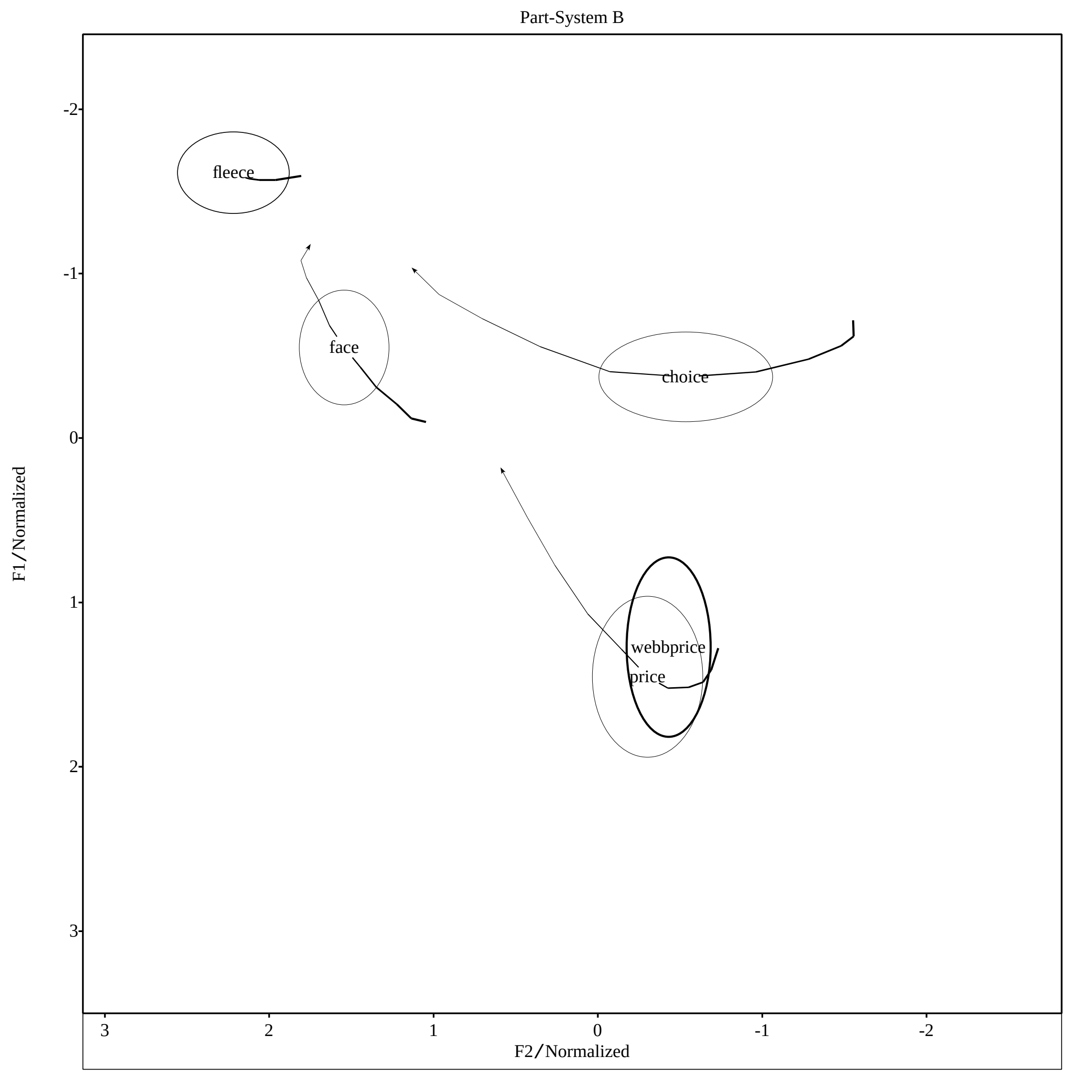


Figure 6.5: PRICE: Comparison with data from Webb (1983)

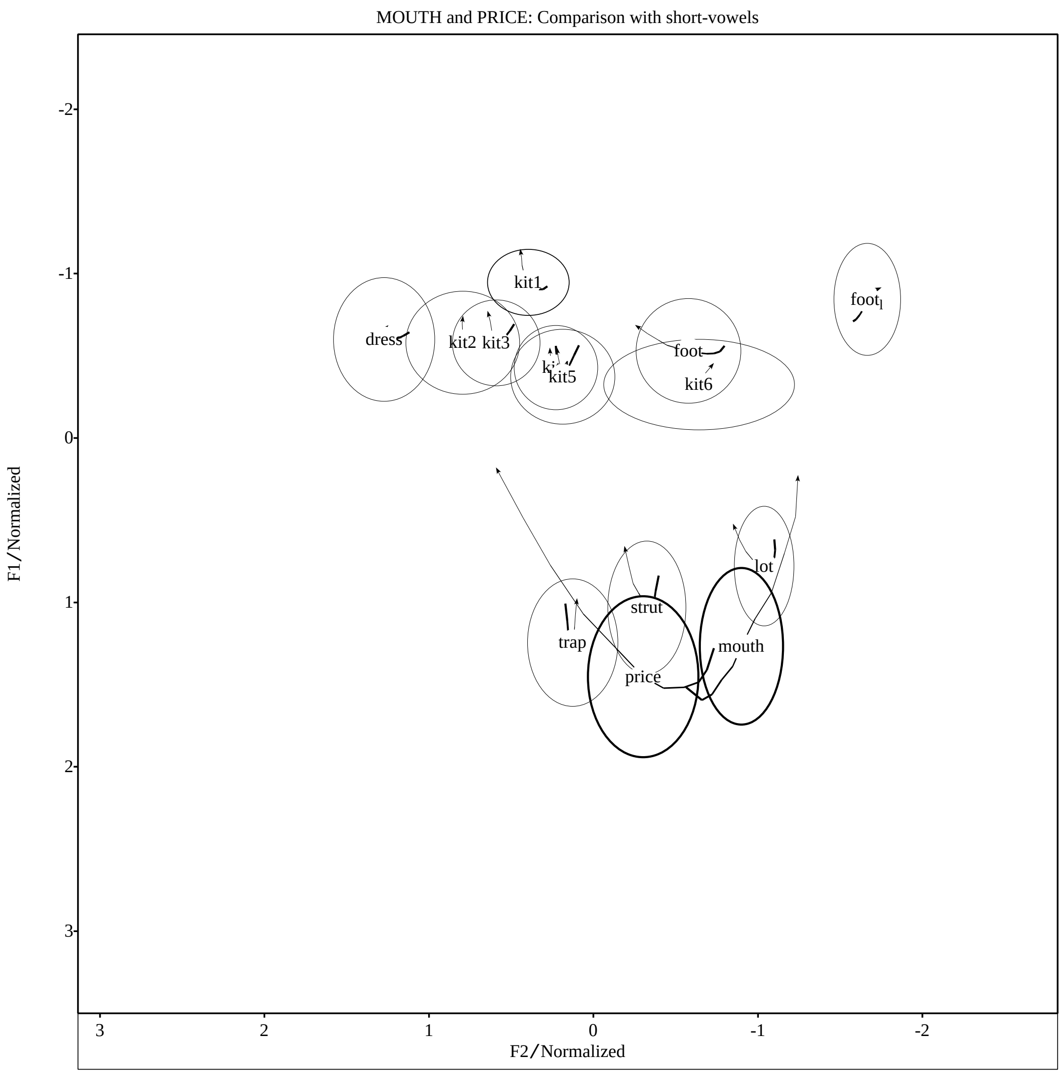


Figure 6.6: MOUTH and PRICE: Comparison with the short-vowels

vis-à-vis the Diphthong and the Southern Shift.

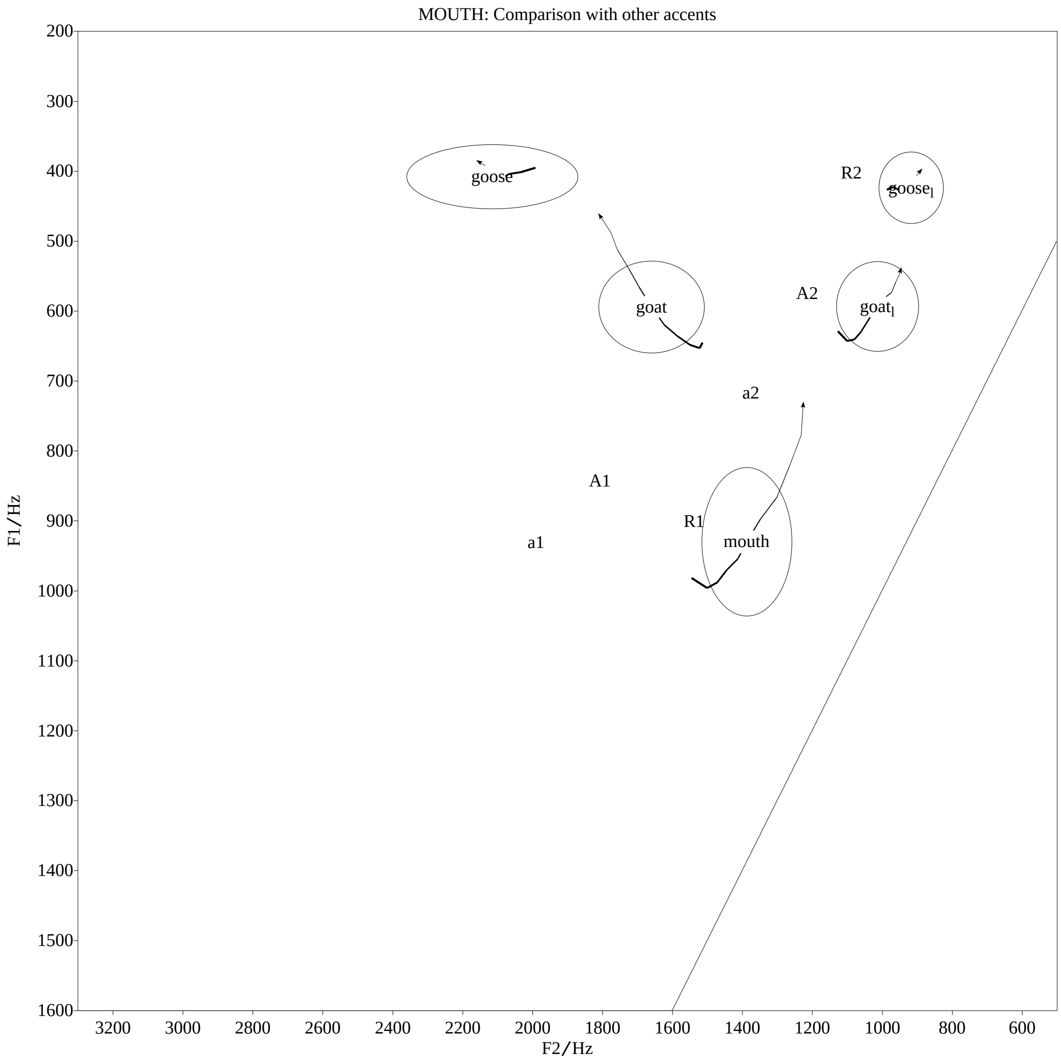


Figure 6.7: SAE MOUTH vs. other English accents

This basic observation is confirmed when we compare the SAE data with that of other accents of English, as provided in Figure 6.7. This representation draws on Cruttenden's (2001) RP data (**R** in the figure) as well as Butcher's (n.d.) and Cox's (2006) AusE data (**A** and **a** respectively). The reader is referred back to §4.6.2 for brief descriptions of each of these sources. In all cases the first diphthongal target has been labeled with a 1, the second target with a 2[28].

[28]The reader is reminded that when the SAE acoustic data is compared with that derived from research on

As can be seen, Figure 6.7 provides clear evidence of the glide-weakened nature of MOUTH in the SAE data compared to other accents of English, particularly RP. We notice too that the first target of SAE MOUTH is even more retracted than the RP citation-form one. In contrast, speakers of AusE show *substantial* fronting of this vowel, even in citation-form. As mentioned in §6.2.2, Wells (1982) has, for example, pointed out that SAE has been less susceptible to Diphthong Shift patterns than the other Southern Hemisphere Englishes.

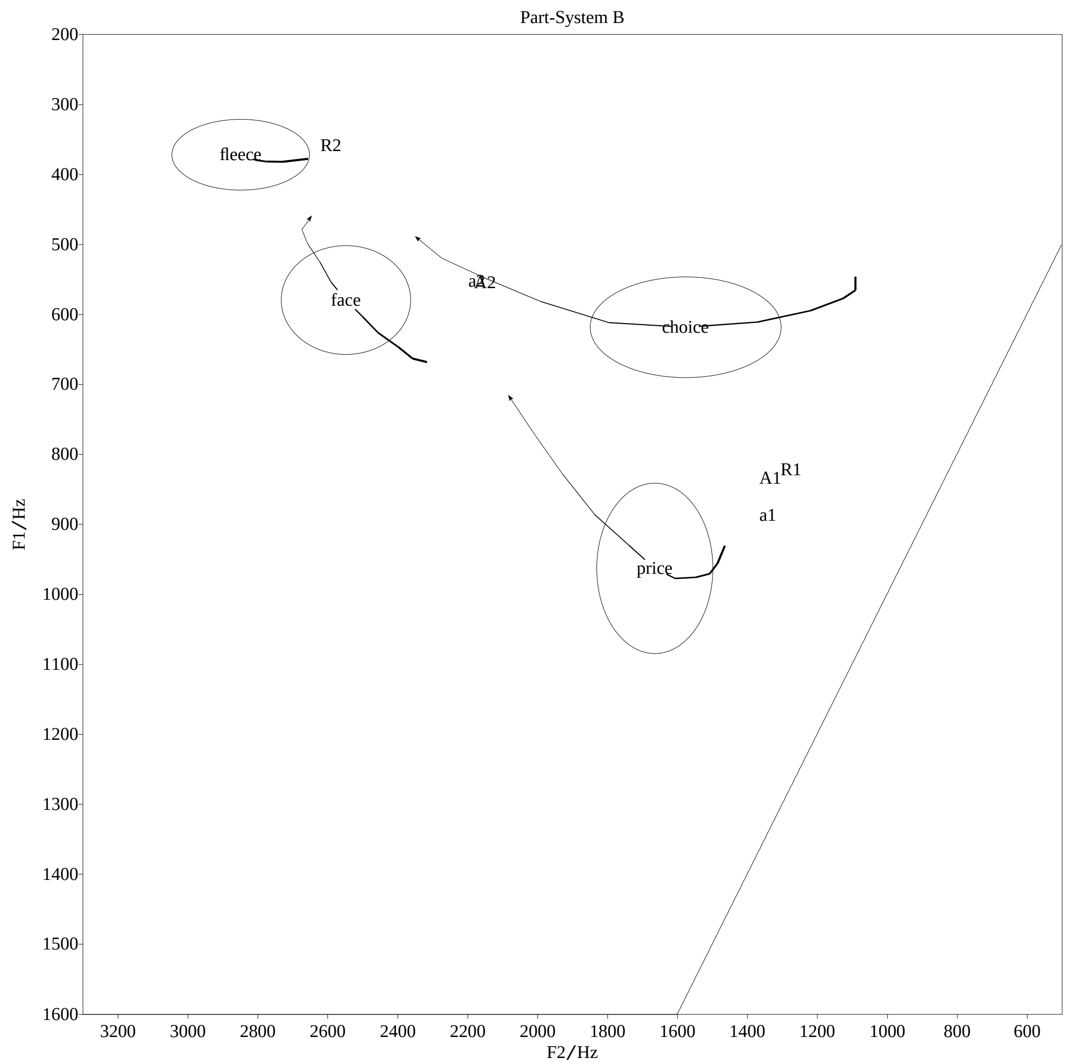


Figure 6.8: PRICE: Comparison with other accents

other accents of English, the 'raw' Hz values are used, and only from the speech of females. See §4.5.3 for more on this.

Figure 6.8 on the preceding page provides the comparison between the SAE PRICE data with that collected from other accents of English. The data sources are the same as for MOUTH[29]. Beginning with the first element of this diphthong, we note the particularly fronted nature of PRICE in SAE. It must be confessed that the RP value for the PRICE nucleus is a little surprising, although, as mentioned in §6.2.1, backed versions of PRICE are attested to in modern RP and this vowel has, in general, been undergoing a process of retraction. It is worthwhile noting, in this regard, that Cruttenden (2001:101), in characterizing the three females on which his RP diphthong data is based, does mention that one subject showed "influence from the London region". Returning to Figure 6.8, we note too, that in comparison with RP, and as expected, both AusE and (particularly) SAE are subject to a degree of glide-weakening; although hardly monophthongisation.

Figure 6.9 on the next page provides a comparison of CHOICE in SAE with the SAE short vowels as well as THOUGHT. As can be seen from this figure, CHOICE begins at a position just in front of THOUGHT (more accurately, from a position that overlaps FOOT before final /l/) and glides towards a position a little closer than DRESS. Since the first and last 15% of the vowel has been excluded from analysis (in order to exclude co-articulatory effects as far as possible), it is, of course, possible that small portions of the glide have been excluded from representation; still there is some suggestion here that the nucleus of CHOICE has moved onto the non-peripheral track. We note in this regard that CHOICE's 'point of inflection' (§2.2.4.2) is a F_1-maximum and not a F_2-minimum. A similar point was made *vis-à-vis* Webb's (1983) CHOICE vowel, as illustrate in Figure 5.5 on page 152, which also appears to have a slightly centralized value. Figure 6.10 on page 184 provides a comparison, with respect to CHOICE, between the current data and that of Webb (1983). As can be seen, Webb's (1983) CHOICE vowel is a little lower than CHOICE in the current data. This is also reflected in Figure 5.5 on page 152 in the sense that CHOICE is discernibly lower than THOUGHT. Whether this reflects an actual difference, though, is difficult to determine given that Webb (1983) bases his analysis on only three tokens, two of which are exemplars of the same word (*boys*, *boys* and *point*).

Figure 6.11 on page 185 provides a comparison between SAE CHOICE and CHOICE in other accents, the same data sources having been used as above for MOUTH and PRICE. It is interesting to note that the difference in first-target closeness between the two Australian data-sets has been commented upon by Butcher (n.d.:451) as possibly reflecting regional differences in AusE, essentially New South Wales (Sydney) vs South Australia. SAE CHOICE comes closest to the former, basically Standard Australian English. The position of RP *vis-à-vis* AusE and SAE is, however, surprising and runs against expectations. As far

[29]Note that in Figure 6.8 the second targets for the two AusE data sets overlap.

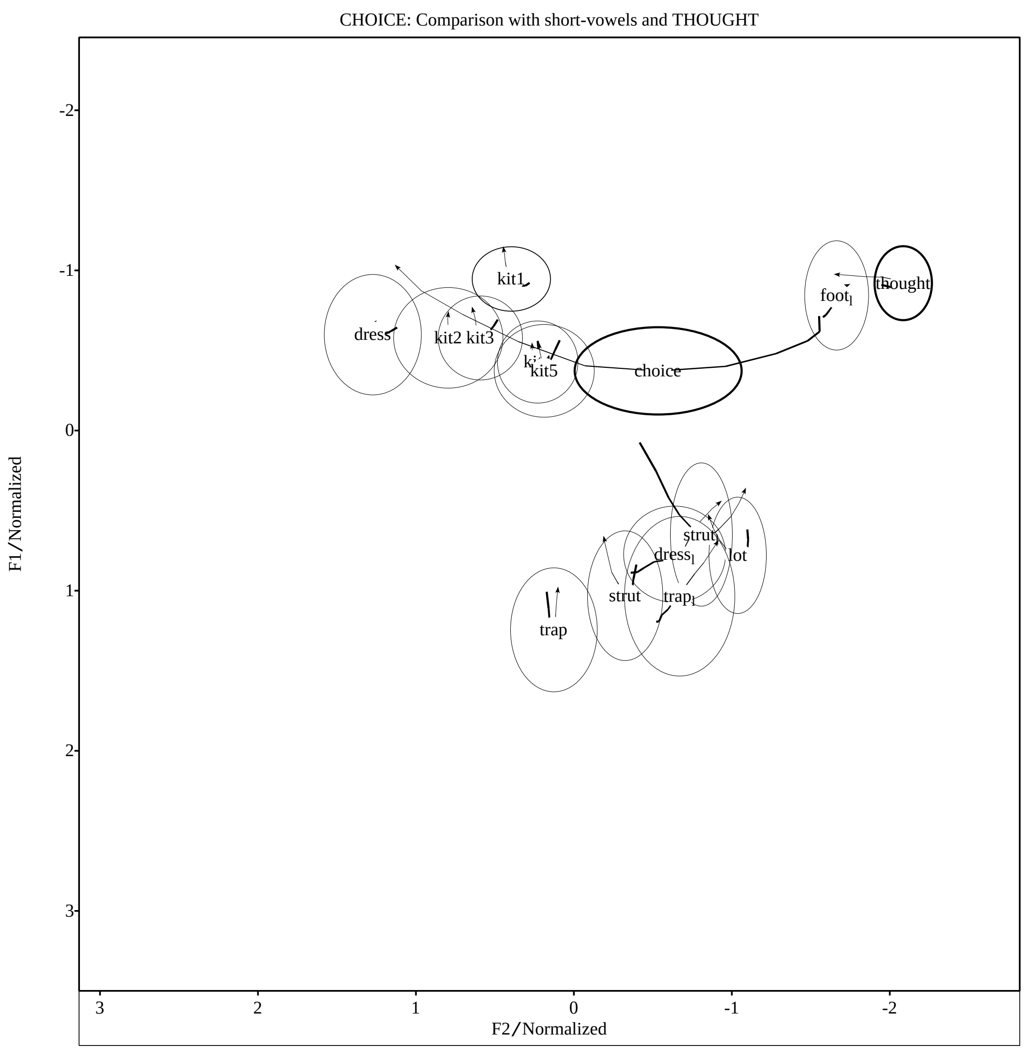


Figure 6.9: CHOICE: Comparison with Short Vowels and THOUGHT

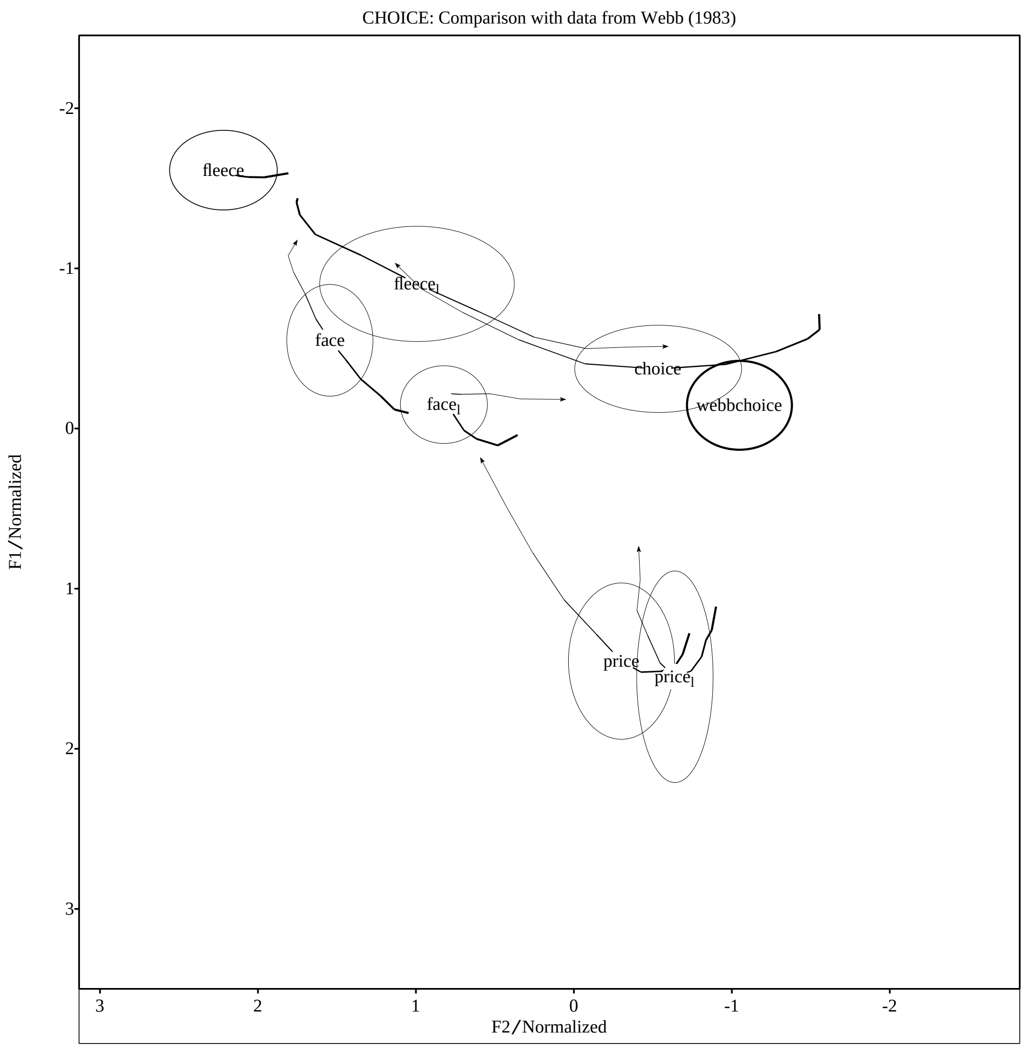


Figure 6.10: CHOICE: Comparison with data from Webb (1983)

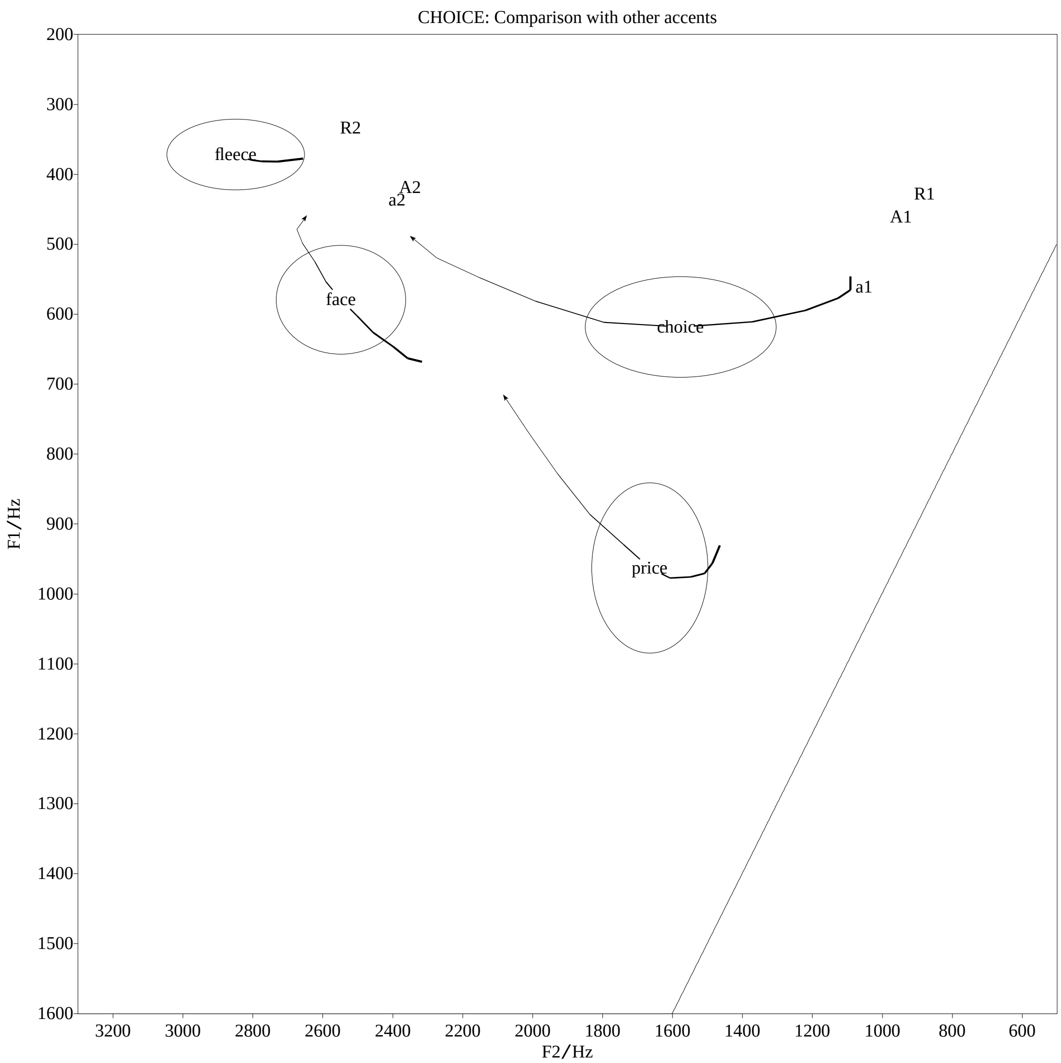


Figure 6.11: CHOICE: Comparison with other accents

as the offglide is concerned, there is also some suggestion that in SAE (and AusE) CHOICE is subject to minimal glide-weakening compared with RP.

The target values for both Cox (2006) and Butcher (n.d.)[30] were determined on the basis of targets constituting "the point of least formant change". Assuming that the excision of the first and last 15% of the CHOICE vowel in the SAE data was successful in ridding the analysis of coarticulatory effects but successful in retaining vowel-target data, then it appears, from Figure 6.11 on the preceding page, that the point of least formant change is far more centralized than in the other accents. This is further (tentative) evidence for a lax, non-peripheral CHOICE-vowel nucleus.

6.2.4 MPC: Synopsis

As suggested by Wells (1982), the fronting and raising of MOUTH, as well as the backing and raising of PRICE, as part of the broader Diphthong Shift, has not become as integral a part of SAE as it has, for example, in AusE. This is reflected very clearly in the analyses provided above. Furthermore, if we accept Lanham and Macdonald's (1979) thesis that there has been a *retreat* from a fronted MOUTH (as opposed to stable sociolinguistic variation) then from the analysis as a whole there appears to be strong evidence to suggest that fronted MOUTH in SAE has undergone an arrestment and retreat, with subsequent diffusion of the backed GenSAE variant at its expense, probably due to its stereotyped nature[31]. While Lass (1995) claims that a PRICE–MOUTH Crossover is found among some GenSAE male subjects, Webb's (1983) vowels overlap rather than cross over and the current data shows a retracted RP-like position for the nucleus of this vowel, although with glide-weakening. While MOUTH is not fronted it is, however, along with PRICE, a peripheral vowel in Labov's (1994) sense of this word.

With respect to PRICE, it appears that while fronted PRICE monophthongisation might

[30]The methodology employed in Butcher (n.d.) for vowel target determination was based on that provided in Cox (2006).

[31]While the slight difference between the current data and that of Webb (1983), as represented in Figure 6.4 on page 177, suggests that there might, in addition, be an incremental reverse shift of MOUTH from a more front to a more back position, there is not enough evidence to make any strong claims in this regard. Related to this is the subtle, but important, distinction between, for example, fronted MOUTH and *fronting* of MOUTH. While there has certainly been, in terms of overall usage in the SAE speech-community, a progressive decrease in the use of the former, this does not necessarily imply that endogenous MOUTH-fronting has completely ceased in those SAE speech-communities which identity with and consistently use BrSAE. The same considerations apply to PRICE and a number of other vowels. I believe, in this regard, that it is important to distinguish, along with Labov (2007), and as briefly touched on in §4.2.3.1, between diffusion (the spread of certain variants at the expense of others) and transmission (often leading to gradual, endogenously-driven, incremental change). As we will see, one possible case of shift or drift in current SAE, as discovered during the course of this analysis, is that of TRAP-Lowering and STRUT and LOT-Fronting; all possible participants in a SECS-like shift *a là* Torgersen and Kerswill (2004).

have constituted a prestige value in Lanham's (1978) and Lanham and Macdonald's (1979) time, this is perhaps no longer the case; in support of Lass' (1990) analysis, PRICE in this data (although subject to a degree of glide-weakening) shows definite diphthongal movement. PRICE-*fronting* does appear, however, to have remained a prestige feature and remains reasonably categorical in GenSAE i.e. seemingly regardless of style. We also note that as with MOUTH, Lanham and Macdonald (1979) claim a *retreat* from the BrSAE value. In this light a lowered, non-peripheral CHOICE vowel makes perfect sense and as seen above there is some tentative evidence to suggest that just such a phenomenon appears in at least some variants of (Gen)SAE.

Overall, therefore, we have evidence for a *reversal* of the PRICE–MOUTH Crossover in GenSAE, although glide-weakened variants have been retained. In the case of MOUTH, there has been a retreat from a front value, while a backed PRICE remains a stereotype of BrSAE and is in retreat from a fronted (whether monophthongised or not) PRICE vowel.

As pointed out in §3.4.3, fronted-PRICE monophthongisation/glide-weakening is not considered, by Trudgill (2004), to be part of the more general phenomenon of Diphthong-Shift-induced Glide Weakening, since, quite simply, the main condition for Glide Weakening (i.e. nucleus-glide differentiation), has not been met. This raises the question of why, in some variants of SAE, fronted PRICE has monophthongised or glide-weakened in the first place.

Some possible light is shed on this question if we view the development of GenSAE, and SAE more generally, in the light of the proposals contained in §3.3, in particular the notion of the 'birth' of modern SAE constituting a relatively protracted three-stage koinéization process culminating in the production, in early Johannesburg, of the sociolects that we know today. In particular, and in the light of this process, it would seem that more than any other variant in SAE, a fronted, monophthongised PRICE is possibly a vestige of north-of-England influence. Firstly, in the formation of GenSAE as a local standard, an avoidance of the CE fronted MOUTH and backed PRICE would have made sense, as would the adoption of the north-of-England fronted and monophthongised varieties of PRICE brought through NE, an adoption, perhaps, further influenced by lower-class north-of-England/Cornish variants of PRICE brought by the British-born immigrants to late 19th-century Johannesburg. Secondly, evidence has been provided above that an older RP variant of PRICE had a fronted nucleus, although it remained diphthongal. All in all, there appears to be enough historical evidence to suggest that current GenSAE values for PRICE are not the end-products of an endogenous chain-shift movement, but have been inherited from the linguistic feature-pool which constituted the linguistic 'mix' of early Johannesburg. These values have, according to this view, and in accordance with Lanham and Macdonald (1979), since diffused more broadly across the SAE speech-community.

More generally, however, one would be foolish to discount possible structural influences in the development-phase. As reported on in §2.3.2.1, a fronted, monophthongised PRICE vowel is viewed as one option for PRICE under a Pattern-4 chain-shift. The best solution is, therefore, to accept Mufwene's (2008) conceptualisation of the 'ecology' of a variety as including a variety of exogenous as well as endogenous (structural) constraints and influences. Under such a conception the 'forces' that led to the adoption of a fronted, monophthongised PRICE in early Johannesburg constitute the following:

1. An older, *fronted* (but diphthongal) RP value, in the region of [aɪ].

2. The influence of the north-of-England through NE, noting that NE speakers with a broader, local accent would, in many cases, have contributed a fronted, glide-weakened or even fully monophthongal PRICE, while NE speakers with a more standard accent would, presumably, have 'offered' the RP PRICE value given above.

3. The influence of north-of-England and Cornish miners.

4. An avoidance of the backed CE value, especially among certain Eastern European L2-speakers of English and others bent on social advancement.

5. A structural 'pressure' on the development of PRICE toward the front, open position, given that PRICE backing and raising was not (for many) a viable sociolinguistic option.

It is thus not surprising that the eventual GenSAE value was settled upon.

In addition, it would appear, in accordance with much of the literature analyzed above, that the L1-Yiddish speakers of early Johannesburg were particularly 'energetic' in their selection of a fronted and monophthongised PRICE vowel. As this value diffused socially and geographically across the SAE landscape and in the process moved from carrying first-order to third-order indexicality (§1.3.3), it may have become stereotyped and associated with Johannesburg and/or the Jewish ethnic group. Such stereotyping provides an at-least plausible explanation for the differences, as reported on in §6.2.2, between Lanham and Macdonald (1979) and Lass (1990) with respect to the 'direction' of PRICE style-shifting i.e. less glide-weakening in casual speech, more in formal contexts for Lanham and Macdonald (1979), more glide-weakening in casual speech, but less in formal contexts for Lass (1990).

With respect to MOUTH, on the other hand, there appears to be less evidence for a backed variant in the early-Johannesburg mix. While Wells (1982) provides some evidence for a monophthongal, backed variant in the north of England, Northern varieties would generally have had uncompleted-GVS variants, such as [u:t] for *out*, or diphthongs with front nuclei i.e. [aʊ] (Britain 2008) and neither is it clear that RP-variants that were brought to Johannesburg would have had the necessary influence, given that an earlier version of RP MOUTH appears to have been fronted. Perhaps the only real ecological factor was the need to adopt a value clearly different from the substantially fronted CE variant[32], a need which led to a hypercorrect retraction, a process which, it appears, led to a similar development in RP itself (although without glide-weakening). This is in alignment with Wells (1982:614; my emphasis), who refers to the reversal of the PRICE–MOUTH Crossover in SAE as a process of "counter-*correction*".

Thus, assuming, and as argued for in §3.3.1, that the Founder effect was substantially-weakened in the early-Johannesburg setting, the above reconstruction of the status of SAE MOUTH provides compelling evidence for the inadequacy of Trudgill's (2004) model of new-dialect formation which proposes that this formation proceeds, in tabula rasa contexts, on a developmental course that is insensitive to indexical forces. The adoption of a relatively-backed MOUTH by GenSAE cannot be accounted for in terms of simple demographics or in terms of endogenous development, but perhaps solely in terms of an indexicality-based need to avoid the stigmatized Diphthong-Shifted and fronted variant. In the case of a backed MOUTH, therefore, we have a variant that was, for all intents and purposes, not present in any numerically significant way in the historical input and is, in addition, a 'movement' *away* from some well-known endogenous shifts and at least one general principle i.e. the PRICE-MOUTH Crossover, the Diphthong-Shift and Labov's (1994) Principle I. With respect to the dating of a backed MOUTH in GenSAE, things are, admittedly, a little less convincing, but we note that while Hopwood (1928) claims an Afrikaans origin for backed [ɐʊ], there is no reason why this might not have constituted an early hypercorrection away from the CE-derived [æu], targeting but overreaching Hopwood's (1928) RP [a̠ʊ]. It is clear that by Hooper's (1945) time a fully-backed [ɑʊ]-variant was a prestige-norm in SAE. While the existence in the linguistic 'feature pool' of, for example, two relatively equally represented alternatives can lead to a reallocation of each variant to a different sociolect[33], in the case of MOUTH the backed variant was, for all intents and purposes, not present in the linguistic mix at all and certainly did not constitute a numerically 'competitive' feature in Trudgill's (2004) terms. In certain cases, such as the prevalence of the Weak Vowel Merger in Southern Hemisphere Englishes (§9.3), Trudgill

[32]We note in this regard Hopwood's (1928) value for RP MOUTH as [a̠ʊ], while his Cockney-derived value is the even fronter [æu] or [aˑ(u)].

[33]For example, FACE in SAE, with a 'narrow' variant being targeted by GenSAE and the open variant by BrSAE, a fact reflected in style-shifting in GenSAE as well, as reported on in §6.4.

(2004) makes recourse to the unmarked status of a 'minority' feature in order to account for its eventual adoption into the new dialect. It is unclear, in this case, how a backed MOUTH would, in any revealing sense, constitute a less marked variant than a front one.

Lastly, whether a lax, non-peripheral and lower CHOICE vowel exists in GenSAE, and, if so, whether it was selected as part of the koinéization process mentioned above, or as part of a subsequent, perhaps structurally-motivated shift initiated by the diffusion of fronted PRICE, is unclear at this stage and requires further research and clarification.

6.3 TRAP

While TRAP is not represented in Figure 2.13 on page 61, it is, as pointed out in §2.3.2.1, an integral part of the Southern Shift in the sense that a raised TRAP vowel is often found 'alongside' the raised DRESS and KIT vowels that form a more central role in the Southern Shift. From this perspective, a raised TRAP vowel should be common to all 'Southern' accents including those of the south of England and the southern USA. Importantly, however, there is some evidence to suggest that the raised short vowels are, in most of the Southern Hemisphere Englishes, in actual fact 'relics' of the historical input and have not participated in any post-settlement raising. The exception to this observation is NZE, which *has*, it appears, undergone 'secondary', endogenous raising of both DRESS and TRAP. In addition, and as was pointed out in §2.14, a *lowered* TRAP vowel appears to be the initiating movement of SECS, a shift of the short vowel system which appears to be a contradiction of, or at least a subsequent development from, Labov's (1994) Pattern 4.

With specific reference to American accents, the status of TRAP is complicated by the fact that, in general, these accents have not been subject to the more advanced Lengthening I processes discussed in §3.4.1 and §3.4.3 i.e. *trap* and *bath* are often distinguished by length alone: [tɹæp] and [bæːθ]. As pointed out in §2.3.2.1, however, in some American accents TRAP and BATH have essentially merged into one phoneme, /æh/, now subject to tensing and raising in accordance with Labov's (1994) Pattern 2; what is known as the 'Northern Cities Shift'; a process which also ultimately involves the lowering and/or centralization of the other front short vowels, among them KIT and DRESS. As explained further in §6.3.1, other U.S. accents, on the other hand, maintain a lexical split between a lax /æ/ and a tense /æh/, both of which do not correlate completely with the TRAP and BATH lexical sets.

6.3.1 TRAP: Other Accents of English

Received Pronunciation: For RP, Lanham (1967:33) gives [æ] with a slightly retracted and lowered [ä̞] when followed by /l/. Lass and Wright (1985:138) provide [æ] as the typical RP value. A reading of Wells (1982:291–2) and Cruttenden (2001:111) indicates, however, that this vowel has undergone change in RP i.e. a progressive *lowering*. Thus Wells (1982:291; my parenthesis) refers to "the newly current [a] [which] is perceptually very similar to fronted realizations of /ʌ/ which have been around in RP for much longer". The same author mentions a chain-shift involving the gradual lowering of KIT, DRESS and TRAP and claims that this process "is a change which promises to carry RP further away from both American and southern-hemisphere accents of English" (Wells 1982:292). Cruttenden (2001:111) explains as follows:

> "This vowel has become more open recently, previously being nearer to C[ɛ] where now it is now [sic] close to C[a]. Only tradition justifies the continuing use of the symbol 'æ' for this phoneme. Since the vowel /ʌ/ has had a tendency over a somewhat longer period to move forward towards C[a], this may occasionally result in a neutralisation of /æ/ and /ʌ/. More often, however, the lowering of /æ/ has resulted in a retreat of /ʌ/ towards the central region".

We note, therefore, that the value for TRAP in RP is [æ] to [a], depending on the age of the speaker. Of interest in this regard is a recent study by Fabricius (2007*a*), who attempts to trace the changes of TRAP and STRUT (and their relationship) across several generations of RP speakers[34]. This author shows how the positions of TRAP and STRUT *vis-à-vis* each other have changed in a rotation movement such that "while TRAP lowered and backed during the 20$^{\text{th}}$ century, STRUT rose and fronted" (Fabricius 2007*b*:1477). According to Fabricius (2007*b*:1480) it appears that TRAP lowering in RP was already underway in the early 20$^{\text{th}}$-century.

It is important to note in this regard that, according to Cruttenden (2001:111), "speakers of Refined RP and older speakers of General RP generally have a closer variety of /æ/ almost at the level of C[ɛ] which may also be diphthongized to [ɛ^{ə}]".

London (Cockney) and Other Accents of England: It appears from a reading of Torgersen and Kerswill (2004) that TRAP-lowering has also played a role in London i.e.

[34]STRUT is dealt with further in §8.4.

Cockney. In fact, for the greater south-east of England, lowered TRAP variants appear to be part of Torgersen and Kerswill's (2004) SECS-Shift, as illustrated in Figure 2.14 on page 64. It is worth repeating that TRAP-lowering is regarded as the initiating move of this chain-shift. Trudgill (2004:37–48) essentially confirms this analysis of the accents of the southeast of England, but distinguishes Modern Cockney from other accents of the southeast, claiming that in the former only TRAP and DRESS have lowered (with KIT remaining at [i]) while in the latter accents all three front vowels have been affected by this drag-chain shift. This is, furthermore, in contrast with, for example, "current East Anglian Traditional-dialects [which] demonstrate the first stage, where only /æ/ has been lowered" (Trudgill 2004:43; my parenthesis). As emphasized in §3.4.3, the contrast between these lowered values as opposed to the close values characteristic of the various Southern Hemisphere Englishes, including SAE, has been advanced by Trudgill (2004) as an essential argument for the conclusion that Southern Hemisphere values of TRAP etc. represent relics of the original values found in the 19th-century accents of the southeastern parts of England. This is in explicit contrast to the view that the Southern Hemisphere accents all underwent an innovatory raising of these vowels[35]. It should also be noted in passing that north-of-England accents are characterized by a [a] or a slightly more central value for TRAP (Wells 1982:356).

AusE, NZE and Falklands Island English: As mentioned above in §3.4.3, the two main non-SAE Southern Hemisphere Englishes, AusE and NZE, have a raised TRAP vowel, usually viewed as being a member of a chain-shift involving the raising of DRESS and TRAP as well as the raising (or centralization) of KIT. Thus, Gordon, Campbell, Hay, Maclagan, Sudbury and Trudgill (2004:26) give [ɛ̝] as the most common value for NZE, with [e̞] and [ɛ̞] representing the most advanced and conservative variants respectively. As is the case with DRESS (§6.4), Lass (2004:374) claims a (further) raising of TRAP in NZE as being a recent innovation and not due to the original input[36]. As confirmed in Evans and Watson (2004:195), TRAP is lower in AusE than in NZE, a fact which appears to be explained by both secondary raising of NZE TRAP and (recent) *lowering* of AusE TRAP (Evans and Watson 2004:195). This more recent lowering of AusE TRAP is of interest given the abovementioned view, found most prominently in Trudgill (2004), that Southern Hemisphere Englishes are more conservative than their 'mainland' cousins e.g. Cockney. From this perspective, the

[35] Although, as we will see in the next section, this is at least partly true of NZE.
[36] Also see Torgersen and Kerswill (2004:29) in this regard.

lowering currently evident in AusE simply mirrors (or rather initiates) a process that modern accents of London and southeast of England have already gone through. It is interesting to note in this regard that, according to Trudgill (2004:43), Falklands Island English has also undergone the initiating move of Torgersen and Kerswill's (2004) SECS-shift i.e. TRAP-lowering.

U.S. Accents: As mentioned in §3.4.1, most U.S. English accents generally only have a length distinction between the TRAP and BATH lexical sets; the two vowels are generally transcribed as /æ/ and /æh/, although it is important to note that TRAP and BATH are not the equivalent of /æ/ and /æh/. In those accents which still maintain the lexical split, many TRAP words (particularly before voiced consonants as in *bad*) have 'fallen in' with /æh/ and its general tendency to tense and develop an in-glide. This raising of /æh/ is, of course, a critical movement in Labov's (1994) Pattern 2 (§2.3.2.1) and the so-called Northern-Cities Shift. Thus, for example, in Detroit "the *entire* short **a** word class is tensed to /æh/ and undergoes fronting and raising" (Labov 1994:99; my emphasis). In the Southern U.S.A. it appears that the lexical split between /æ/ and /æh/ has been maintained with the short /æ/ subject to the general Pattern 4 movement.

6.3.2 TRAP in SAE: The Impressionistic Data

Hopwood (1928), Breckwoldt (1961): For SAEP TRAP, Hopwood (1928:32) claims "**æ(˙)** > **ɛ(˙)** ... a tensifying of the vowel, found in Cockney EP, and in AEP[37], owing to substitution of the nearest articulatory position; *e.g.* ... cat **kæt** > **kɛt**". Breckwoldt (1961:7) gives a similar [ɛ] value.

Lanham and Traill (1962), Lanham (1965): Lanham and Traill (1962:22), surprisingly, claim [æ][61], basically [æ], for SARP TRAP, and posit little or no variation. In contrast to Hopwood (1928), the authors assert that this vowel "has a *laxer* articulation that in RP" (Lanham and Traill 1962:22; my emphasis), although admittedly Hopwood (1928) deals, in effect, with BrSAE, while Lanham and Traill (1962) are focussed on GenSAE and CulSAE. The analysis in Lanham (1965:94) is more conventional: the author mentions the raised tongue position for the DRESS and TRAP vowels, although claims that a raised TRAP is not as widespread as a raised DRESS vowel.

Lanham (1967; 1978), Lanham and Macdonald (1979), Lanham (1982): Lanham (1967:66; 67) gives [ɛ][51/61]{[ɛ̞]}, sometimes as high as [ɛ][51]{[ɛ]}, as the unmarked allophone

[37] Afrikaans English Pronunciation.

of TRAP in SAE, with [æ]^{63}{[ɐ̝]} being the value of the vowel before tautosyllabic /l/. The author claims that "in some idiolects /æ/ replaces /e/ before /l/" (Lanham 1967:81). An even higher [ɛˆ]{[ɛ̝]} is given by Lanham (1978:152) as the most advanced or Broad value for this vowel: "High index scores correlate strongly with lower class in English cities" (Lanham 1978:152). Lanham and Macdonald (1979:46) confirm a raised TRAP as a GenSAE variable and Lanham (1982:339) emphasizes the link between a raised DRESS and a raised TRAP, presumably by means of a chain-shift.

Bailey (1984), Jeffery (1982): As already discussed in §3.4.2, Bailey (1984:11–3) spends some time discussing lengthening of TRAP in certain contexts and, more intriguingly, the possible breaking of TRAP before [ŋ]. While I have certain reservations regarding the accuracy of the relevant analyses, it does highlight the intriguing possibility that TRAP raising (and tensing?) constitutes an element in a Pattern 2-type shift in SAE (or in one of the main inputs into SAE), as originally suggested by Jeffery (1982).

Lass and Wright (1985; 1986), Lass (1990; 1995; 2004): Lass and Wright (1985:137) confirm the considerably raised status of TRAP in SAE, giving the characteristic value as [ɛ̜]{[ɛ̞]}, the lowering and centralization of this vowel before final /l/ (Lass and Wright 1985:155) as well as a less-raised variant in CulSAE (Lass and Wright 1985:156). With respect to the possible origins of a raised TRAP vowel, according to Lass and Wright (1986:209), "raising of ME /a/ is well-attested from the 15th century onwards . . . and reports of raised qualities in the London area occur as early as the late 18th century. We find it . . . represented in spellings like *keb* 'cab' as a stereotypical Cockney feature by 19th century writers".

While Lass (1990:276) claims that the normal value for SAE TRAP is "rather more than half-way between CVs 2 and 3, and somewhat retracted" (thus [ɛ̝̈]), the more posh varieties (and, presumably, styles) are claimed by the author to have an [æ̣]{æ̝}-like value. Before dark /l/(i.e. [ɫ]) this vowel is normally the centralised and low [æ̜̈][æ̞̈] e.g. in *pal*. As mentioned in §3.4, this vowel is subject to lengthening before voiced stops and nasals (except /ŋ/). The result is a long, centralised [æ̈ː], "often with a slight [ə]-like off-glide" (Lass 1990:276)[38].

Lass (1995:98) confirms TRAP as an important social variable in SAE and claims that while it is sometimes slightly higher than RP /æ/, it never approaches [ɛ] in CulSAE or GenSAE. The latter is, of course, the BrSAE value and is the focus of much

[38]Note that one of the characteristic features of vowels which are in the process of being raised along Labov's (1994) peripheral track is the development of an in-glide.

imitative stereotyping. As far as TRAP before final-/l/ is concerned, Lass (1995:98) confirms its retracted and lowered status, “except, curiously, in some Respectable pronunciations of the word *Natal*, where it raised ([nætɛɫ])”.

In Lass (2004:376), TRAP in SAE is characterised as highly variable and between [a] and [ɛ] for CulSAE and GenSAE, although usually closer to the later, with [æ] being the most typical value for GenSAE speakers. Any value above [ɛ] is likely to be part of a Broad accent and is the subject of much stereotyping. A raised TRAP is usually “unsurprisingly in conjunction with a closer DRESS (a secondary partial chain shift as it were)” (Lass 2004:376).

Wells (1982), Branford (1994), Bowerman (2004), Da Silva (2007): Wells (1982:613) gives [æ] for conservative SAE accents; it is only in BrSAE that it becomes “cardinal 3 or closer” (Wells 1982:613). Branford (1994:474) confirms this assessment, as well as the link between the Broader raised values and “London usage of the early nineteenth century” (Branford 1994:477). Bowerman (2004:937) gives a vowel slightly raised above [æ] as being the canonical value for non-Broad SAE accents. BrSAE TRAP is raised to [ɛ]. Da Silva (2007:86), in her review of the available SAE literature, stressed that “the raised [æ] may not necessarily be an innovation but rather an antiquation” i.e. a relic of the original input into SAEP.

Bekker and Eley (2007): More recent (acoustic) evidence (Bekker and Eley 2007) suggests that the TRAP vowel might currently be subject to lowering, especially among female speakers. As hinted at by Lass (1990:277–8; my parenthesis), “the lower [æ̣] {[æ̞]} types in more standard speakers suggest a movement away from the vernacular norm [i.e. [ɛ]], and we probably have here a case of ‘secondary’ (British-influenced) lowering”. As mentioned above, a similar process has been noted in AusE (Evans and Watson 2004:199), although the most accepted value in AusE, as in Cockney, is around C[ɛ] (Cruttenden 2001). Lass’ (1990) claim regarding British influence seems unlikely, however, particularly in the light of the SECS data (§2.14) and Trudgill’s (2004) analysis regarding colonial lag. From this perspective, at least, any lowering would have a decidedly endogenous basis.

6.3.3 TRAP: A Comparative Acoustic Analysis

Webb’s (1983) acoustic data, as given in Figure 5.4 on page 151, shows a TRAP vowel with a fair degree of internal variability. Webb (1983:145) concludes as follows:

"/æ/ has 4 allophones, viz. a sharply retracted [æ] before dark /l/, a tensed [æ] before nasals, a slightly raised [æ] before $\left[\begin{smallmatrix}-son\\+cor\end{smallmatrix}\right]$, and the [æ] elsewhere".

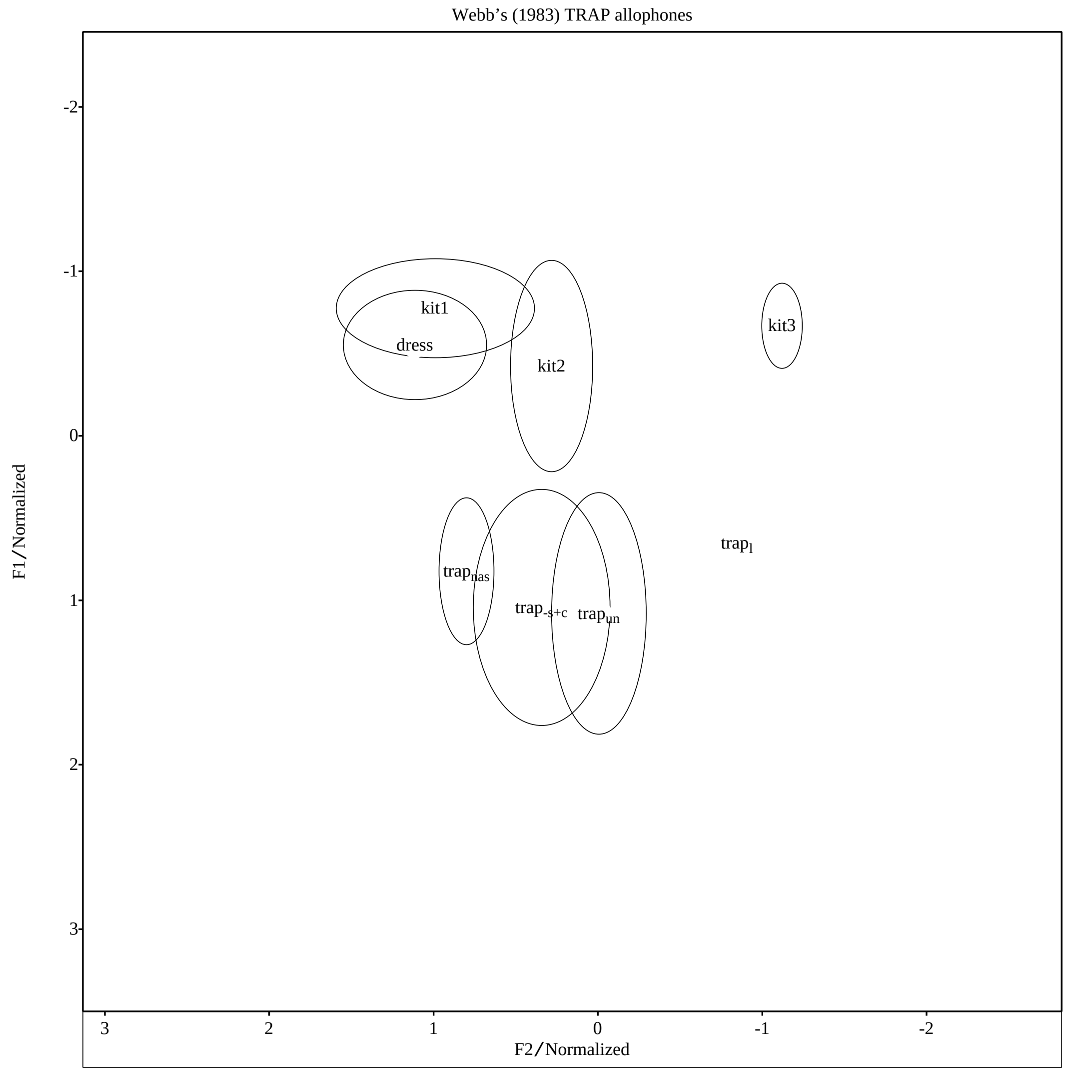


Figure 6.12: Webb's (1983) TRAP allophones

Figure 6.12 provides a visual representation of Webb's (1983) four TRAP 'allophones'. While there is certainly evidence for retraction before final /l/, any conclusion relating to the other three putative allophones remains highly speculative given the degree of overlap, degree of variability and number of tokens analyzed (23 for the remaining three categories). The possibility of a tensed, peripheral TRAP vowel before nasals remains suggestive, however.

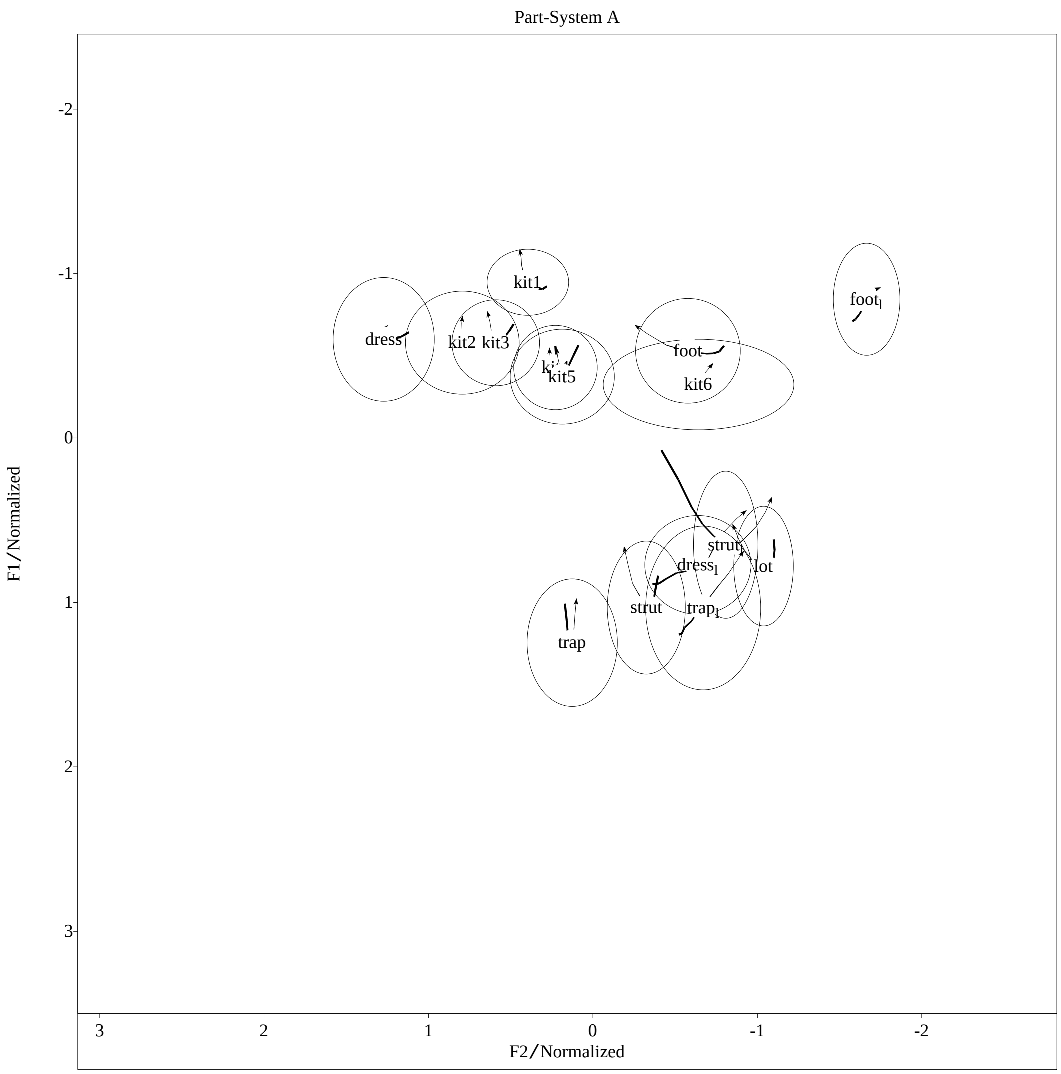


Figure 6.13: Overall Results: Part-System A (27 subjects)

Lexical Set	**No. of Tokens**	**No. of Measurements**	**N-G Tokens**	**O-G Tokens**
TRAP	301	6622	*had*, *lass*, *has*, *back*, *bat*, *bad*, *ban*, *cap*, ***ra**bbit*, ***a**bbot*	*bat*, ***ta**xes*, ***ta**xis*, *bad*, *van*, *back*, *bag*, *bang*, *hat*, *had*, *mad*, *ham*, *lap*, ***sa**lly*
TRAP before /l/	27	594	*pal*	*pal*

Table 6.2: Summary of TRAP data

With respect to the data collected for the current project, Figure 6.13 on the preceding page provides a clear picture of the position of TRAP vs. the other SAE short vowels. This 'picture' of TRAP is based on data provided in Table 6.2.

From Figure 6.13 on the preceding page, we note that any spectral movement appears to be co-articulatory in nature. There is no evidence, in this figure, for diphthongal movement proper, i.e. VISC. Of more interest, however, is the slight overlap between the standard deviations of STRUT and TRAP.

Before dealing with the relative positions of TRAP and STRUT however, it is worth noting any differences between the current data and that of TRAP in Webb (1983), as originally depicted in Figures 5.4 on page 151 and 6.12 on page 196. In Figure 6.14 on the next page the two data sets are compared. While there is some overlap between these two sets of data, the raised position of Webb's (1983) data compared with the current data presents no surprise, given the fact that the former was produced by a male and in spontaneous speech. It is only natural that the data should present a somewhat more raised and thus Broad value. The larger standard deviation of Webb's (1983) data also needs to be noted, however. As discussed in §6.3, and as depicted in Figure 6.12 on page 196, Webb (1983) posits four separate allophones for TRAP in SAE: TRAP before final /l/, TRAP before nasals, TRAP before $\left[\begin{smallmatrix}-son\\+cor\end{smallmatrix}\right]$ and an unmarked TRAP allophone; all of which accounts, presumably, for the large degree of variation present. The fronted value for TRAP before nasals is, presumably, also meant to indicate the possibility of a tensed and peripheral TRAP vowel and thus evidence for 'short **a**'-raising *à la* Labov; a theme also taken up by Jeffery (1982) and Bailey (1984), and as discussed in §6.3.

The standard deviation for TRAP in the present data is, though, demonstrably less;

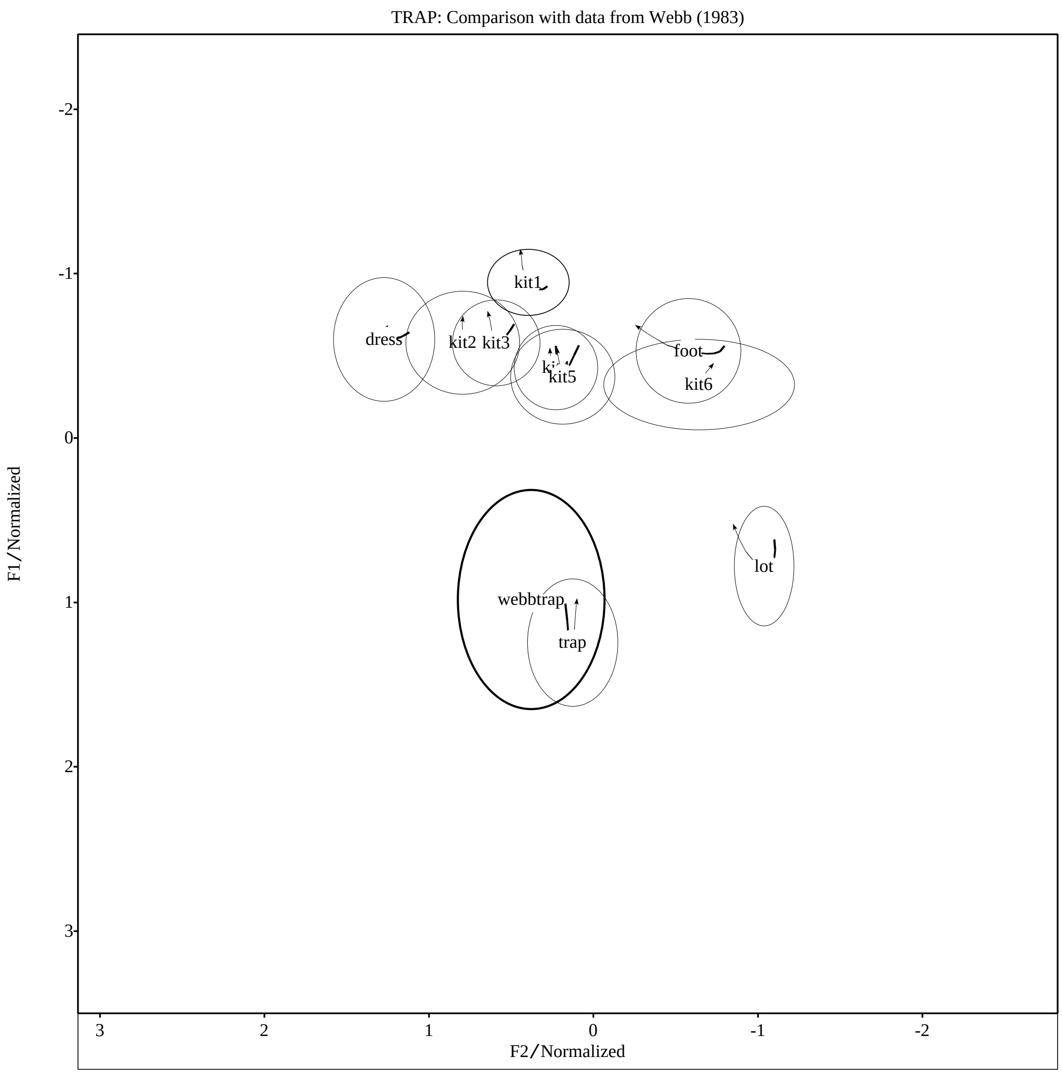


Figure 6.14: TRAP: Comparison with Webb (1983)

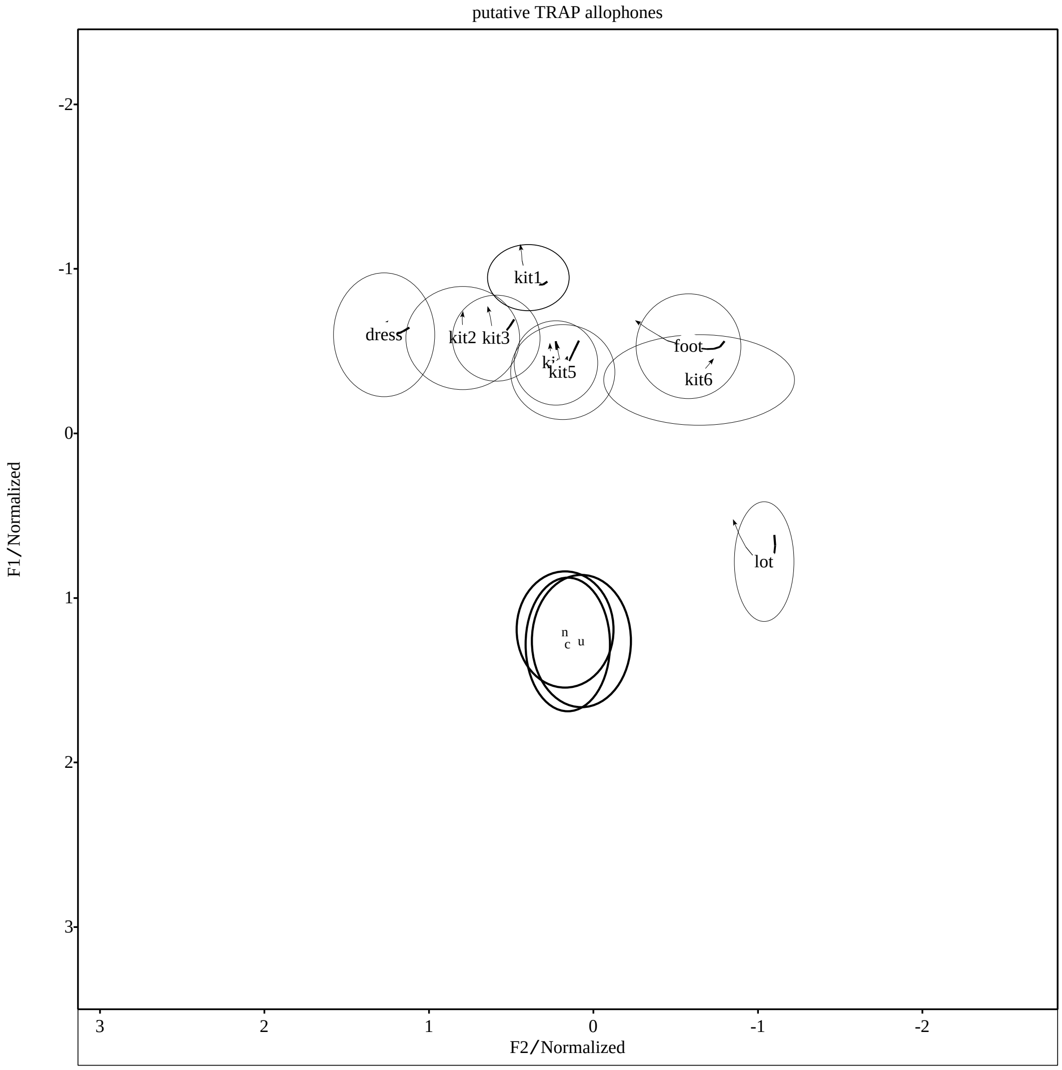


Figure 6.15: TRAP: Putative allophones

‘Allophone’	Tokens	No. of Tokens
Unmarked	***a**bbot*, *back*, *bag*, *cap*, *lap*, ***ra**bbit*, ***sa**lly*, ***ta**xes*, ***ta**xis*	123
$\begin{bmatrix}-son \\ +cor\end{bmatrix}$	*bad*, *bat*, *had*, *has*, *hat*, *lass*, *mad*	133
Before nasals	*ban*, *bang*, *ham*, *van*	44

Table 6.3: Putative TRAP Allophones

thus indicating, perhaps, less allophonic variation. This is confirmed in Figure 6.15 on the preceding page where there is obviously no evidence for any allophonic variation in the current data. In Figure 6.15 ‘u’ indicates unmarked TRAP, ‘c’ TRAP before non-sonorant coronals and ‘n’ TRAP before nasals. The graphic representations are based on data derived from the midpoint of each vowel. The various ‘putative’ allophones have been based on the data outlined in Table 6.3.

In §3.4.2, and also briefly in §6.3, Bailey’s (1984) observations about possible tensing of TRAP, particularly before a final /ŋ/, are mentioned. Fortuitously, at least one token in the current research met the required condition i.e. *bang*. Unfortunately, however, only O-G subjects were asked to read this token, leaving only 9 tokens of this word. As Figure 6.16 on the next page clearly shows, there is not much difference between TRAP generally and TRAP before final /ŋ/, at least in terms of the position of its mid-point. What little raising there is easily attributable to chance especially given the low number of tokens. What *is* interesting, however, is the direction of the spectral movement i.e. for most TRAP tokens the spectral movement is from ‘out’ to ‘in’; whereas in the case of TRAP before final /ŋ/ there is movement (and larger movement than for TRAP in general) from a central position to a more ‘peripheral’ and close one. This is no doubt attributable to the influence of the following velar sound, as pointed out by Bailey (1984:13) himself; whether one would want to call this ‘breaking’, however, is another story.

Moving on to the comparison between SAE TRAP and TRAP found in RP and the other Southern Hemisphere Englishes, Figure 6.17 on page 203 provides the necessary comparison. In Figure 6.17 one notes the lowered position of RP TRAP (‘R’ and ‘r’), the closer position of AusE TRAP (‘A’, ‘a’ and ‘*a*’) and the dramatically close position of NZE TRAP (‘N’ and ‘n’)[39]; all of which makes perfect sense given a reading of the relevant literature.

[39] The RP citation-form (R) and connected-speech (r) female-speaker data is from Deterding (1990), the AusE data is from Watson, Harrington and Evans (1998), Cox (2006) and Butcher (n.d.) (A, a and *a* respectively), while the NZE data is from Watson, Harrington and Evans (1998) and Easton and Bauer (2000) (N and n). The reader is reminded that all the AusE and NZE data is from females and in citation-form.

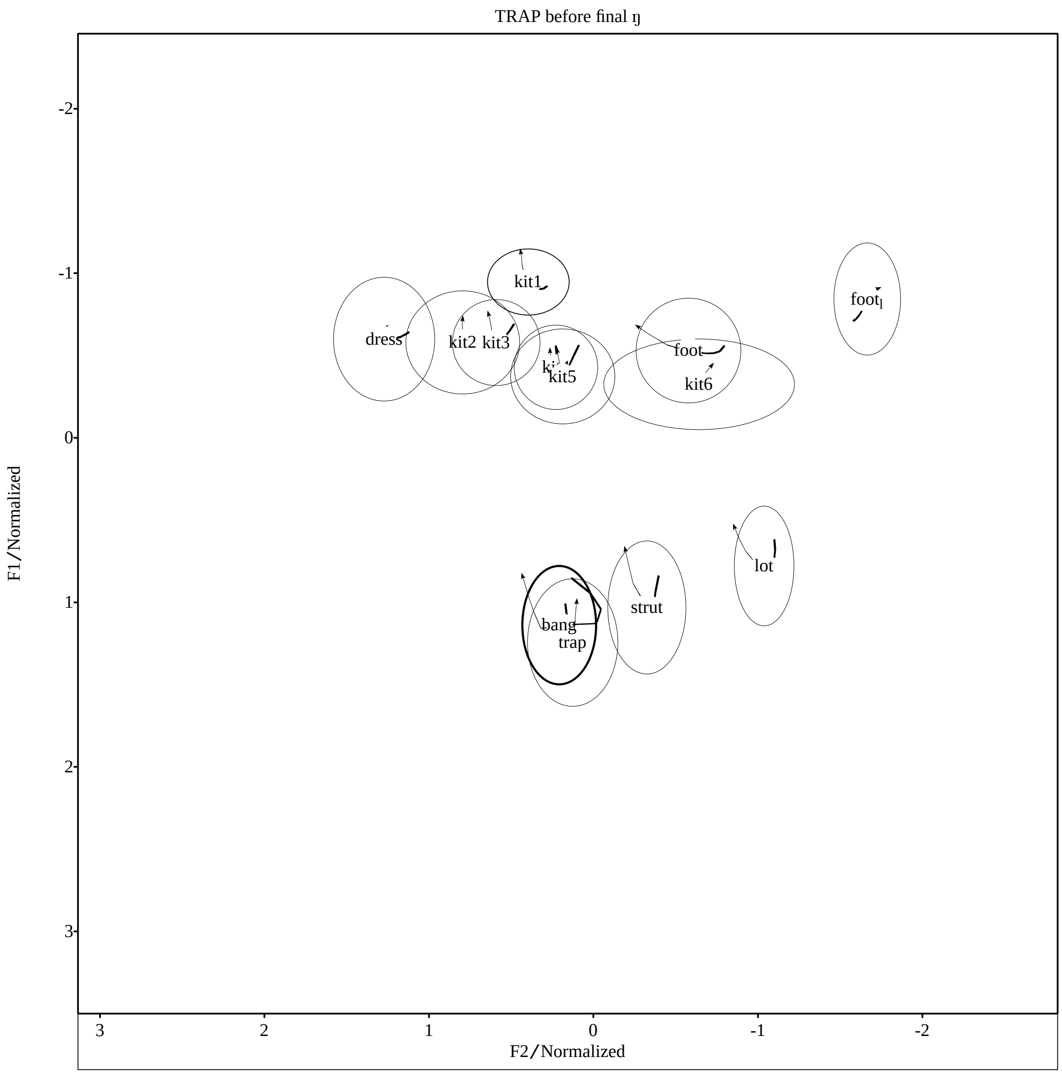


Figure 6.16: TRAP before final /ŋ/

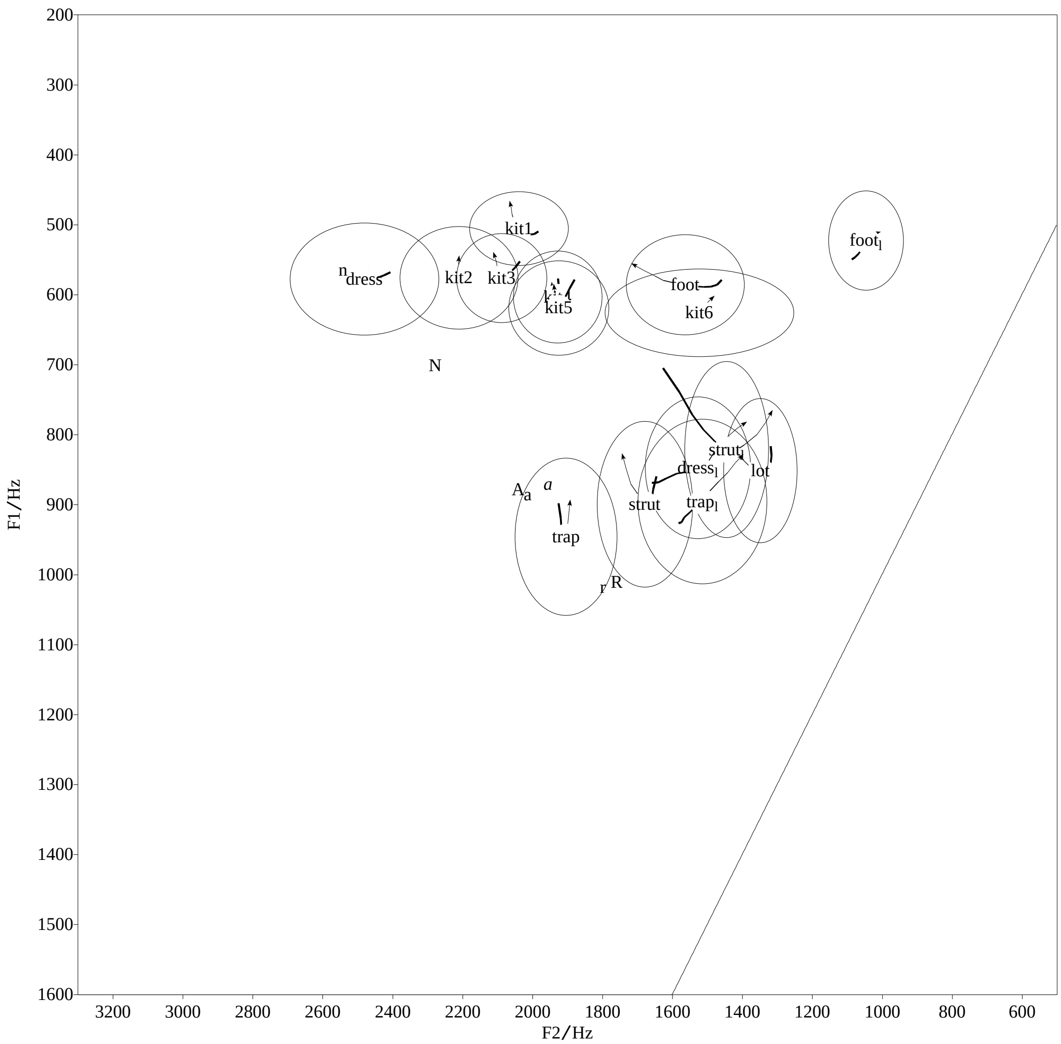


Figure 6.17: SAE TRAP vs. other English accents

We note too that SAE TRAP finds itself solidly between the AusE and RP values. What this comparison seems to suggest is that SAE, similar to what is described by Fabricius (2007*a*) for RP, is in the process of undergoing a re-positioning of its STRUT and TRAP vowels i.e. with the later (eventually) lower, as opposed to (formerly) higher, than STRUT; or, alternatively, in the process of merging these two vowels. While it is *perhaps* the case that a lowered TRAP in SAE is a case of (British-influenced?) hypercorrection, it is, however, more than certain that it has not undergone the degree of secondary raising of TRAP found in NZE. This makes sense given the stereotyped value of a close TRAP in SAE. While a raised TRAP was in all likelihood an input into the development of SAE, its indexical value has prevented further raising and perhaps lent itself, in Labov's (1994) terms, to the reclassification of this vowel as lax rather than tense. Within this framework a further lowering of TRAP would make perfect sense. From this perspective it is probable, too, that any evidence of TRAP tensing and 'breaking' (as in TRAP before /ŋ/), is as much a *relic* of a Pattern 4-type shift of the short front vowels as TRAP raising is. It this regard it is apposite to repeat the observation by Evans and Watson (2004:199) that a similar process is underway in AusE; in these authors' words, "our data combined with the data displayed by Cox (1996) . . . suggests that AE /æ/ has been lowering from a much higher position over a 50 year period at least"; it appears that SAE *might* be undergoing a similar process.

Figure 6.18 on the next page compares the current SAE data with that from Torgersen and Kerswill's (2004) Ashford data. The bullets show mean values for each older-generation female, while the diamonds show data for the younger females. We note immediately that SAE TRAP is as lowered as that of Ashford[40]. This provides clear evidence for TRAP-lowering in (recent) SAE, whether viewed as a hypercorrection or as an endogenous movement. As has been mentioned in §2.14, TRAP-lowering appears to be the impetus for SECS; which in turn suggests that SAE has undergone this lowering and has initiated a process that might eventually lead to developments similar to those seen in the southeast of England. As seen in §8.4, as well as in Figure 6.13 on page 197, STRUT is, however, fully-front and has not yet begun to retract under the pressure of the fronted TRAP vowel.

6.3.4 TRAP: Synopsis

By all accounts, GenSAE, along with AusE, is following RP and the southeast of England more generally, in lowering its TRAP vowel. Whether this is an endogenous change or simply a hypercorrection is, as yet, unclear; although given that, as far as the author can

[40]The somewhat more centralized position of TRAP in the Ashford data is almost certainly due to the admixture of citation-form and connected-speech data.

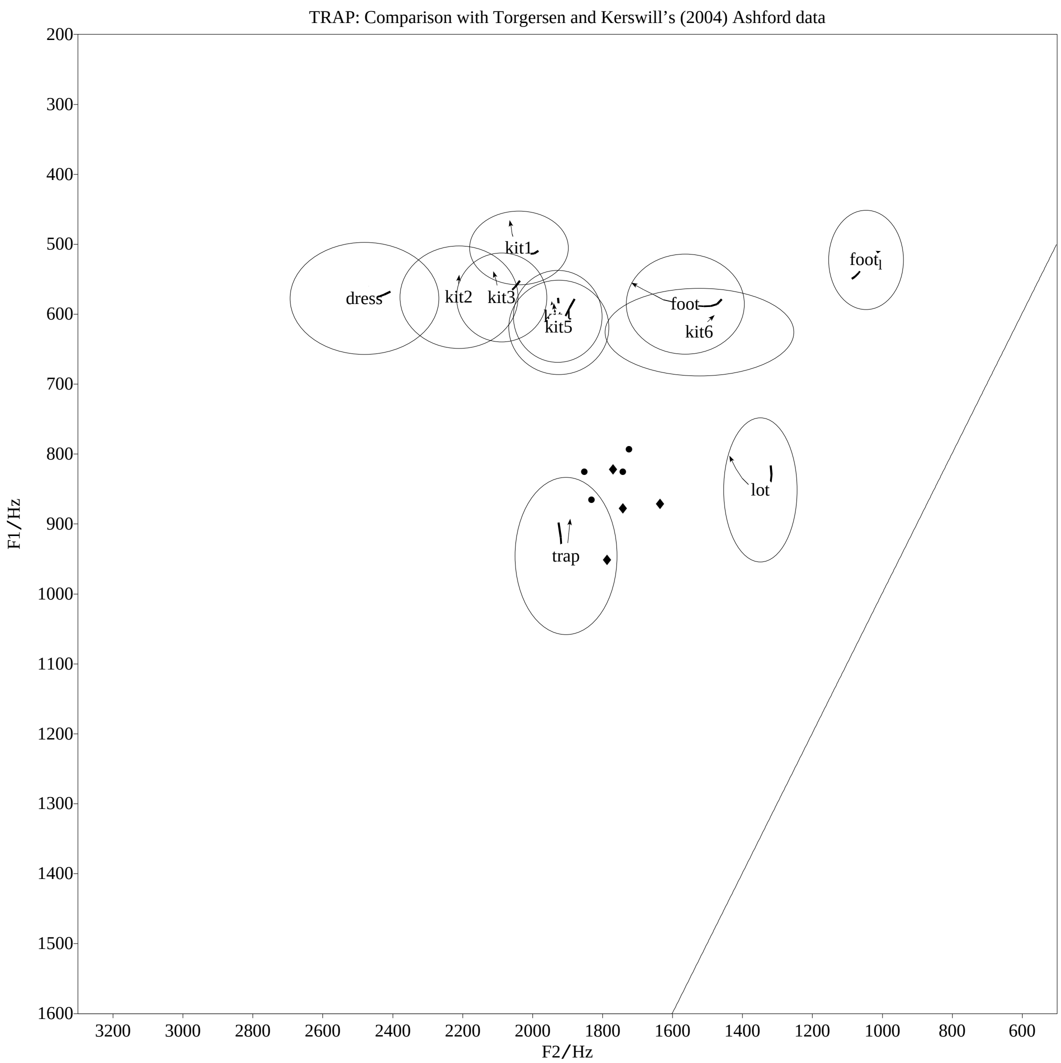


Figure 6.18: SAE TRAP: Comparison with Torgersen and Kerswill's (2004) Ashford data

ascertain, there is no overt consciousness of such lowering among SAE speakers, it is more likely to be the former. What is clear, though, is that both these accents differ in this respect from NZE which has shown substantial (secondary) TRAP raising, which of course raises the question why NZE English has broken off from the other two main Southern Hemisphere accents in this way.

More generally, the lowering of the short front vowels across these various ‘Southern accents’ constitutes a challenge to Labov’s (1994) Pattern 4 and the Southern Shift. In particular, what needs to be explained, is the mechanism by which these supposedly peripheral and tense vowels were reanalyzed as lax and non-peripheral, thus allowing them to ‘fall’. If, in addition, it can be shown that TRAP-Lowering in SAE is, indeed, an endogenous, structurally-motivated change, and not simply a hypercorrection, then there is evidence, given the fact that this movement appears to be a ‘prestige’-driven, anti-BrSAE one, for an imminent SECS-like chain-shift, but one driven from ‘above’. This would seem to correlate in some important, but as yet unclear, fashion with Labov’s (1994) observations, briefly dealt with in §2.3.2.1, concerning reverse chain-shifts and their ‘source’ in the upper end of the socioeconomic spectrum.

On a diachronic level, the source of the traditionally raised TRAP vowel in SAE can, obviously, not be the accents brought over from the north of England, given that such accents have predominantly had an open [a]-like value. One of the primary sources, therefore, must have been the Cockney-like raised TRAP brought through CE. Perhaps more importantly, we have noted above in §6.3.1, that a pre-20th-century version of RP also had a raised TRAP. The confluence of these two values in, for example, early Johannesburg, coupled, perhaps, with some late-19th-century raised variants brought in by working-class speakers from the southeast of England, was presumably enough to secure it a place in the SAE vowel repertoire. The effect of north-of-England values on the linguistic mix was further reduced by the fact that at least some of the NE speakers would have endeavored to use the more prestigious RP value.

6.4 FACE, DRESS and FLEECE (FDF)

In Labov (1994:160), FACE and FLEECE are transcribed as /ey/ and /iy/ respectively, emphasizing the glided nature of these vowels. Labov (1994:160; 162) notes, however, that there are a number of dialects in Scotland, Ireland and the north-of-England, as well as a number of American dialects with a “strong non-English substratum” in which either (or both) of these vowels are not subject to diphthongization. Diphthongization of FLEECE is clearly linked, historically, to the south-of-England and thus common in many ETEs, particularly those involved in the Southern Shift. As emphasized by Labov (1994:211), FACE is perhaps the most well-known variable in the accents of Southern England, with an open version constituting “a well-recognized and long-established social stereotype”. It, along with DRESS and FLEECE, forms as essential component of Labov’s (1994) Pattern 4, with a lowered FACE often viewed as the initiating force of this pattern. As illustrated in Figure 2.13 on page 61, in terms of this pattern FACE and DRESS ‘swop’ positions, with the nucleus of FACE appearing on the nonperipheral track and subject, in terms of Principle IIa, to lowering. DRESS, on the other hand, finds itself on the peripheral track and is thus subject, in terms of Principle I, to raising. The lowering of FACE often leads, pull-chain fashion, to the diphthongization of FLEECE. Thus, for example, Labov (1994:173) shows that in Cockney a diphthongized FLEECE vowel (with the nucleus on the nonperipheral track) is an essential step in a Pattern 4 (Southern)-Diphthong Shift. It is clear, however, that in some accents displaying other Pattern 4 features, the nucleus of FLEECE, on occasion, remains in high-front position; as pointed out in §2.3.2.1 this is is the case in Norwich.

6.4.1 FDF: Other Accents of English

Received Pronunciation: For RP FACE, Lanham (1967:41) gives [ɜ]^{33}{[ə̟]} “or slightly lower” for the nucleus and [ɪ]^{22}{[ɪ]} for the off-glide, “which is seldom lost”. Interestingly, according to this author, the nucleus has a clearly mid-central quality and does “not have the quality of [e] . . . assigned . . . by Jones and Ward” (Lanham 1967:39) i.e. evidence for a Pattern 4 movement in RP. Wells (1982:293) stresses the variability of this vowel in RP and for the nucleus defines an area of the traditional vowel chart bounded by the following three qualities: [e, ɛ, ɛ̈]. According to this author, the closest and most monophthongized variants are characteristic of U-RP, with an older variant having a nucleus close to Cardinal 2 and a more recent variant having “an almost monophthongal [ɛ̣e̜ ∼ ɛ̣ː]{[ɛ̞e̝ ∼ ɛ̞ː]}” (Wells 1982:293). Cruttenden (2001:130) concurs with Wells (1982) and gives the values for General RP as [e̝ɪ ∼ eɪ], with Refined RP having a diphthongal variant with a starting point close to C[ɛ] as well as a monophthongal [ɛː].

For RP DRESS, Lanham (1967:32) claims two main allophones:

1. The unconditioned one which he gives as [E]$^{41/51}${[ɛ̞]} in words such as *beggar* and *rent*
2. DRESS before /l/, which is lowered and retracted to [ɛ]^{52}{[ɛ̈]} as in *well*.

As with KIT (Chapter 7), in the case of DRESS, the closer and more peripheral the vowel generally the older the speaker of RP: “an /e/ with a quality approaching C[e] is typical of Refined RP” (Cruttenden 2001:110); more open and central qualities are generally associated with younger speakers (Wells 1982:291). This change in RP DRESS has been identified by Hawkins and Midgley (2005:193) as part of a more general chain-shift of the DRESS, TRAP, GOOSE and FOOT vowels, the former two becoming more open, with the latter two becoming more fronted and unrounded and, as discussed in §2.14, a more open DRESS vowel is a part of the more general SECS-Shift. Cruttenden (2001:110) gives [e̞ ∼ ɛ̝] while Lass and Wright (1985:138) give [ɛ̝]{[ɛ̝]} for General RP DRESS.

For RP FLEECE, Lanham (1967:41) provides [ɪ]^{22}{[ɪ]} for the nucleus and [i]^{11}{[i]} for the off-glide, but does claim that the glide is sometimes lost, with the vowel simply being lengthened. As confirmed by Cruttenden (2001:105), the FLEECE in RP is diphthongized: “The vowel is often noticeably diphthongized, especially in final positions. A slight glide from a position near to [ɪ] is common among RP speakers, being more usual than a pure vowel”.

The Other Accents of England: As stressed by Cruttenden (2001:130), an even more open starting point for FACE is characteristic of a whole range of English English accents, mostly as a result “of a more general change known as the ‘Southern diphthong shift’, typical of London, Birmingham and the south and midlands of England”. Thus Cockney has [æɪ] according to Cruttenden (2001:130), while Wells (1982:307) gives [ɛ̞̈ ∼ ɐɪ]{[ɛ̞̈ ∼ ɐɪ]} for “popular London” and [æɪ ∼ aɪ] for “broad Cockney”. Of interest, though, is Trudgill’s (2004) observation that, with respect to the broader southeast of England, a diphthongal FACE vowel, let alone a Diphthong-Shifted FACE vowel, was still somewhat of a rarity in the nineteenth-century (Trudgill 2004:52–5); in other words, many regional accents of the southeast of England (and of course elsewhere in Britain) had, well into the 19th-century, still not undergone the pre-Diphthong-Shift Long Mid Diphthongization of [eː] to [eɪ]. This applied to an even greater extent to the accents of the north of England.

For DRESS, Cruttenden (2001:110) points out that traditionally Cockney has a relatively close variety which “may additionally involve a glide towards [ɪ]”. However, for London and surrounding areas, Torgersen and Kerswill (2004) mention a lowered DRESS vowel as being part of the broader, more recent, SECS-Shift. As pointed out in §6.3, such lowering is conceived of as being a response to TRAP-lowering, the initiating move of the SECS-Shift. Trudgill (2004:46–7) argues for closer values of DRESS in mid-nineteenth century “RP and other southeastern English English accents more generally". For the north of England, Wells (1982:356) implies a value for DRESS that is traditionally lower than the RP one.

In the case of FLEECE, glides with starting points more centralised or open than [ɪ] are unlikely in (General) RP and are more characteristic of regional British varieties such as Cockney [əi] or Liverpool [ɨi]. Interestingly enough, Refined RP shares a centralised nucleus with Cockney. On the other hand, Estuary English[41] is likely to have inherited an extra degree of diphthongisation from Cockney itself i.e. above and beyond that found in General RP[42]. The historical status of FLEECE is of particular relevance, given claims by Trudgill (2004:59) that the diphthongisation of this vowel was hardly underway in the southeast of England by the mid-nineteenth century. On this analysis, a diphthongal FLEECE vowel was unlikely to constitute a major input into SAE as well as other Southern Hemisphere Englishes, and its subsequent diphthongisation in AusE and NZE was thus, in all likelihood, the result of an endogenous 'drift'-like development. For the north-of-England, and according to Wells (1982:357), "The FLEECE Merger ... has not been carried through everywhere in the north ... thus in a broad swathe through the middle north *meet* is [miːt] but *meat* may be [mɪət]". Historically, this implies the input of at least some 'pure' tokens of FLEECE into the early SAE mix.

The Southern Hemisphere Englishes: For FACE in Cultivated, General and Broad AusE, Wells (1982:597) gives [ɛɪ], [ʌɪ] and [ʌːɪ ∼ a-ːɪ]{[ʌːɪ ∼ a̠ːɪ]} respectively. Harrington et al. (1997:174) provide acoustic evidence for a lowered first target for FACE in Broad AusE.

For DRESS, Cruttenden (2001:110) points out that AusE has a closer variety which "may additionally involve a glide towards [ɪ]". A more peripheral DRESS vowel (i.e. higher F_2) appears to be associated with Broad as opposed to General or Cultivated speakers of AusE (Harrington et al. 1997:163).

NZE has a particularly close quality for DRESS (Wells 1982:607), with Gordon et al. (2004:26) giving [e̝] as the general value, with an even closer [i̞] and slightly opener [e̞] being the more advanced and Cultivated values respectively. According to Evans and Watson (2004:195), "the NZE /e/ vowel may now be of a similar height and fronting to NZE /iː/". The two extremes of RP and NZE can be appreciated by comparing the acoustic data represented in Figure 6.19 and Figure 6.20 on the following page, which provide data for RP and NZE respectively. It should be noted again

[41] London Regional RP - see Cruttenden (2001:81).
[42] See Rosewarne (1984) and Rosewarne (1994:6) in this regard.

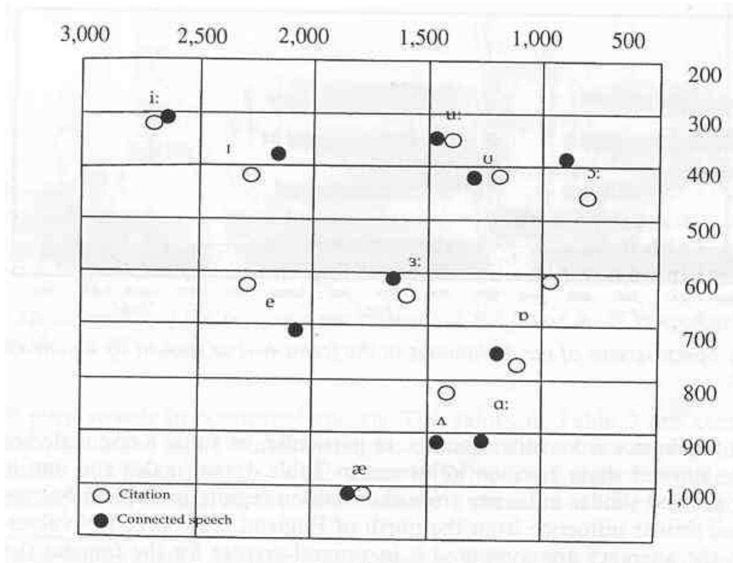


Figure 6.19: Formant frequencies for RP pure vowels said in citation form and in connected speech (female speakers); taken from Cruttenden (2001:102)

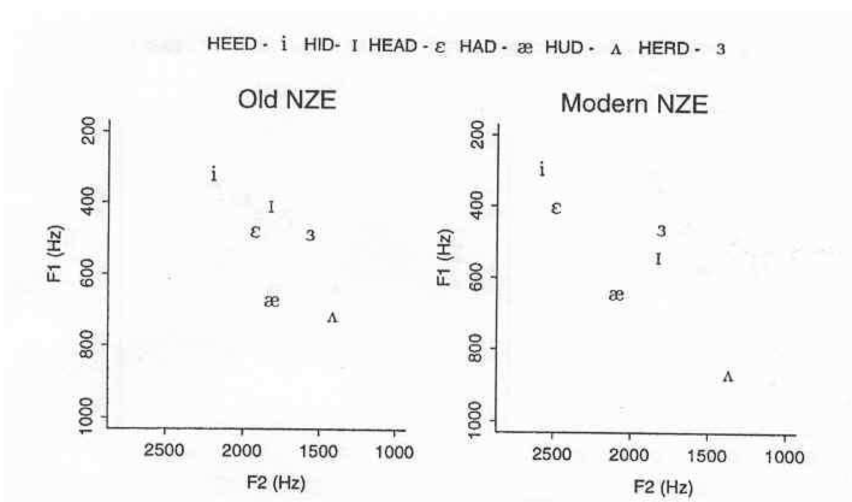


Figure 6.20: Old New Zealand versus Modern New Zealand short vowels; from Watson et al. (2000:59)

that, in the case of NZE, the raised DRESS vowel is, according to, for example Lass (2004:374), not attributable to the original input, but a more recent innovation, a fact also discernable from inspection of Figure 6.20, which compares old and new NZE data[43]. As confirmed by Evans and Watson (2004:195), DRESS in AusE is lower than in NZE, with the former having remained relatively stable over the years.

NZE and AusE FLEECE show a degree of diphthongisation (Lass 1995:98) irrespective of the social background of the speaker. Thus for AusE, Wells (1982:597) provides a cline of values, [ɪi ∼ ɨi ∼ əːi][44], the first characteristic of Cultivated AusE, the second of General AusE and the third of Broad AusE. Harrington et al. (1997) provide acoustic evidence for diphthongization of FLEECE in AusE, with Broad accents showing the greatest movement. Interestingly, these authors also provide suggestive (though tentative) evidence for a movement away from such diphthongization: "it may therefore be that . . . there is less of a tendency for young male talkers to produce more extensive /i/ onglides than older male talkers" (Harrington et al. 1997:168). For NZE, Wells (1982:607) mentions a degree of diphthongization, although not as dramatic as that found in AusE – [ɨ]{[ï]} is given as the most retracted value for the nucleus. Interestingly enough, the English of Tristan da Cunha has a completely monophthongal FLEECE vowel, while Falklands Islands English displays the typical diphthongal Southern Hemispehre variety, with a value in the region of [ɪi] (Trudgill 2004:51).

The U.S.A.: In the US South, FACE swops position with DRESS, a central movement in Labov's (1994) Pattern 4. In Philadelphia, in contrast, we have the "fronting and raising of the nucleus of checked /ey/" (Labov 1994:60), often to the point of overlapping with FLEECE and contrary to the Southern Shift, but, on the other hand, the standard lowering of the nucleus in FACE in open syllables e.g. in *day* (Labov 1994:81). As mentioned in §2.3.2.1, this behaviour appears to be part of a retrograde movement away from the Southern Shift.

For the US South, DRESS forms a prominent role in "the reversal in acoustic position of the /ey/ (*bait*) and /e/ (*bet*) class" (Fridland 2008:69). As confirmed by Wolfram et al. (2004:345), a common Southern feature is a partial merger of DRESS and KIT such that *pin* and *pen* are pronounced identically i.e. [pɪn]. In Philadelphia, a lowered and backed DRESS vowel appears to be an incipient change, along with a similar movement for KIT as well as STRUT-raising (Labov 1994:82).

[43] Also see Torgersen and Kerswill (2004:29) in this regard.
[44] Wells (1982:xvii) gives [ɨ] as being equivalent to [ï].

As confirmed in Fridland (2008:69), the lowering movement of FLEECE onto the non-peripheral track (i.e. its diphthongisation) is also a common feature in the US South, although in the case of Philadelphia (traditionally a Southern accent) there appears to be little evidence of FLEECE diphthongization (Labov 1994:211).

6.4.2 FDF in SAE: The Impressionistic Data

Hopwood (1928): According to Hopwood (1928:13), "**˔ɛɪ** > … **ëi** or **ɛi**" for SAEP FACE. The first value above, i.e. [ëi], is, according to this author, prevalent in AE and is a result of substituting the Afrikaans diphthong in words such as *eier* 'egg', *my* 'my' etc. The more open (Broad) value, i.e. [ɛi], is seen as being mainly Cockney-influenced. Note that, in general, Hopwood (1928:33) claims a more tense articulation of the second element of diphthongs in SAEP, thus accounting for the **i** element in the SAEP values for FACE given above[45].

Interestingly enough, Hopwood (1928:24–5) also gives two possibilities for the DRESS vowel in SAEP. Firstly, he claims that "**˔ɛ** > **ɛ(ˑ)** in stressed syllables, due to A. [Afrikaans] spelling and sound substitutions" (Hopwood 1928:24; my parenthesis). The author, however, provides an alternative which is that [ɛ̝] becomes [ë]. Hopwood's (1928) explanation of the origin of this latter value is as follows: firstly, he points out that in the case of weak vowels in non-final position (§9.3), SAEP often follows the Scottish English example of using a stressed vowel e.g. RP [ɡlædnɪs] 'gladness' becomes [ɡlɛdnës] in SAEP. According to Hopwood (1928), SAE–speakers then extrapolate this value to stressed syllables; thus RP [dɹɛ̝s] 'dress' becomes SAE **drës** (note the BrSAE obstruent /r/). The author then continues by stating that "this change is considerably aided by the Cockney change of **ɛ** to **ɪ** (*e.g.* "Get out of it" becomes Cockney E.: "Gitahtuvit"…)" (Hopwood 1928:25).

For FLEECE, Hopwood (1928:31) claims "**ìˑj**{[ɪ̆ˑi]} > **ìˑ**{[ɪ̆ˑ]} or **i**, a shortening, purifying, and sometimes tensing of the E. sound". The **i** value almost certainly belongs to BrSAE or AE, particularly in terms of its shortness.

Hooper (1945; 1951): Hooper (1945) illustrates quite clearly that an open and retracted first element in FACE[46] is associated with a relatively lower social-class status as well as with being male and younger. Hooper (1951:84) is unambiguous about the

[45] In fact, the **i** should rather be transcribed, in Hopwood's (1928) terms, as **ì** i.e. as modern [ɪ̆] instead of modern [i] – in off–glide position **j** would be used for the latter. This analysis is, in fact, confirmed later by Hopwood (1928:33) where he transcribes SAEP 'lay' as **lëì**.

[46] Hooper (1945) gives [əɪ].

supposed source of the FACE value in SAEP, and is referring in what follows to the rather open Cockney-like onset to this vowel:

> "...one or two characteristic South African pronunciations are very similar to pronunciations regarded as inferior socially in England. For example, words like *gate*, *pale*, *train*, tend to be pronounced in a way that sounds to English ears very like a Cockney pronunciation. In fact, it is almost certainly another case of Afrikaans influence and not Cockney at all".

Breckwoldt (1961): Breckwoldt (1961:7) gives raised values for the DRESS vowel in SAE: [e ∼ ë ∼ ɩ]. Peculiarly enough, Breckwoldt (1961:7) provides a more open value for FLEECE giving, by way of example, "riːd > reːd".

Lanham and Traill (1962), Lanham (1964?; 1965): Unlike Hopwood (1928), Lanham and Traill (1962:29) claim that in non–SARP SAE FACE vowels are characterized by "the weakening or complete loss of up–glides after the low vowels of complex nuclei" (Lanham and Traill 1962:29; footnote 29). With respect to FACE in 'SARP A', Lanham and Traill (1962:29) claim [ə]^{43} or [ə]$^{43/53}$ for the main vowel, with [ɪ]$^{22/33}$ as the endpoint of the up–glide; thus [ə̟ë̞ ∼ ə̟˔ë̞]. On the other hand, SARP 'B' has [ə]^{44} for the vowel and a slightly more fronted up-glide terminus at [ɪ]^{22}; simply [əɪ]. Non-SARP SAE takes the retraction of the vowel further to [ə]$^{44/54}${[ə̠˔]} or even [ɔ]$^{54/64}${[ɜ̠˔]} with a very weak and short glide following. This general picture is confirmed in Lanham (1965:93), who predicts a gradual lowering of the onset with an eventual value for SAEP in the region of the STRUT vowel.

For SAEP DRESS, Lanham and Traill (1962:22) give [E]$^{41/51}${[ɛ̞]} or [E]^{41}{e̝}, with the former being more characteristic of their SARP 'A' group and the latter more common in their SARP 'B' subjects. Even for the authors' SARP 'B' group these values seem surprisingly open, and very close in fact to the typical RP [ɛ̞] value. Still, the authors do mention that "the tendency to raise the representative of /e/ is a trend in SAE," (Lanham and Traill 1962:22), referring here no doubt mainly to non-SARP SAE. This is confirmed further in Lanham (1965:94) who claims that, for example "/wen/ *when* and /red/ *red* are pronounced with [e] or raised [ɛˆ]{[ɛ̞]} by the majority of English-speaking South Africans". Lanham and Traill (1962) also mention the influence of a following /l/ on this vowel. For words such as *tell* the value [ɛ]$^{51/62}$ is given, thus approximately [ɛ̝̈]: "*sell* sounds like *Sal* and the acoustic distance between these words is noticeably narrower than between *pet* and *bat*" (Lanham and Traill 1962:22). A similar value for DRESS before /l/ is given in Lanham

(1964?:30): "the vowel written e . . . sounds like (æ) in hat when followed by l". Lanham (1965:93) gives [æ̝] as the SAE value, and claims that "the contrast /e/: /æ/ is neutralized before /l/. Many children can no longer distinguish *elf* from *Alf*".

Lanham and Traill (1962:8) claim that in SARP the FLEECE vowel receives a short up–glide and does not have a single vowel quality, although "the glide is very short and easily lost . . . in environments conditioning 'short' allophones . . . [and] . . . under tertiary stress" (Lanham and Traill 1962:29; my parenthesis). The main component of the vowel is given by these authors as [ɪ]^{22}{[ɪ]} or [ɪ]$^{21/22}${[ɪ̞]}, reaching [i]^{11}{[i]} on the up-glide. Lanham and Traill (1962) claim a higher, fronter and purer vowel for non-SARP SAE.

Lanham (1967): For FACE, Lanham (1967:86) gives [ʌ]$^{54/64}$ {[ɜ̠]} as the value for the onset of this vowel in non-conservative SAE. According to Lanham (1967:92), the glide reaches [ï]^{23}{[ï]}. For conservative SAE the author gives [ə]^{44}{[ə]} for the onset and [ɪ]^{22}{[ɪ]} as the termination of the glide, although in non-CV syllables the glide is shorter.

Lanham (1967:66) provides somewhat more retracted values for DRESS before /l/ in SAE, the most retracted value being [æ]^{63}{[ɐ̝]}. Later, Lanham (1967:81) provides an even closer value for DRESS in non-conservative SAE: "the norm is [e]^{31}{[e]} or somewhat lower with [æ]$^{52/63}${[ɐ̝˖]} occurring before /l/".

According to Lanham (1967:89), for SAEP FLEECE, monophthongal variants are also attested to, the author giving [ɪˑ]^{21}{[e̝ː]} or ^{22}{[ɪː]} as possibilities even in *conservative* SAE[47]. For non-conservative SAE FLEECE, Lanham (1967:64) only gives a high, front value at [i]$^{11/22}${[i̞ː]} for the most extreme forms of non-conservative SAE i.e. those under Afrikaans influence. The more common value is claimed to be [ɪˑ]^{21}{[e̝ː]} (Lanham 1967:89), thus overlapping with conservative SAE to some degree. This author also mentions the possibility of a retracted allophone before final /l/: [ï]^{23}{[ï]} or ^{24}{[ɨ̞]} (Lanham 1967:66;89).

Lanham (1978), Lanham and Macdonald (1979), Lanham (1982): Lanham (1978:151) identifies a backed and lowered FACE vowel, with glide-weakening especially before /l/, as a defining value of BrSAE but also mentions that it is found in GenSAE as

[47]Note that the use of the half-length symbol /ˑ/ in Lanham's (1967) transcription system appears to be a *phonemic* device and does not reflect phonetic fact. This is borne out, for example, by the author's transcription of *leader* in non-conservative SAE as /liˑdə/ phonemically, but [lɪːdə] phonetically. Half-long vowels in SAE are mentioned however as being characteristic of "certain extreme forms of SAE, particularly in the Eastern Province [Eastern Cape Province in modern terms]." (Lanham 1967:90; my parenthesis).

well, particularly in less formal styles. This value for FACE is “often corrected and therefore idiolectally involved in prominent stylistic variation . . . [and] . . . is above the level of ‘social consciousness’ in English cities” (Lanham 1978:151l; my parenthesis). The author also links this trend to similar developments in Britain (e.g. Cockney) and Australasian English. Lanham and Macdonald (1979:46) add little new, giving [ʌ^{e}] as the most advanced variant, with more advanced variants linked to a greater degree of ‘Broadness’. While involved in style-shifting and correction, the variable appears to be stable with no evidence of receding. A backed, lowered and glide-weakened FACE vowel is often correlated with a similar GOAT vowel (§8.3) and is identified as a GenSAE variable.

For DRESS, Lanham (1978:152) confirms the general picture provided above, giving a C[e]-value (or higher) for the most advanced value of DRESS. “Prominent in Ext SAE [BrSAE] and pervasive in general SAE. Limited evidence of social consciousness, but not usually corrected” (Lanham 1978:152; my parenthesis). Vowel retraction before /l/ towards an [æ]-like value is confirmed, with Lanham and Macdonald (1979:43) giving [æ̈] as the most advanced GenSAE variant. Lanham and Macdonald (1979:46) identify a close DRESS vowel as a GenSAE variable with a cardinal 2 quality or even closer. The authors imply an implicational link between a close DRESS vowel and a close and glideless SQUARE vowel (§9.2) i.e. if a close DRESS vowel exists in a particular idiolect it is likely that a close and glideless SQUARE vowel will exist in that same idiolect. Lanham (1982:339) claims that “in general, the movement toward cardinal vowel [e] . . . has not come to a conscious level of social awareness”.

Lanham (1978:153) mentions a raised and fronted FLEECE vowel as being characteristic of BrSAE, and gives Afrikaans as the origin for this value.

Lass and Wright (1985; 1986): Lass and Wright (1985:137) confirm the considerably raised status of DRESS in SAE, giving the characteristic values as [e $\sim$ ë]. Notice that a slightly centralized Cardinal 2 for DRESS overlaps with Lass and Wright’s (1985) value for RP KIT; see Chapter 7. These authors also confirm the centralization and lowering of DRESS before final /l/ (Lass and Wright 1985:155) as well a less-raised version in CulSAE (Lass and Wright 1985:156).

Lass and Wright (1986:209) confirm that the DRESS vowel is one of the three vowels involved in the ‘chain-shift’ mentioned in §3.4.3, the others being KIT and TRAP. More specifically, in relation to the hypothesis that the values for these three vowels are due to the original input brought by the original settlers, it is worthwhile mentioning that "raising of ME /e/ is attested as early as the 13th century ... and is stigmatized in the 17th century as vulgar" (Lass and Wright 1986:209).

Lass (1990; 1995; 2004): For his middle-class Cape Town subjects, Lass (1990:278–9) gives an onset in the region of [e] or [e̝] for FACE. While [eɪ] appears to be common, the diphthong is in other cases often extremely narrow in the sense that there is not much difference in quality between the first and second morae of the diphthong, creating a near-monophthong e.g. [ee̞] or [e̝e]. The impression of monophthongisation is, according to Lass (1990:278), hightened by the fact that the first mora is longer than the second i.e. [eˑe̞ ∼ e̝ˑe]. Lass (1990:279) also states that it is his impression "(not yet subject to quantitative analysis) that the narrower diphthongs are more typical of female speech than of male". Given the monophthongal nature of SQUARE in SAE, the potential for confusion between FACE and SQUARE is obvious – see §9.2 below. This author confirms that the more extreme the variety the more open and retracted the onset. Thus Lass (1990:273;279), for example, gives [äɪ ∼ ʌ̈ɪ] as extreme variants.

According to Lass (1995:99), the value of the FACE vowel in SAE is an important social marker. In general, "the closer the onset is to STRUT, the more Extreme" (Lass 1995:99). For this author, CulSAE and GenSAE have an RP-like [eɪ] diphthong as the norm, although "in some younger (and especially) female Respectable speakers the second element is very short and rather peripheral and open, so that the nucleus is almost monophthongal". Male GenSAE speakers may, on the other hand, verge towards a slighly more open value such as [ɛɪ ∼ æɪ], while BrSAE speakers have even further retracted and open values, such as [ɛɪ ∼ äɪ ∼ ʌɪ][48].

Lass (2004:377) confirms the socially-marked nature of this vowel also emphasising the degree of variation. The earlier analyses are confirmed; the most typical value for GenSAE speakers is given as [eɪ]; these speakers rarely go "lower than [æ], except among older Jewish speakers" (Lass 2004:377).

According to Lass (1990:276), DRESS is a half-close front vowel that is often centralised i.e. [e ∼ ë]. It is also, in GenSAE, often slightly raised above C[e] position in the speech of women: thus [e̝ ∼ ë̝]. This vowel is thus "close enough to RP or similar [ɪ] to cause perceptual problems for outsiders" (Lass 1990:276). As discussed in Chapter 7, RP KIT is, according to Cruttenden (2001:107), pronounced with a quality of "a centralised C[e] = [ë]". Lass (1990:276) gives [ɛ̈] as the value of DRESS before

[48]In the original, the first option is given as [ɛi]. I have assumed that this was a typological error.

final /l/, “a natural response to uvular co-articulation”.

Lass (1995:97–8) claims that the DRESS vowel is not an important social indicator, giving its value at around Cardinal 2, and confirming that it is slightly higher for female speakers in non-CulSAE varieties. It is also often centralized, “approaching the fronter KIT allophones” (Lass 1995:98). As far as DRESS before final- /l/ is concerned, Lass (1995:98) has the following to say:

> “The pre-/l/ allophones in Extreme SAE often have a preceding [j], especially initially and after /h/: *help* [(h)jɛɫp]. In some Respectable and Extreme, this vowel lowers and retracts before dark /l/, to around [ɛ] or even [æ]”.

Lass (2004:376) adds that, in BrSAE, DRESS is raised above [e] (although more in Coloured than White speakers). “It is never so high as to merge (as innovative NZE) with shorter realizations of FLEECE” (Lass 2004:376). The author also mentions, in the case of CulSAE and GenSAE, the (occasional) hypercorrection to a more [ɛ]-like quality.

According to Lass (1990:277), FLEECE is ‘pure’ in all varieties of SAE, with no glide to [i] (e.g. [ɪi]), a fact which this author reports as surprising given the “London-vernacular-like forms” prevalent in SAE, particularly BrSAE. For Lass (1990:277), the phonetic quality is either nearly Cardinal 1, or a slightly more retracted and lowered value. According to Lass (1995:98), this vowel is invariably [iː]: a “long close vowel [iː] in all varieties, with no social variation”. Likewise Lass (2004:376) claims that, with respect to FLEECE in SAE, “values like [ei, əi] do not occur in any varieties I am aware of”.

The remaining sources: Wells (1982:614), while confirming the correlation between a more open and more back starting point for FACE and ‘broadness’ in SAE, gives rather close starting points for GenSAE i.e, [ëɪ] or [əɪ]. Advanced (Broad) values reach [ʌɪ] or, with glide-weakening, [ʌe]. The author confirms extensive style-shifting and sociolinguistic variation with respect to this vowel. Branford (1994:481) generally concurs with the above description, linking the CulSAE variant with a front, close and narrow RP-like value and, drawing on Lanham and Macdonald (1979), characterizing BrSAE as having a retracted and more open, glide-weakened FACE vowel with [ʌ^{e}] as the most advanced variant[49].

[49] This is, in fact, a reconstruction of what Branford (1994:481) *should* have written. The original is as follows:

Da Silva (2007:117) mentions possible values of [eɪ], [ɛɪ] and [aɪ] for FACE, making a general connection between the lower values and BrSAE and Afrikaans English and, with respect to the Greater Johannesburg region in particular, noting that "a lowered first vowel is characteristic of residents of the East Rand, of areas such as Boksburg, Benoni and Edenvale"[50]. The inherent variability of this vowel is, furthermore, borne out in her results which, in the case of her primarily 'white' lect 1, shows an almost equal distribution of her [aɪ] and [eɪ] variants; understandable given the conversational Labovian-style nature of her interviews (Da Silva 2007:159).

For DRESS, Wells (1982) generally agrees with the foregoing assessments, giving [æ̈] as the value of DRESS before final /l/ and confirming the neutralization of DRESS and TRAP in this environment in SAE (Wells 1982:617). Branford (1994:474) gives the range of DRESS in SAE as CulSAE [ɛ] to cardinal vowel no. 2 or even higher at the Broader end of the scale. According to this author a final /l/ conditions a [æ̈]-quality. With respect to the non-CulSAE variants, Branford (1994:477) adds that "The DRESS vowel was similarly higher, at any rate for some speakers, in southern British lower-class speech in the 1830s". According to Bowerman (2004:936–7), the realization of DRESS is around cardinal 2. The author adds the following comments, however:

> "... though it is lowered to [ɛ] in Broad, sometimes approaching [æ], especially before [ɫ]. Some varieties of Broad and General WSAfE place this vowel higher, around raised [e] or lowered [ɪ]"

This seems incorrect on two levels. Firstly, it implies that outside of the context of a final /l/, a lowered DRESS vowel can be an indicator of a Broad accent. The opposite would, surely, be the case i.e. a lowered DRESS would, if anything, be an indicator of a CulSAE accent. Secondly, the author, in all likelihood, also meant a lowered [i], as opposed to a lowered [ɪ], for the raised variant.

Interestingly enough, Bailey (1984:16) gives traditionally monosyllabic words such as *feel*, *real* etc., which have FLEECE before final /l/, disyllabic status in SAE

> "Lanham & MacDonald (1979: 46) describe this glide for broader SAE as 'backed, lowered, glide-weakened [aɩ]', with [ʌ̈ᵊ] as its most advanced variant" (Branford 1994:481).

This is a misquote (and, in certain respects, a misanalysis) of the original, which is as follows:

> 'Backed, lowered, glide-weakened əi ... most advanced variant [ʌᵉ]" (Lanham and Macdonald 1979:46).

[50] *The* prototypical 'white' lower and lower-middle class areas of Johannesburg.

i.e. /fiy.əl/ and /riy.əl/. A process similar to this appears to be common in AusE (Palethorpe and Cox 2003). Taylor (1991:79) mentions a "extremely close and fronted monophthong," for SAE FLEECE although he does not indicate its sociolinguistic value. A high-front [iː] is given by Branford (1994:481) for all sociolects of SAE, although "in more extreme accents it may be higher and more fronted than in RP" although given that RP is often diphthongized (see above), Branford (1994) must be referring to the terminus of the glide for this to make sense. It should also be noted that the symbol used by Branford (1994) for FLEECE in SAE is [ɩː]. As is explained further in Chapter 7, Branford (1994) appears to be using [ɩ] as an equivalent for [i] and *not* [ɪ].

6.4.3 FDF: A Comparative Acoustic Analysis

Data for Webb's (1983) FACE and FLEECE vowel is included in Figure 5.5 on page 152. The position of this nucleus is commented upon by the author as being surprisingly low (Webb 1983:161; footnote 9), and it does, indeed, appear to have a Broad (AusE-like) value. As far as the glide of FACE in Webb's (1983) data is concerned, he notes "that glidelessness is not generally characteristic of the glides presented" (Webb 1983:155) in the data:

> "The glide in /ey/ is strongly present, being roughly 133Hz higher, and 557Hz in "front" of the vowel nucleus in the environments __$ and __# and 253Hz in "front" in the environment __ [-son]. Relatively speaking there is some weakening before nasals, the glide being only 185Hz in "front".

In other words, there is little evidence of the expected glide-weakening mentioned in §3.4.3. In a footnote, the author speculates that the absence of glide-weakening in the data might be "due to the fact that the author is Afrikaans-speaking" (Webb 1983:161; footnote 8). As already mentioned in §5.1.3, at another point in the article, however, the author states clearly that "Natal English . . . is the variety of English acquired by the author" (Webb 1983:154). If this is the case it seems unlikely that the lack of glide-weakening can be accounted for by the bilingualism of the author.

Webb (1983:153–4) believes, furthermore, that the data justifies the recognition of two FACE allophones, one in word or syllable-final position, one elsewhere. In Figure 6.21 on the following page, FACE in final position is indicated by *face*$_{fin}$ while the unmarked allophone is marked with *face*$_{un}$. While this figure does provide some evidence for the existence of the two allophones, the matter requires further exploration, particularly because the conclusion is based on formant data from only 12 tokens (from one individual), with FACE in final position being represented by four of these i.e. *way*, *may*, *places* and *major*.

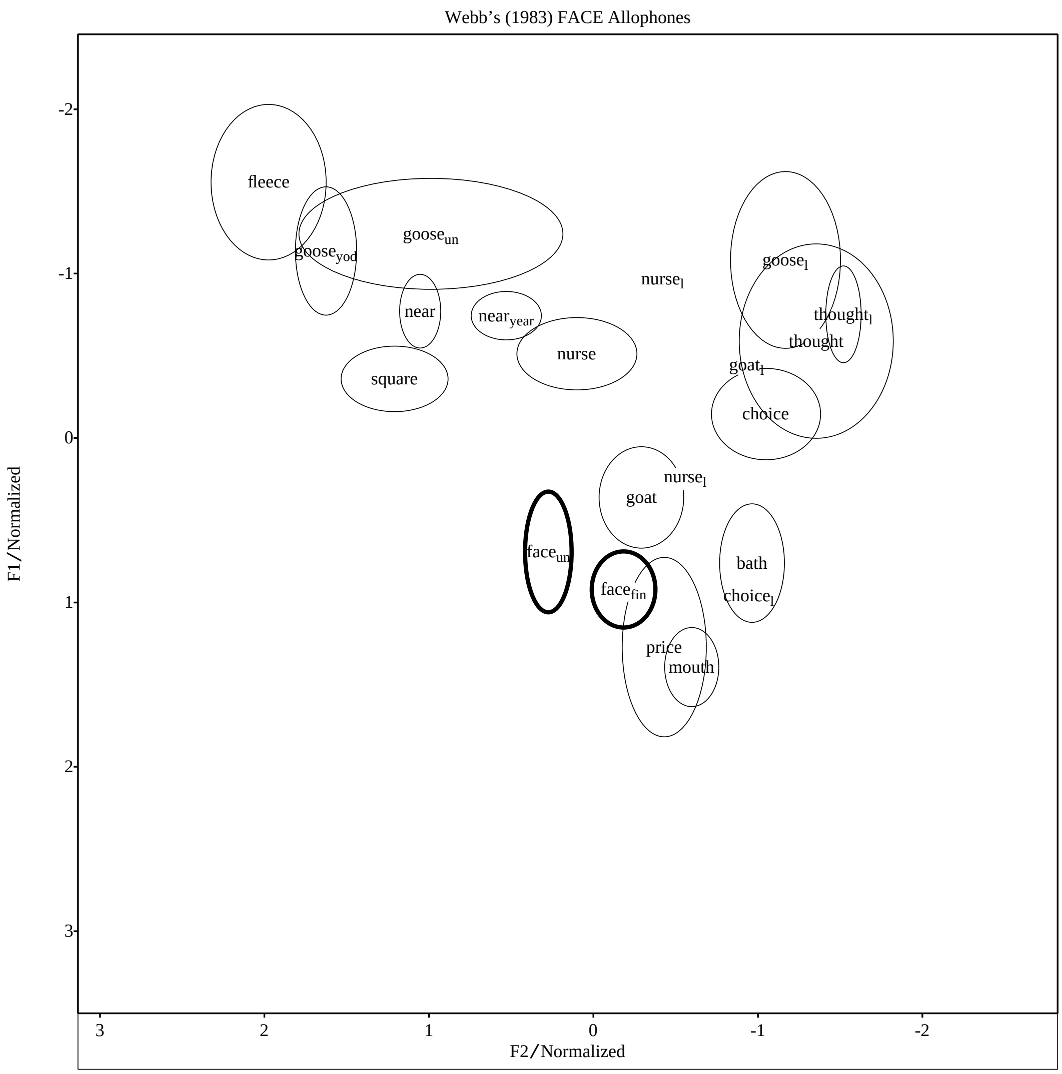


Figure 6.21: Webb's (1983) FACE Allophones

Contrary to expectations, Webb (1983:149) claims diphthongal status for FLEECE in his data:

> "Of the twelve tokens in the data eight undergo formant movement, this is they glide forward, by an average of 200Hz. In three of these cases the glide is also slightly higher than the vowel nucleus. The remaining four tokens remain stable with respect to both formant values".

Interesting too is that FLEECE before final /l/ (the one token is *field*) shows no sign of retraction[51].

In Figure 5.4 on page 151, which gives the lax vowels of Webb (1983), we have ample evidence of the raised status of the main allophone of DRESS, as well as dramatic evidence for the retraction of this vowel before a final /l/, the two tokens in question being *selves* and *well*. What is surprising here is the *degree* of retraction with the quality being closer to LOT than any other short vowel, a far greater degree of retraction than generally predicted by the impressionistic literature, in which the default value given is usually a centralized [æ̈]. Impressionistically, the overlap between the DRESS vowel before /l/ and LOT has to do both with the degree of retraction of DRESS in this context as well as the fact that LOT in many South African idiolects appears to have very slight lip-rounding, if at all – see §8.4. The lowering of F_2 that is usually caused by lip-rounding is thus absent.

Moving on to the current data, Figure 6.2 on page 173 provides a graphic representation of SAE FACE and FLEECE in the context of the other SAE Part-System B vowels, while Figure 6.13 on page 197 shows DRESS in the context of the other short stressed vowels. The data on which these various graphic representations are based is summarized in Table 6.4 on page 223.

Beginning with FACE, as one can plainly see from Figure 6.2 on page 173, this vowel has a relatively close onset coupled with a weak glide to an even closer position.

The position of this vowel becomes clearer when placed in the context of the Part-System A short vowels, as depicted in Figure 6.22 on the next page. As can be seen the onset is a little lower than DRESS and the glide 'travels' to a position just above DRESS. Note that the position of the terminus of the glide is in itself of interest given that it does not overlap at all with the fronter allophones of KIT, which would be expected from a [eɪ]-like understanding of the trajectory of this vowel. From this figure it also appears that the nucleus of FACE in the current data remains mostly on the periphery; there is little evidence here for the laxing and lowering of the nucleus so characteristic of accents of English undergoing

[51] Note that in Figure 5.5 on page 152 *fleece$_l$* is 'hiding' behind *fleece*.

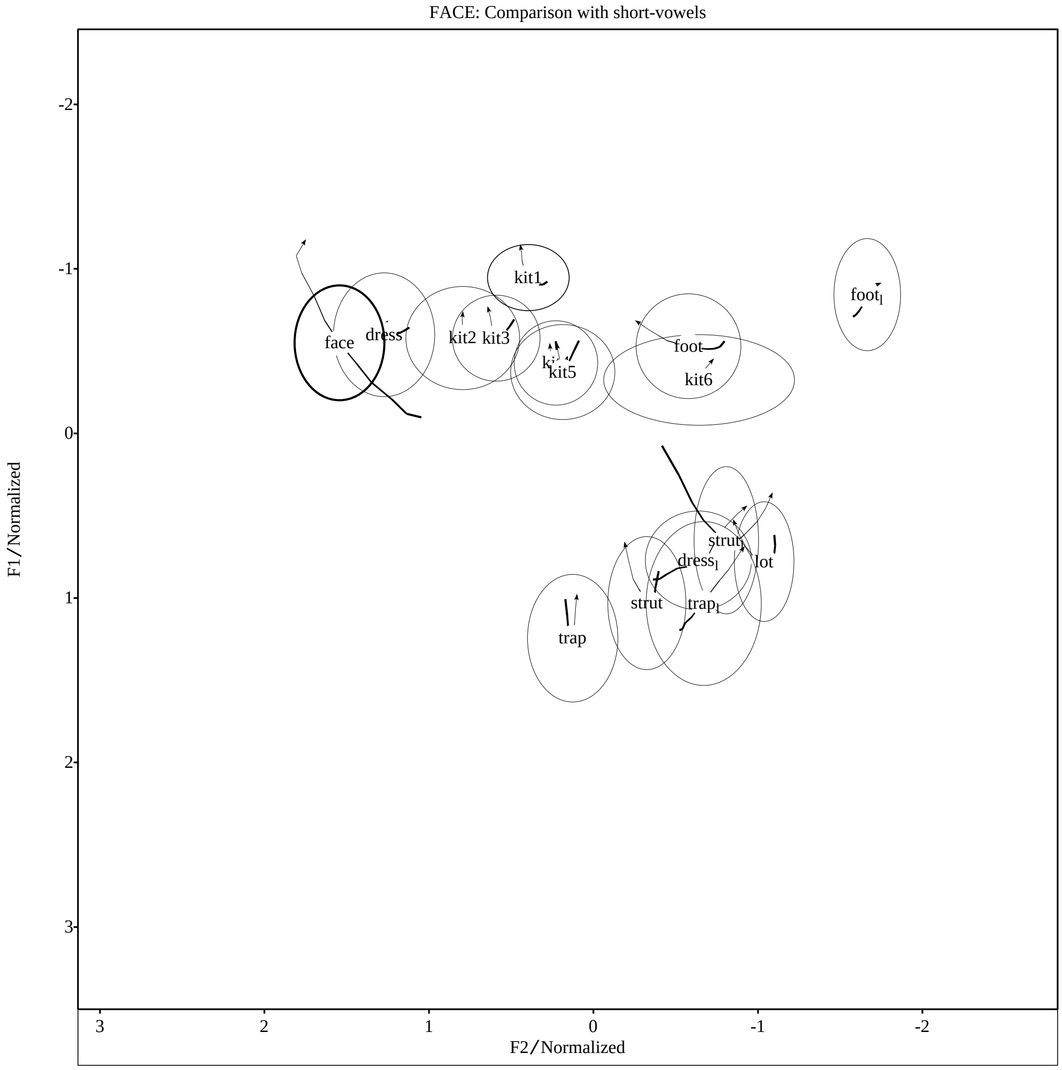


Figure 6.22: FACE: Comparison with Short Vowels

Lexical Set	No. of Tokens	No. of Measurements	N-G Tokens	O-G Tokens
FACE	82	1782	*hay*, *maid*, *mate*, ***pa***_per_	*eight*, *mate*
FACE before /l/	18	396	*fail*	
DRESS	94	2064	*bet*, *head*	*bet*, *peck*, *beck*, *speck*, *let*, *red*, ***te***_lly_
DRESS before /l/	27	594	*tell*	*tell*
FLEECE	58	1276	***are***_na_, *heed*, *beat*	*beat*
FLEECE before /l/	18	396	*field*	

Table 6.4: Summary of FACE, DRESS and FLEECE data

the Southern Shift. All in all, the data provides support for Lass' (1990) transcription of this vowel for his (female) subjects as [ee̝] or [e̞e]. It also, perhaps more importantly, provides evidence for the non-peripheral, lax status of the DRESS vowel in SAE.

The FACE data collected as part of this project is in dramatic contrast with Webb's (1983) FACE vowel, as clearly depicted in Figure 6.23 on the following page. This difference thoroughly confirms the status of FACE as a sociolinguistic marker in SAE. It is interesting to note, in this regard, that Webb's (1983) FACE vowel comes very close to overlapping with PRICE in the current data, as well as PRICE in Webb's (1983) own data – see Figure 5.5 on page 152 i.e. either PRICE has not undergone much of a compensatory retraction or, as intimated in §6.2, has undergone a 'reversal' *vis-à-vis* an original and 'relic' Pattern 4 shift. It also suggests that PRICE in SAE is far less susceptible to style-shifting than FACE, in all likelihood due to the stereotyped role that a backed and rounded PRICE has traditionally played in SAE.

As was discussed above, and as represented in Figure 6.21 on page 220, Webb (1983) hypothesizes an allophonic distinction between FACE finally and FACE non-finally. In order to test this hypothesis for the current data, all the tokens of the word *hay* were compared with the overall FACE values. The results are represented in Figure 6.24 on page 225. As can be seen, there is no significant difference.

Figure 6.25 on page 226 provides a comparison of SAE FACE with other accents of English. As can be clearly seen, citation-form AusE has a far more open nucleus than the SAE equivalent. The terminus of the glide is roughly equal to that of SAE though. It should be stressed here that the AusE data represented here is word-list data; it is thus quite

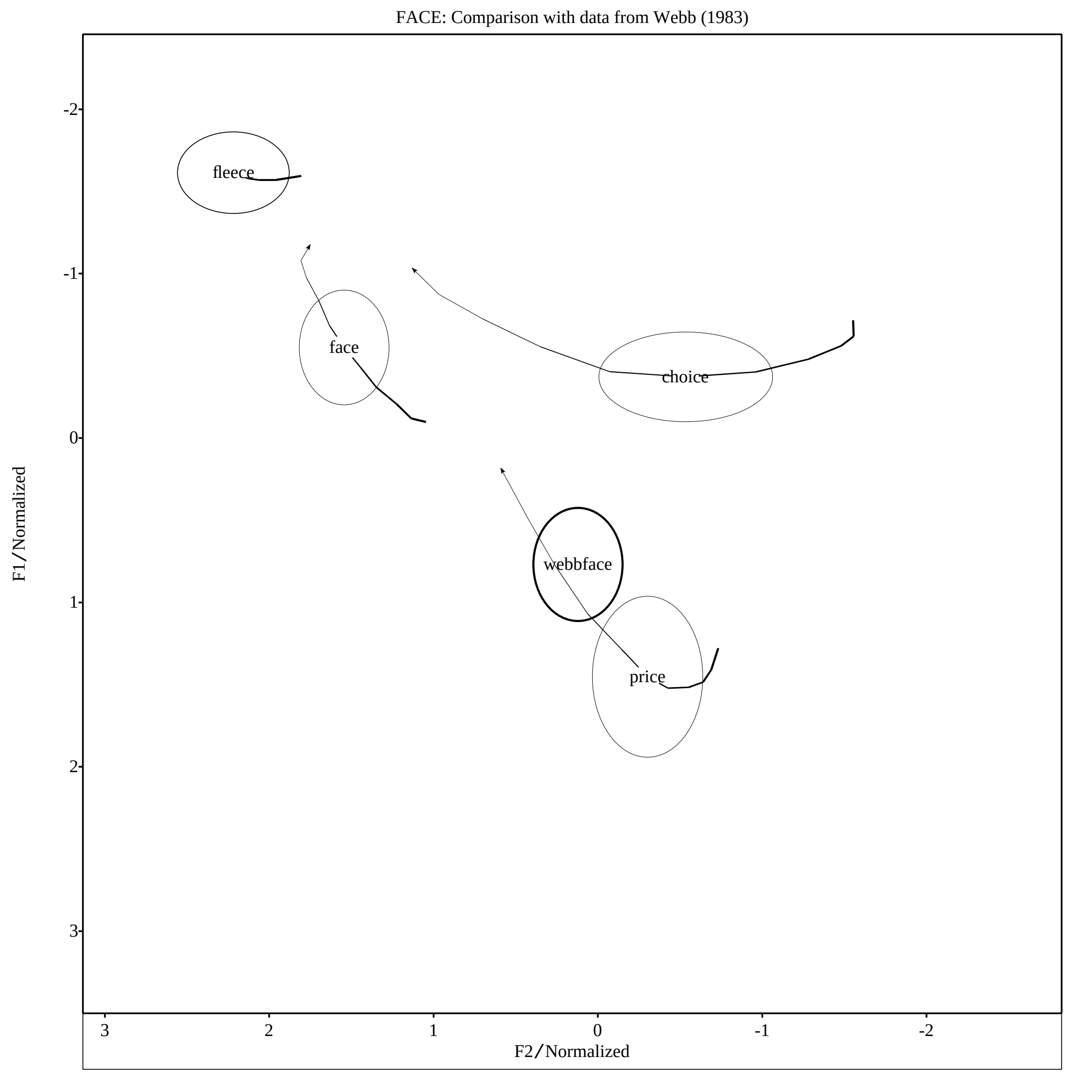


Figure 6.23: FACE: Comparison with data from Webb (1983)

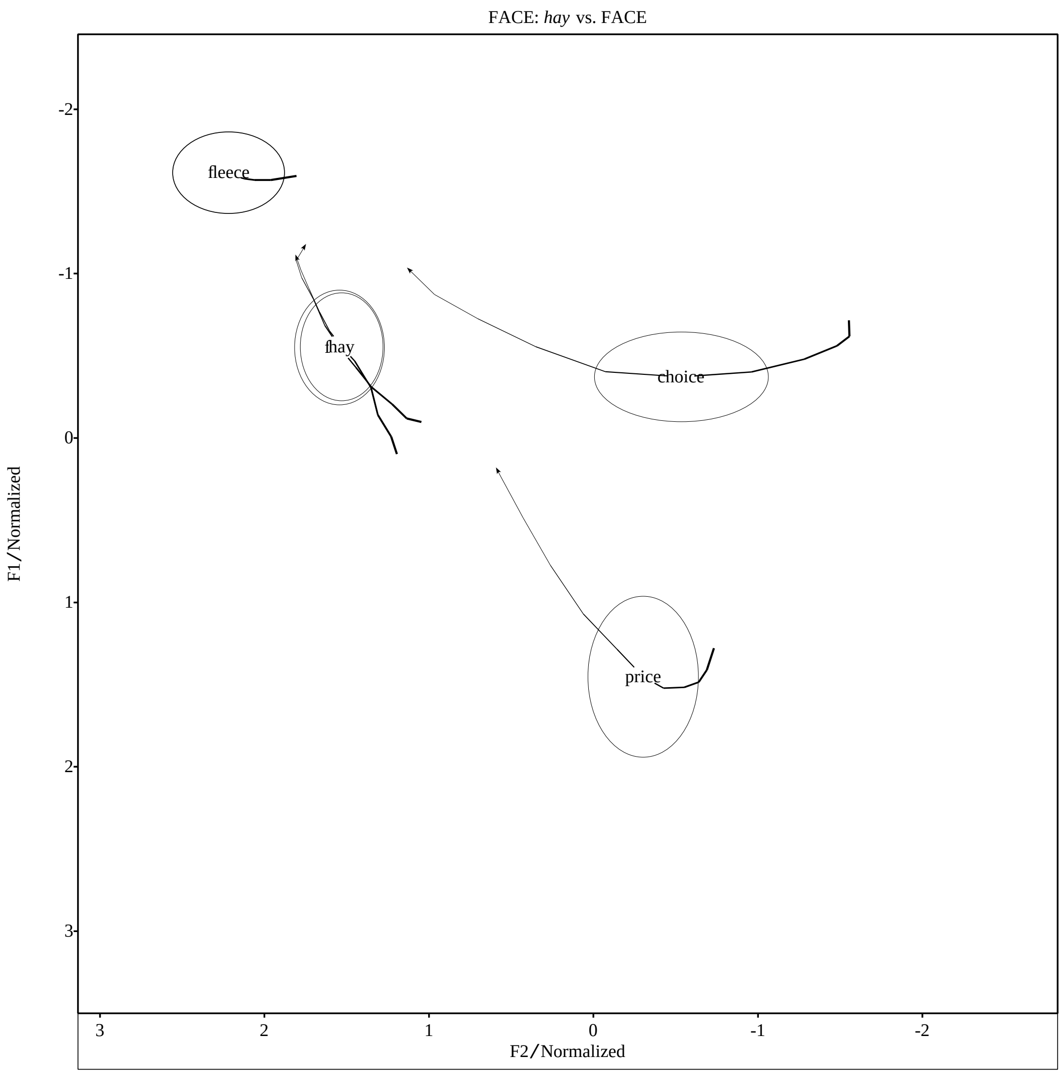


Figure 6.24: FACE: *hay* vs. FACE

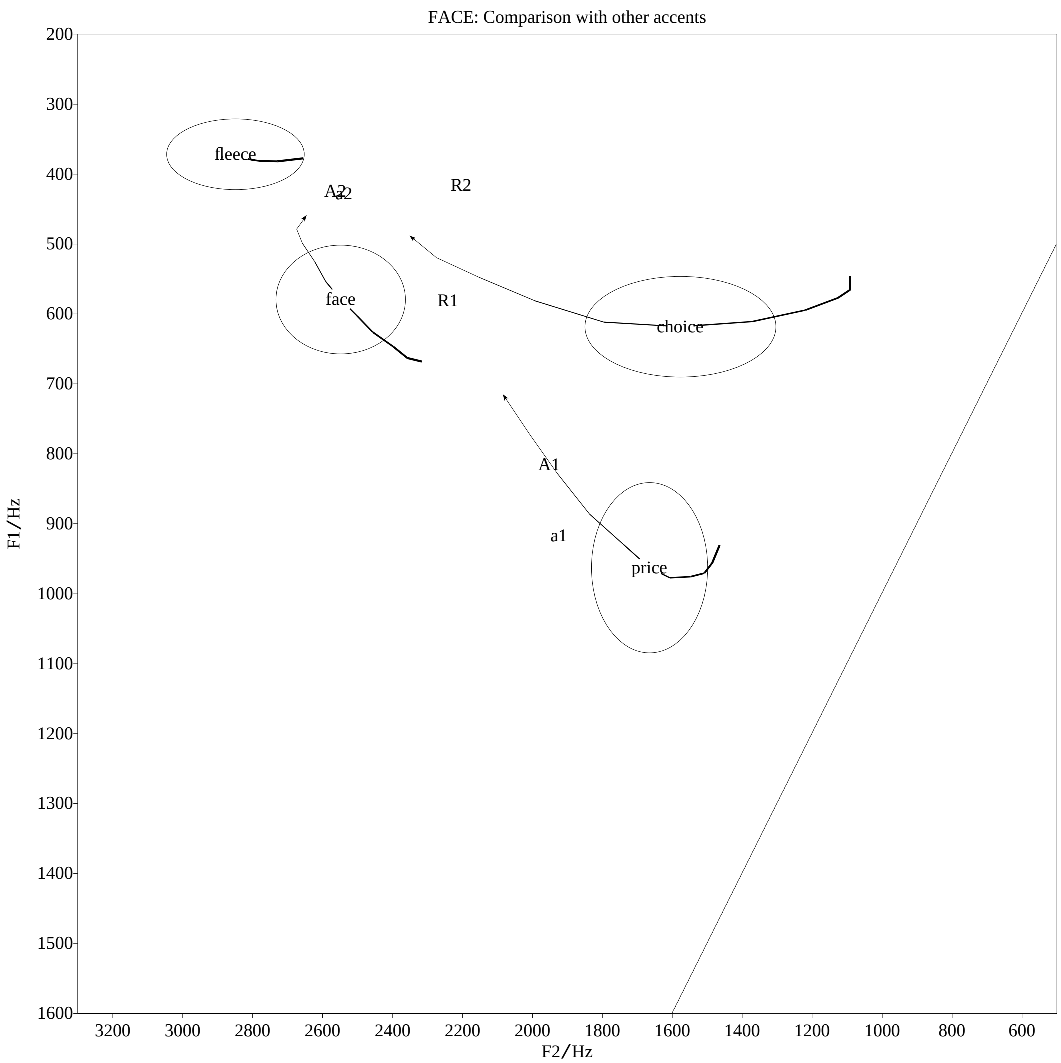


Figure 6.25: FACE: Comparison with other accents

apparent that FACE lowering has taken a far greater 'hold' of AusE than SAE. We note too the more centralized values for RP, a fact which seems to indicate that while AusE and RP appear to maintain a laxed status for FACE (in word-list style), SAEP appears to have (reclaimed?) a peripheral status for this vowel, at least in the more formal styles.

As expected, the position of DRESS in Figure 6.13 on page 197, is a front and close one. Of interest, too, is the degree of overlap between KIT2 (as in *it*, *hit*, *kit* etc.) and DRESS, a fact which accounts for the confusion that speakers of other accents of English often experience in this regard when listening to SAE speakers use the DRESS vowel i.e. *sex* and *six*, for example, sound virtually the same. Note too that there is no evidence of a diphthongal quality to this vowel; it is, by all accounts, steady-state.

Figure 6.26 on the following page provides a comparison of the DRESS-vowel data provided in Webb (1983) and the data collected for the same vowel for the current research[52]. As is clear from this figure, the height of the vowel across the two data-sets is virtually equivalent. It would appear, therefore, that DRESS in SAE is subject to very little style-shifting, and does not vary dramatically across important social parameters. Since Webb's (1983) analysis was only based on 17 tokens (from one speaker) it is unclear whether anything should be read into the higher degree of centralization present in his data. Such centralization to [ë] has been reported in the impressionistic literature, particularly by Lass and Wright (1985) and we have seen above that the relative positions of FACE and DRESS provide some evidence for the peripheral status of the former and the non-peripheral status of the latter.

Figure 6.27 on page 229 compares SAE DRESS with DRESS found in RP and the other Southern Hemisphere Englishes. In this figure we note the dramatically raised DRESS vowel of NZE and the slightly lowered DRESS vowel of RP. SAE DRESS, along with AusE DRESS, finds itself between these two extremes.

Figure 6.28 on page 230 provides a comparison between the SAE data and that for Torgersen and Kerswill's (2004) Ashford. As before, diamonds indicate the young-female data. As in the case of TRAP (§6.3), the more centralized values are, in all likelihood, due to the use of both citation-form and connected-speech data. As can be clearly seen, SAE (and by implication AusE) has not been subject to the same degree of lowering, if any at all. This makes perfect sense if we view the southern-hemisphere Englishes as being behind the southeast of England in terms of the actualization of SECS. While both these accents show evidence of recent TRAP-lowering, the gap left by this lowering, has, for the time being,

[52] Webb's (1983) DRESS data is given as a solid oval, while the data from the current research is given as dotted ovals. The use of such dotted lines in this and later figures is simply a means of making the data easier to read. Note that in the current figure, the DRESS value for the current data is 'hiding' behind that of Webb (1983).

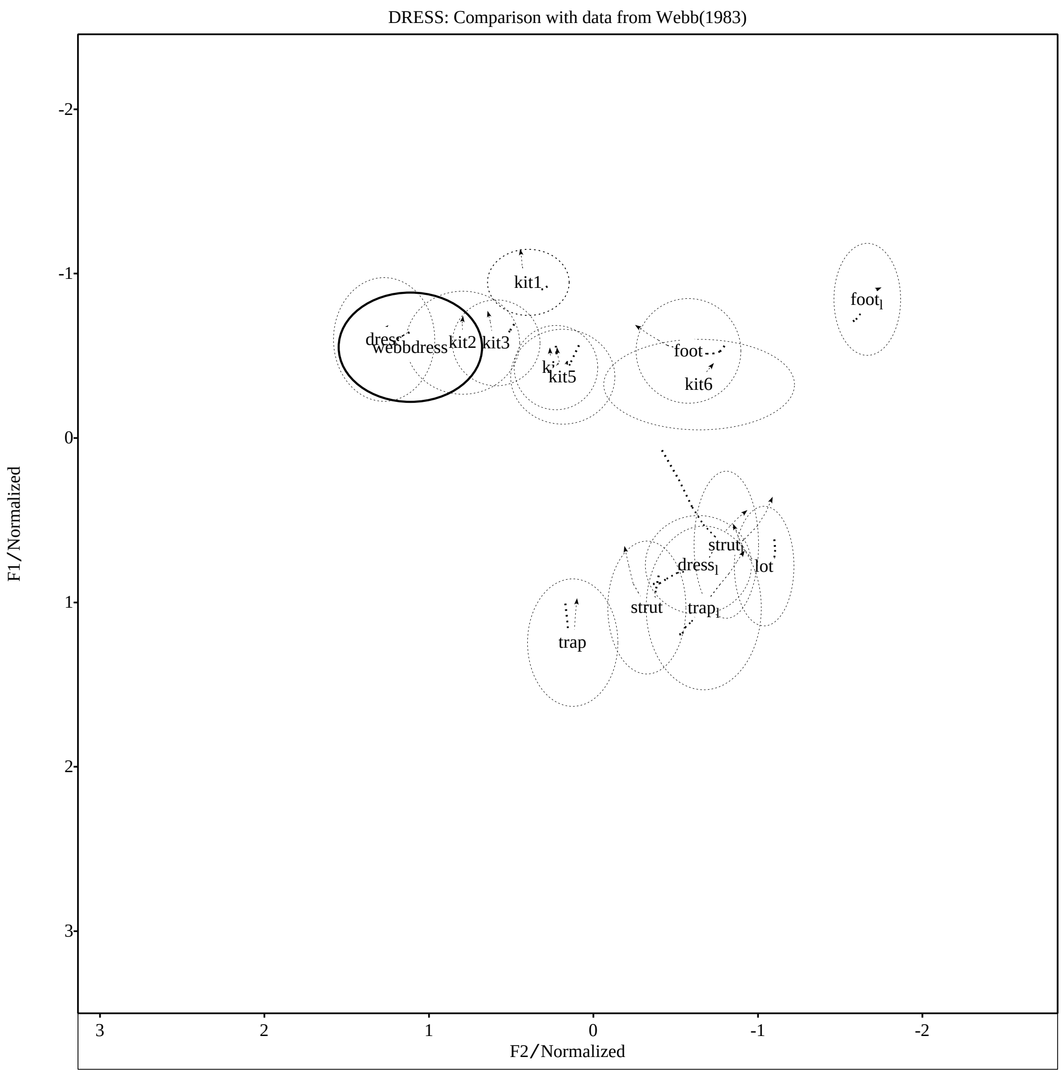


Figure 6.26: DRESS: Comparison with Webb (1983)

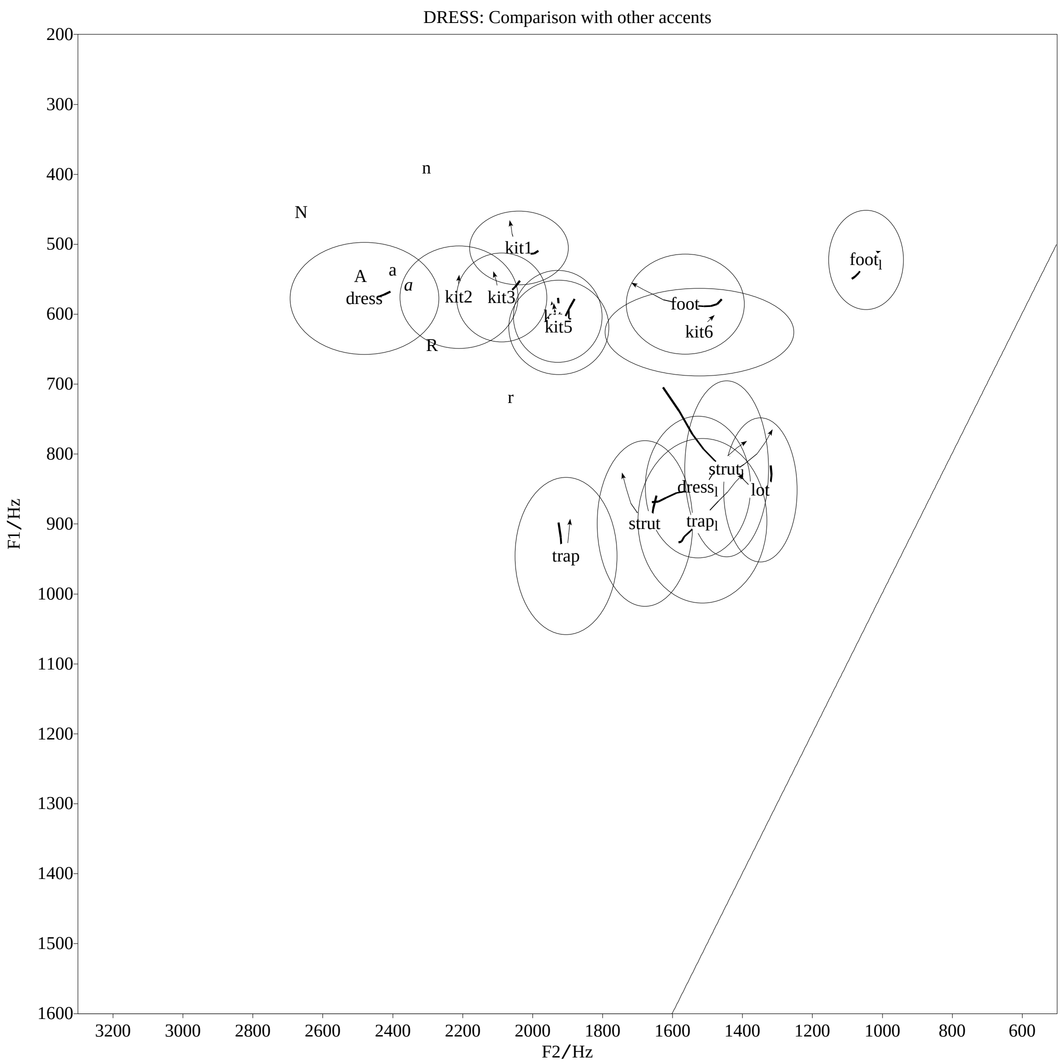


Figure 6.27: SAE DRESS vs. other English accents

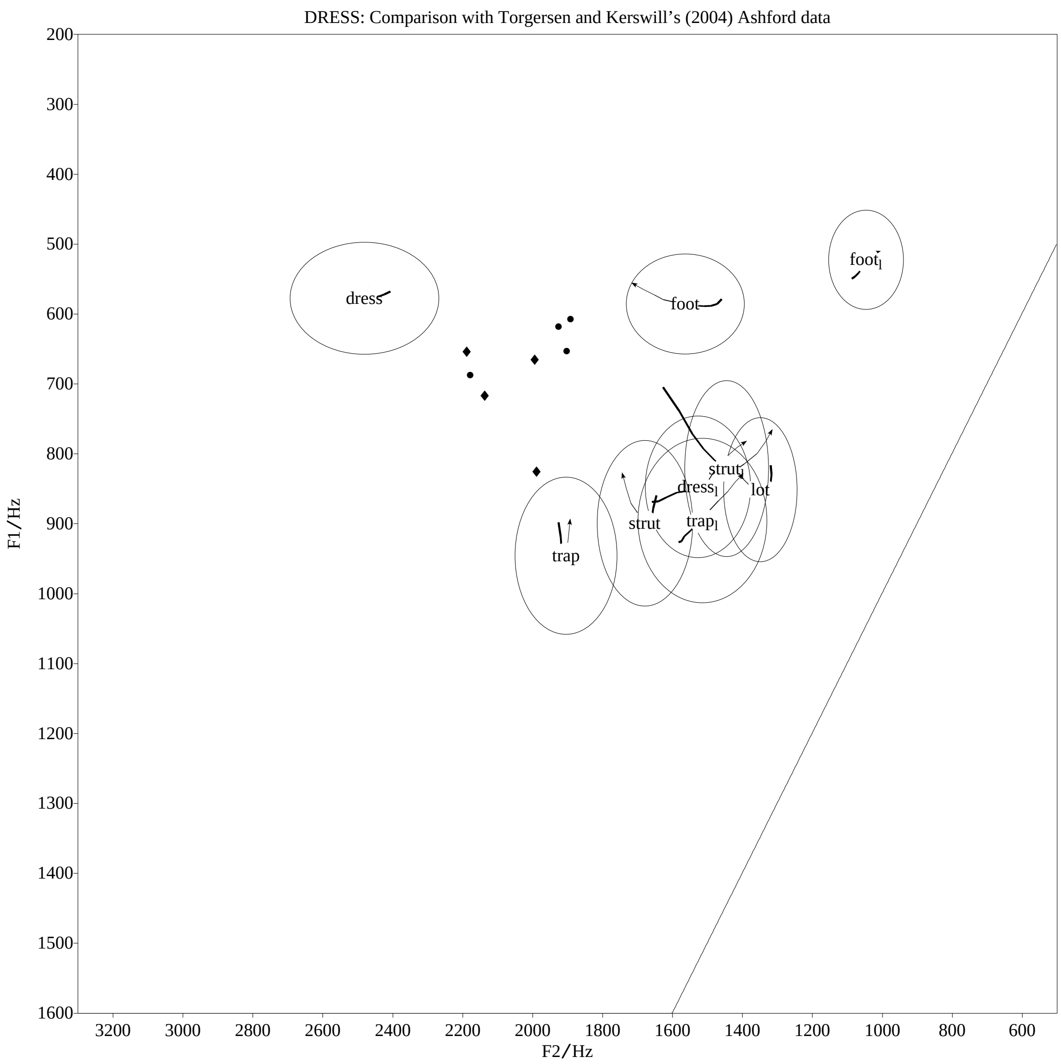


Figure 6.28: SAE DRESS: Comparison with Torgersen and Kerswill's (2004) Ashford data

not yet led to the relevant pull-chain and consequent DRESS-lowering. While RP (and the southeast of England more generally) has undergone a gradual lowering of both the TRAP and the DRESS vowel, it would appear that while SAE (and AusE) have undergone some lowering of the former, the latter has remained relatively stable. NZE, on the other hand, has shown secondary tensing of both these vowels.

The reader is referred again to Figure 6.2 on page 173, which places FLEECE in the context of the other SAE Part-System B vowels. As can be seen from this figure, there is only slight evidence for spectral movement, and this is more than likely due to the presence of *arena* as one of the tokens analyzed i.e. the effect of the preceding /r/. By all accounts, therefore, FLEECE in this data is steady-state; 'pure'. This is contrast to Webb's (1983) claims regarding his own accent, which appears to show formant movement for this vowel. It should be reiterated that most impressionistic accounts of FLEECE in SAE emphasize a lack of diphthongization for this vowel in contrast to its behaviour in other Southern Hemisphere Englishes.

Figure 6.29 on the following page provides a comparison between the current data and that of Webb (1983). The latter author does not provide Hz-values for more than the 'point of inflection'; and one assumes that in this case one is dealing with the F_1-maximum. The overlap between Webb's (1983) data and the current data is apparent, although one should keep Webb's (1983) claims regarding spectral movement in mind.

Figure 6.30 on page 233 compares SAE FLEECE with FLEECE found in RP and the other Southern Hemisphere Englishes. Unfortunately, none of these studies provide more than one target value for FLEECE so the presence of any diphthongisation has been obscured. It is suspected, moreover, that the retracted value for 'N' is a typographic error in Easton and Bauer (2000:98), where the F_2 is given as 2022Hz. The only 'trend' visible is a slight lowering *vis-à-vis* RP. Easton and Bauer (2000:98) pick up on this when they claim a "slight lowering in the Australasian speakers in comparison with their RP counterparts". SAE is perhaps 'playing along' in this regard.

6.4.4 FDF: Synopsis

The data analyzed in the previous few sections confirms the fact that FACE is a marker in SAE, with second-order indexicality; but the lowered and retracted variants have probably not yet been stereotyped, given the reported and illustrated high-degree of stylistic (and broadly sociolinguistic) variation FACE is subject to. Thus while the presence of an extremely narrow diphthong with a front, close nucleus in citation-form GenSAE, as well as the difference between AusE and SAE citation-form values for FACE, might lead one

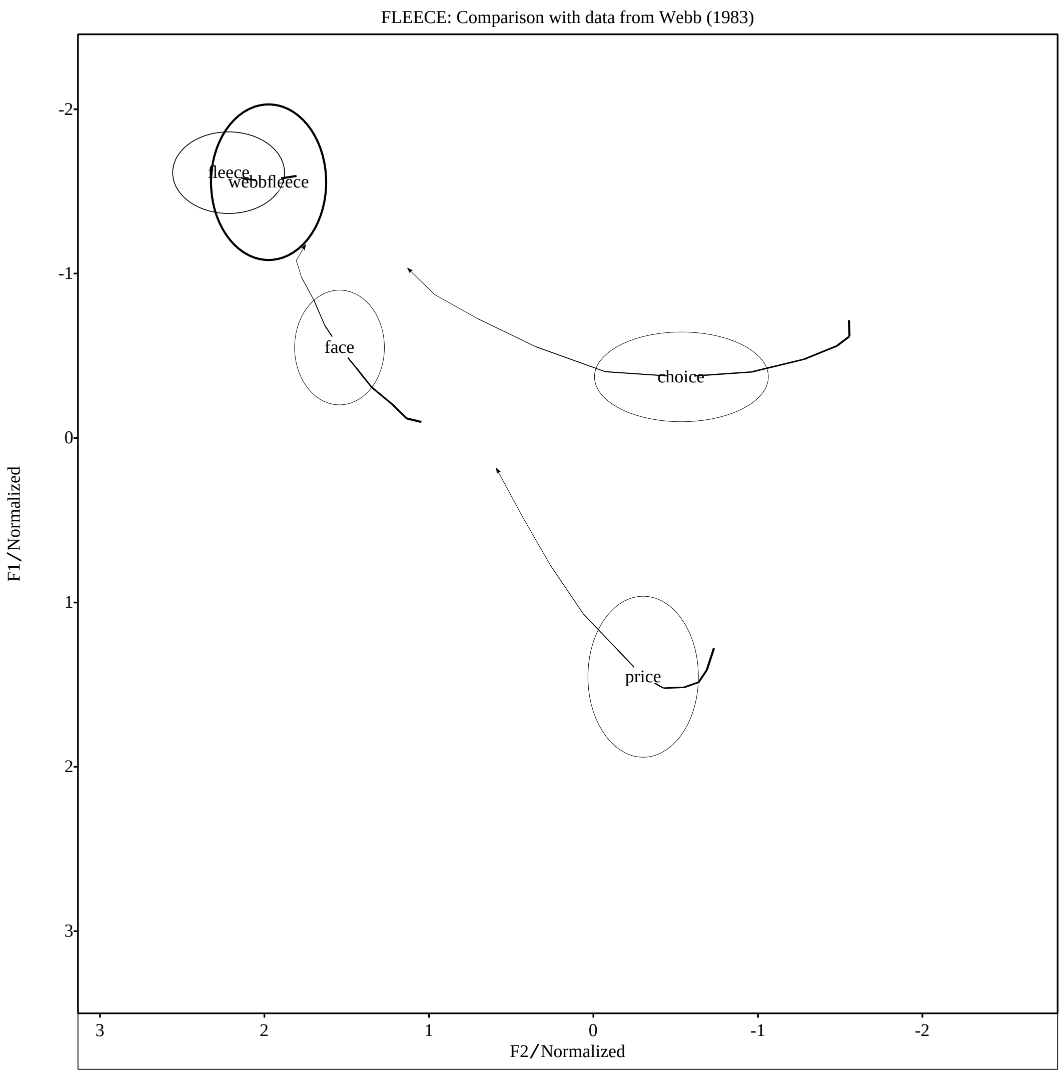


Figure 6.29: FLEECE: Comparison with data from Webb (1983)

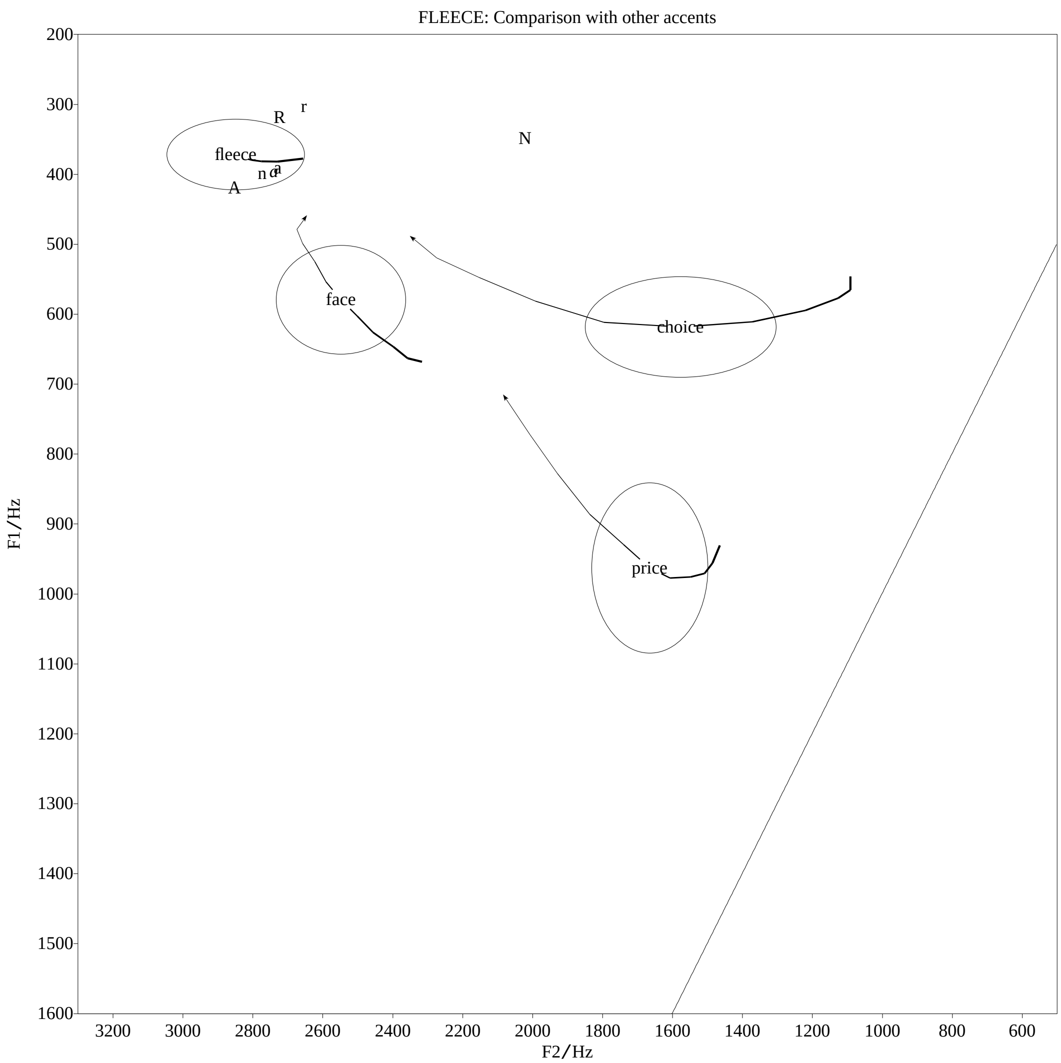


Figure 6.30: SAE FLEECE vs. other English accents

to suggest that these features provide even further evidence for a retreat from a Pattern-4 (or Diphthong Shift) chain-shift pattern in GenSAE, FACE is unlike PRICE or MOUTH in one important respect. In particular, PRICE and MOUTH appear to show very little style-shifting in *GenSAE*, reflected by the fact that this variety has been characterized by a growing avoidance of the stereotyped Broad variants, both in its inception and subsequent diffusion at the expense of BrSAE. In contrast, Lanham and Macdonald (1979) observe that the 'wide'-variants of FACE are not receding and that the variable is, as a whole, in stable sociolinguistic distribution.

This is, I believe, directly related to the likely origins of FACE-lowering in SAE. It would appear that Hooper's (1951) claim for an Afrikaans origin gains support in the light of Trudgill's (2004) observation that many southeast of England varieties had not even undergone diphthongisation, let alone Diphthong-Shifting, of FACE during the period of the first settlement at the Cape[53]. In addition, the Natal settlement, bringing with it a preponderance of north-of-England accents, was likely to be characterized by Pre-Long-Mid-Diphthongization monophthongal values or, in some cases, an RP-like value that was certainly narrower than the values usually associated with Cockney etc., and perhaps even narrower than even modern RP, given the observation that "older speakers [of RP] may have a starting point nearer to C[e]" (Cruttenden 2001:130; my parenthesis). On the other hand, the requisite Afrikaans-English contact and extensive bilingualism that would have led to the adoption of FACE-lowering from Afrikaans into SAE, was, of course, available in both the early Eastern Cape settlement as well as in early Johannesburg[54].

On the other hand, the rather narrow starting point for FACE in older versions of RP, along with the monophthongal north-of-England variants brought via NE and other immigrants to early Johannesburg have *arguably* provided the formal-style GenSAE values. The stable variation that this variable is characterised by seems to provide support for the notion that, in comparison with MOUTH, the adoption of the more prestigious variant in the case of FACE was less motivated by a hypercorrective reaction against the CE/Afrikaans-derived value. What is unclear, though, if this reconstruction is correct, is why this difference arose i.e. why have CE/Broad values of PRICE and MOUTH been subject to stereotyping and not, to the same degree, those of FACE.

The seemingly-peripheral value of GenSAE citation-form FACE, as reported on above,

[53] Although it has to be recognized that, left alone, CE would, in all likelihood, have developed a lower FACE, in accordance with Trudgill's (2004) notion of drift and like AusE and NZE. The influence of Afrikaans was perhaps, therefore, just an added impetus.

[54] This proposed reconstruction will, of course, be substantially bolstered by research on 19th-century varieties of Afrikaans. In this thesis I have simply taken Hooper's (1951) implicit 'word' for lowered FACE-variants in early Afrikaans. Further research is, thus, required in this regard.

also seems to provide some evidence for the lax, non-peripheral status of DRESS in GenSAE, an observation which ties in with previous transcriptions implying a somewhat centralized value (Lass 1990), its acoustic and perceptual overlap with the fronter allophones of KIT, as well as the fact that the current citation-form data did not differ substantially from that provided by Webb (1983) i.e. DRESS appears to be reasonably categorical in its positioning, an analysis backed up by the fact that the impressionistic literature provides little evidence for a discernable distinction between BrSAE and GenSAE DRESS; it is, it would appear, to be a broad indicator of non-Cultivated SAE. A non-peripheral DRESS is contrary to Labov's (1994) 'Southern Shift', but in-line with Torgersen and Kerswill's (2004) SECS-shift, although, as is the case with AusE, it appears that it has not yet been subject to change in response to what might be a SECS-initiating lowering of TRAP in the two relevant accents. This reflects Trudgill's (2004) notion of colonial lag and the conservative nature of the Southern Hemisphere Englishes, although in this particular case, as with TRAP, NZE has diverged from the other two main Southern Hemisphere Englishes by raising its DRESS vowel even further.

The fact that DRESS does not constitute a marker (or stereotype) in SAE is easily linked to the various inputs into SAE. In particular, a raised DRESS was to be found in:

1. CE; transplanted directly from the southeast of England.
2. a late-19th-century version of RP i.e. that to which early Johannesburg was exposed.
3. The accents of the more hypercorrective of the NE immigrants to early Johannesburg.

The more open north-of-England varieties were thus 'swamped' by the above-mentioned inputs and the equivalence between the CE and RP variants precluded the possibility of firm indexical associations.

Turning to FLEECE, the synchronic acoustic data as well as the generally impressionistic data provides little to surprise; by all accounts SAE has always had and still has a pure 'monophthongal' variant of this vowel. The reconstruction of its history is however, of import, not least because it distinguishes SAE from AusE and NZE. Lanham's (1978) contention that monophthongal FLEECE has an Afrikaans origin is bolstered by Trudgill's (2004) contention that a diphthongal FLEECE was unlikely to be an input into the English of early-19th-century settlements. The influence of Afrikaans on CE is of important explanatory value given that, unlike in the other two varieties where the endogenous Diphthong-Shifting of FLEECE (and FACE) was left to continue unhampered, FLEECE-diphthongisation in SAE was instead 'hampered' by the importation of Afrikaans values, both in CE and later in early-Johannesburg. An added influence came perhaps from the various north-of-England accents, which, in some tokens of FLEECE, have a pure monophthongal form. Another possible influence was the presence of some FLEECE tokens of 'pure' quality in the RP of the time; still, as pointed out in §6.4.1 Refined RP, presumably reflecting an older version of this dialect, has Cockney-like diphthongal variants, rendering this possibility speculative at best.

6.5 Conclusion

This chapter has provided a comparative analysis of SAE MOUTH, all of Wells' (1982) Part-System B vowels, as well as two of the front short-vowels of SAE. In the case of the diphthongs it has shown, I believe relatively convincingly, that SAE's 'half-hearted' participation in the Diphthong-Shift, so characteristic of the other two main Southern Hemisphere Englishes, AusE and NZE, is, in at least the case of GenSAE, due to the fact that the development of this variety involved contact, in early Johannesburg, between a variety of English dialects as well as contact between English, Afrikaans and Yiddish speakers. While the influence of the later language is still unclear, the possible influence of Afrikaans has explanatory value both in terms of its influence on early CE as well as on the koinéization process that took place in Johannesburg. This is particularly true in the case of FLEECE where it could have had a restraining influence on the endogenous development of diphthongisation in CE as well as in early-Johannesburg, where it was perhaps 'assisted' by monophthongal north-of-England variants. In the case of FACE, it also perhaps had a role to play in assisting structurally-driven FACE-lowering in CE and certainly added the relevant feature to the early-Johannesburg feature-pool; in this latter context, it had to compete with narrowly-diphthongal or even monophthongal variants from north-of-England sources as well as the RP of the time. Instead of one of these main trends winning out over the other, both were reallocated and reanalyzed as stylistic variants of the same variety, accounting for the present-day situation in which the two variants of FACE are in stable sociolinguistic variation.

In terms of the synchronic status of FLEECE and FACE, the former remains clearly monophthongal, while the latter is characterised by the stable stylistic and more generally sociolinguistic variation mentioned already. In its word-list 'form' it also provides one kind of evidence for the non-peripheral status of DRESS in SAE, a possibility which, in turn, links to SAE's possible participation in a SECS-like shift.

In the case of MOUTH and PRICE, on the other hand, a reaction against Cockney-like CE values was, in the case of the former perhaps the main (even only) basis for the adoption of a virtually categorical backed MOUTH in GenSAE, and, in the case of the latter, one of many influences that led to a similarly categorical avoidance of a backed PRICE in GenSAE and the adoption of a fronted variant. The possible lowering of CHOICE along the non-peripheral track in modern GenSAE would be a logical continuation of the subsequent diffusion of fronted PRICE. The synchronic data also provides an indication that a fully monophthongal fronted-PRICE no longer carries the prestige it might have had a few decades ago, and a change in indexical value has been proposed to account for the diphthongal (although still glide-weakened) nature of this vowel in citation-form style.

Most importantly, the case of backed MOUTH appears to provide evidence against Trudgill's (2004) model of new-dialect formation in tabula rasa contexts, as dealt with briefly in §1.3.3, in particular his emphasis on an indexicality-'free' dynamic that drives this process. In essence, it appears that only an appeal to indexicality can explain the adoption of this variant into GenSAE.

In the case of the two front short vowels that were investigated, TRAP and DRESS, the characteristic adoption of close values in SAE for both these vowels was due to the fortunate 'coming-together', in the late 19th-century, of CE and the RP of the time. In short, both provided similar values. These two particular cases provide, I believe, an additional, although perhaps less convincing, challenge to Trudgill's (2004) thesis of new-dialect formation in tabula rasa contexts. I think it quite possible, given the facts outlined in §3.3 regarding the confluence of linguistic variants brought to early-Johannesburg, that the CE/RP values were in the minority. If the above historical reconstruction is in any way accurate it is therefore likely that the only factor that eventually settled the adoption of these close variants was the prestige they carried, via RP, for almost all concerned[55].

Lastly, and particularly as a result of the comparisons with data from other accents, the SAE data provides some evidence that SAE has begun with the initiating move of a SECS-like chain-shift i.e. TRAP lowering. Additional evidence derives from the fact that DRESS might be on the non-peripheral track. Furthermore, if SAE is indeed shown to be participating in such a SECS-like, structurally-driven, shift, this raises interesting questions regarding the usual correlations, in the variationist literature, between changes-from-above vs. changes-from-below and social class (see §4.2.3.1), given that this *shift* would appear to have been initiated from 'above'. This, in turn, links to the notion of a reverse chain-shift, as briefly mentioned in §2.3.2.1.

The next chapter will focus exclusively on the remaining short front vowel i.e. KIT.

[55]We note here that some of the NE immigrants to Johannesburg probably used these close variants and/or were aware of their indexical value.

CHAPTER 7

The KIT Vowel

7.1 Introduction

Although KIT, sometimes referred to as "short I" is, in terms of Labov's (1994) Southern-Shift-related Pattern 4, intimately connected with the lowering and laxing of FLEECE, and thus, more properly belongs under the previous chapter, it has been selected for individual treatment mainly as a result of the extensive analysis it has received in the extant SAE literature; far more than any other vowel in the SAE system.

7.2 KIT: Other Accents of English

Received Pronunciation: Some striking realizations of KIT mentioned by Wells (1982:127–8) include a closer and fronter [i]-quality for the West Midlands of England (e.g. in Birmingham), an open [ɐ]-like quality for Scottish English and, more importantly, the prevalence of what he calls "shading" in both Southern Hemisphere Englishes and the English used in the southern states of the USA. In these accents, "KIT is particularly subject to . . . the development of different allophones conditioned by the place of articulation of the following consonant" (Wells 1982:128). It is interesting to note that Lanham (1967:31–2) implies such shading as a feature of RP, although the distribution of different qualities is, according to him at least, interphonemic as opposed to allophonic. Thus Lanham (1967) claims that in RP a central [ɪ̈][24] {[ɨ̞]} quality is contained in words such as *sin*, *pit*, *will*, *ship* and *dinner*, while a more front (and conventional) [ɪ] is found in the context of an adjacent velar, word-initially, after /h/ and "before syllable margins with palatal consonants" (Lanham 1967:31). Lanham (1967:114; footnote 5) also claims, in fact, that some RP-speakers have the slightly lower [ɜ][34]{[ɘ]} for the central allophone of KIT.

Readers with prior knowledge of the basic 'design' of the so-called 'KIN-PIN Split' in SAEP will find the implied parallels between RP and SAEP suggestive. This particular issue is taken up in more depth in §7.3. For the time being, it is worthwhile noting that such shading in RP is hardly surprising since, as confirmed by, for example, Schmitt (2007:322), "there are two major allophonic varieties of all front checked vowels in English . . . a raised variety before velars . . . and a lowered one before dark /l/". Note the implication here that there are other more *minor* forms of allophonic variation as well. It is of interest, though, that Lanham (1967) does not refer specifically to a retracted allophone of RP KIT before final /l/.

Wells (1982:291) claims that the realization of KIT in RP very often differs with the age of the speaker: "relatively close and peripheral qualities are associated particularly, but not exclusively, with old-fashioned RP; relatively open and central qualities are common with younger speakers".

According to Cruttenden (2001:107), the KIT vowel is pronounced in General RP with a quality of "a centralised C[e] = [ë]". Lass and Wright (1985:138) concur with this assessement. Refined RP often diphthongizes this vowel, giving it a [ɪᵊ]-like quality. As far as available acoustic data on the KIT vowel in RP is concerned, Figure 6.19 on page 210 provides data on "formant frequencies for RP pure vowels said in citation form and in connected speech" (Cruttenden 2001:102) for female speakers. The figure is taken from Cruttenden (2001:102), who bases the figure on data derived from Deterding (1990) and Deterding (1997).

Other Accents of Britain: Some non-RP British values for KIT have already been provided in the previous section. For the southeast of England, Torgersen and Kerswill (2004) seem, interestingly enough, to find evidence for KIT-fronting – see Figure 2.14 on page 64. This is, to a degree, somewhat surprising given that one would expect a lowered TRAP, followed by a lowered DRESS, to initiate KIT-lowering as well, as is evident, by all accounts, in younger RP speakers. One possible explanation of this is that RP (as a conservative variety) is somewhat behind more vernacular accents of the south-east of England and has thus not undergone the full impact of SECS. For one, it appears that FOOT-fronting would have an obvious role to play in KIT-fronting and while this appears to have begun in RP (see §8.4), it appears to not be as advanced as in the vernacular accents. One might wish to argue, therefore, that while RP has undergone some of the more classic SECS changes (e.g. lowered TRAP and DRESS) it has still not experienced the full force of FOOT-fronting which appears, in other accents, to be, in turn, pushing KIT to the front. In the absence of this pressure, KIT is therefore allowed to lower a little. Thus, Trudgill (2004:43) claims KIT-lowering to be a feature of the southeast of England, and, in fact, contrasts these accents with Cockney which, he claims, remains at [i].

Importantly, Wells (1982:356) also implies a more *open* value for KIT in many accents of the north of England. For Scotland, Wells (1982:404) deserves to be quoted in full:

> "In KIT words, the quality of /ɪ/ is – in an educated Scottish accent – much the same as in RP. In more popular accents it may be considerably opener and/or more retracted. In the north-east, it is often pretty open and [ɛ]-like …even among educated speakers. In Glasgow it ranges from [ɪ] to [ʌ̣]{[ʌ̝]}, including various intermediate qualities".

We have already noted, in §3.3.1, that the British settlement of Natal was characterised by a sizeable Scottish contingent particularly from Glasgow and the Highlands.

U.S. Accents: For the US South, as confirmed in Fridland (2008:69), FLEECE and KIT often reverse positions, in accordance with Labov's (1994) Pattern 4 i.e. FLEECE is diphthongised, while KIT moves onto the peripheral track and becomes more front and close as a consequence. In Philadelphia, in contrast, we appear to have the backing and falling of KIT as an incipient change (Labov 1994:82), connected, it appears, with a similar falling and backing of DRESS as well as a raising of STRUT.

NZE and AusE: It is highly instructive to compare the KIT vowel in SAE with the same vowel in NZE, the latter being another Southern Hemisphere English, which, in its modern version, displays KIT retraction; it (and other NZE vowels) has also been subject to intensive acoustic analysis over the last few years. What is particularly important to emphasize is that KIT-centralization in NZE appears to be a relatively recent phenomenon, and in all likelihood linked to the secondary raising of TRAP and DRESS reported for this dialect (Evans and Watson 2004:195). These latter authors also point out that over the last thirty years or so KIT in AusE has undergone fronting.

Lastly, it is perhaps pertinent to point out that one possible thesis advanced by, for example, Trudgill (1986), for the centralization of KIT in NZE (and thus indirectly for the similar phenomenon in SAE), but a thesis subsequently rejected in Trudgill (2004:110), was that KIT-centralization "was the result of Scottish English input into the dialect mixtures that eventually led to the development of this feature," Scottish English being well-known for its low and central variants.

NZE is often characterised by the wholesale retraction of this vowel as opposed to the 'split' characteristic of SAE. Gordon et al. (2004:26) give [ɘ] with the most advanced values being in the region of [ə] or [ɜ] and with Cultivated speakers having [ɪ̞]. In contrast, the AusE KIT vowel has increased its peripherality relative to RP (Watson et al. 2000), with Broader speakers generally showing a more peripheral (higher F_2) KIT vowel than General or Cultivated speakers of AusE (Harrington et al. 1997:163). According to Watson et al. (2000:64), this peripheral KIT vowel is in turn related to two other features of AusE: diphthongization of FLEECE "and the loss of the /ɪ/ - /ə/ contrast in unstressed position". The first of these features links AusE unambiguously to Labov's (1991) Southern Shift, which is characterised, among other things, by an 'exchange' between the two short high front vowels and their long counterparts. According to Watson et al. (2000:64), the pattern shown in AusE is one way for the system to accommodate for possible confusion between the high short front vowels so characteristic of Southern Hemisphere Englishes. As Labov (1994:138) puts it, "if /i/ had not moved to an upper . . . position, its margin of security with /e/ would have been seriously endangered". Another option is for the KIT vowel to centralise as in NZE and SAE. In their study of vowel shifting in NZE, Watson et al. (2000) provide the vowel charts in Figure 6.20 on page 210 for Old versus Modern NZE. As can be seen, the modern NZE KIT vowel is far more centralised than the same vowel in Old NZE, and is in fact lower than both DRESS and NURSE. Of interest, though, is the following comment by Taylor (1991:79) with reference to the historical behaviour of the KIT vowel in SAE:

> "But this impingement of raised /ɪ/ on /iː/ could be avoided by a raising of /iː/. This latter development may in fact be observed in SAfE, where /iː/ is frequently realized as an extremely close and fronted monophthong. In this respect, SAfE contrasts with NZE, where /iː/ is realized by an Australian-like dipthong [ɪ̈i]. It has thus been possible for SAfE /ɪ/ to undergo both centralization *and* raising, without the possibility of either variant merging with neighbouring sounds. Interestingly, a similar split appears to have occurred in NZE also, although on a much

> more limited scale. Bauer (1986:236–7) notes that alongside the dominant realization of NZE /ɪ/ as "very central open", there exists a much closer allophone, confined, it would seem, to velar environments, especially before /ŋ/".

The diphthongisation of FLEECE in NZE of course raises the question of why KIT centralized in the first place and, secondly, the existence of a fronted allophone in velar contexts points to the possibility of there being less of a difference between SAE and NZE than is commonly assumed. This is of particular relevance to Natal English which, as will be discussed in §7.3, is characterized in the literature as being centralized in all contexts but the velar and thus potentially very similar in this respect to certain varieties of NZE.

According to Lass (2002:114), the KIN-PIN Split in SAE is possibly a change-in-progress. Firstly, the "split of fronter and backer allophones" has not completely resolved itself and "there are therefore lexical exceptions" (Lass 2002:114). Secondly, there is evidence that the KIT vowel "is continuing to centralise and lower" (Lass and Wright 1985:154). There is similar evidence for the centralisation of NZE being a change-in-progress. So, for example, Easton and Bauer (2000:98), in their acoustic analysis of NZE vowels, provide evidence that the younger the speaker the lower and (interpreted from a certain perspective) more retracted the KIT vowel.

It was mentioned above that one of the possibilities for KIT vowel movement was towards the periphery (with the related diphthongisation of FLEECE), as in AusE. The motivation behind this shift was to maintain a functional difference between the DRESS and KIT vowels. The other option (i.e. centralisation) is claimed by Watson et al. (2000) to be similarly motivated and to provide a similar solution, just in a different direction[1]. This process of centralisation does, however, pose a problem in the sense that the KIT vowel encroaches on schwa, [ə], the vowel quality usually reserved for vowels in unstressed, weak position. On the one hand "some studies of NZE have suggested that the NZE HID has the same quality as a schwa" (Watson et al. 2000:64); on the other hand "considering the high functional load in English of HID and schwa, and keeping in mind the principal of maximal perceptual contrast . . . it does not make sense for HID to take on the quality of schwa, since this would decrease perceptual contrast and lead to a considerable loss in comprehension".

Watson et al.'s (2000) theoretical objection to KIT having a schwa quality is com-

[1]Of course a more serious problem for all these 'functional' accounts is that, as shown convincingly in Labov (1994), languages abound in mergers.

plicated though by the fact that their own data provide evidence for such a quality[2], and an earlier study (Watson, Harrington and Evans 1998:203) in fact advances [ə] as the proper transcription for the KIT vowel in NZE. A 'solution' to this problem is advanced in Watson, Harrington and Palethorpe (1998), which provides "a kinematic and acoustic comparison of the vowel spaces of New Zealand English . . . and Australian English" (Watson, Harrington and Palethorpe 1998:2363). In this study, a number of physiological measurements are reported on which seem to indicate that the KIT vowel remains a high vowel; centralised though not lowered. So, for example, KIT is found to have a higher jaw position than DRESS while the tongue height for KIT, NURSE and DRESS are the same. Furthermore, the degree of lip aperture for KIT also seems to place it higher than DRESS i.e. KIT "was produced with more compressed lips" (Watson, Harrington and Palethorpe 1998:2364). Their conclusion is that "it is perhaps the case that NZE HID has backed but not fallen . . . further, the results suggest that a phonemic transcription for the NZE HID with a barred [i] (cardinal vowel 17) might be more appropriate than schwa". In other words NZE KIT should be transcribed as [ɨ]. While the disparity between the acoustic and the physiological data is intriguing, no explanation for the disparity itself is advanced. Since the lowering of the KIT vowel on the F_1 plane must have some physiological basis, the results of Watson, Harrington and Palethorpe (1998) simply leave the question posed by the acoustic data unanswered.

7.3 KIT in SAE: The Impressionistic Data

Of all the vowels in SAE, the KIT vowel has been subject to the most scrutiny. It has also been the subject of controversy, both in terms of the historical origins of its unique qualities and, relatedly, its phonemic status. As such, it has been separated out for individual treatment. SAE is characterised by, indeed (in)famous for, its 'KIN-PIN Split': SAE is the only variety of English in which the tokens of these subsets (like *kin* and *pin* themselves) do not rhyme. The term has, however, been placed in scare–quotes because the actual phonemic status of the split is in question, as will become clearer below.

Hopwood (1928), Hooper (1945), Breckwoldt (1961): The 'split' itself is not mentioned in Hopwood (1928:17; my emphasis), who does, however, claim mid and central values for this vowel i.e. that "**ɪ > ë** or **ə** in *stressed* and weakly stressed syllables". Hopwood (1928) gives, as an exemplification of the above values, [nëbz] for *nibs*

[2] See figure 6.20 on page 210.

and [ˈspərət] for *spirit*. As can be seen from the use of the obstruent [r] in the latter transcription, the emphasis in Hopwood (1928) is on a very Broad, if not Afrikaans, accent; [ˈspɪɹɪt] is given for RP. Hopwood's (1928) explanation of the origin of the SAEP value is as follows:

> "The . . . change is due in the first place to the AE. spelling pronunciation of words with written *i* in closed syllables, pronounced [turned-i] in A., and ë before nasals and sometimes *r* and *l*. It was then extended in AEP to StEP [i.e. RP] **ɪ**-pronunciations not written with *i* in closed syllables" (Hopwood 1928:17; my parenthesis).

Hopwood (1928:18) also claims the occasional [ə] pronunciation for Afrikaans *i*, "which would also help this change to take place in SAEP". He also mentions an alternative pronunciation of KIT in SAEP i.e. the closer and fronter [i], again based on substitution with an Afrikaans sound, this time the *ie* in words like *diep* and *skiet* [trans.: *deep* and *shoot*]. Three examples of SAEP words that, according to this author, use this quality are *history*, *pin* and *is*.

Hooper (1945) notes the the variable use of schwa as opposed to [ɪ] in words such as *swim* across social-class and gender. As expected, the relatively lower-class as well as male respondents all made more use of the former.

Breckwoldt (1961:7) follows pretty much in Hopwood's (1928) footsteps by claiming [ə] and [iː] qualities for SAE KIT thus "pɩn > pən . . . wɩʃ > wiːʃ".

Lanham and Traill (1962), Lanham (1964?; 1965; 1967; 1982): In their analysis of SARP and non-SARP SAE, Lanham and Traill's (1962) description of this vowel is substantially more sophisticated that the previous authors' and needs to be understood in the context of *their* understanding of its phonemic status. This, in turn, relies upon their notion of 'marginal contrast', which relates to "sound units which only rarely act as the sole differential between words. The term 'marginal' is applied to such units: 'marginal phonemes' partaking of 'marginal contrasts"'(Lanham and Traill 1962:12). In particular, the authors make use of this notion to describe and explain the KIT vowel; in short, it is not one vowel, but two or more. A phonemic split has occurred, although it is, at best, a marginal one. By way of contrast, it should perhaps be noted that the latter Lass (1990:276) argues, for example, against the use of marginal contrasts in this sense. He, for example, claims that "it would be best to take even occasional minimal pairs rather as the results of failed lexical diffusion of a relatively recent change than as part of an oppositional system". The same author later advances this position against the possibility of a phonemic split in KIT in SAE (Lass 1990:284; footnote 12). Returning to Lanham and Traill (1962:12–3), however, these authors mention two kinds of evidence that point to such a marginal contrast and which exist in the pronunciation of some of their subjects:

1. Firstly, there are cases of unconditioned vowel qualities residing *consistently* in only a few words. The example provided by these authors is that of the KIT vowel in *children* in many SARP 'A' pronunciations i.e. [ï]^{24}{[ɨ̞]}, phonemicized as /ï/. The consistency of the phonetic value is important here, given that free variation could be argued for if there was relatively random fluctuation between phonetic values. It is, thus, troublesome that Lanham and Traill (1962:23; my emphasis) offer the following comment with respect to their putative /ï/ phoneme:

   > "The tenuous position of this vowel within the vowel pattern is reflected in the unusual degree of fluctuation connected with strongly stressed /ï/, which is partly free fluctuation within the general area [ï]^{24}{[ɨ̞]}, *and partly fluctuation with neighbouring phonemes*. SARP 'A' speakers carry this fluctuation into the articulatory areas of /u/{/ʊ/} . . . others move forward to [ɜ]$^{33/43}${[ɘ̟˕]}, i.e. /ɜ/, some fluctuate between all three. The word most consistently presenting /ï/ is *chíldrèn*."

   In general, the phonemic status of /ï/ appears tenuous, not just systematically, but theoretically too.

2. The second kind of evidence involves vowel qualities which are, in the main, in complementary distribution, but where a few words do not adhere to said distribution. The existence of any minimal pairs is, of course, an added incentive for recognising vowel qualities as representative of separate phonemes. Again, Lanham and Traill (1962) use KIT–vowel tokens as the basis of exemplification. Thus the first author claims for his own pronunciation[3] the use of a conditioned mid–central [ɜ]$^{33/43}${[ɘ̟˕]} for words such as *tin* and *pit*, and a similarly conditioned high–front variant for words such as *brick* and *pig*. There are, however, a number of words which do not fit the pattern: the author mentions "*pin*, *finish*, *wind* (noun), *window* and a few others . . . containing an environment in which [ɜ]{[ɘ]} is predicted"(Lanham and Traill 1962:12), but which have the higher and fronter alternative. The 'clincher' minimal pair provided by Lanham

[3]It is perhaps worth mentioning that Lanham was South-African born.

and Traill (1962:13) is *finish* and *Finnish*, the first word's first vowel having the high-front variant, with the latter displaying a mid-central value for the first vowel.

It is on the basis of the above considerations that Lanham and Traill (1962:18) distinguish between three different (marginal) KIT-vowel phonemes for the SARP 'A' system: /i/, /ï/ and /ɜ/. Of particular interest is the authors' claim that "marginal contrasts often co–occur with a fluctuation between phonemes in the representation of particular words" (Lanham and Traill 1962:13). According to them, this is often a reflection of a change–in–progress, and the authors mention SARP 'A' speakers who maintain a marginal contrast in their formal register, but who 'collapse' this contrast in colloquial style in favour of a fully–predictable distribution of elements. In general, a higher degree of predictability (i.e. complementation) is claimed for SARP 'B' and non-SARP SAE as well as for younger speakers. The authors summarize as follows:

> "Our reading of this evidence is that SAE is moving towards the establishment of one phoneme /i/{/ɪ/} with allophones ranging from a high front [ɪ] to approximately [ɜ]$^{33/43}${[ə̟˕]} or even a lower vowel. *Pin*, *wind*, etc. will then be [pɜn], [wɜnd], etc. in all circumstances, for all speakers." (Lanham and Traill 1962:13).

Thus for SARP 'B', Lanham and Traill (1962:18) claim a fully predictable relationship between "the allophones corresponding to SARP 'A' /i/ and /ɜ/". They constitute one phoneme, which the authors label /i/. This phoneme also includes SARP 'A' /ï/ – the /l/ in *children* being the relevant conditioning environment (see below).

With respect to the various conditioning environments that determine the quality of KIT tokens in SARP 'B' (and *most* of the KIT tokens in SARP 'A'), the authors provide the following list:

1. A following velar consonant e.g. in *sick*.
2. A preceding velar consonant e.g. in *kiss*.
3. Initial position or post-/h/ e.g. *it* and *hit*.
4. A following palatal consonant (*fish*); "or the sequence /n/ plus palatal consonant (e.g. *dingy* . . .)" (Lanham and Traill 1962:20).
5. a following tautosyllabic /l/ e.g. in *milk*.

The high and front allophones are conditioned by environments 1–4, a syllable-final /l/ has a further retracting influence on KIT, while the the unconditioned allophone is the lower and more central one. In accordance with Lanham and Traill (1962:20), the first four environments above have been listed in their order of conditioning strength. Thus, in SARP, an immediately following velar consonant (/g, k, ŋ/) as well as an immediately preceding one, usually conditions an allophone in the region of [ɪ]^{21}{[e̞]} or [ɪ]$^{21/22}${[ɪ̟]}, although a preceding velar "is often weaker and produces allophones in the range of [ɪ]$^{21/22}${[ɪ̟]} to [ɪ]^{22}{[ɪ]}" (Lanham and Traill 1962:20). KIT in initial position or after /h/ is produced, according to Lanham and Traill (1962:20), as a slightly more retracted allophone, in the region of [ɪ]^{22}{[ɪ]} to [ɪ]$^{22/23}${[ɪ̠]} i.e. approaching but not quite [ɪ̈]. Lanham (1967:80), however, gives [ɪ̈]^{23}{[[ɪ̈]]} as a possibility for conservative SAE.

The influence of a following palatal is variable and differs across the SARP 'A'/SARP 'B' boundary. In SARP 'A' the relevant allophone is the same as that conditioned by KIT in word-initial position or following an /h/ i.e. [ɪ] or a slightly–more retracted [ɪ̠]. In SARP 'B' however the value of the allophone depends, firstly, on the nature of the palatal consonant: voiceless palatals (/ʃ, ʧ/) generally conditioning higher and fronter allophones than voiced palatals (/ʒ, ʤ/). Secondly, the value is usually more retracted and lowered than for SARP 'A' i.e. "[ɜ]^{33}{[ɘ̟]} or even uninfluenced [ɜ]$^{33/43}${[ɘ̟˕]}" (Lanham and Traill 1962:20). Thus, for example, Lanham (1965:94) asks us to "notice …how *bridge*, *dingy* pronounced with [ɪ] in British English is already [brəǰ] and [dənǰɪ] in S.A.E.".

For non-SARP SAE, Lanham and Traill (1962:20–1) mention the possibility of an even higher and fronter allophone conditioned by all the environments given above, bar post-vocalic /l/ i.e. [ɪ]^{21}{[e̞]} to [i]$^{11/21}${[i̞]}. From a modern perspective these values appear particularly Broad, if not indicative of AE, an analysis which finds support in Lanham (1967:80; my parenthesis) where the author qualifies as follows:

> "…abnormally high-front [i]^{11} or $^{11/21}${[i ∼ i̞]} is heard …in regions where Afrikaans influence is strongest. In Natal and the English speaking cities, allophones in [non-conservative SAE] may be similar to those in [conservative SAE]".

The main exception to this similarity appears to be KIT before final-/l/. According to Lanham and Traill (1962), this has a substantive retracting influence in SARP 'B', reflecting a more general trend in non–SARP SAE. SARP 'B' has values in the region of [ɜ]^{34}{[ɘ]} or [ɜ]^{35}{[ɘ̠]}, while non–SARP SAE has [ɜ]$^{35/46}${[ɘ̠˕]}. Lanham (1965:93) goes further by claiming a partly-unrounded [ʊ] as the value. The author also claims that FOOT and KIT will possibly undergo neutralization before /l/ “and *pill* will be pronounced as *pull* is in present-day Standard English” (Lanham 1965:93). In contrast, SARP ‘A’ KIT is immune to a final tautosyllabic /l/ and remains at [ɜ]$^{33/43}${[ɘ̟˕]}. This, however, contradicts latter comments by the first author regarding CulSAE. In addition, and surprisingly, the authors claim that the effect of a final–/l/ overrides that of an initial velar consonant, initial /h/, etc: words such as *kill* and *hill* have “well centralised allophones” (Lanham and Traill 1962:21; footnote).

Up until this point, the authors’ arguments are relatively easy to follow. Yet Lanham and Traill (1962:18) then obfuscate matters somewhat by including /ə/ as a contrastive unit within the SARP ‘B’ system. Lanham and Traill (1962:19) point out that an important difference between SARP ‘A’ and SARP ‘B’ is the high frequency of [ə] in stressed syllables in the latter, contrasting with its absence in the same environment in the former. They provide the following explanation:

> “The unconditioned representative of /i/{/ɪ/} in SARP ‘B’ is generally ɜ$^{33/43}${[ɘ̟˕]} ... A stronger or weaker tendency for [ə] to replace [ɜ] in SARP ‘B’ ... is part of an important trend in SAE. Examples quoted ... may well occur with */ə/ and not /i/*{/ɪ/} therefore” (Lanham and Traill 1962:21; my emphasis)

A phonemic /ə/ is also evident from the chart provided on page 18 of Lanham and Traill (1962) as well as from other comments. Lanham and Traill’s (1962) SARP ‘A’ has, thus, /i/, /ï/ and /ɜ/ while their SARP ‘B’ has /i/ and /ə/. According to Lanham and Traill (1962:24), SARP ‘B’ /ə/ is realised as [ə]^{44}{[ə]}. Compared to non-SARP SAE there are relatively few words in SARP ‘B’ which choose /ə/ over /i/ – examples provided by the authors are “*vívìd* /vəvəd/, *prétty* /prəti/, *mínùte* /mənət/, *activity* /æktəvəti/, *límitèd* /ləmətəd/” (Lanham and Traill 1962:24). These words do, however, have a [ɜ]{[ɘ]}-value in more formal registers.

The phonemic status of SARP ‘B’ (and non–SARP SAE) /ə/ seems to be based on the fact that although certain words may fluctuate (depending on style) between [ɜ]{[ɘ]} and [ə], “there are certain words which only have [ɜ]{[ɘ]}” (Lanham and Traill 1962:24; footnote) – a situation very similar to /ï/ in SARP ‘A’. There is also a suggestion that the existence of contrast between [ə] and other values in weakly-stressed syllables is an added argument for the existence of a /ə/ phoneme. Thus the authors’ state that “in a substitution frame such as: *The drawer – the table* ... a vowel in the region of [ɜ][33]{[ə̟]}, or possibly a slightly higher position, is differentiated from [ə]. Thus [ɜn]{[ə̟n]} is usually heard in this frame as *in*, and [ən] as the reduced form of *and*" (Lanham and Traill 1962:25).

In non-SARP SAE the authors claim that [ə] occurs in almost all unconditioned cases (Lanham and Traill 1962:24) i.e. non–SARP SAE [ə] is pretty much equivalent to SARP 'A' [ə̟˕] in terms of its distribution through the lexicon.

In general, though, the argument for a separate /ə/ phoneme for SARP 'B' and non-SARP SAE seems tenuous at best. This is particularly true in the light of arguments provided by Lass (1986) against the use of schwa in the transcription of SAE KIT.

In Lanham (1964?:26), the greater predictability of the typically SAEP with respect to the KIT vowel is emphasised again. The author begins by emphasizing an important similarity that SAEP shares with many other dialects of English:

> "…in most English dialects the vowel written <u>i</u> is pronounced in two different ways according to the position in which it occurs".

The conditioning environments are for, all intents and purposes, identical to the ones already provided above (i.e. velar context, word-initial position etc.). Lanham (1964?) is, of course, referring to the shading alluded to in §7.2 above, a feature which, on this analysis, SAE appears to share with most other varieties of English. However, in SAE the unconditioned allophone has a different *quality* to that found in other accents, "which is the result of borrowing from Afrikaans …S.A. English tends to a vowel with a much lower tongue position like that of Afrikaans <u>dik</u>, <u>wit</u>" (Lanham 1964?:25) (trans.: *thick* and *white*).

In Lanham (1967), the source of the so-called KIN-PIN Split (as a *phonemic* split) is identified as RP. Lanham (1967:31–2) offers the following description of RP:

> "[ɪ] and [ï]{[ɨ̞]} complement each other only partially, however. There are numerous words in RP where [ɪ] occurs apart from the environments listed above, e.g. <u>wind</u> /wind/ [wɪnd] …<u>antedate</u> [æntɪdɜit]. RP speakers doubtless are not very sensitive to the difference and there may well be no minimal-pair contrasts, complementation is, nevertheless, partial and we therefore accept a marginal contrast between /i/ [ɪ] and /ɜ/ [ï]{[ɨ̞]} and the fact that there is often alternation between the two in certain words".

The putative marginal phonemic split in SARP 'A', as advanced in Lanham and Traill (1962), and the devolution of this split into simple allophony in SAE proper, makes perfect sense from this perspective and it would be natural to conclude that SAEP has been characterized, in some sense, by a gradual dephonemicization and centralization of the KIT vowel. This perspective is all the more surprising, however, given that SAE is usually identified as the *source* of the so-called KIN-PIN Split. So, for example, Wells (1982:612) claims that "historical /ɪ/, the KIT vowel, has undergone a phonemic split in South African English". From the previously mentioned perspective, therefore, the identification of the KIN-PIN split with SAE is incorrect on two levels:

1. The general tendency in SAE is for there to be no split as such, but for any variation to resolve itself into full complementation. The further away from RP the idiolect in question the less the contrast (or split) holds.
2. The origin of the split is not SAE, but RP. Lanham (1967:114) explains one possible source of this last point of confusion in the following manner:

   > "Standard descriptions of RP often ignore the fact that i in sin is represented by a central vocoid, and i in sick, hit, it by a front vocoid under identical conditions of style and stress. The probable reason is that strong emphasis on word or phrase, or a highly formal style in certain contexts, does indeed produce [ɪ] instead of [ɪ̈] in words such as sin, pit, will".

With respect to the analysis of KIT in SAEP, Lanham (1967) differs little from that of Lanham and Traill (1962). There are, however, three points that merit attention:

1. Lanham (1967:66) provides the following analysis of the effect of /l/ on KIT:

   > "/i/ [ɪ]^{21}{[e̞]} becomes [ɪ̈]^{23}{[ɪ̈]} in a word like kill in forms of SAE where the influence of /l/ is strongest. To a certain extent /l/ influence is regional … and not all forms of … SAE have an /l/-influenced allophone of /i/. The influence of /k/ … may be stronger than that of /l/ …
   >
   > The /l/-influenced allophone of /ɜ ∼ ə/ is so distinctive as to receive popular recognition among South Africans as a hallmark of SAE. The influenced allophone is unrounded [o]^{36}{[o̟]} (or slightly lower) and is so like /u/ [ʊ] in words like pill, build, milk that children are admonished for saying pull for pill, mulk for milk".

   This is different from the analysis provided by Lanham and Traill (1962) in that it alludes to the stronger influence of an initial velar such as /k/ on the following /i/ than the final /l/. In this regard it should be stressed that Lanham (1967) distinguishes between the effect of a final tautosyllabic /l/ on /i/ and on /ə/. Thus for Natal and other English speaking cities, Lanham (1967:80–1) gives /i/ as [ɪ̈]^{24}{[ɨ̞]}, while for /ɜ/ in conservative SAE the author gives [ɜ]^{35}{[ɘ̠]} with unrounded [ɵ]^{46}{[ɤ̞̈]} as the realization of /ə/ in the non-conservative variety[4].

2. Lanham (1967:72) argues that in the more extreme forms of SAEP a following palatal consonant looses its conditioning influence, except in the case of a final [ʃ]. Thus words such as *bridge* would have a schwa-like quality[5].

3. Lanham (1967:73) places much emphasis on the use of schwa in stressed syllables in SAE:

> "Collectively the ... differences from RP ... operate to increase the frequency of occurrence of low schwa to a degree probably unsurpassed by any other dialect of English. In RP low schwa as a simple nucleus is confined to a low percentage of weakly stressed syllables; in extreme SAE it predominates in these syllables and also occurs commonly with stronger stresses. The high frequency of low [ə] at all levels of stress clearly stamps SAE – it is reported from Australia that the pronunciation of <u>pin</u> as [pən] is popularly recognized as a means of identifying South African immigrants".

Lanham (1978), Lanham and Macdonald (1979): Lanham (1978:152) generally confirms the analysis provided up till this point. He confirms the partial complementation of fronter allophones with more centralized allophones of KIT in RP and that, under the influence of Afrikaans, SAE has a trend towards a stricter level of complementation and the greater seperation of the allophones of KIT in phonetic space. In its most advanced (BrSAE) form the conditioned fronted allophones approach C[1] and the centralized unmarked values are labelled "low schwa" (Lanham 1978:152). According to this author the prevalence of low schwa, not only with respect to KIT but in relation to unstressed syllables as well, "is above the level of social consciousness, at least in the English cities, and is subject to correction, even hypercorrection" (Lanham 1978:152). Lanham (1978:152) also mentions the effect of final tautosyllabic /l/ on the central allophones of the KIT vowel and gives [˂ɤ ∼ ˂ˇɤ]{[ɤ̈ ∼ ɤ̞̈]} as the values of KIT before /l/. A further contribution of Lanham (1978:158) is his identification of certain typically Natalian centralized values for KIT in contexts where ordinarily

[4] With respect to the variant of /i/ before /l/, however, it is difficult to reconcile the value Lanham (1967) gives on page 66. (i.e. [ɪ̈]^{23}{[ɪ̈]}) with that provided on page 80 (i.e. [ɪ̈]^{24}{[ɨ̞]}), although the values are, admittedly, not far apart.

[5] See also Wells (1982:612) in this regard.

one would expect a conditioned, fronted value: "before syllable margin with palatal C ('which, fish, ridge'); after h ('Hilton'); word initial after juncture. These features have associations with the North of England". It is not clear, though, whether Lanham (1978) views this as a true regionalism or simply as a Natal variable i.e. in the case of the latter as a feature of GenSAE which has its origin in Natal but has subsequently diffused throughout much of the SAE speech community.

The issue is resolved in Lanham and Macdonald (1979), however, who overtly confirm the existence of regional differences with respect to the KIT vowel:

> Natal English at the present time has other variables with clear local correlations and strong suggestions of an origin in the northern counties of England or in Scotland; for example, high schwa as allophone of /i/ in contexts h_ , #_ , _ C palatal, in *Hilton*, *industry*, *fish*. Except for the latter, this variable is uncommon outside Natal".

By "latter" Lanham and Macdonald (1979:45) presumably mean KIT in pre-palatal context as opposed to just the token *fish*: this interpretation is borne out in a related footnote in which the authors discuss "schwa-quality vocoids realizing /i/{/ï/} in the context _ C palatal". This interpretation is strengthened by noting Taylor's (1991) later comments regards the pre-palatal context: "more 'respectable' varieties of SAfE tend to have a non-centralized vowel before the palato-alveolar fricatives and affricates ... while broader accents typically have a range of vowels of varying degrees of centrality" i.e. pre-palatal [ï ∼ ə] occurs outside the Natal region. In general, if the above-mentioned analysis of NE is correct then NE should show evidence of [ɪ] only in a velar context.

These authors further confirm the "stricter complementation" (Lanham and Macdonald 1979:46) of KIT allophones in SAE and confirm that correction (or even hypercorrection) to an [ɪ]-quality does take place. This feature of SAE "originates in Afrikaans which phonemically opposes /i/ : /ə/ in stressed syllables" (Lanham and Macdonald 1979:46). The prevalence of "low schwa" in SAE is identified as a GenSAE variable, a theme taken up again in Lanham (1982:339) who, furthermore, confirms "a polarization of the allophones of /ɪ/":

> "Under the influence of Afrikaans, South African English has expanded the phonetic space between /ɪ/ and /ə/. The result is that some allophones approach cardinal vowel [i] (e.g., [kis] *kiss* and [wiɡ] *wig*) while others are realized as low schwa (e.g., [pən] *pin*)".

Jeffery (1982): It is no doubt comments such as these that led Jeffery (1982:258), as already touched on in §3.4.2, to collapse the SAE KIT-system into a polarization of [i] and [ə][6]. His overall conclusion, as discussed in §3.4.2, is that the centralization and polarization of the KIT vowel is the final product of a Pattern 2, Northern Cities shift-like push-chain (see §2.3.2.1), beginning with the fronting and glide-weakening of PRICE and MOUTH, the effect of which is to initiate a raising and tensing of the front vowels, including TRAP, SQUARE, DRESS and KIT. KIT tokens not in a velar context were raised first to [i] and then fell down along the non-peripheral track to [ə]. In the meantime tokens with a velar context followed up to [i].

Wells (1982): Wells (1982:612-3) dedicates a full section to the so-called "KIT Split" in SAE, and appears to be the originator of this unfortunate term. His description of the phonemic status of this lexical set is as follows:

> "Historical /ɪ/ . . . has undergone a phonemic split in South African English. The difference between the two resultant sound types is clearly heard if a word such as *sing* [sɪŋ ∼ siŋ] is compared with one such as *limb* [lɨ̈m ∼ ləm]. Description of this phenomenon is complicated by the fact that both the vowel qualities used and the conditioning environments are variable. It is possible to adopt various views on the phonemicization of the sound types in question: Lanham at one time (1967) argued for the recognition of three distinct phonemes, though he now (e.g. 1978b) considers that only one phoneme is involved. We shall here tentatively adopt the view that there are two" (Wells 1982:612).

By way of orientation, Wells (1982:612) explains that "Lanham refers to what I write [ɨ̈], [ə] as 'high schwa' and 'low schwa' respectively". While this analysis does justice to Lanham's penchant for creating phonemic contrasts, it appears deficient in two main respects:

1. It is incorrect in claiming that Lanham (1967) posits three phonemes, insofar as three phonemes is solely posited for RP and the near-RP SAE variety, CulSAE.
2. It does little justice to the *trend* in SAE towards a greater degree of complementation *vis-à-vis* RP. It, in short, and as discussed above, creates the impression that the relevant split originated in SAE, whereas in actual fact (or, at least, in

[6]This author's summary of the allophony of KIT is incorrect, since he provides [i] in the context of a velar, [ə] elsewhere. As noted above, preceding-[h] and word-initial contexts are at least two other environments that condition a fronted allophone.

terms of Lanham's analysis of RP and SAE) the trend is *away* from the marginal contrast in question, with the origin of the split being, in fact, RP.

That Wells (1982) recognizes the problematic status of the split in SAE is further confirmed when, in reference to BrSAE KIT allophones ([i] and [ə]), he remarks that "in an accent of this type a pair such as *kit* ... and *bit* ... clearly do not rhyme, *and it is this which makes one reluctant to regard their vowels as still belonging to the same phoneme*" (Wells 1982:612–3; my emphasis). Implicit in this statement is the recognition that, in fact, KIT in SAE generally paints a picture of complementary distribution. Thus, after providing the contexts in which the fronter allophone [ɪ] occurs, he confirms that "a centralized or central vowel, [ɪ̈], is used in the complementary set of environments" (Wells 1982:612).

As far as Wells' (1982) analysis of the *phonetic* value of the various KIT allophones is concerned, it agrees substantially with what has been provided above. Particular mention is made of the BrSAE [i] and [ə] values which replace [ɪ] and [ɪ̈] respectively. Wells (1982:617) gives the value of KIT before final /l/ as being [ɤ̈]. However, Wells (1982:612) adds to the confusion even further by making the following claim:

> "In conservative accents [ɪ̈] may contrast with [ə] in weak syllables, as *haggis* [-gɪ̈s] vs. *bogus* [-əs]. But most South Africans have no such contrast. For them, too, pairs such as *illusion* and *allusion*, *except* and *accept* are homophonous; and so are *scented–centred* [ˈsentəd], *sources–saucers*, *Lenin-Lennon* etc. This means that it is generally possible to regard [ɪ̈] and [ə] as allophones of the same phoneme, /ə/, in South African English. This approach is supported by the fact that in broad accents [ɪ̈] tends to be replaced by an opener [ə] quality even in stressed syllables, as [ˈdənə] *dinner*, [ˈləmətəd] *limited*" (Wells 1982:612; my endnote).

This appears to be Wells' (1982) own analysis and differs from, for example that of Lanham and Traill (1962) who, as discussed above, make a distinction between /i/ and /ə/ for SARP 'B' and non-SARP SAE, with /i/ including both [ɪ] and [ɪ̈] qualities (but not [ə] qualities which belong to /ə/). Secondly, it is unclear why a neutralization of contrast in weak-syllables coupled with the replacement of one phonetic value with another in BrSAE should lead inexorably to the conclusion of a phonemic split between /ɪ/ and /ə/ as advanced by Wells (1982). Thus, Wells (1982:613) continues: "it seems more satisfactory to regard [ɪ ∼ i] as /ɪ/ and [ɪ̈ ∼ ə] as /ə/, the latter having become one of the stressable vowels in South African

English”. The arguments for the existence of a separate /ə/ in SAE seems, however, to me, as tenuous as those proposed by Lanham and Traill (1962).

Bailey (1984): Bailey (1984:13–4) begins by positing a regional difference between SAE more generally and Cape Town English with specific reference to the KIT vowel: “it is reasonably well-known in Transvaal and Natal that SAE speakers in Cape Town pronounce ‘pin, window, winter, finish’ with a short high vowel [ɪ] in stressed position while more typical SAE has some type of central vowel [ə]” (Bailey 1984:13). Bailey (1984:14) attempts to trace this feature to Afrikaans-English bilingualism as well as hypercorrective attempts to emulate RP-norms. It is interesting to note (given that the author appears to come from Natal) that *finish* and, later on, *bitch*, *ditch*, *witch*, *bridge*, *religion* and *prestigious* are implied or given as having [ə]. *Finn*, *Finland* and *Finnish* are all given as having [ə].

The author, focussing on his own (Natalian) pronunciation, gives only two allophones (i.e. no KIT before final /l/): [ɪ] and [ə], with the later being the unmarked one. This author’s characterisation of the allophonic conditioning of [ɪ], is, firstly, that it occurs when followed by a “hushing sibilant” i.e. /ʃ/ or /ʒ/. A following /s/, /z/ and, more importantly, /ʧ/ or /ʤ/, do not (by themselves) condition the fronted allophone. This is in line with what is a (presumably) typical Natalian accent. Furthermore, the *onset* conditions that lead to a fronted [ɪ] are characterised broadly as velar or glottal, recognising in the process the use of the glottal stop in word-initial position, as in *in* [ʔɪn]. There are two extra points worth mentioning here. Firstly, the author, given his Natalian accent, surprisingly gives *dish* as having a fronted allophone. A central [ə] in *dish* is a well-known shibboleth of Natalian English, at least in GenSAE. Second, many of the examples provided seem to miss the point that a following segment can be either tautosyllabic or non-tautosyllabic, and that this could make a substantial difference to the conditioning. In general this distinction seems to be missed in the literature. In Bailey (1984:17), for example, an example of /ʒ/ conditioning the fronted allophone of KIT, is given as *vision*. While I have no reason to doubt that the fronted allophone is used (at least in the author’s idiolect) in this word, this fact might be equally attributable to the fact that KIT is syllable-final.

Bailey (1984:20) also provides an interesting analysis of KIT in words like *kill* and *gill*, which he claims begin with the phonological sequence /ky/, although it is unclear whether he means for English or SAE in particular. The author first distinguishes between the (variable) use of a true palatal stop, i.e. [c] and [ɟ], in words such as *kill* and *gill* in SAE, as opposed to the use of a fronted velar, [k̟] and [g̟], in words such as *kit* and *give*. The following comments are of interest with respect to *kill* and *gill*:

> "it seems that historically the vowel underwent breaking as a result of the strain in the tug-of-war between the preceding and succeeding consonants. On the one hand the initial velars select a high front allophone for this short stressed vowel. On the other the velarized /l/ causes retraction of the body of the tongue producing a central vowel. The first part of this hypothetical diphthong then loses its syllabicity becoming a glide /y/ which merges with the preceding velar" (Bailey 1984:20).

The author then gives [ˈcʰəɫ] for *kill* but [k̟ʰɪlə] for *killer* as a result of the fact that the /l/ in the latter word belongs to the onset of the second syllable and is not subject to velarization. I should add, however, that the proposed pronunciation of *kill* seems highly improbable to me (although it might be a Natal-ism); I would pronounce the word as something like [cʰɪᵊɫ] i.e. retaining the diphthongal quality of the vowel, but targeting the front vowel allophone of KIT. While the pronunciation of *killer* seems correct it does alternate with [cʰɪᵊlə]] (in my idiolect, for example). A similar process of palatalization has occurred in what Bailey (1984:21) refers to as /hy/ sequences, as in *huge* and *hue*. A word such as *hill* is thus pronounced in my idiolect and I suspect in many South African's speech as [çʰɪᵊɫ], with slight aspiration if any.

Lass and Wright (1985; 1986): As was briefly mentioned in §7.2, both Cruttenden (2001:107) and Lass and Wright (1985:138) assign a [ë] value for General RP KIT, more commonly transcribed, of course, as [ɪ]. As mentioned as well in §6.4, the latter authors also give [ë] as one characteristic value of DRESS in SAE, the implication being that there is a substantial overlap between RP KIT and SAE DRESS. The reflex of this state-of-affairs is that, according to Lass and Wright (1985), KIT in SAE is an even more centralized [ï] (with respect to the *fronter* values) while a [ɘ̜]{[ɘ̝]}value is commonly used for the even more centralized variants. Lass and Wright (1985:138) define [ï] as follows:

> "a further centralized [ë], but not yet fully centralized ... the auditory quality we intend is pretty much that of Afrikaans *i* in *sit, is* or the Trager-Smith [ɨ] in some U.S. *just*, *twenty*, *finish* . . . or the more centralized varieties of RP /ɪ/ before syllable-final /l/ (*milk*, *children*)".

On the other hand, [ɘ] is defined as "a fully central half-close nucleus . . . not [ə] 'schwa"' (Lass and Wright 1985:138). As such, these authors essentially challenge

Lanham's and Jeffery's (1982) conceptualization of the polarization of KIT allophones in SAE, since the [ɪ̈]-allophone is viewed as the most front value, with the more centralized [ɘ]-value being used in other contexts. Thus, in relation to one of their two subjects, a middle-class, female, General (Respectable) SAE speaker, the authors' claim the following:

> "[she] has two main variants of the *bit* vowel. The fronter is [ɪ̈]; the other is the lower, fully central [ɘ̞]{[ɘ̞]} ... At first glance, [ɪ̈] seems to occur after the velars /k g/, and /h/; [ɘ̞]{[ɘ̞]} in all other environments ... she shows no sign whatever of 'polarization': the front allophone [i], so basic to Lanham & Macdonald's scenario, does not occur at all even in high tongue-body environments. Instead she has the centralized [ɪ̈] there, after /h/, and initially. Thus [ɪ̈] is in fact the frontest allophone" (Lass and Wright 1985:150; my parenthesis).

It should be noted that the study involved the elicitation of citation-form KIT tokens (204 per subject, to be precise) (Lass and Wright 1985:150). Other, more minor, features of this speaker's KIT vowel include the following:

- Surprisingly, a retracted value preceding velars (as opposed to following them – see above), the values being [ɘ̞]{[ɘ̞]} to [ɤ̽̈]; thus [bɤ̈̽g] for *big* and [pɤ̈̽k] for *pick*. The authors do admit, however, that this might be an idiolectal idiosyncrasy (Lass and Wright 1985:151). The same set of values is given for KIT following /w/, /f,v/ and, inconsistently following /r,l/ (Lass and Wright 1985:150–1).
- A slightly fronter [ɪ]-value in a highly restricted set of contexts, in particular proper names (*Tim*, *Lynn* etc.) and "in disyllabic words in [-ŋgl], e.g. *single*, *tingle*, *mingle*. The historically older [ɪ] remains only in one restricted lexical class, and one very specific phonetic environment" (Lass and Wright 1985:151).
- A retracted KIT before final /l/, which Lass and Wright (1985:151) give as [ɤ̈]. "If a velar or /h/ precedes and [ɫ] follows, the vowel has a slight onglide from [ɪ̈] through to [ɤ̈], as in *skill* [skᶦɤ̈ɫ], also *hill, kill*" (Lass and Wright 1985:151).
- The existence of at least one minimal pair: [hɘ̞m] *him* vs. [hɪ̈m] *hymn*.

Instead of polarization, we have thus fine changes in phonetic quality from [ɪ̈] 'backwards', mainly based on phonetic environment, but with a few lexical exceptions; "the overall picture ... is one of the general entrenchment of a 'basic' quality no fronter than [ɪ̈]" (Lass and Wright 1985:151), with a few exceptional cases of [ɪ] as mentioned above.

Turning to Lass and Wright's (1985) other subject, a male BrSAE speaker, the authors also claim [ɪ̈] and [ə̝] as the basic qualities, but when "adjacent to the velars ... following /h/, and word-initially, a closer and fronter [ɪ̜̄]{[i̠˔]}" ... Thus *thing* [θɪ̜̄ŋ]{[θi̠˔ŋ]}, *Kitsch* [kʰɪ̜̄ʧ]{[kʰi̠˔ʧ]} . . . , *hit* [hɪ̜̄t]{[hi̠˔t]} . . . , *ink* [ɪ̜̄ŋk]{[i̠˔ŋk]}" (Lass and Wright 1985:152). Further features of this idiolect's KIT vowel include the following:

- No consistency after /w/, i.e. the value is not categorically [ə̜]{[ə̝]} to [ɤ̟̈] as in the case of the GenSAE speaker reported on above. So, for example, a following velar (as in *wing*) conditions the fronted [i]-value: [wɪ̜̄ŋ]{wi̠˔ŋ} (Lass and Wright 1985:152).
- Similarly, when KIT is followed by a final /l/, but preceded by a velar or /h/ there is an offglide beginning with a [ɪ̜̄]{[i̠˔]}-quality and ending at [ɤ̈]; thus *skill* [skɪ̜̄$^{\text{ɤ̈}}$ɫ]{[ski̠˔$^{\text{ɤ̈}}$ɫ]} (Lass and Wright 1985:152).

While the existence of the [ɪ̜̄]{[i̠˔]}-quality is, of course, *prima facie* evidence of polarization for at least BrSAE, the authors claim that from a historical perspective the [i̠˔]-value is recessive. They base this conclusion on the fact that in the speech of their second subject there are a number of exceptions to the assimilatory force underlying the existence of [i̠˔]: "the systematically marginal status of [i] ... is ... suggested by the fact that it often fails to occur in places that ought to predict it ... [the speaker] shows 'basic' [ɪ̈] in newer or more colloquial items e.g. *Bic* 'pen, razor' [bɪ̈k], *flick* 'movie"' (Lass and Wright 1985:153; my parenthesis). In essence, the [i]-quality is viewed as situated outside the main force of a chain shift, the leading edge of which is the centralization of the KIT vowel. The centralization of the KIT vowel, has, in its turn, been the result of pressure from the other front vowels i.e. DRESS and TRAP. This is summarized diagrammatically in Figure 7.1 on the following page, which is a reconstruction of the figure in Lass and Wright (1985:156).

According to Lass and Wright (1985:156), "the boxed area represents the range of categorical inputs and ouputs", while the values outside of the box are either "continuations of the original shift" (in the case of [ɘ]), or "a supervening development" (in the case of [i]). The underlined values (i.e. [æ] and [ɪ]) do not, bar a few marginal exceptions, occur in either GenSAE or BrSAE.

The authors summarize their position in the following words:

> "The high front reflexes in [the BrSAE speaker's] dialect can be seen as representing a supervening and largely irrelevant development ... While the /ɪ/ >

[ɪ̈]/[ɘ] change was contextually determined in its earlier stages, its results are now categorial, with traces of lexical and phonetic conditioning for [both speakers]. But the shift to [i] is in the main phonetically conditioned – with a few lexical interferences. And it shows no strong evidence of productivity: it is … more or less a relic" (Lass and Wright 1985:155–6; my parenthesis).

In short, if this analysis is the correct one, then one would predict, for GenSAE speakers, minute phonetic variation ranging from a roughly [ɪ̈] position to a [ɘ]-value, with perhaps the exception of KIT before final /l/ which has an even further-retracted value i.e. [ɤ̈]. Lastly, while Lass and Wright (1985) provide no overt position on the so-called KIN-PIN Split, it is clear that they recognize the existence of marginal minimal pairs, such as the *him*:*hymn* example already mentioned above. They, however, explain such phenomena in the following terms:

> "If we visualize the typical trajectory of a sound change in time as a sinusoid … then pairs like … *him*; *hymn* represent the tail-off of the exponential or 'log' phase of the curve … Such residues are the garbage left behind – at least in the early stages – by the asymptotic nature of change-curves. Curves that complete – i.e. 'Neogrammarian' changes – normally require much more time than has been available for our shift. 'Exceptionlessness' develops by patch-up operations on the irregular detritus left by the tail-off of the curve … Assuming a mid-to-late 19th century date for the large-scale centralization of ME /i/, the elapsed time is

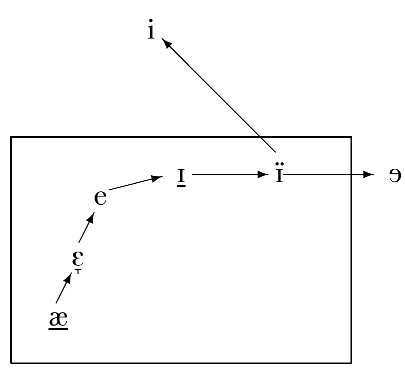


Figure 7.1: Lass and Wright's (1985) chain-shift

...rather small" (Lass and Wright 1985:155).[7]

From this perspective, therefore, there is no real KIN-PIN Split. Any evidence of minimal pairs is simply the result of a sound change not having reached full completion. Change with regard to the KIT vowel in SAE should, thus, involve the gradual 'mopping up' of lexical exceptions to general allophonic rules. Thus while Lass and Wright (1985) reject the notion of polarization, they embrace the notion of a stricter degree of complementation as a trend in SAE. "The presence of marginal minimal pairs assignable to failure of an allophonic rule" (Lass and Wright 1985:155) should thus gradually disappear from SAE. Secondly, and as is implied in Figure 7.1, it appears that SAE KIT has not yet 'settled' and that further centralization and lowering is likely, even with respect to environments which currently condition the fronter allophones (Lass and Wright 1985:154).

Although not a focus of the study, Lass and Wright (1985:156) tentatively claim, for CulSAE, "an allophonic range for /ɪ̈/ encompassing both unshifted [ɪ] (especially after /h/) and centralized [ɪ̈] in many (most?) other environments".

Lass and Wright (1986:207–11) essentially confirm the analysis provided by Lass and Wright (1985). Both Lass and Wright (1985:210-11; 12; my parenthesis) argue, thus, for an essentially endogenous (non-Afrikaans influenced) etiology for SAE KIT.

Lass (1990): Lass (1990:273) confirms that CulSAE is characterised by "a levelling out of the marked allophonic distinctions in KIT (e.g. rhyming of *it* and *sit* ...)" found in the more local varieties of SAE. With respect to GenSAE, Lass' (1990) description is unfortunate in a sense, given that he, by his own admission, provides an oversimplication of the qualities of this vowel i.e. in comparison to Lass and Wright (1985). He posits three main allophones for his particular variety of GenSAE: [ɪ], [ɪ̈ ∼ ə] and [ɤ̈]:

1. [ɪ] intially, after /h/, in the environment of velars, "usually before palato-alveolars ... but rarely after them" (Lass 1990:275). According to this author the [i]-like qualities only occur in the less-standard varieties.
2. [ɪ̈] in other environments, with the further retracted [ɘ] in the "vicinity of labials (*limb*, *miss*), and after /r l/" (Lass 1990:275). Lass (1990:275) identifies this as being "virtually identical to the famous American [ɨ]".

[7]See Lass (1981) for a (highly entertaining) exposition of other English examples of this phenomenon. Lass (1984:34–6) provides a more gentle introduction to the notion of the 'failure' of allophonic rules.

3. [ɤ̈]: before final /l/ and after /w/ as in *fill* and *wish*. The author mentions the possible neutralisation of KIT and FOOT in this context, but claims that in most (standard) varieties the distinction is maintained through a difference in lip-rounding i.e. weak lip-rounding in FOOT vs. no lip-rounding in KIT before final /l/. In reading-style this allophone is often hypercorrected to [ɪ] "followed by a [ɤ̈]-like off-glide to /l/" (Lass 1990:276).

This description appears, however, to be more than an oversimplification of Lass and Wright (1985). It is incompatible, in certain subtle ways, with this earlier analysis. The following points need to be made:

- In Lass and Wright (1985), [ɪ̈] is given as the frontest allophone with [ə̝] being reserved for the more centralized values. In Lass (1990) we have [ɪ] for the former and [ɪ̈ ∼ ə] for the latter.
- In both Lass and Wright (1985) and Lass (1990), [ɨ] is identified as being equivalent to [ɪ̈].

It thus appears that Lass (1990) and Lass and Wright (1985) provide two seperate analyses. The former is not just an oversimplification of the latter. While Lass (1990:275) does hedge somewhat by stating that "[ɪ] is as front as KIT gets in the posher varieties", the (overt) discrepency between these two analyses remains.

Taylor (1991): Taylor (1991:75), which has as its sole primary focus the KIT vowel in SAE, has as subsidiary foci "(a) the distribution and phonological status of the different vowel qualities, and (b) the origin of the present-day SAfE situation". Although the article refers the reader to Lass and Wright (1985) for "a detailed account of the range of phonetic variants in the speech of two individuals" (Taylor 1991:84; footnote 2), its author's description of the distribution of the allophones of KIT is essentially 'Lanham-esque' i.e. a fronted allophone in *kit*, *fig*, *sing*, *him* and *id*, "a vowel which resembles the RP equivalent, or which may even be slightly more front and more close than the RP equivalent" (Taylor 1991:75); and a centralized value in *pin*, *fit*, *sin*, *dim*, *lid* "which is centralized with respect to the RP pronunciation, with values ranging from around [ɪ̈] for speakers of 'respectable' SAfE, to values approaching schwa in more extreme varieties" (Taylor 1991:75). Thus, unlike in Lass and Wright (1985), where the most front value for KIT *overall* is [ɪ̈], in Taylor (1991) the frontest value for the *centralized* variant is [ɪ̈]. Furthermore, polarization of the variants is implied, which is, again, contrary to the position taken in Lass and Wright (1985).

Taylor (1991:75) refers to the two relevant qualities, *à la* Wells (1982), as KIN and PIN.

As far as the conditioning environments regulating KIN vs. PIN are concerned, there is little to surprise. In particular, a velar context, a preceding /h/ and a preceding word-boundary condition the fronted variant[8]. In general, velar and laryngeal environments tend to be exceptionless in their conditioning of KIN. Palato-alveolar and labial environments have a more variable influence, however. The former seem to condition KIN in Respectable variants but PIN (or at least approaching PIN) in less Respectable (i.e. Broader) variants. Labial environments, according to Taylor (1991:76) seem to create an even more erratic pattern of distribution:

> "About half the informants had a fronted stressed vowel in *finish*, while only two had the fronted vowel in *Finnish*. Other words which were spoken with a fronted vowel include *fin* (five informants), *Finn* (2), *wind* (10), *limb* (2) and *lip* (1)".

This lack of strict complementation is considered by Taylor (1991:76) to be "a dilemma for phonemic analysis". On the one hand positing two separate phonemes misses the obvious fact that in the vast majority of cases there is predictibility. On the other hand, positing KIN and PIN as two allophones of a single phoneme misses the fact that there are a number of lexical exceptions. Taylor (1991:77) also mentions the existence of at least a few minimal pairs in the idiolects of some speakers e.g. *finish* and *Finnish* or Lass and Wright's (1985) *him* and *hymn*. With respect to this marginal contrast, the author concludes that "the potential phonemic contrast between the KIN and PIN vowels in SAfE is of marginal status, and is highly unstable. Though perhaps real enough for an individual speaker, it has no functional value within the SAfE community as whole" (Taylor 1991:77). This conclusion is based on a subsidiary set of experiments[9] on the *finish–Finnish* minimal pair. Without entering into the details, it was basically established that the distribution of the KIN and PIN vowel qualities *vis-à-vis* these two words is random for the SAfE speech community, though on an idiolectal level certain individuals seem to maintain a contrast between these two words, mainly in terms of KIN for *finish* and PIN for *Finnish*. The irony of all this is, of course, that the earlier Lanham (1967) identified *RP* as the source of this lack

[8] Or, strictly-speaking, in the case of the last-mentioned environment, a preceding glottal stop i.e. *it* is, in SAE, generally [ʔit] and not [it]; Taylor (1991:76) concludes thus that "preceding /h/ and preceding word boundary may thus be subsumed under the characterization 'laryngeal environment"'.

[9] Taylor (1991) elicited data from 25 informants, most of whom were students at the University of the Witwatersrand.

of full complementation and stressed that SAE was characterized by an *increasing* degree of complementation. Taylor (1991) presents the *lack* of complementation as a feature of SAE.

With respect to Taylor's (1991) arguments relating to the origin of the KIT 'Split', it is interesting to note the connection this author makes between FLEECE, which he characterizes as "an extremely close and fronted monophthong" (Taylor 1991:79) and KIN, the implication being that the more (than RP) front and close KIN value in SAE has led to a raising of FLEECE; in contrast, for example, with Australian English where FLEECE has been subject to diphthongization: [ïi]. The author also reminds the reader that NZE is characterized by gross centralization of KIT, but points out that apart from the central values there is also a much closer allophone, "confined, it would seem, to velar environments, especially before /ŋ/" (Taylor 1991:97).

As far as Taylor's (1991) solution for the origin of KIT values in SAE is concerned, the claim is that "the centralization of /ɪ/, in certain environments, has its origin in mechanically conditioned perturbations in the acoustic quality of the /ɪ/-vowel". In basic terms, Taylor (1991) exploits the notion of consonantal acoustic locus, as originally developed by Delattre et al. (1955), to explain how a velar context allows easier coarticulation to a high, front vowel than a non-velar one. Using spectrographic data from RP, and focussing specifically on the second formant, he shows how in a non-velar context subjects 'undershoot' the vowel target. Similar arguments are provided for /h/-initial and /ʔ/-initial environments. In short, these differences that occur as a result of the processes of coarticulation and 'undershoot', have then, according to Taylor (1991:81–2; my parenthesis)

> " . . . been exaggerated [in SAE] to such an extent that the /ɪ/-vowel split into two perceptually quite distinct classes of sound . . . this process was undoubtedly facilitated by . . . the presumed presence of sporadically centralized instances of /ɪ/ amongst the early English-speaking settlers, internal pressure to restructure the vowel system, as well as, perhaps, the presence of a stressed centralized vowel in Afrikaans, especially in Afrikaans cognates of English words".

In the case of palato-alveolars, which according to Taylor (1991:81) have a variable influence on SAE KIT, the RP data "are less clear ... a following [ʃ] and [ʧ] cause no noticeable raising of F_2". Interestingly enough, Taylor (1991:81) also notes one case of RP *pin* having a high F_2 i.e. a fronted articulation. This accords with Lanham and Traill's (1962) and Lanham's (1967) earlier comments about the lack of full complementarity in RP – these authors actually mention *pin* as an example of a lexical item that doesn't fit the general RP KIT pattern.

Thus, it is interesting to note that the RP acoustic data that Taylor (1991) draws upon provides some evidence for the earlier position provided in Lanham and Traill (1962) and Lanham (1967) i.e. that RP shows the same general pattern as found in SAE e.g. fronting in the context of a velar, lowering in the context of a bilabial. From this perspective, it would appear that SAE 'took' an already-existing co-articulatory process and simply exaggerated it; or perhaps a better way of putting it would be to say that what was a form of accommodatory allophony in RP became non-accommodatory (i.e. part of the grammar) in SAE – a case of phonologization as described in §1.4, coupled with a gradual 'mopping up' of lexical exceptions.

Branford (1994), Lass (1995; 2004), Bowerman (2004): Branford's (1994) analysis of KIT is essentially Lanham-esque as well, in that he posits, for 'Respectable' speakers, a [ɪ]-value for velar contexts, while [ɪ̈] occurs in the complementary set of environments, retracting to [ə] in broader lects. It is odd, however, that Branford (1994:476), in providing a summary of the short-vowels of SAE, gives /ɪ/ɩ/ for the two keywords BIT and HIS. Firstly, the two symbols have usually been viewed as equivalent to each other, with the IPA having replaced [ɩ] with [ɪ] in 1989[10]. That Branford (1994) does not imply equivalence is clear from other examples, like TRAP, which he renders as /æ/ɛ/ i.e. as capturing the range of phonetic values from Cultivated to Broad SAE. This is also evident from a latter discussion of the happY vowel (Branford 1994:477), in which the author claims that a [ɩ]-value is common in the Southern Hemisphere Englishes and quotes Wells (1982) in this regard. The value Wells (1982:166) provides in the relevant quote is [i], which appears, then, to be what Branford (1994) means by [ɩ]. It is interesting in this regard to note that Wells (1982:xx) himself stipulates an equivalence between [ɪ] and [ɩ], which renders Branford's (1994) use of this symbol as all the more confusing[11]. Returning to Branford's (1994) use of /ɪ/ɩ/ for BIT, HIS, and accepting that this author is using [ɩ] as an equivalent for [i], the attribution of values seems to relate to his more general goal of contrasting "'Conservative' with 'Broad' or 'Extreme' renderings" (Branford 1994:473). Thus RP-like [ɪ] is the CulSAE value, while [ɩ] (i.e. [i]) reflects the Broad value for HIT. What's missing from

[10]See Pullum and Ladusaw (1996:89) and International Phonetic Association (1999:167–8).

[11]Another use of the [ɩ]-symbol in Branford (1994), is in relation to a peripheral discussion concerning the status of Afrikaans loanwords such as *dik* 'full' and *skrik* 'fright'. The author states that "DIK [dək] does not rhyme in SAEP with *Dick* [dɪk] in conservative speech, often [dɩk] elsewhere," (Branford 1994:475) implying (perhaps) the use of [i] for *Dick* in non-CulSAE accents, which is true enough for many BrSAE speakers, although untrue of GenSAE.

this picture is, of course, the Broad [ə] for BIT.

Lass (1995:96–7) generally paints a familiar picture. He emphasizes the role played by KIT in the general SAE front-vowel chain shift, emphasizing that the shift to the center as opposed to ‘up’, was based on the likelihood that “the original shift seems to be of the type that prohibited mergers of categories, and raised KIT would intersect the short allophones of FLEECE (e.g. *sit* would become a near-homophone of SEAT)” (Lass 1995:105; footnote 11). Later, in BrSAE in particular, fronter and closer variants of KIT do, in fact, appear and this sociolect is generally characterized by [i] in *it* vs. [ɪ̈] in *sit*. CulSAE[12] has [ɪ] for both KIN and PIN, while GenSAE has [ɪ] for the former and [ɪ̈] for the latter. Further vowel retraction before final /l/ and after /w/ is mentioned as typical of both GenSAE and BrSAE speakers, while in the later there is also further considerable retraction is the context of /f/ e.g. *fifty* and *fit*. This even further retracted version of KIT leads to near-homophony with FOOT (e.g. *will*/*wool*, *bill*/*bull*), although the KIT vowel in this context usually has slightly weaker lip-rounding. According to Lass (1995:97), as far as the “KIT split (as Wells calls it),” more generally is concerned, while the distribution is close to being fully predictable there are lexical exceptions. Lass (2004:375) adds nothing new to his earlier Lass (1995) ‘picture’.

Bowerman (2004:936) claims that alternation between [ɪ̈] and [ə] in the unmarked allophone (with [i] as the fronter allophone) is characteristic of BrSAE, while the sole unmarked variant of GenSAE is [ɪ̈], with [ɪ] being the fronter and marked allophone, the latter “occurring in velar and palatal environments”. While Bowerman (2004:936) refers to the KIT Split, he acknowledges that the “split is an allophonic variation”. According to the author the ‘split’ does not occur in CulSAE. Before a final /l/, “the vowel may be as far back as [ʊ̈]” (Bowerman 2004:936).

7.4 KIT: A Comparative Acoustic Analysis

With respect to Webb’s (1983) tokens in particular, KIT1 in Figure 5.4 on page 151 shows the mean value and standard deviation for tokens with a velar context or with KIT word-initially (*if*, *it*, *kick*, *think*, *kids*, *kid*, *give*, *is*), KIT2 provides values for the unconditioned allophone, *did*(x3), *sister*, *this*(x3), while KIT3 provides values for KIT before final /l/(*until*, *will*). The retraction of KIT3 is apparent and it is interesting to note that it is substan-

[12] As well as, according to Lass (1995:97), in “more formal styles for some Respectable speakers, especially female”.

Allophone	Environment(s)	N-G Tokens	O-G Tokens	No. of Tokens
KIT1	disyllabic	***ci**ty*	***ci**ty*, ***si**lly*	36
KIT2	ʔ_ ; h_ ; in context of velars	*sing*, *hid*, *kit*, *it*	*sing*, *hit*, *it*, *kit*	108
KIT3	before palato-alveolars	*bitch*, *dish*, *fish*	*fish*	63
KIT4	Unmarked	*chin*, *sit*	*sit*	45
KIT5	l_; ɹ_; in context of labials	*rid*, *bit*, *lit*	*limp*, *rim*, *bit*, *lip*	89
KIT6	w_; _ɫ	*till*, *with*	*fill*, *pill*	54

Table 7.1: KIT allophones

tially more retracted than FOOT. The relationship between KIT and DRESS is particularly interesting and draws comment from Webb (1983:145; my parenthesis):

> "The second striking feature of the chart . . . concerns the lowering of /ɪ/ (or raising of /e/) . . . The "margin of error" . . . seems to be very small. [The figure] shows that these two phonemes . . . overlap quite considerably . . . /ɪ/ and /e/ thus seem candidates for merging".

From Figure 5.4 it would appear that the front allophone of KIT finds itself in a position similar to, or perhaps even slightly more central than, the KIT of modern RP i.e. slightly centralized from cardinal 2 i.e. [ë] or slightly more centralised[13]. While there is some evidence of 'polarization' what strikes me about this data is the lack of high-front or dramatically lowered (BrSAE) values.

Turning to the current data, Figure 7.2 on the following page[14] provides a clear picture of the position of KIT vs. the other SAE short vowels. Note that, in this figure, *kit4* is 'hiding behind' *kit5*. Table 7.1 provides an outline of the various allophones of KIT as represented in Figure 7.2, along with the label used, a description of the relevant environments, the tokens analyzed for each allophone (both N-G and O-G) and the number of tokens analyzed for each variant. Overall the representation is based on 8690 measurements over 395 tokens over 27 subjects at an average of 14.63 tokens per subject. The distribution of allophones is largely based on Lass (1990).

[13]The exact analysis depends, of course, on whether we wish to accept that DRESS in SAE has become non-peripheral and lax or, alternatively, remains on the periphery – see §6.4 for more on this.

[14]The same diagram as Figure 6.13 on page 197, repeated here for the reader's convenience.

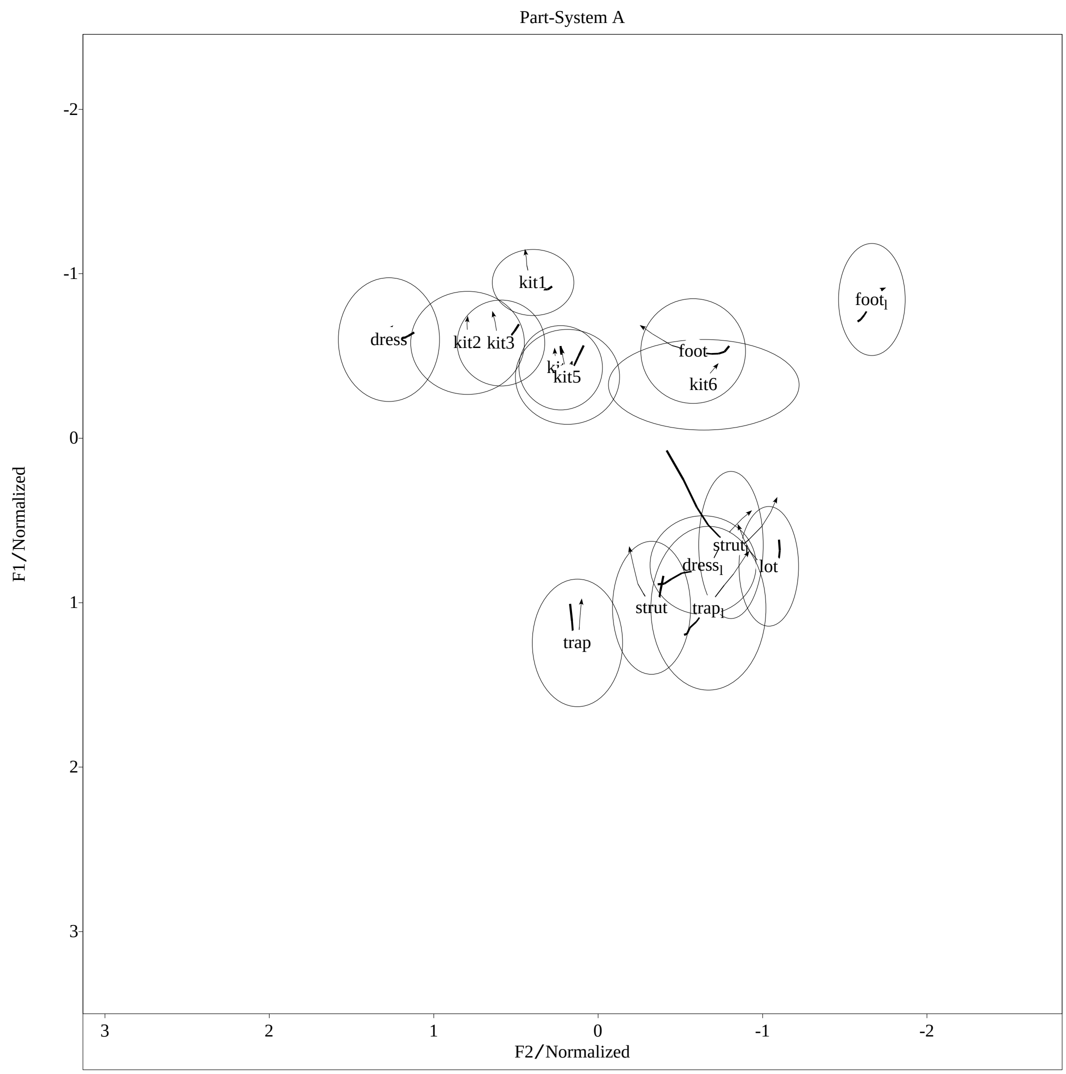


Figure 7.2: Overall Results: Part-System A (27 subjects)

Beginning with KIT1 (i.e. in *silly* and *city*) we note, in Figure 7.2 on the preceding page, its relatively high position *vis-à-vis* the other KIT allophones. It is not completely clear what to make of this, not least because KIT in this environment is not dealt with in the impressionistic literature; that there is a degree of vowel-target undershoot seems likely, however. In terms of the front-back dimension there is clear evidence of centralization.

The overlap between KIT2 and SAE DRESS has already been noted in §6.4, in the discussion of DRESS. There is also some evidence for a greater degree of centralization of KIT3 i.e. KIT before palato-alveolars. The same does not, however, apply to the putative distinction between KIT4 and KIT5. Note that in all these cases (i.e. KIT1 to KIT5) any evidence of spectral movement is easily attributable to co-articulatory effects. In the case of KIT6, however, the seeming lack of spectral movement in Figure 7.2 is, in all likelihood, due to the 'lumping' together of KIT after /w/ and KIT before /l/; the co-articulatory effects of these two environments are, after all, likely to cancel each other out i.e. gliding from high-back to a more front position in the case of the former, while gliding from a more central position to a position further back in the case of the latter.

Figure 7.3 on the following page provides details of each individual word used under the KIT6 label i.e. *till*, *pill*, *fill* and *with*[15]. As is clear from this figure, KIT after /w/ involves a glide from a FOOT-like position to a more fronted KIT4/5 position. KIT before tautosyllabic dark /l/, on the other hand, shows a distinct retraction. The effect of /l/ will be discussed more generally in §9.4.3, but it is worth noting here that this data provides evidence to suggest that in modern SAE there is little or no hypercorrection of KIT before /l/ in word-list context; the value of KIT in this context has, undoubtedly, become the product of a phonologization process and constitutes a broad indicator of SAE, a supposition strengthened by the position of KIT before /l/ in Webb's (1983) data – *kit 3* in Figure 5.4 on page 151.

Figure 7.4 on page 271 provides a comparison of the SAE KIT data collected for this research project, with that collected and analyzed by Webb (1983). In order to render the data easier to assimilate, the various labels have been shortened by replacing KIT1, KIT2 etc. with 1, 2 etc.; Webb's (1983) two KIT allophones have been labelled (in italics) as *W2* and *W4*. The labels *W<u>2</u>* and *W<u>4</u>* were chosen because all the tokens analyzed by Webb (1983) fall into the KIT2 or KIT4 categories.

The first thing that strikes one about this comparison, is the high degree of overlap between the two sets of data. It is true that Webb's (1983) data has a slightly higher degree of variation. Thus, for one, Webb's (1983) *W2* category overlaps substantially with DRESS; both DRESS in the current data as well as DRESS in Webb (1983) – see Figure 5.4 on

[15]With *fill* being obscured by *pill* in Figure 7.3.

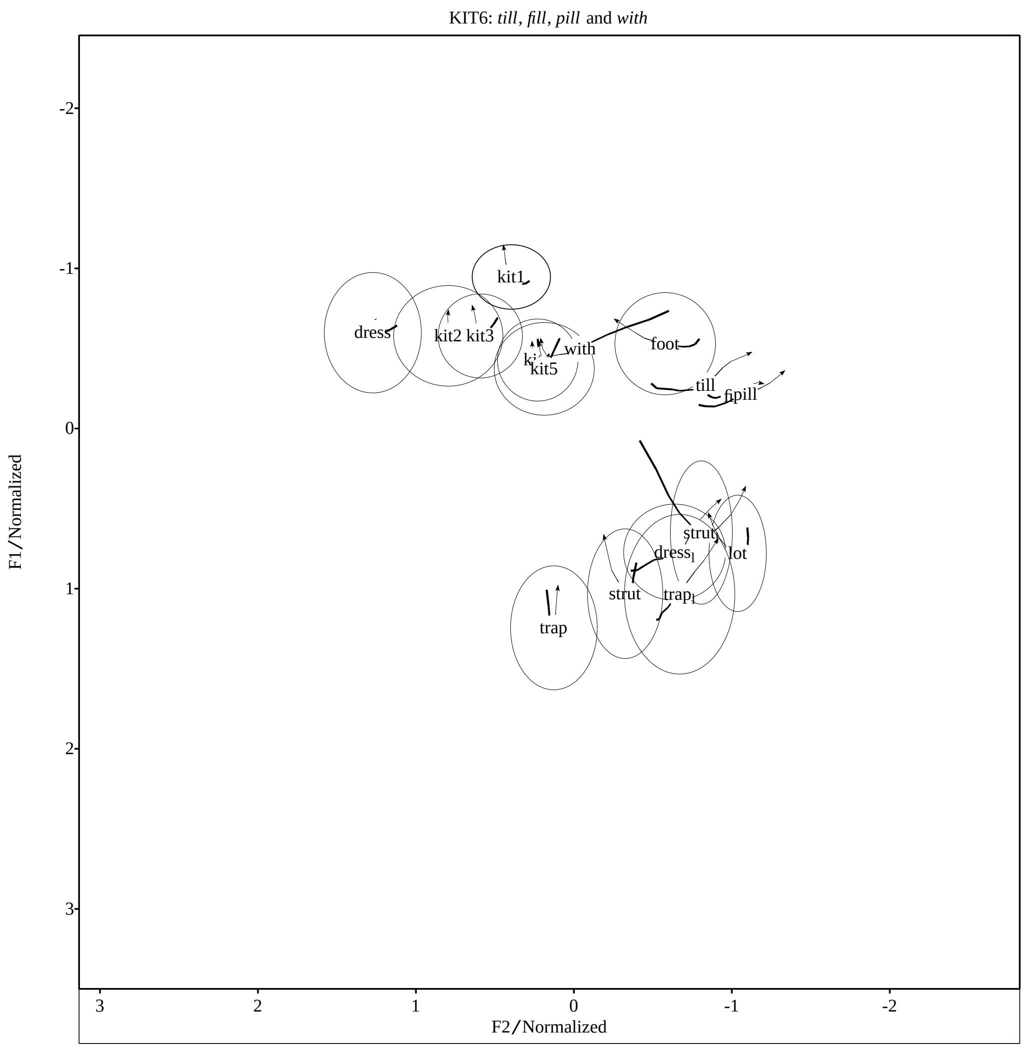


Figure 7.3: KIT: *till*, *fill*, *pill* and *with*

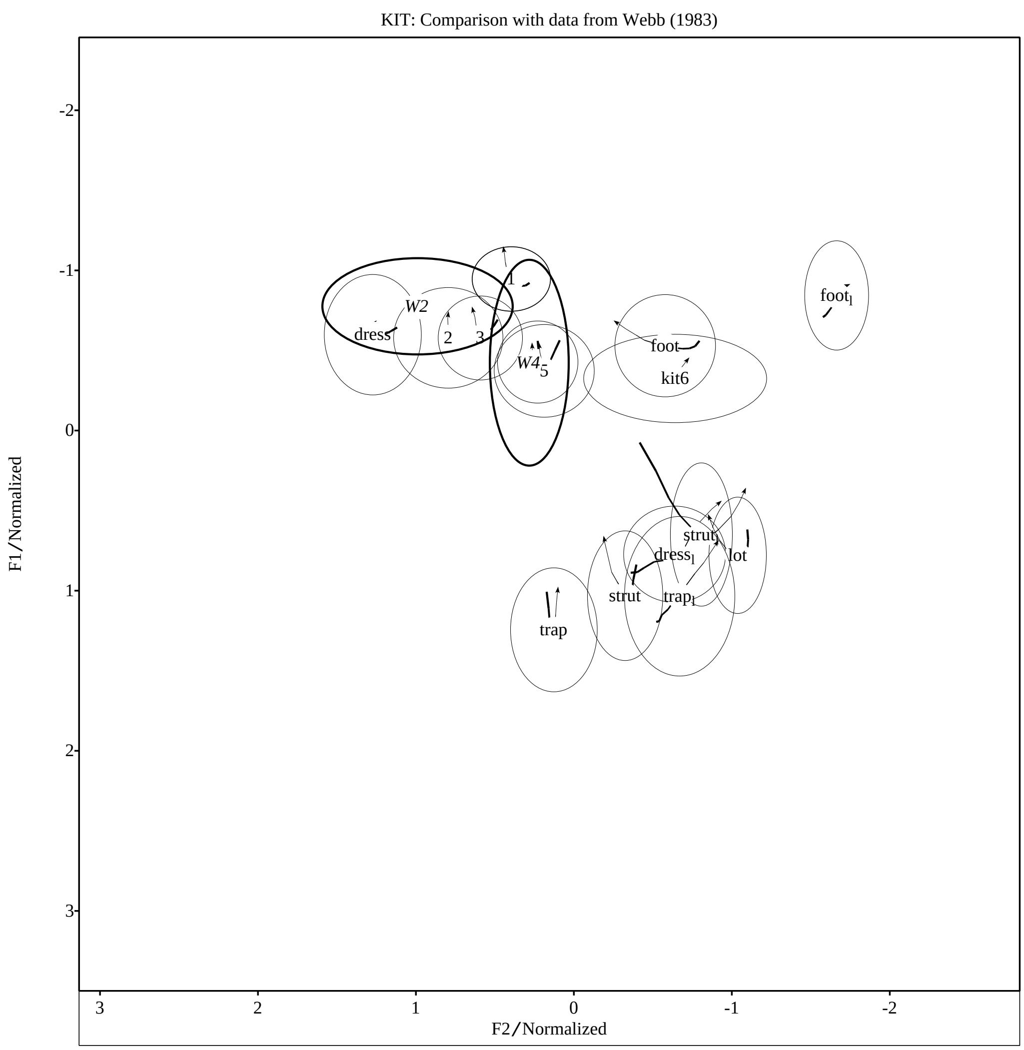


Figure 7.4: KIT: Comparison with Webb (1983)

page 151. Webb's (1983) *W4* category also appears to indicate the slightly greater prevalence of more open tokens. These difference are, however, easily attributable to the small number of tokens (15 across the two categories) used by Webb (1983) and, overall, there seems to be little evidence of substantial style-shifting or differences across generations and gender.

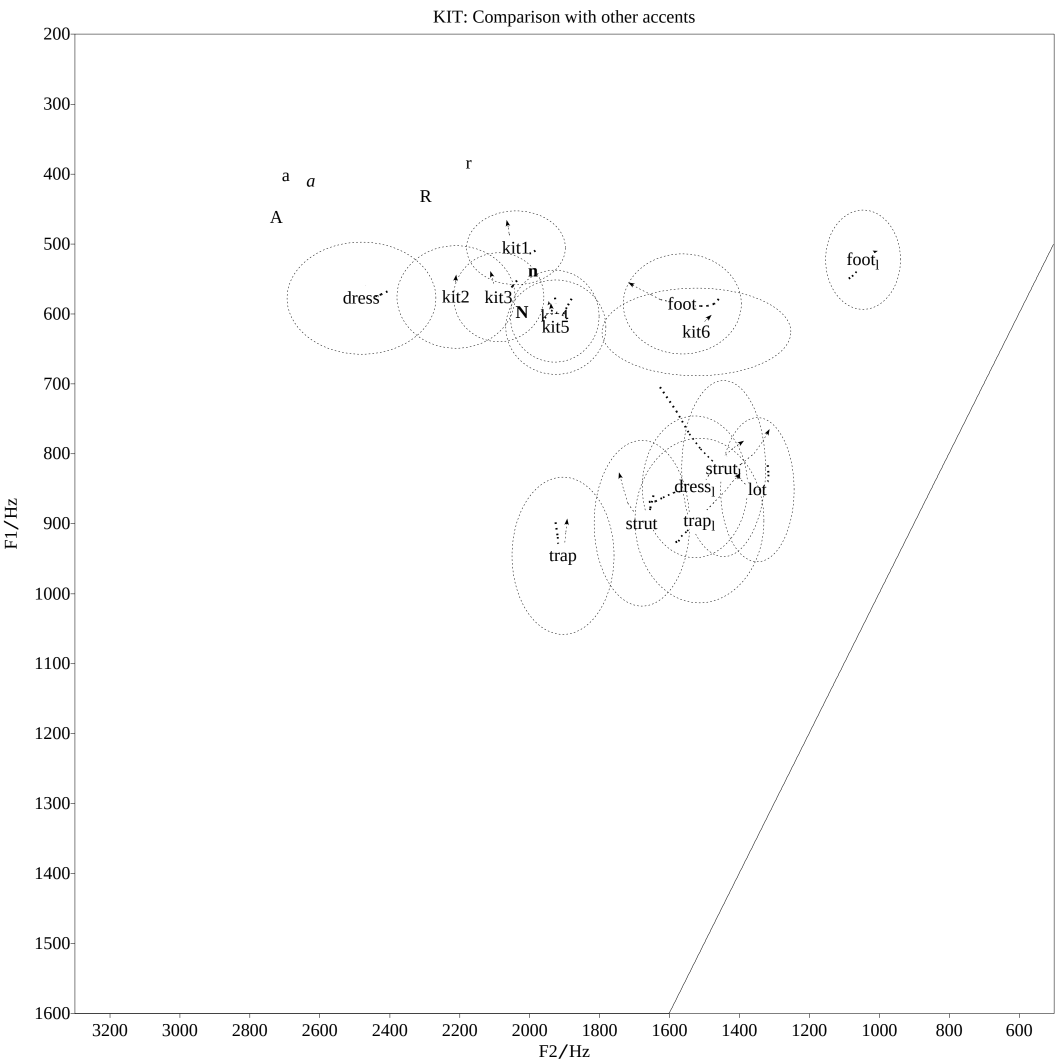


Figure 7.5: SAE KIT vs. other English accents

Figure 7.5 compares SAE KIT with KIT found in RP and the other Southern Hemisphere Englishes. The data represented in this figure creates some problems for interpretation mainly due to the lowered position of SAE KIT1–3 *vis-à-vis* the RP values. As reported on in §7.2, General RP KIT is usually given as a slightly centralized cardinal 2 i.e. [ë]. Older versions of RP have, however, closer values for this vowel. Let us assume that DRESS in SAE is cardinal 2 on the basis of its identity with AusE DRESS and its slightly higher position in comparison with RP DRESS, as depicted in Figure 6.27 on page 229. One might wish to argue that on the basis of these two facts there is little reason to doubt that SAE DRESS is anything lower than C[2]. This also seems to make sense with respect to the AusE KIT values which are often reported as having a more fronted (and close) value – a fact which, of course, connects with the diphthongisation of FLEECE in this accent. The problems remains, then, of accounting for the somewhat unexpected position of RP KIT. It is, furthermore, impossible to make recourse to the fact that the close values reflect citation-form speech only (and thus an older, more conservative form of RP), since as can be seen in Figure 7.5, the connected-speech data for KIT (labelled 'r') shows an equally (if not more) close value, although somewhat more centralized. Furthermore, the relevant female subjects were recorded in 1990, with an average age in that year of 36 (Deterding n.d.). Already in Wells (1982:291), there is evidence that "relatively close and peripheral qualities are associated particularly, but not exclusively, with old-fashioned RP; relatively open and central qualities are common with younger speakers".

Whatever one wishes to make of the RP values contained in Figure 7.5, a clearer picture of SAE KIT is provided on consideration of the NZE values. A close inspection of Figure 7.5, will reveal that the NZE KIT data finds itself in a position half way between *kit3* and *kit4/5* (represented by *N*), or slightly higher and just below *kit1* (represented as *n*); essentially in the 'middle' of the allophonic distribution of KIT in SAE. If one is, thus, willing to assume that KIT in NZE is also subject to 'shading' (and there is at least some evidence for a fronter allophone before velars in NZE) then the difference between NZE and ***Gen***SAE seems slight indeed.

Figure 7.6 on the following page provides a comparison between the current SAE data and that of Torgersen and Kerswill's (2004) Ashford data; the position of the diamonds (young-female) compared with the bullets (older-female) indicate that a process of KIT-fronting is underway; presumably in response to a fronted FOOT – see Figure 2.14 on page 64. What is clear, however, is that there is substantial overlap between the RP-data and the Ashford data and that *both* show evidence of being in a closer position than the SAE (and NZE) data.

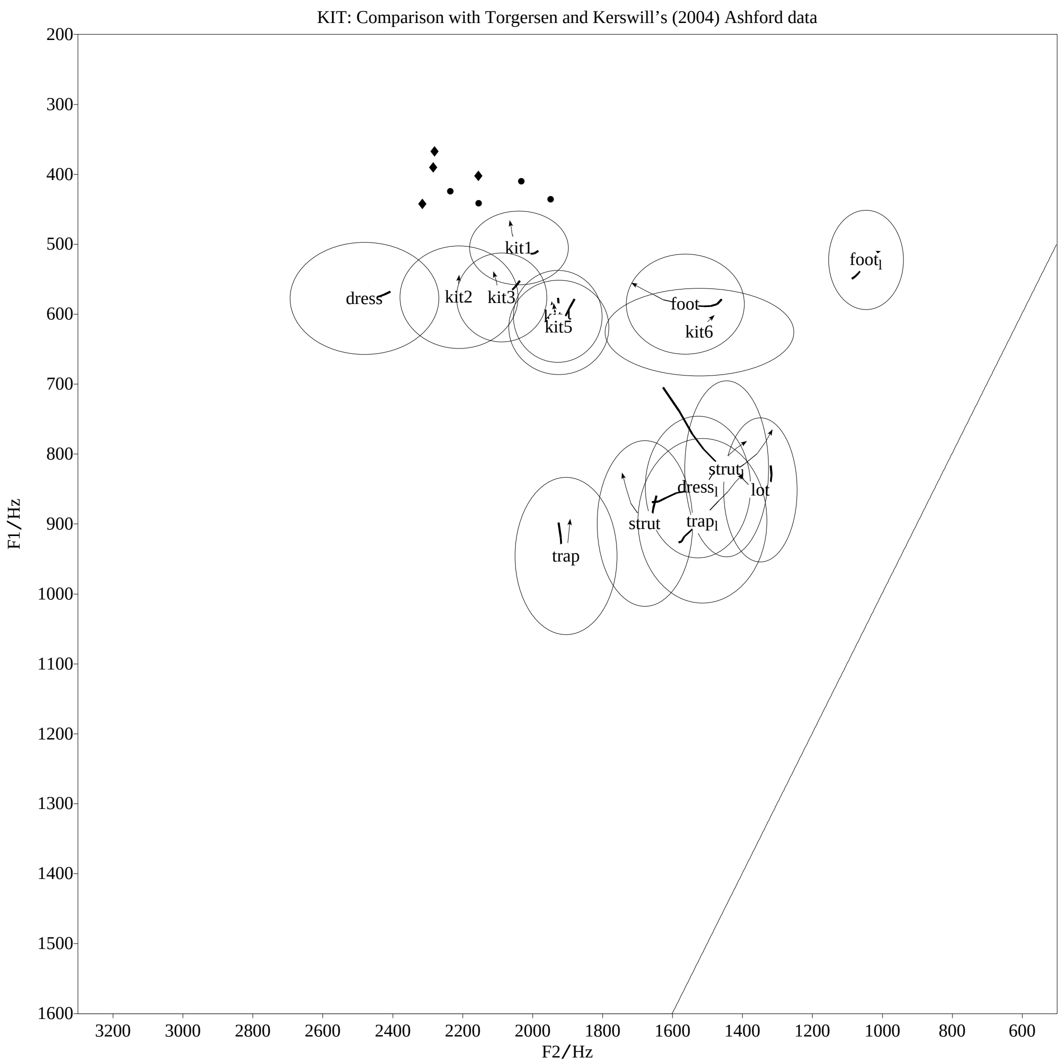


Figure 7.6: SAE KIT: Comparison with Torgersen and Kerswill's (2004) Ashford data

7.5 Conclusion and Synopsis

Much has been said about KIT in the impressionistic literature, with a variety of possible hypotheses offered to explain its synchronic and diachronic status. What I would like to offer here is a selective *synthesis* that brings together a number of 'strands' in the analysis provided above, but with the added merit that it takes seriously both endogenous forces *and* the various koinéization processes that, I believe, took place during the development of SAE. In this regard I take seriously Branford's (1994) 'warning', briefly mentioned in §3.4.3, that "in the case of a variety or variety cluster with as complex a history as that of SAEP, we are unlikely to be able to construct a single monolithic explanation that will fit the observable facts". The reconstruction which follows also has the advantage that it is in accordance with the notion, repeated at various places by Lanham, that GenSAE is expanding at the expense of BrSAE, as well as with other reconstructions already provided in this thesis, in particular that of PRICE, which, as seen in §6.2, is amenable to explanation only if we take into account the influence of multiple factors on its development.

Trudgill (2004:47) confirms that "there is some evidence to suggest that . . . KIT . . . may also have had a closer quality in some forms of nineteenth–century English English than that which is most current today" a fact reflected in the traditionally high-values for Cockney TRAP, DRESS and *KIT* as well as closer values for KIT in an older version of RP. Lass and Wright (1985:145–6; my emphasis), on the other hand, claim that "a tendency to *lower* /i/ was prevalent from the 15th century onward, especially in the East Midlands; forms seem to have entered London English early" and provide *some* additional evidence for centralized values being brought over from England during the 1820-settlement while at the same time providing reasonably convincing evidence that the [i]-value of BrSAE could not have been sourced from Afrikaans. Lass and Wright (1985:153), in fact, then 'smuggle' an English-working-class derived [i]-value into the early-Johannesburg mix to account for the polarization of [i] and [ə] values in BrSAE and AE. Whether the high-front values arrived during the 1820 settlement or via British immigration during the gold rush is, however, to some degree irrelevant to the main thrust of this argument; essentially an [i] was brought to Johannesburg, whether from early 19th-century England (and perhaps Afrikaans) via CE, or from late-19th-century working-class southeast England[16].

[16]Note, however, that there are indications that the north-of-England (with a generally low KIT) had a greater representation in early Johannesburg than the south, at least on a working-class level. Further research would, however, be needed to resolve this issue, one possibility being research on older informants in the Eastern Province area in order to ascertain whether they display high and front KIT variants or not. Note too that, on the basis of Lass and Wright's (1985) argument, the [i] of Afrikaans English is derived from working-class English and not Afrikaans; an interesting hypothesis that would seem to imply that Afrikaans-English is a fully-fledged ethnolect and not simply the variable result of L2-learning. The potential parallels here between AE and BSAE

Either way, a further impetus for the development of centralized values was the retention of a monophthongal FLEECE vowel in SAE, which could easily be seen as ‘repelling’ at least the wholesale development of an [i]-value for SAE KIT (as has been the case in AusE). One possible explanation for the similarity between NZE and SAE (regardless of the fact that the former has a diphthongised FLEECE, while the latter has a monophthongal FLEECE) is that while a monophthongal FLEECE played a constraining role against KIT-raising in SAE, *secondary* DRESS-raising played a similar role in NZE[17].

As far as the NE ‘influence’ is concerned, this was likely to have contributed relatively open values from both north-of-England accents as well as Scottish English, which, as reported on in §3.3.1, was well-represented in the Natal settlement. This is perhaps reflected in the fact that *modern* NE is still often characterized as having a greater degree of centralization than elsewhere. A further north-of-England influence would have been felt in early-Johannesburg as a result of the accents brought over by working-class miners from this area.

As reported on above, older forms of RP had closer values for KIT, although presumably not as close as the stigmatized working-class ‘Cockney’-[i]; the ‘shading’ reported on by Lanham in his various publications would also, presumably, have characterized this version of RP. Furthermore, it is not inconceivable that normative-minded Natal immigrants attempting to emulate this shading would have overgeneralized the pattern (thus assigning ‘incorrect’ values to those few lexical items like *pin* which did not follow strict complementation) while retaining a relatively open value[18]. L1-Yiddish speakers would, given their predilection for NE, have possibly ‘picked up’ on this pattern as well.

Thus arose the GenSAE version of KIT, motivated by a strict aversion towards the polarized variants (particularly [i]) characteristic of CE (or southeast England working-class speech) and eventually characteristic of BrSAE and AE. The adoption of a more centralized (NE-based) KIT, but with RP-derived ‘shading’, thus, like PRICE, became categorical in GenSAE and has since spread throughout South Africa at the expense of BrSAE (as well as CulSAE) values.

At the same time, the development of centralized KIT variants was, no doubt, motivated by a *structural* ‘pressure’ to avoid merger with a substantially raised DRESS vowel. Staying briefly with endogenous considerations, one of the advantages of Lass and Wright’s (1985) analysis is that the proposed absence of [i]-variants in 1820-settler English means that one need not account for the endogenous raising of KIT (whether *in toto* or only in certain

are intriguing.

[17]As shown in §6.4, DRESS in NZE has, in certain lects, risen as far as [i̞].

[18]This, of course, assumes that the north-of-England/Scottish open values did not display shading of their own. Either way, however, the shading was likely to display a stricter degree of complementation than RP.

contexts) in the face of a monophthongal FLEECE vowel.

Turning to the acoustic data, one notes that the position of *Gen*SAE KIT in comparison with RP and Torgersen and Kerswill's (2004) data, as well as its substantial overlap with the NZE data, provides evidence for the virtual wholesale centralization of KIT in this sociolect; thus providing evidence for Lass and Wright's (1985) analysis of this vowel i.e. as having an [ï]-like value for the fronter allophones and an [ɘ]-like value for the more central allophones[19]. In GenSAE, therefore, there is little evidence for a polarization, or 'split' in the loose sense, of KIT, but rather for a form of shading reported for other accents of English as well. In fact, one wonders, in this regard, if there is currently much difference between, for example, NZE and GenSAE, and thus whether pronouncements to the contrary are not based on the difference between *Br*SAE and NZE. Given a continued expansion of GenSAE at the expense of BrSAE, one would predict the similarity between the two accents to become, over time, more apparent.

This possibility aside, we also note that the comparison between the current data and that of Webb (1983) provides some evidence for the categorical status of KIT in GenSAE[20] and thus its status as an indicator of this sociolect. We note, furthermore, the overlap between DRESS and the fronter allophones of KIT, as well as the fact that KIT before final /l/ does not undergo hypercorrection in word-list style, a fact that will later, in §9.4.3, be linked to the more general phenomenon of vowel-retraction before final /l/ and an historical process of phonologization that might explain this feature of SAE.

In addition, the acoustic representation of SAE KIT appears to provide evidence for a value higher than schwa for the central allophones. In particular, these allophones are only slightly more open than the fronter ones. While further evidence for the non-schwa position of central KIT will be provided later, particularly with regard to the analysis of weak vowels as contained in §9.3, it can be claimed, for now, that there is at least suggestive evidence to support a [ɘ]- or [ɨ]-like transcription for the central allophones of KIT, at least in GenSAE. This fact, coupled with the substantial overlap between the SAE and NZE acoustic data, provides additional although indirect support for similar proposals, as outlined in §7.2, for the transcription of KIT in the latter accent.

The thesis now turns to an analysis of BATH, THOUGHT and GOOSE (§8.2), GOAT and CURE (§8.3) as well as the remaining short vowels, STRUT, LOT and FOOT (§8.4).

[19] I am not sure, however, whether I would be willing to extrapolate this 'pattern' to BrSAE.

[20] Complicated, admittedly, by the fact that Webb is a NE-speaker; one possibility therefore is that the centralized values shown in his connected-speech are a reflection of this regional bias and that, in other areas of South Africa, GenSAE KIT undergoes a greater degree of polarization in less-monitored styles.

CHAPTER 8

Labov's (1994) Pattern 3 and SECS

8.1 Introduction

This chapter is focussed, firstly, in §8.2, on those vowels which participate in Labov's (1994) Pattern 3 chain-shift, which is, in turn, part of his broader Southern Shift i.e. BATH, THOUGHT and GOOSE[1]. Secondly, the focus is turned, in §8.3, on GOAT and CURE, the first of which plays a 'peripheral' role in Labov's (1994) Pattern 3 but a central role, along with GOOSE, in the more broadly-conceived Diphthong Shift, while CURE is included at this point because of its potential relationship with THOUGHT via the Second FORCE Merger, THOUGHT being dealt with in the immediately preceding section. Lastly, §8.4 deals with the remaining 'SECS'-vowels i.e. those vowels which participate in Torgersen and Kerswill's (2004) SECS-Shift. These are, of course, the short vowels which have not been dealt with in the preceding two chapters: STRUT, LOT and FOOT.

8.2 BATH, THOUGHT and GOOSE (BTG)

As outlined in §2.3.2.1, and as illustrated in Figure 2.12 on page 59, Pattern 3, a component pattern of Labov's (1994) Southern Shift, involves the raising and then fronting of the tense, long back vowels along the peripheral track. In SAE this mainly involves /ɑː/, /ɔː/ and /uː/.

[1]The reader is reminded that, unless otherwise indicated, BATH is used as a 'cover' term for BATH, PALM and START, while THOUGHT is used to refer to THOUGHT, NORTH and FORCE.

With respect to Wells' (1982) lexical sets, BATH refers to the standard lexical set definable "as comprising those words whose citation form contains the stressed vowel /æ/ in GenAm, but /ɑː/ in RP" (Wells 1982:134). In RP (and SAE), BATH is, in general, phonetically equivalent to two other lexical sets, START and PALM, whereas, in GenAm, BATH 'belongs' with TRAP. START has /ɑː/ in RP "and the sequence /ɑr/ in GenAm" (Wells 1982:157–8). Lastly, RP has /ɑː/ for PALM and /ɑ/ in GemAm. In Labov (1994), /ah/ is given for the vowel in *father* in a non-rhotic *or* rhotic USA accent i.e. for PALM, but for START in a non-rhotic USA accent only.

On the other hand, Wells (1982:144) defines THOUGHT as "comprising those words whose citation form in RP and GenAm has the stressed vowel /ɔ(ː)/, excluding cases where this vowel is followed by non-prevocalic /r/ in GenAm". In modern RP (and SAE), THOUGHT, NORTH and FORCE are mostly realized by the same quality and the distinction is thus mainly as a result of differences in GenAm, where NORTH refers to the "sequence /ɔr/ in GenAm, or rather in that variety of GenAm which retains the opposition between /ɔr/ and /or/" (Wells 1982:159), while FORCE refers to /or/. Transcribed as /oh/ in non-rhotic USA accents for THOUGHT, NORTH and FORCE, but only for THOUGHT in rhotic accents. Note, however, that many USA accents have merged THOUGHT with LOT.

8.2.1 BTG: Other Accents of English

Received Pronunciation: For RP /ɑː/, Lanham (1967:42; my parenthesis) provides the following description:

> "Vocoid is usually central [a]^{73} {[ɐ̹]} but tends to a more retracted position in advanced RP [Refined RP]. Many speakers have only extra length in this nucleus, others fluctuate between a long vocoid and a vocoid followed by a centring glide which is usually short and moves toward, but seldom reaches, [ə]^{44}{[ə]}".

Cruttenden (2001:115) provides somewhat more retracted values, with [ɑ̈ː] for mainstream RP, whereas "a variety of /ɑː/ retracted near to the quality of C[ɑ] is typical of Refined RP".

In his description of RP /ɔː/, Lanham (1967:43) does not distinguish between the three lexical sets (THOUGHT, NORTH and FORCE) and provides the phoneme /oᵊ/, the nucleus of which varies between a rounded [ɔ]^{56}{[ɔ]} and "a higher, more central, almost unrounded [ɔ]$^{46/55}$ {[ɔ̜̈˔]}" (Lanham 1967:43). The author adds that there is variation in terms of the existence of a centring glide to [ə], which, if present, "is most clearly heard in open syllables with strong stress" (Lanham 1967:43). Wells (1982) generally concurs with the above assessment giving a value between cardinals 6 and 7 as the norm for RP of the day and mentioning that an older RP had an opener quality. Wells (1982:293) differs from Lanham (1967), though, in claiming a greater degree of rounding for the closer variants. With respect to the FORCE lexical set in particular, Wells (1982:161) mentions the existence of a quality separate from THOUGHT and NORTH i.e. /ɔə/ in older versions of RP. The so-called FORCE-NORTH Merger has, however, mostly completed itself in modern-day RP, with FORCE taking on the same value as THOUGHT and NORTH. Wells (1982:293) also mentions that some RP speakers "have a centring diphthong [ɔə] as a positional allophone for /ɔː/, used in final position, as [sɔə] *saw–sore–soar*" (i.e. applicable to all three lexical sets) but that this is most probably a "near-RP Londonism" reflective of the behaviour of THOUGHT etc. in popular London and Cockney accents. Bailey (1984:16) also briefly mentions the absence (or progressive disappearance) of an [ɔə]-like quality in RP.

For RP GOOSE, Lanham (1967:42) gives the value of the first element of this long vowel as an open rounded [υ]^{26}{[ʊ]}, "although a lower, more central allophone with little lip-rounding is heard". The short-off glide is given as reaching [u]^{17}{[u]}, "but may be lost leaving only extra length". Wells (1982:294) mentions a fully-back GOOSE as characteristic of Refined RP, with General RP having more central values and Cruttenden (2001:123) concurs, though mentioning a slight tendency towards unrounding. For General RP, Cruttenden (2001) gives [üː ∼ ʉː ∼ ṻː ∼ ʊu ∼ ʉ̈ː]. Hawkins and Midgley (2005:193) mention a growing tendency for GOOSE in RP to become more fronted as well as unrounded. Harrington et al. (2008) provide additional support for the hypothesis of GOOSE-fronting in RP over the last few decades.

Overall, therefore, there is evidence that a Pattern 3 movement is operational in advanced varieties of RP.

London (Cockney): For Cockney /ɑː/, Wells (1982:305) claims a quality similar to RP, though in the broadest varieties there is a tendency to retract to Cardinal 5.

For Cockney and popular London speech in general, Wells (1982:310-1) proposes a THOUGHT Split to account for minimal pairs such as *bored*[bɔəd] vs. *board*[boːd]. The basis for this split does not concern us here, but it is worth pointing out that the split is a post-FORCE-NORTH-Merger one and is not, therefore, a residue of the state-of-affairs that existed before said merger. More importantly for the present discussion is the close quality of at least one member of this split. As emphasized by Wells (1982:310), "the vowel of THOUGHT (merged of course with NORTH and FORCE) is closer in London speech than in perhaps any other accent of English. Qualities around cardinal 7, [oː], are by no means unusual". Somewhat confusingly, Labov (1994:202) mentions the high back position of /ohr/ (*nor*, *or*, *four*, *fort*, *cord*), "which develops an upglide, joining the word classes of /oh/ in *law*, *cawed*, etc. so that all these words have [o˔ʊ]{o̝ʊ}". This raised vowel then pushes GOOSE into a more front position, in accordance with Labov's (1994) Pattern 3.

Increased fronting and diphthongization of GOOSE is, of course, characteristic of those accents taking part in the Southern Shift. Thus, according to Wells (1982:306), Cockney characteristically has [ʊ] to [ə] for its starting point, with a centralized to central off-glide. Cruttenden (2001:123) gives [əɯ̈ ∼ ïɯ̈] indicating a general lack of lip-rounding, while Labov (1994:202) characterizes /uw/ as a "front rounded vowel, with a front rounded upglide, [üᶣ]".

Other Accents of Britain: With specific reference to the possible origins of GOOSE-Fronting in SAE, Branford (1994:481) mentions that "fronted [ʉ] occurs in several British dialect areas, including rural Devon . . . parts of Scotland . . . and . . . East Anglia". For Norwich, Labov (1994:204) gives [ü] with a glide to rounded [ɥ]. Trudgill (2004:55-8) provides evidence of GOOSE-Fronting and/or diphthongisation in many areas of England and Britain more generally, and it seems clear that "this fronting and/or diphthongisation had certainly started in some areas in the 1800s" (Trudgill 2004:58), although it should be noted that in some of these areas, GOOSE-before-/l/ is also fronted, "unlike in London and modern RP".

Wells (1982:349–63) claims that the north-of-England generally has flat-BATH accents (i.e. merged with TRAP, generally at [a]), while a complex distribution for the THOUGHT, NORTH and FORCE sets pertains and one which is characterized a number of diphthongal and/or rather open values. GOOSE, generally, has a backed quality, although there are some fronted (e.g. [ʏː]) and diphthongal variants as well. Generally, the north-of-England does not appear to be a likely source for modern SAE values.

The 'other' Southern Hemisphere Englishes: The quality of BATH etc. is one of the main differences between SAEP and the two other main Southern Hemisphere Englishes. AusE and NZE use a vowel perhaps best transcribed as /aː/; Harrington et al. (1997:178) imply [ɐː] as being a common quality. Tristan da Cunha English, on the other hand, does show START-Backing, i.e. a [ɑː]-like value, while, intriguingly enough, Falklands Island English, "which probably developed later than South African English" (Trudgill 2004:63) does not.

In addition, in certain varieties of AusE there are a number of BATH words that are pronounced with the same vowel as TRAP, particularly "where the vowel is followed by a nasal plus another consonant, e.g. *advantage, chance* ..." (Wells 1982:599).

The quality of THOUGHT etc. in AusE and NZE is not given much attention to in Wells (1982) or Cruttenden (2001), with the former author giving [ɔː] for AusE and both [ɔː] and [oː] for NZE English. For AusE, Wells (1982:600) does, however, mention a "tautosyllabic [ɔə] ... restricted to the Cultivated accent". Harrington et al. (1997:178) propose [oː] as being the most suitable symbol for transcribing AusE THOUGHT etc. According to Evans and Watson (2004:197), /oː/ "has shown very little movement in AE over the last 30 years"; they also claim little variation for this vowel across AusE, NZ and "British English". Trudgill (2004:66–7) confirms the phonological merger of THOUGHT, NORTH and FORCE in the Southern Hemisphere Englishes, although adds that in New Zealand English, the merged vowel /ɔː/ has two main allophones, [ɔə ∼ ɔɐ] and a monophthongal version, the former occurring word-finally and the later pre-consonantally i.e. *flaw* [flɔə] but *caught* [cɔːt]. In the English of Tristan da Cunha, the phonetic realization of the merged vowel is [oə]. Generally, and as implied in Trudgill (2004:145), AusE and NZE appear to have a somewhat lower, [ɔː] value, for THOUGHT compared with SAE, "which ... along with Falklands Islands English and the English of southeastern England, notably London, is experiencing the movement of the NORTH-FORCE-THOUGHT vowel upwards from [ɔː] to [oː] or, in Tristanian, [oə]".

For Cultivated, General and Broad sociolects of AusE, Wells (1982:597) gives [ʊu], [ɨɯ ∼ ʊʉ] and [əːʉ] respectively for GOOSE. Harrington et al. (1997:163) provide acoustic evidence for an association between fronted GOOSE and 'broadness' in AusE. Their acoustic evidence also provides tentative support for the notion that the broader the accent in AusE, the greater the diphthongal movement.

For NZE, Wells (1982:607) comments as follows:

> "New Zealand /uː/ (GOOSE) is usually a close central [ʉː], or a lip-diphthong [ɨʉ] ... in broader accents, noticeable diphthonging may occur, but with rather less lowering of the first element than in Australian".

From Trudgill (2004:51), it would appear that while Falklands Islands English has a (slightly) diphthongal GOOSE vowel, [ʉ̩ʉ], the English of Tristan da Cunha has a monophthongal variant with back values being typical of older speakers.

U.S. Accents: In US English BATH, along with a number of TRAP words (collectively transcribed as /æh/), has developed a tendency to tense and develop an in-glide. This forms part of Labov's (1994) Pattern 2 and is an integral part of his Northern-Cities Shift. For New York, for example, Labov (1994:75) outlines a change from late-19th century [ɛːᵊ], to 1940s [eːᵊ] to 1960s [iːᵊ]. For Philadelphia, "the tensing and raising of (æh) is the most marked stereotype of the ... system" (Labov 1994:80). In the case of START, Philadelphia /ahr/ has completed a sound change whereby this vowel backed from [aɚ] to [ɒɚ] (Labov 1994:79). This, in turn, led to a raising of /ohr/ i.e. NORTH/FORCE. A raised /oh/ is typical of non-rhotic New-York speech, as discussed in Labov (1994:48–9). The relevant variable includes words such as *law* and *awful* but also LOT words such as *coffee* and *lost* and is also characterized by a discernable in-glide: thus [koəfɪ] or an even higher nucleus in *coffee*. In the latter case this leads to homophony between, for example, *law*, *lore* and *lure*. There is a tendency to correct to a low back position in careful speech (Labov 1994:101; footnote). The raising of /oh/ in New York City speech leads to an overlap (with respect to the relevant nuclei) with /uw/. From (Labov 1994:203–4) it appears, however, that this does not lead to GOOSE-fronting given the different directions of the glides. For Philadelphia, the backed START vowel discussed above has led, in its turn, to a raising of NORTH-FORCE, /ohr/, creating, for most speakers, a merger between /ohr/ and /uhr/ i.e. CURE which, of course, is not a separate phoneme in rhotic dialects but a sequence of vowel plus /r/. Thus *moor*, *more* and *boor*, *bore* are homonyms (Labov 1994:80). According to Labov (1994:202), THOUGHT (/oh/) is also raised in Philadelphian speech, but not /ah/ (PALM). A similar pattern is evident in Texas (Labov 1994:202–3). For the Southern USA more generally, a typical (rural) feature is the use, for THOUGHT, of a back, upgliding diphthong e.g. [baɒt] for *bought* (Wolfram et al. 2004:345). Labov (1994:216) gives [ɑʊ] as a possible value.

For Philadelphia, GOOSE-Fronting is a prominent sociolinguistic marker (Labov 1994:80–1), but retains a substantially backed allophone before final /l/. It would appear, in its most advanced values, to have an unrounded nucleus with a glide to back [u] (Labov 1994:202). Texas also shows prominent GOOSE-fronting (Labov 1994:203). GOOSE-Fronting in California (Los Angeles), as reported in Labov (1994:219), is an example of a non-Southern US dialect showing this feature, and it appears that similar fronting is taking place in the dialects participating in the Northern Cities Shift. As such, and as confirmed by Fridland (2008:69), GOOSE-Fronting is, in fact, a feature of all regional dialects of the US, not just the South.

As noted in §2.3.2, GOOSE-Fronting is in line with Labov's (1994) Principle III of chain-shifting, as well as an integral part of Pattern 3 and, thus of his Southern Shift. As already mentioned briefly in §1.4, Harrington et al. (2008:2825) attempt to explain such GOOSE-Fronting (in RP, but more broadly too) in terms of a process by which new language learners fail to "compensate for coarticulation". The gradient effects predicted by such a process seem to sit well with the facts of GOOSE-Fronting as well as with the notion that GOOSE-Fronting is a system-driven, endogenous change. Note, however, that the mechanism for change appears to entail the existence of "the prevalence with which RP /u/ follows consonants with a high F2 locus, both after alveolars in words like *noon*, *soon*, and after /j/" (Harrington et al. 2008:2825). As pointed out by these authors, the various Southern Hemisphere Englishes have been equally frugal in terms of yod-dropping and have retained at least an affricated reflex of yod in contexts such as *tune* and *duty* i.e. while RP often retains an allophone of /j/ in words such as *tune* ([tj̊ʉːn], SAE, for example, has [tʃʉːn][2]; regardless of the realization, a high F2 locus is maintained. However, since GOOSE-Fronting, as pointed out above, is extensive in US English, and since American dialects often drop their yod ([tʉːn]), this creates a potential problem for Harrington et al.'s (2008) analysis.

8.2.2 BTG in SAE: The Impressionistic Data

Hopwood (1928): For SAEP /ɑː/, Hopwood (1928) gives the (Broad) value of this vowel as the half-long [ɒˑ], although the full-length [ɒː] is also implied: "The E. Province surname *Larsen* ˈ**lɑːs(ə)n** is there pronounced as if it were spelled *Lawson* ˈ**lɒːsn**" (Hopwood 1928:12). The author claims a Cockney origin for this value and that in AE it is often further raised to [ɔː]. Hopwood (1928:13) also mentions the possibility of rhoticization in the case of START vowels in particular (e.g. **pɑrt**) as well as shortening of the vowel more generally (e.g. **plɑnt**). Interestingly, the author claims that Scottish English Pronunciation "aː is prevalent in Natal and the east of the Eastern Province, and is due either to Scotch teachers, or to the loan of a Native [i.e. Nguni] "clear" aː for E. "dark" ɑː, or both" (Hopwood 1928:13; my parenthesis).

For SAEP THOUGHT, Hopwood (1928:15) claims the following: "**ɒː** > **ɔː**, or even **oː** . . . This is also a Cockney E. change from **ɒː** to **oː**; a raising and tensifying of the vowel, which obtains a good start from the usual AE substitution of **ɔː** for **ɒː**".

[2] See Limbrick (1989) for more on this.

In this regard, Hopwood's (1928) [ɒː] for RP THOUGHT is an old-fashioned one, even for his time. As stated succinctly by Wells (1982:235), "this vowel had become less open, i.e. [ɔː], by the nineteenth century". For the possible RP norm during Hopwood's (1928) time, it is worthwhile quoting Wells (1982:293; my parenthesis) in full:

> "The vowel /ɔː/ has been getting less open over the last half–century. Newsreels from the thirties [i.e. at the least two years after Hopwood (1928)] often evidence a cardinal–6–like quality [[ɔː]] which now seems dated.".

With respect to the so-called NORTH and FORCE vowels, Hopwood (1928), as with SQUARE (§9.2) etc., claims a degree of rhoticity. So, for example, he claims that StEP (i.e. RP) "*short* **ʃɒː⁽ᵊ⁾t** > "sho(r)t" **ʃɔrt** or **ʃoˑ(ɹ)t**". A similar analysis is provided for *warm* and *score*. Again this is undoubtedly, a BrSAE feature; as mentioned in §3.4.3 all other varieties of SAE are non-rhotic[3]. It should also be mentioned that the RP values provided by Hopwood (1928:19) for *short*, *warm* and *score* are suspect i.e. **ʃɒː⁽ᵊ⁾t**, **wɒː⁽ᵊ⁾m** and **skɒː⁽ᵊ⁾**. The first two belong to the NORTH lexical set[4], while *score* belongs to FORCE (Wells 1982:162). The homophony of THOUGHT, NORTH and FORCE in modern-day RP is, historically, mainly a result of R-Dropping and the First FORCE Merger, the following summary of which follows that provided in Wells (1982).

Firstly, in the 18th-century /r/ disappeared in non-prevocalic position. By this time the original [ɒː] in THOUGHT had been replaced by [ɔː], NORTH had changed from [ɒːr] to [ɔˑr] and FORCE was [ɔər]. As a result of R-Dropping NORTH then became [ɔː] and FORCE became [ɔə]. An [ɒ] value in early 20th century RP for *short*, *warm* and *score* is thus highly unlikely. Secondly, the First Force Merger involved the monophthongisation of FORCE vowels i.e. [ɔə] becomes [ɔː]. As mentioned in Wells (1982:235), "the First FORCE Merger was not complete in RP until the current century". Thus while a diphthong in *score*, [skɔə], is possible for RP during Hopwood's (1928) time, given that it qualifies as a FORCE vowel, both *warm* and *short* qualify as NORTH vowels and were thus never diphthongised in the first place. It can

[3] A more general rule is provided by Hopwood (1928:19) to cover the behaviour of the SQUARE vowel, as well as all the other vowels with historical 'r' i.e. NORTH, FORCE, NEAR, CURE etc.: "A vowel (or diphthong) followed by historical *r* still written, is pronounced shorter (or monophthongised, raised and tensified, and followed by a semi-rolled **r**, or coronal **ɹ̣** especially when final . . .)". To be frank, it is somewhat difficult to reconcile this rule with some of the examples provided by Hopwood (1928:19).

[4] see Wells (1982:160).

only be concluded that Hopwood's (1928) values for *warm* and *short* are hypercorrective (and artificial). A final piece of evidence in support of the above analysis is that in Jones (1924), the first edition (revised) of *An English Pronouncing Dictionary*, *warm* is transcribed as [wɔːm], *short* as [ʃɔːt] and *score* as [skɔː], with [skɔə] as the less–common alternative.

For SAEP GOOSE, Hopwood (1928:22) gives a fronted and rounded vowel: "**ùˑw**{[uː]} > **ÿü(w)**{[ÿü(ː)]}". Note that this transcription implies a (diphthongal) retraction from front to back while rounding is maintained. There is also the option of extra length. Hopwood (1928) claims a Cockney source for GOOSE-Fronting[5], helped by the presence of a high, front, rounded vowel in Afrikaans i.e. /y/ and "pairs of words such as . . . E. (noun) excuse **ɛkˈskjùˑws**{[ɛkˈskjuːs]}, and A. ekskuus **ɛ̈ksˈkys**". Note that Afrikaans /y/ is also lengthened in some contexts; thus Lass (1990:282) provides [fɐn ˈfyːrən] for the Afrikaans surname '*Van Vuuren*'.

Hooper (1945), Breckwoldt (1961): Hooper (1945) shows a correlation between raised and more rounded variants of BATH and having a relatively lower social-class status or being male. The Broadest value that Hooper (1945:479; 482) provides is a "vowel near that in *on*". Breckwoldt (1961:7) also gives rounded and raised variants of this vowel in SAE i.e. [ɒ] and [ɔː]. For THOUGHT etc., Breckwoldt (1961:7) notes a raised variant in SAE i.e. [oː]. His other claim, that [ɔː] becomes [ʊː] seems unlikely, however. For GOOSE, Breckwoldt (1961:7) gives both [ɷ]{[ʊ]} and [ÿ] as possible values for this vowel. As an aside, the substitution of /ʊ/ for /uː/ is a characteristic of my own speech. It was only recently that I discovered that *room* is, prescriptively, [ruːm] and not [rʊm].

Lanham and Traill (1962), Lanham (1964?; 1965): For non-SARP varieties of SAEP, Lanham and Traill (1962:31) claim a shortening of the BATH-PALM-START nucleus, as does Lanham (1964?:27). These varieties as well as SARP 'B' show backing, raising and slight rounding as well. The combination of all of these features naturally leads to near–homophony between BATH and LOT (§8.4). The more common SARP 'B' as well as SARP 'A' value is, however, given by Lanham and Traill (1962) as the comparatively fronted [*a*]$^{64/73}${[ɐ̠]}. Lanham (1965:92) confirms this picture when he claims that "/ɒ/:/ɒˑ/ is for many South Africans already established as a contrast depending only on length, e.g. in *pot* /pɒt/: *part* /pɒˑt/.

For SARP THOUGHT etc., Lanham and Traill (1962:31) give "long [θ]$^{47/56}${[ɔ̝]}

[5] Thus Hopwood (1928:22) says, for Cockney, that "Prof. Jones writes it **iü**".

with weak slit rounding” for this vowel, with a trend towards a higher vowel with more lip–rounding, especially in the context of an immediately preceding /w/ or /y/, the authors giving [θ]^{47}{o̝} as one possible value. The authors claim RP THOUGHT “is generally represented by a more central, slightly lower vowel which lacks lip–rounding” (Lanham and Traill 1962:31). According to Lanham (1964?:27) the NORTH and FORCE vowels are often shortened in SAEP, and in the particular case of FORCE, Lanham (1965:92) emphasizes the monophthongal nature of this vowel.

For SARP GOOSE, Lanham and Traill (1962:8) mention the possibility of a brief backwards up-glide (i.e. diphthongisation), particularly when it is “under primary stress and immediately before a voiced plosive or fricative, or finally, e.g. in *booze*, *boo*”(Lanham and Traill 1962:29). As the diphthong progresses so does lip–rounding. For the onset, these authors provide the values [ʋ]^{26}{[ʊ]} or [ʋ]$^{26/35}${ʊ̟˔} while the glide is at [u]^{17}{[u]}. Common in non-SARP SAE (but frequent enough in SARP nonetheless) is, however, a tendency to front this vowel and the more front the more extreme the variety of SAE. The slighty-less extreme value is given as [ü]$^{14/25}${[ʉ̠˔]}, “which is followed by a slight glide towards the back of the mouth or is merely ‘long’,” while the most extreme value provided by Lanham and Traill (1962:30) is [ï]^{14}{[ʉ̟]} with [ʋ]^{25}{[ʊ]} being the value of the terminus of the glide[6].

Lanham (1967): Lanham (1967:64) gives rounded [ɒː]^{65}{[ɒ̞̈ː]} as the non-conservative value in SAE for BATH and claims that it is a long vocoid “without any trace of a centring glide” (Lanham 1967:64). The overlap between BATH and LOT in terms of quality, although not length, is again emphasized. The conservative SAE values are given as unrounded [ɒː]^{74}{[äː]} or [aˑ]^{73}{[ɐ̝ː]}[7].

For THOUGHT etc., Lanham (1967:64; 68) gives rounded [oː]$^{36/46}$ and [ɵ]^{46} (both values translatable into modern IPA terms as {[ö̝ː]}) for this vowel in non-conservative SAE and also confirms its generally monophthongal nature. Conservative SAE is described by Lanham (1967:92) as displaying a “somewhat centralized, rounded, long [ɔː]$^{46/56}${[ɔ̞̈]} with no following glide”.

For GOOSE, Lanham (1967:60) concurs with his earlier assessments, but also mentions a trend towards loss of lip-rounding as the variety in question becomes closer to the Broad end of the sociolinguistic cline. Thus, “in extreme SAE . . . /uu/ is often represented by central unrounded [ï]{[ɨ]} which dialect-sensitive South Africans

[6]Note that I suspect that Lanham and Traill (1962:30) meant [ʋ]^{26} here and thus my translation into [ʊ].

[7]For the first alternative, Lanham (1967:91) does not, in fact, mention that it is unrounded. I suspect he simply forgot to mention this, and thus my transcription.

hear in the SAE pronunciation of a word like school, as the "placing of y before the vowel"" (Lanham 1967:60). Later, Lanham (1967:74) concludes that "in extreme forms of SAE the vocoid is weakly rounded [ü]$^{14/24}${ʉ̹}". In more general terms, Lanham (1967:91) claim a weakly-rounded [ʉ̹] value for non-conservative SAE, with this trend towards centralization having made in-roads into conservative SAE. As in Lanham and Traill (1962), an open-rounded [ʊ]^{26}{[ʊ]} with a glide towards [u]^{17}{[u]} would be the archetypal conservative SAE value, although the author stresses that most SAE speakers lie somewhere between the extreme (usually) monophthongal [ʉ̹] and this archetypal value.

Lanham (1978), Lanham and Macdonald (1979), Lanham (1982): Lanham (1978:153) mentions the possibility of (near)-neutralization between the backed and raised BrSAE version of BATH and its glide-weakened, backed and raised equivalent for PRICE – see §6.2. According to this author the backed raised variant is "stigmatized and corrected to front [aː] and hence involved in stylistic variation . . . originates . . . in speech of lower classes from Home Counties" (Lanham 1978:153).

Lanham and Macdonald (1979:38–40) confirm a "backed, raised, rounded, shortened aa" (Lanham and Macdonald 1979:38) as a Cape English variable and thus associated with BrSAE. The most advanced variant is given as [ɔ̈ˑ], which is corrected to [ä̝ː] or even hypercorrected to [aː]. The variable appears to be above the level of consciousness and would appear to carry covert prestige in the SAE community. "Backed aa correlates with the Eastern Cape as a region, in fact with the Cape generally. Lower class in the cities is a further correlation" (Lanham and Macdonald 1979:39)[8].

Lanham (1978:154) confirms the raised value of THOUGHT in SAE, giving the most advanced variant as [˂oː]{[öː]}. According to this author, the feature is below the level of social consciousness and not subject to stylistic variation. The origin is unknown. Lanham and Macdonald (1979:47) essentially repeat this assessment, identifying a raised THOUGHT as a Gen SAE variable and adding that "a British origin is probable". Interestingly, by the 1980s, Lanham (1982:338) is reporting a "raised /oᵊ/ in, for example, [boːd] *bored* and [poːz] *paws*" for *Cul*SAE. As with GOOSE, it is possible that this reflects the growing influence of local traits on this near-RP ac-

[8]An interesting parallel development in this regard is the use of a backed and raised [ɒ]-like vocoid in place of [ɑː] in some urban forms of Afrikaans – particularly in female speech. Thus Lanham and Macdonald (1979:39; footnote) mention "advanced variants . . . prominent in the speech of at least one female professional voice on the Afrikaans service of SA television" and, more recently, this accentual feature has been parodied in a series of television advertisements for Polka.com, an internet service-provider. The later Branford (1994:480) also mentions this phenomenon and Wissing (2006) provides a study specifically focussed on, and confirming of, this trend.

cent: Lanham (1978:149; 153) identifies a centralized GOOSE vowel as (probably) a Natal variable and as a feature of GenSAE, but as also being prominent in BrSAE. Accordingly, even the most advanced values (a fully central [ʉ̞] especially after a yod e.g. in *mule* /mjul/) are "below social consciousness; high index values are maintained in formal style and little stylistic variation" (Lanham 1978:153–4). Lanham and Macdonald (1979:45) bring tentative support to this assessment. Lanham (1982:338) gives a centralized GOOSE vowel for CulSAE, "especially following /j/ in such words as *new* and *due*". Again, it is unclear whether this is a reassessment of the quality of GOOSE in CulSAE or reflective of the growing impact of "South Africanisms" on this near-RP variety. Given Lanham's (1982) insistence that this trait is below the level of social consciousness it seems likely to be the latter.

Wells (1982): The range of values provided by Wells (1982:615) for BATH is from the more conservative and more formal [ɑ+ː]{[ɑ̟ː]} to Broader and more colloquial [ɒː], with the latter having weak lip-rounding. Wells (1982) comments further:

> "South Africans can often be readily recognized as such by the very back quality of the /ɑː/ of START, BATH, and PALM. In broad speech it may also be weakly rounded, thus becoming qualitatively identical with the /ɒ/ of LOT. Thus *last part* ranges from conservative [ˈlɑːst ˈpɑːt] to broad South African [ˈlɒːst ˈpɒːt]. This broad variant is, however, stigmatized and so subject to stylistic alternation with a less back [ɑ+ː]{[ɑ̟ː]}" (Wells 1982:615).

Wells (1982:615) claims that THOUGHT-FORCE-NORTH in SAE follows "the same trends as in other southern-hemisphere accents . . . [and] . . . in England" i.e. it is becoming closer, essentially reaching cardinal [o]. Echoing the work of Lanham etc., Wells (1982:615–6) mentions the centralization of GOOSE, particularly after yod and its wide-spread (non-conscious) adoption by the SAE speech community. "This agrees with developments in the south of England and Australian-New Zealand, except that South Africa lacks the competing trend towards diphthonging of /uː/ found in the other places mentioned" (Wells 1982:616).

Lass (1990; 1995; 2004): According to Lass (1990:278), for GenSAE in Cape Town, BATH-PALM-START has a range from [ɑ̟ː] to [ɑ̈ː], with fully-back, rounded and raised variants [ɑː ∼ ɒː ∼ ɔː] reserved for BrSAE. Fronter values such as [ɐː] do occur, though rarely. He also states that "like LOT, this is characteristically less than fully open" (Lass 1990:278). A more accurate description for GenSAE BATH might thus be, for example, [ɑ̝̈ː].

Lass (1995:99) confirms the social significance of this variable in SAE, giving [ɑ̈ː] to as front as [a̶ː]{[ɐ̞ː]} for CulSAE, with a backer value in GenSAE and BrSAE, "even fully back [ɑː]" (Lass 1995:99). In GenSAE, men back this vowel more than women and younger speakers more than older. It becomes particularly socially significant (and amenable to stereotyping) when lip-rounding (and sometimes raising) is added to provide BrSAE values such as [ɒ̈ː ∼ ɔː]. Lass (1995:99), however, adds that "there is some evidence that at least weak rounding is beginning to become less stigmatised now". Lass (2004:377) adds nothing really new to his earlier analyses, giving [ɑː] as the most typical value for GenSAE speakers.

Lass (1990:278) gives the value for THOUGHT etc. as [oː], with some speakers having the slightly more open [o̜ː]{[o̞]}. The author adds that the main value, i.e. [o], is "very like the current London nucleus in the same category" (Lass 1990:278). The author claims that values as low as [ɔ] do not occur. As discussed briefly in §8.4, [oː]-like values also occasionally occur in so-called CLOTH words.

Lass (1995:99) provides a similar picture to the one already provided, stressing that the THOUGHT etc. variable has only marginal importance sociolinguistically in SAE: a value approaching [ɔː] in CulSAE, with [oː] in the other sociolects. The use of THOUGHT for CLOTH words (*soft*, *cloth*, *loss*, *Austria*) is mentioned as highly variable, but, according to (Lass 1995:99), a general trend appears to be that "the more conservative the style or lect, the less likely the CLOTH words . . . are to have THOUGHT; though most varieties have it in *off*" . This seems incorrect, however, and, according to Lass (2004:377), SAE is, in fact, to be seen as lagging behind British dialects in the use of LOT for CLOTH: "/oː/ in *off*, *cloth*, *sausage*, *Austria* is not uncommon even in teenagers" (Lass 2004:377). The more conservative the speaker, therefore, the *more* likelihood that THOUGHT will be used.

According to Lass (1990:273), CulSAE is distinguished from GenSAE by a more back GOOSE vowel. For GenSAE this author gives [ʉː] as the most common value, but often even fronter values occur, such as [ʉ̟ː] or [ÿː], "particularly in female speakers and most often in final position under high prominence" (Lass 1990:278). This author emphasizes the fact that a fronted GOOSE vowel lies well beneath the level of social consciousness. Again, before /l/ there is retraction, [üː ∼ u̟ː], and in certain environments (especially before sonorants) there is tendency to diphthongise e.g. *pool* [pʉuɫ] (Lass 1990:278).

According to Lass (1995:98), CulSAE has a [uː] quality, whereas "in all other varieties it is never backer than central [ʉ]" (Lass 1995:98). Among younger Respectable speakers (especially female) there is a tendency to front the vowel even further, even as far as [yː], "with 'compressed' rather than 'pouted' lip-rounding, very like the vowel in Afrikaans *vuur*" (Lass 1995:98). The author denies, however, an Afrikaans origin for this value and claims that, in general, "the higher up the Respectable scale, and the younger the speaker, the fronter the vowel" (Lass 1995:98). He confirms the important role played by this value as a social variable, both within SAE, and *vis-à-vis* other ethnolects of SAE e.g. Coloured English in which GOOSE has a particularly backed value.

A fronted GOOSE vowel is thus particularly 'white'. Thus, for example, Lass (2004:377) confirms that 'black' speakers using a fronted GOOSE vowel have often been "stigmatised by the political left as linguistic turncoats or ethnic/class traitors". As with FLEECE, Lass (2004:376) contrasts the monophthongal status of this vowel in SAE, with its diphthongal status as commonly found in the two Australasian varieties, AusE and NZE. He generally confirms his earlier analyses, stating that "in university-age females it is now often completely fronted to a weakly rounded [yː], or even a vowel with a neutral lip-attitude distinct from [iː] only in that the latter is not neutral but spread" (Lass 2004:377). The most typical value in GenSAE is given as [ʉː].

The Remaining Sources: According to Mesthrie (1993:30), in BrSAE the BATH etc. vowel is typically back and rounded into [ɒ̈ː] or even raised into [ɔː]. The use of [ɔː] is a much stereotyped feature of BrSAE and AE. "Thus to an outsider . . . *park* may sound like RP *pork*, *cart* like RP *court*". This author claims "the influence of lower class Home Counties speech" for the value of BATH etc. in SAE (Mesthrie 1993:30). Branford (1994:480) claims a comparatively retracted "[ɑː] in conservative speech". A rounded and raised variant is confirmed as stereotypical of lower-class and/or rural speech. Bowerman (2004:937) gives [ɑː] as the quintessential GenSAE value, with BrSAE shortening, raising and rounding to [ɒ ∼ ɔ]. CulSAE has a more central version than GenSAE.

For THOUGHT etc. Bailey (1984:16) confirms the absence of an in-glide for this vowel in non-RP-influenced SAE accents. He also mentions the occasional use of a THOUGHT vowel in CLOTH words such as *cross*, *loss*, *wrath* and *salt*, but confirms they are heard in "a laboured, exaggerated RP style in Britain and less frequently in South Africa" (Bailey 1984:19). Branford (1994:480) confirms [ɔː] for CulSAE with the higher [oː] in other lects. Likewise Bowerman (2004:937–8) confirms [ɔː] for CulSAE, while [oː] characterises GenSAE and BrSAE. Very surprisingly, the use of THOUGHT for CLOTH words is identified as a BrSAE feature.

For GOOSE, Branford (1994:480) confirms the fully-back [uː] in CulSAE and a more central variant for non-CulSAE variants of SAE; Bowerman (2004:937) confirms the general picture provided above. Da Silva (2007:117) gives [ʉ(ː)] to [yː] as encapsulating the possible range, although the implied shortened value seems unlikely.

8.2.3 BTG: A Comparative Acoustic Analysis

In Figure 5.5 on page 152, Webb's (1983) BATH vowel is represented among the other SAE long vowels and diphthongs. As is to be expected, its quality appears relatively retracted. While there is obviously no danger of a PRICE–BATH merger in this idiolect, "there is a considerable overlap between the long vowel /ɑ/ . . . and /ɒ/" (Webb 1983:149). This state-of-affairs is discussed further in §8.4, in the discussion on LOT, and is represented clearly in, for example, Figure 8.22 on page 341.

As in the case of THOUGHT etc. (see below), Webb (1983:149) has a rather curious take on the nature of his BATH-START-PALM data i.e. he (tentatively) suggests the possibility of a difference in phonetic quality between BATH-PALM on the one hand and START on the other:

> "There seems to be a slight but possibly significant difference between /ɑ/ immediately followed by orthographic r, and /ɑ/ elsewhere. The evidence is, admittedly, slender but it does bear looking into. If the distinction is found to hold, it will provide additional evidence for underlying /r/ in SAE".

The slender evidence that he refers to is a F_1 and F_2 value of 646Hz and 997Hz for START, as opposed to 622Hz and 1065Hz for BATH-PALM.

Figure 5.5 on page 152 also provides Webb's (1983) acoustic data on his THOUGHT etc. vowel. The close nature of this vowel is apparent. Note that THOUGHT-NORTH-FORCE before final /l/ does not appear to undergo significant retraction. As in the case of BATH-PALM-START, but even more speculatively, Webb (1983) claims some evidence for a quality difference between THOUGHT on the one hand and NORTH-FORCE on the other, i.e. /o/ without following orthographic 'r' and /o/ with following orthographic 'r' respectively. The evidence is, by the author's own admission, slender at best. It should also be noted that the author's two tokens of *sure* have been included under THOUGHT etc. and not CURE, since "the two tokens measured are phonetically [ɔ]". As discussed in §8.3, there is some evidence for the merger of CURE and THOUGHT in SAEP i.e. the so-called Second FORCE Merger.

In addition, Figure 5.5 shows three allophones of the GOOSE vowel i.e. an unmarked fronted allophone (goose$_{un}$), a slightly more fronted allophone after yod (goose$_{yod}$) and a markedly retracted allophone before final /l/ i.e. (goose$_l$). Webb (1983:152) in fact deals with GOOSE after yod as a separate phoneme, /ʉw/, an approach which does not seem warranted in the SAE context. Besides GOOSE after yod, Webb (1983:149; my parenthesis) claims three distinct allophones for GOOSE:

> "First of all there is the retracted (and slightly lowered) ... [allophone] ... before a following /__l/, then there is the central [uw] following the liquids /m/ and /r/, and finally there is the sharply fronted ... [allophone] ... in the other positions".

In fact, Webb's (1983) data also includes a *Luther* token i.e. GOOSE after /l/. Figure § 8.1 on the facing page provides a graphic illustration of the relevant allophonic distribution.

Contrary to Wells (1982) and others' claims above, and as in the case of FLEECE (§6.4), Webb (1983:149) claims a (tentative) diphthongal status for his GOOSE tokens, with an average F_2 retraction of 314Hz. His concession, however, that "more data are ... required before reliable deductions can be made" (Webb 1983:149).

Turning to the current data, Figure 8.2 on page 296 places the BATH vowel in the context of the other Part-System D vowels. It is based on 946 measurements across 43 tokens across 27 subjects, at an average of 1.59 tokens per subject. The N-G words were *heart* and *hard* while the O-G token was *palm*. There were 18 tokens of *snarl* for the analysis of BATH before /l/ and, in this regard, we note a slightly retracted value which also shows an expected 'glide' towards the back.

While there were, in effect, no BATH tokens, there is little to suggest that BATH is different to PALM and START in SAE[9]. The choice of mainly START vowels was, perhaps, fortuitous though, because, as is evident from Figure 8.2 on page 296, there is some evidence of an in-glide.

Figure 8.3 on page 297 separates out the three relevant words; the in-gliding is particularly prevalent in *hard*, where the final voiced obstruent conditions a longer vowel. Note that *palm* is also, in general, somewhat separate from the two START tokens. Curiously enough, and as reported on above, Webb (1983) also provides slight evidence for a distinction between these two categories i.e. START vs. BATH/PALM; although for Webb (1983) the distinction has to do with the average value of the nucleus, not any evidence for an in-glide.

It seems clear, however, that the 'ingliding' could equally well be explained by the fact that both *hard* and *heart* end in an alveolar consonant. As confirmed by Stevens (1998:357),

[9]Except for the (very) occasional Americanism such as in the use of [dæːns] for *dance*.

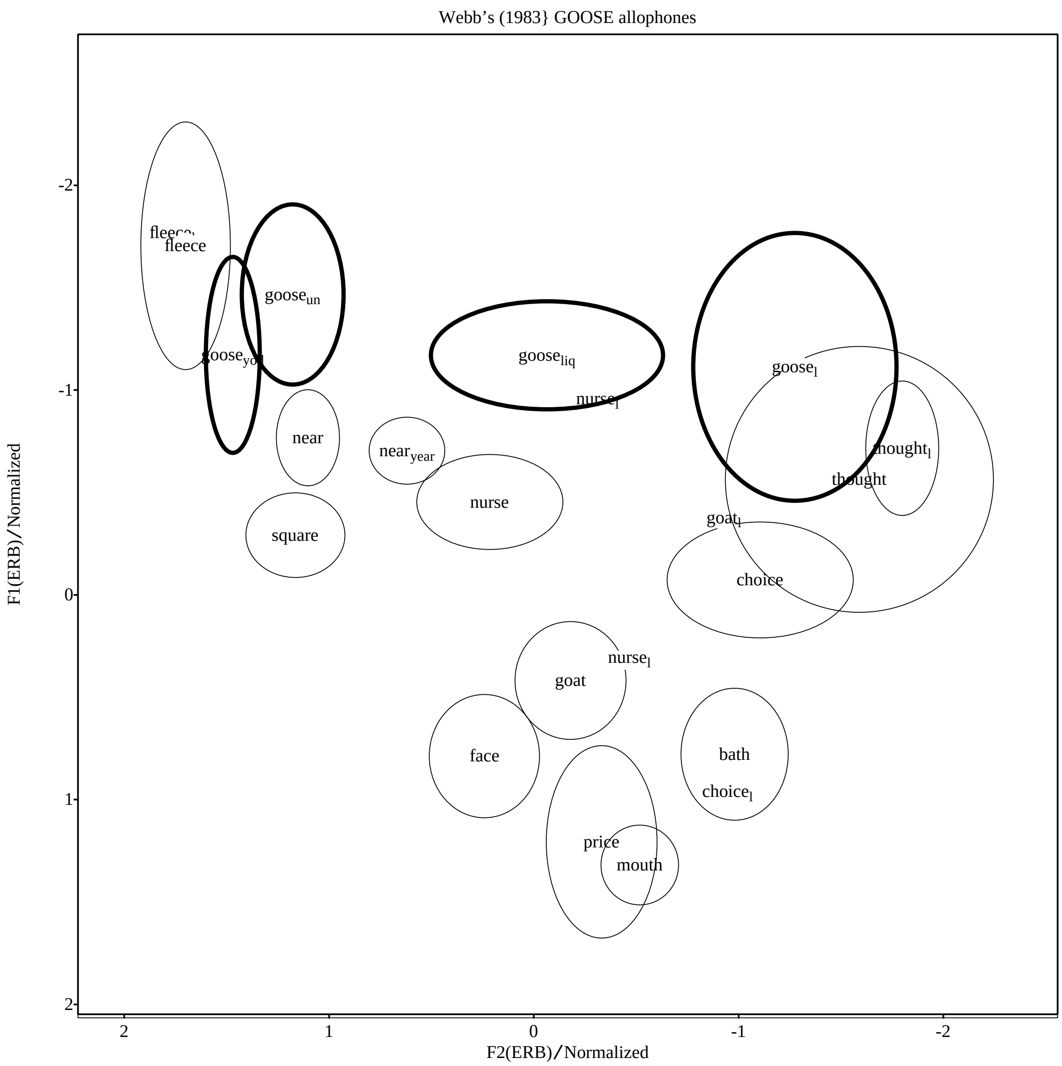


Figure 8.1: Webb’s (1983) GOOSE allophones

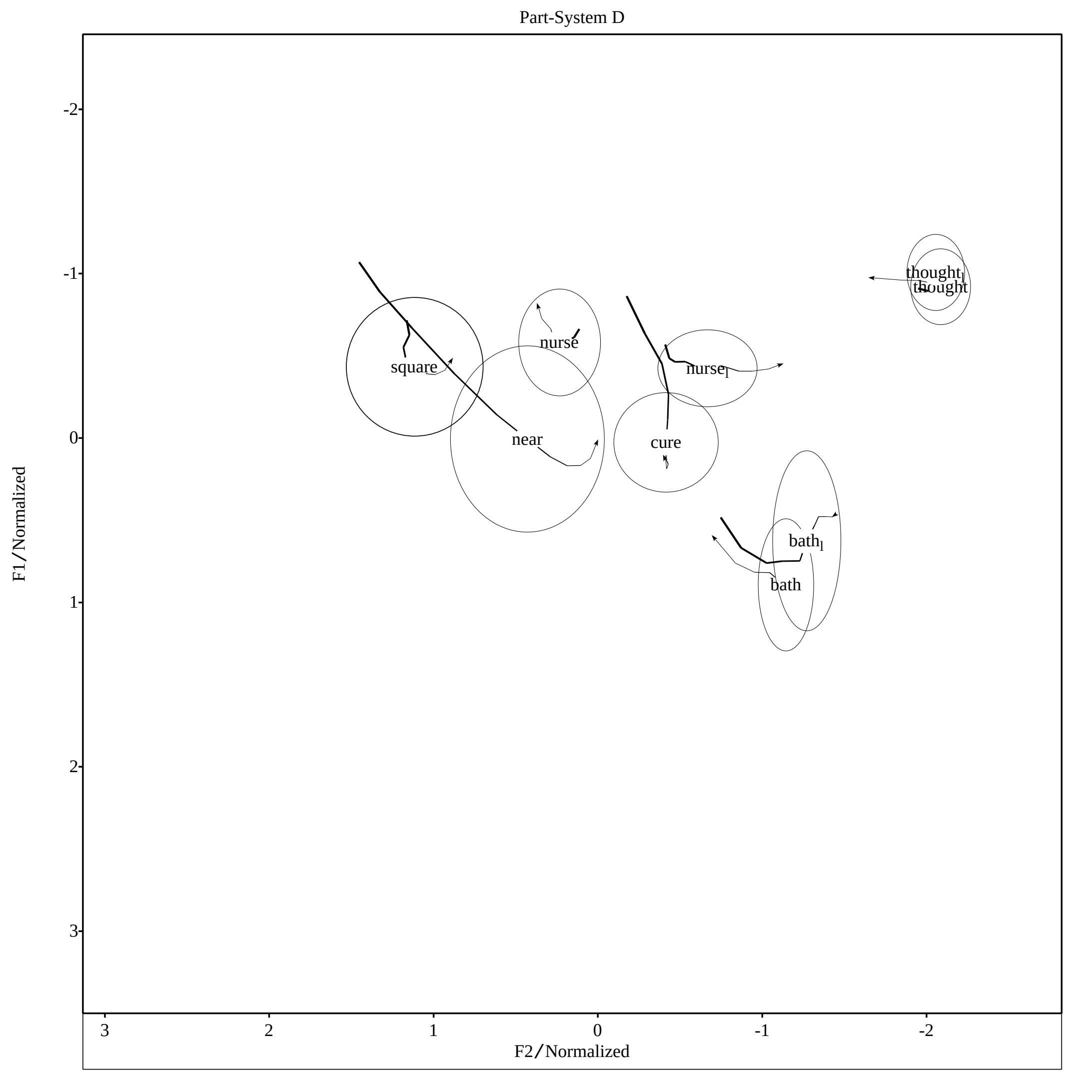


Figure 8.2: Overall Results: Part-System D (27 subjects)

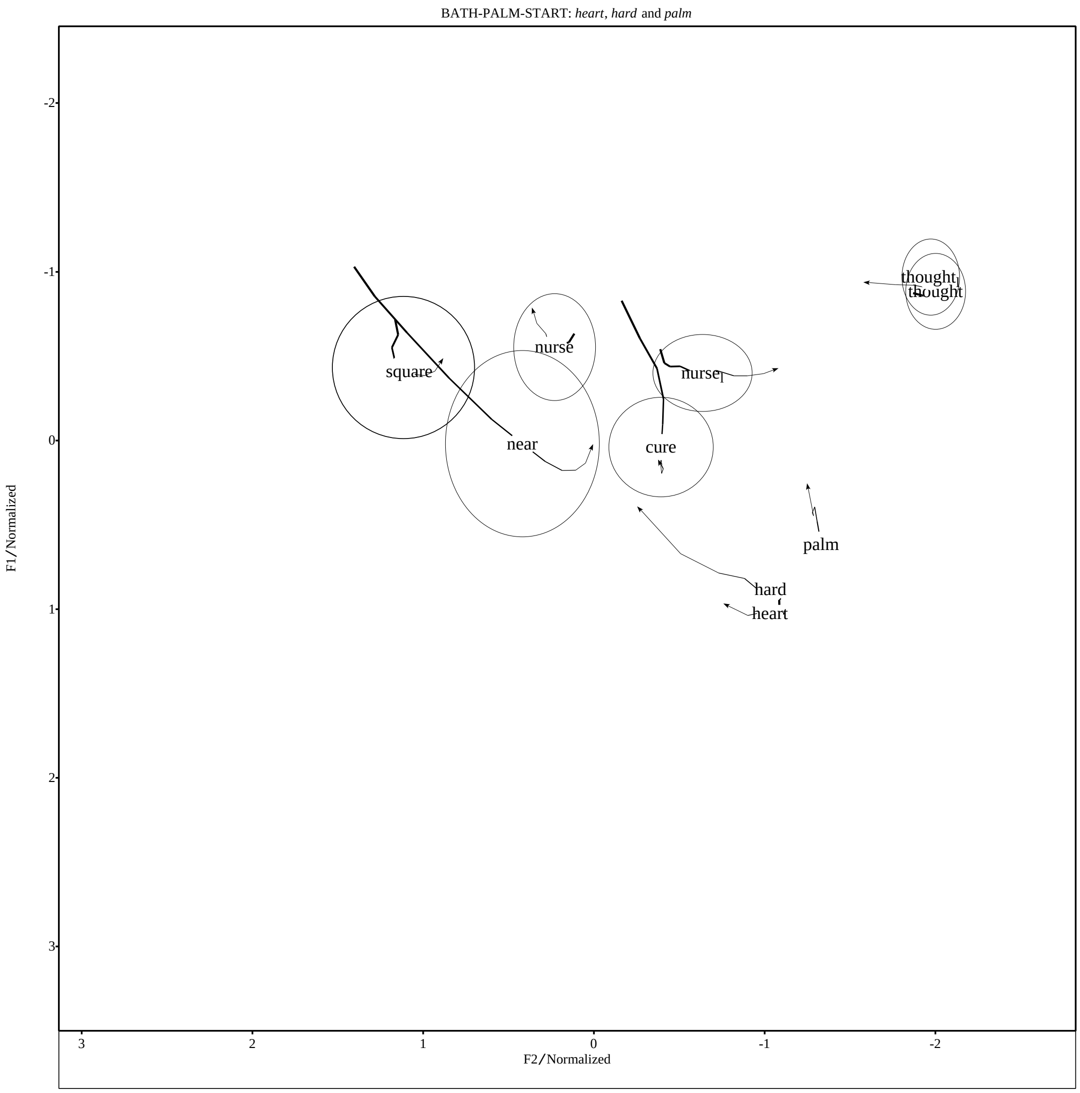


Figure 8.3: BATH-PALM-START: *heart*, *hard* and *palm*

one of the characteristic features of a post-vocalic alveolar stop consonant is that it conditions a rise of F_2 and a fall in F_1. What is a little more difficult to explain in terms of formant transitions, however, is the greater degree of ingliding in the case of a following /d/. Since percentages were used to segment the vowels for formant analysis (i.e. the first 15% and the last 15% were excluded from analysis in the case of long monophthongs and diphthongs) the *relative* length of the area measured for the two sub-types (before /t/ and before /d/) are equivalent. The *absolute* amount of time excluded, however, would be different given the differential length of vowels before voiced as opposed to voiceless consonants. As such, one might expect less evidence of coarticulatory effect in the context of a following voiced consonant. The opposite is in fact the case, and can perhaps be better explained by positing a degree of vowel-internal inglide i.e. the greater length of the vowel before a voiced consonant allows more time for the in-glide target to be reached.

As can be seen in Figure 8.4 on the facing page, there is little to differentiate the two data sets: i.e. the current data and that of Webb (1983); neither is the distinction (in terms of nucleus position) between PALM-BATH ('ah' in Figure 8.4) and START ('ahr') in Webb's (1983) data particularly convincing.

Figure 8.5 on page 300 compares the SAE data with that of other accents of English. One notes, first of all, the expected fronting of the Australasian data. The SAE and RP data seem to be close in terms of quality, although, perhaps surprisingly, the latter (in citation-form) has a slightly more retracted value.

Figure 8.2 on page 296 also places THOUGHT in the context of the other Part-System D vowels. It is based on 1848 measurements across 84 tokens across 27 subjects, at an average of 3.11 tokens per subject. The N-G words are *hoard* and *bought* while the O-G words were *pour*, *paw* and *bought*. Eighteen (18) tokens of *haul* were analyzed for THOUGHT before /l/, but, as can be seen from Figure 8.2, there is not much to report. For THOUGHT more generally, there is very little variation with respect to this vowel and it is in a relatively close position; more intriguingly there is *slight* evidence for in-gliding spectral movement.

Figure 8.6 on page 301 compares the two SAE data-sets. As can be seen Webb's (1983) data shows an uncomfortable degree of variability and a fronter, slightly lower position.

As discussed above, Webb (1983) did advance the possibility of a difference in value between THOUGHT on the one hand and NORTH/FORCE on the other. This putative distinction is represented in Figure 8.7 on page 302, with THOUGHT being labeled 'webboh' and NORTH/FORCE 'webbohr' respectively. While there is some evidence of NORTH/FORCE having a slightly more retracted value, neither of the two values overlap the THOUGHT-NORTH-FORCE values for the current data. Secondly, in both of Webb's (1983) categories, there is still an uncomfortable degree of variance. There are thus two aspects of this situa-

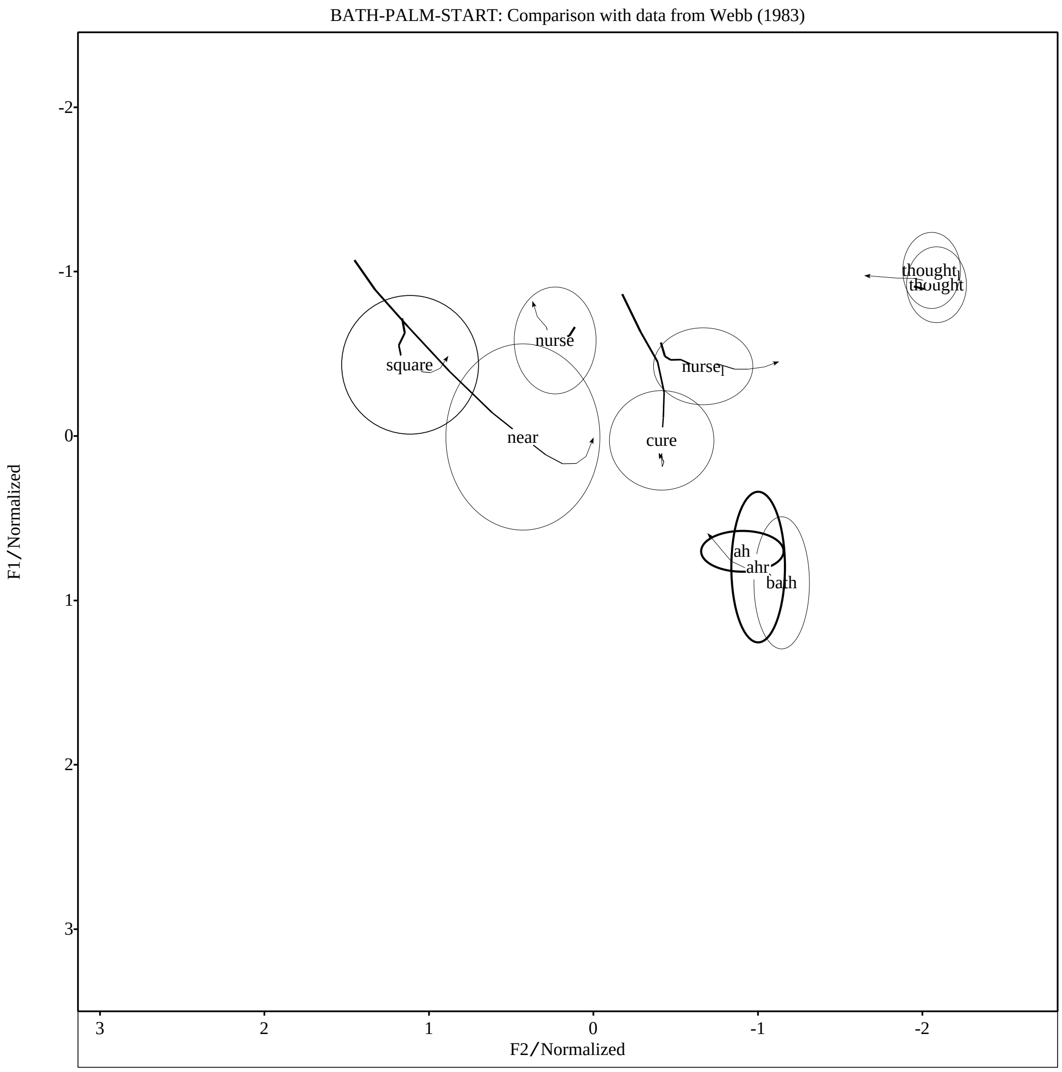


Figure 8.4: START-BATH-PALM: Comparison with data from Webb (1983)

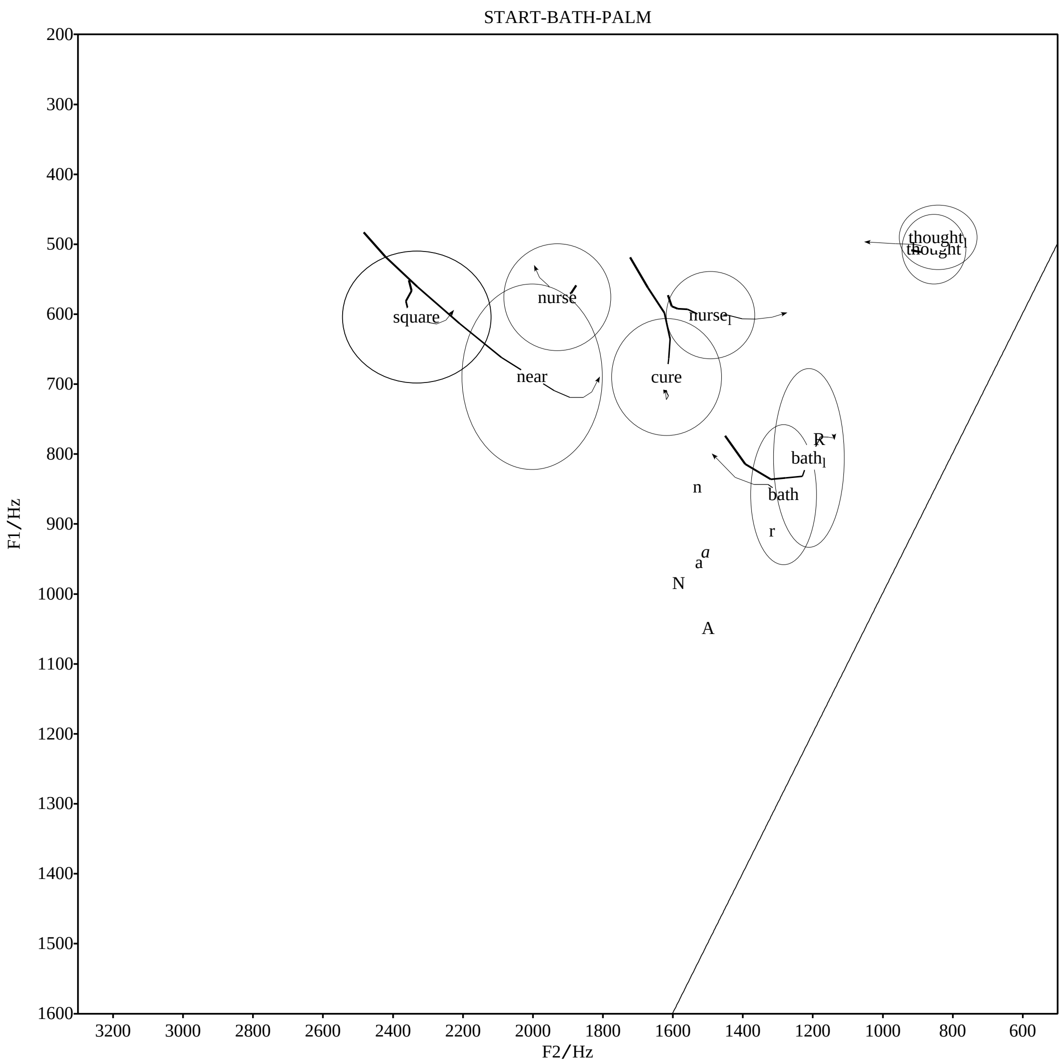


Figure 8.5: SAE START-BATH-PALM vs. other English accents

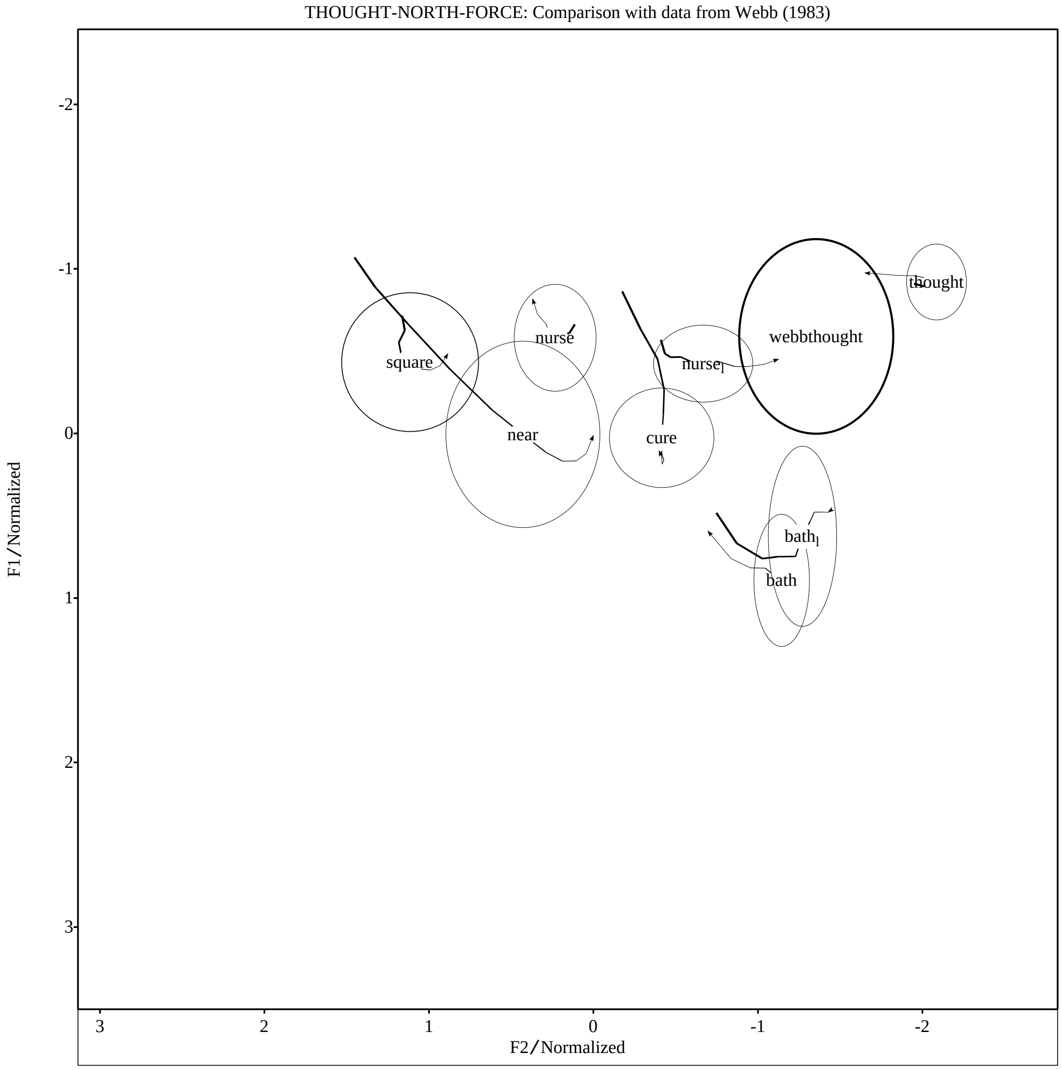


Figure 8.6: THOUGHT-NORTH-FORCE: Comparison with data from Webb (1983)

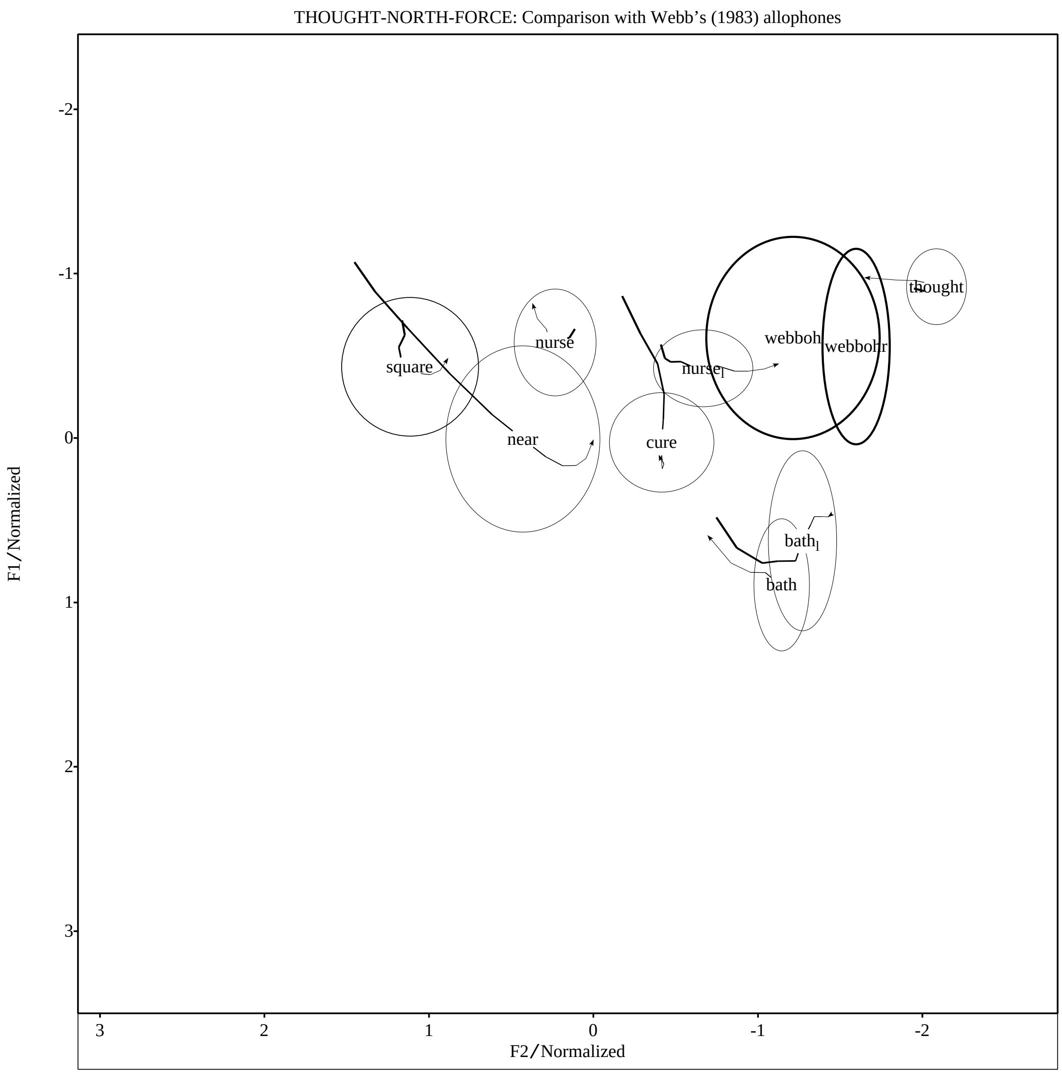


Figure 8.7: THOUGHT-NORTH-FORCE: Comparison with Webb's (1983) allophones

tion that require an explanation; firstly, the high degree of variance in Webb's (1983) data and the difference in quality between Webb's (1983) and the current data.

One possible, and intuitively plausible, explanation which unites these two facts, is that SAEP has undergone (and recently completed) a sound change involving the raising of THOUGHT etc. along the back peripheral track. On this analysis, the first tokens to be affected by this raising were the NORTH and FORCE ones, with the THOUGHT tokens following shortly. This would account for the high degree of variance and, of course, for the differences in value. That the sound-change has completed itself is evident in the lack of variation in the current data. On this account, a raised /oː/ vowel is not simply a relic of the input to SAE, but has risen since the establishment of SAEP as a distinct accent, an analysis which does receive support in the impressionistic literature. However, while the fact of a raised THOUGHT-NORTH-FORCE vowel in SAEP has been confirmed by most commentators in the impressionistic literature, its being variable across sociolinguistic categories (like gender or style) has received little confirmation. Although Lass (1990:278) does mention the possibility of a lower value for some speakers, a lower value for a male using connected-speech strikes one as unlikely, given the conservative connotations of the lowered variants. If THOUGHT-NORTH-FORCE has indeed undergone further raising recently, it has escaped the notice of impressionistic judges. It would, however, fit snugly into a Pattern-3 movement for SAE.

Figure 8.8 on the next page compares SAE THOUGHT-NORTH-FORCE with the same vowel in other accents of English. While it is a little surprising, given a reading of the literature reviewed in §8.2.1, to find SAE THOUGHT etc. in virtually the same position as RP, AusE and NZE, there are enough hints to suggest that all of these varieties are under a structure pressure, in accordance with Labov's (1994) Pattern 3 chain-shift, to raise this vowel.

Moving on to GOOSE, Figure 6.1 on page 172 provides the broader context of this vowel for the current data. The graphical representation is based on 1672 measurements across 76 tokens across 27 subjects at an average of 2.81 tokens per subject. The N-G words were *who'd*, *boot* and *food*, while the O-G tokens were *boot*, ***woun**ded* and *cute*. Seventeen (17) tokens of *fool* were analyzed for GOOSE before /l/ and the dramatic disassociation of this allophone from the main 'group' is obvious enough in Figure 6.1 on page 172. In this figure we also note the expected (relatively) fronted position of GOOSE in SAE. There is also clear evidence of spectral movement (i.e. from back to front) although this is, in all likelihood, due to the presence of a post-vocalic coronal context in all of the tokens (Oh 2008:363).

Figure 8.9 on page 305 places GOOSE in the context of the Part-System B vowels, in-

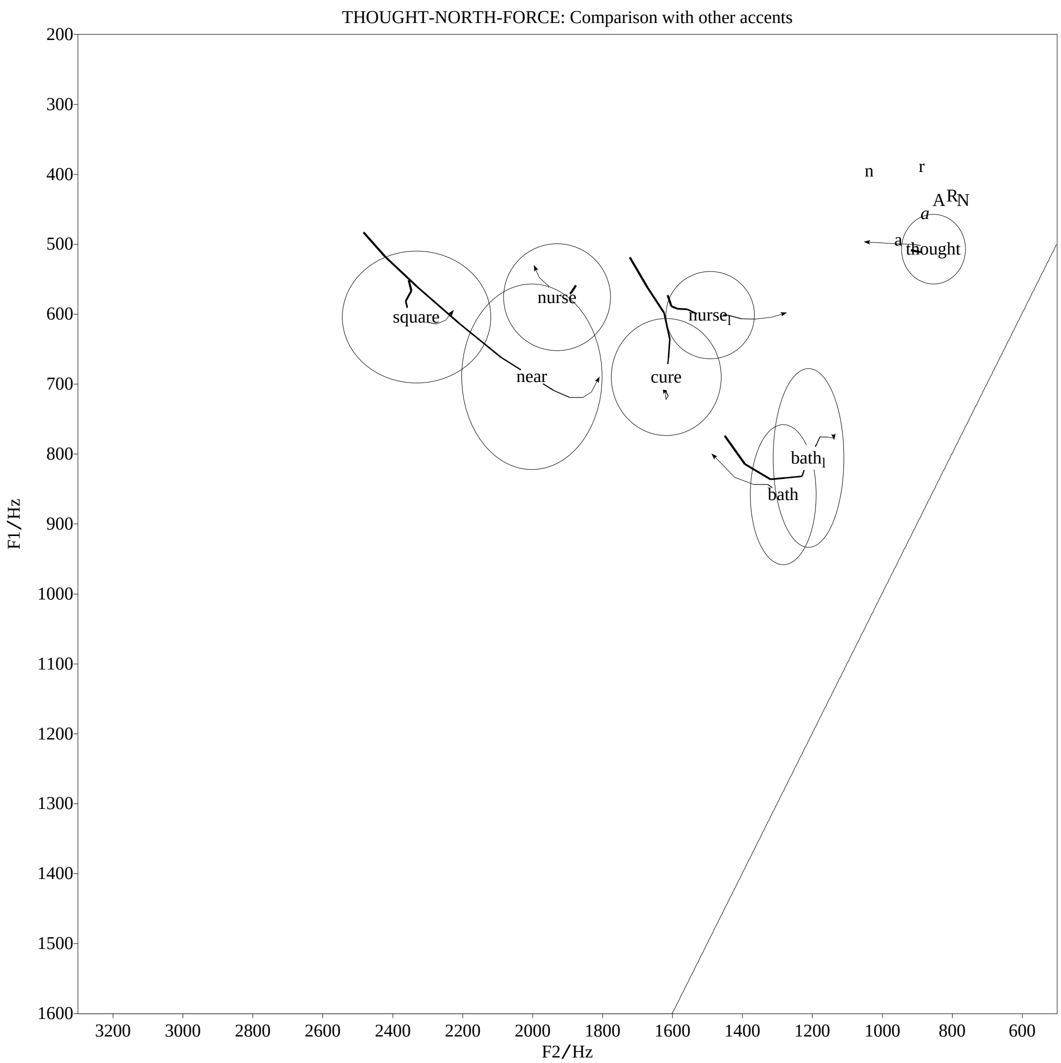


Figure 8.8: SAE THOUGHT-NORTH-FORCE vs. other English accents

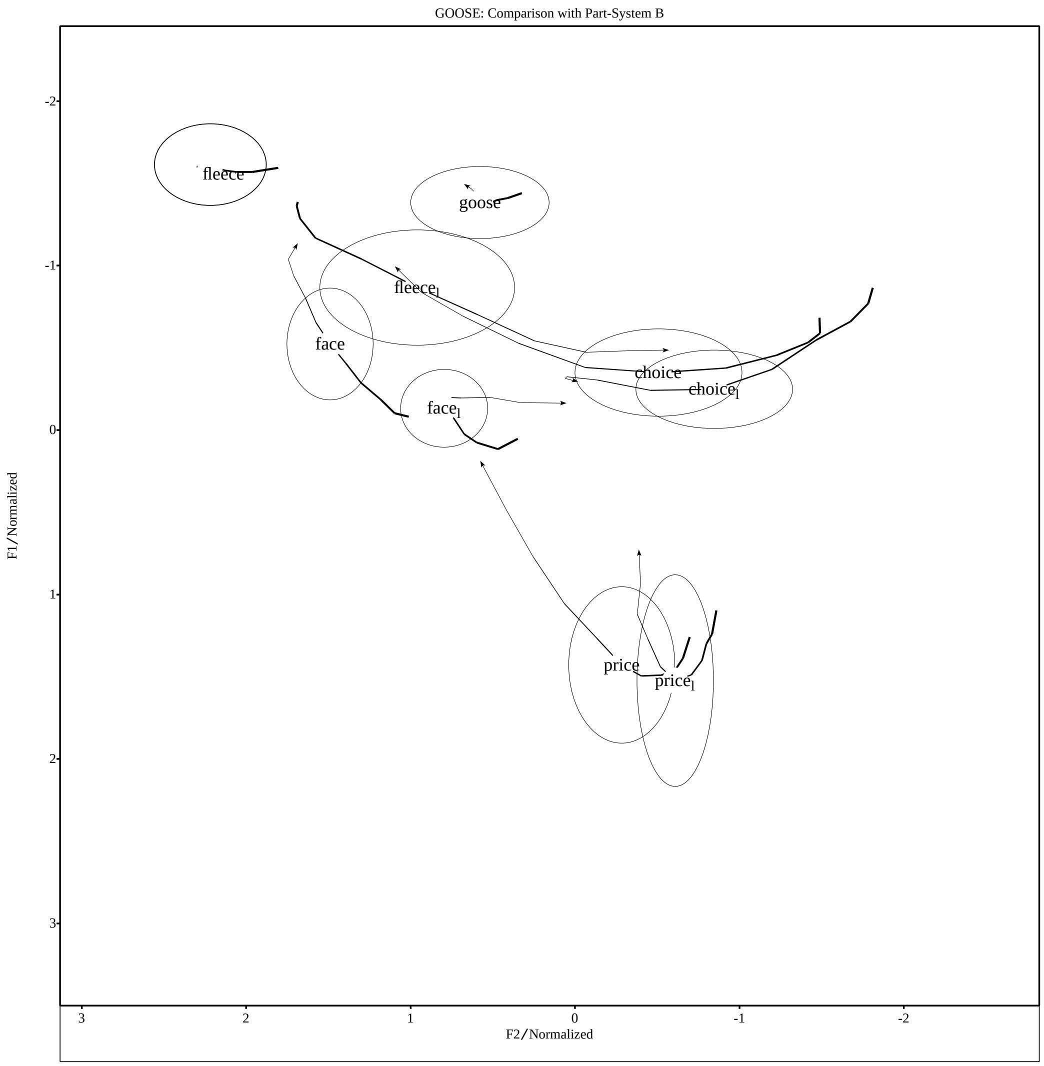


Figure 8.9: GOOSE: Comparison with Part-System B

cluding FLEECE. As can be seen there is still some distance between GOOSE and FLEECE, whether due to lip or tongue configuration (or both); it is also perhaps due to the low number of tokens with a preceding /j/ (see below) and the absence of tokens with a preceding coronal.

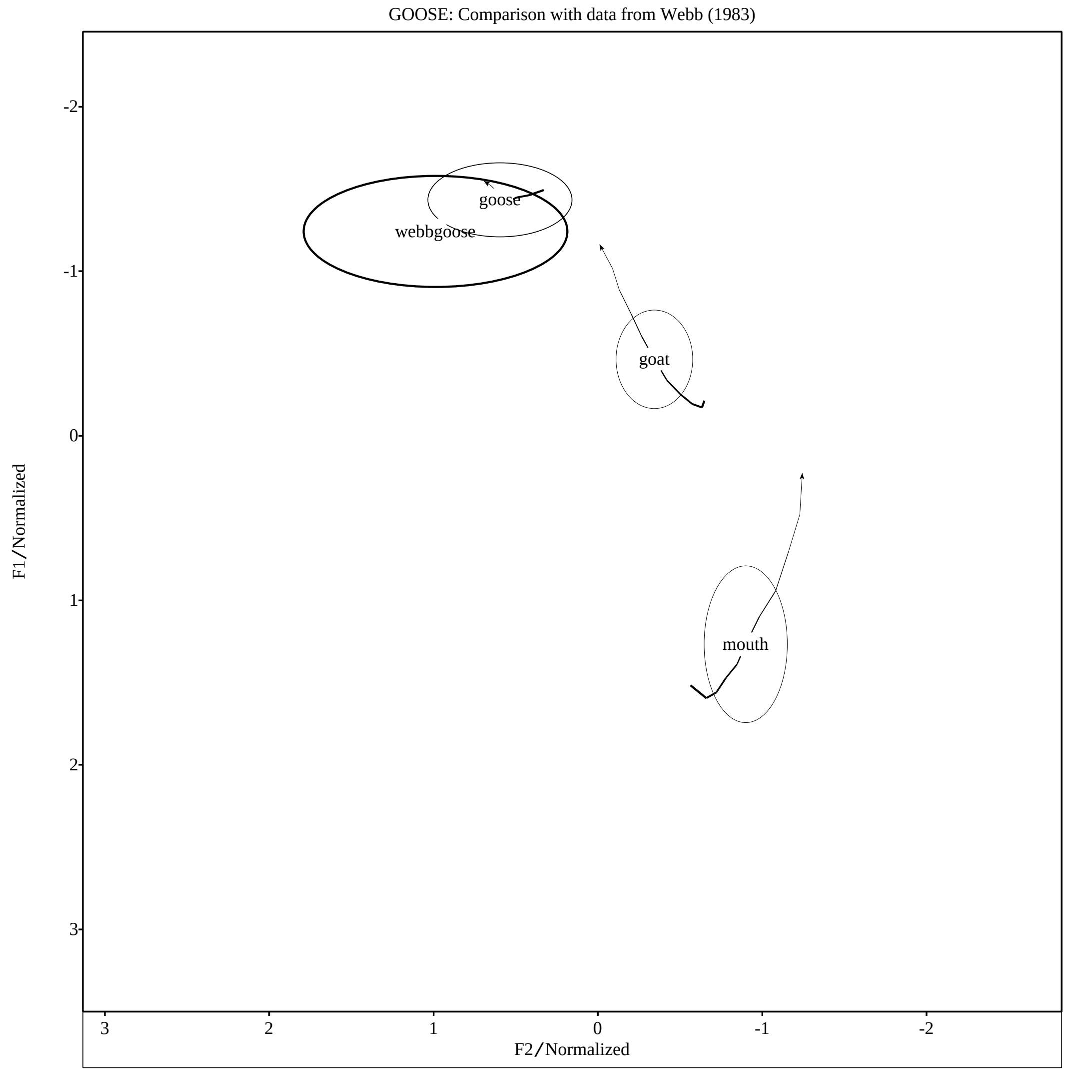


Figure 8.10: GOOSE: Comparison with data from Webb (1983)

Figure 8.10 provides a comparison with Webb's (1983) GOOSE data i.e. including GOOSE after yod, but excluding GOOSE before final /l/. As can be seen there is a fair degree of overlap; the somewhat more fronted value (and the higher standard deviation) for Webb's (1983) GOOSE is more than likely the result of a greater proportion of GOOSE-after-yod tokens: 6 out of 19 (31.5%) compared with 9 out of 76 (11.8%).

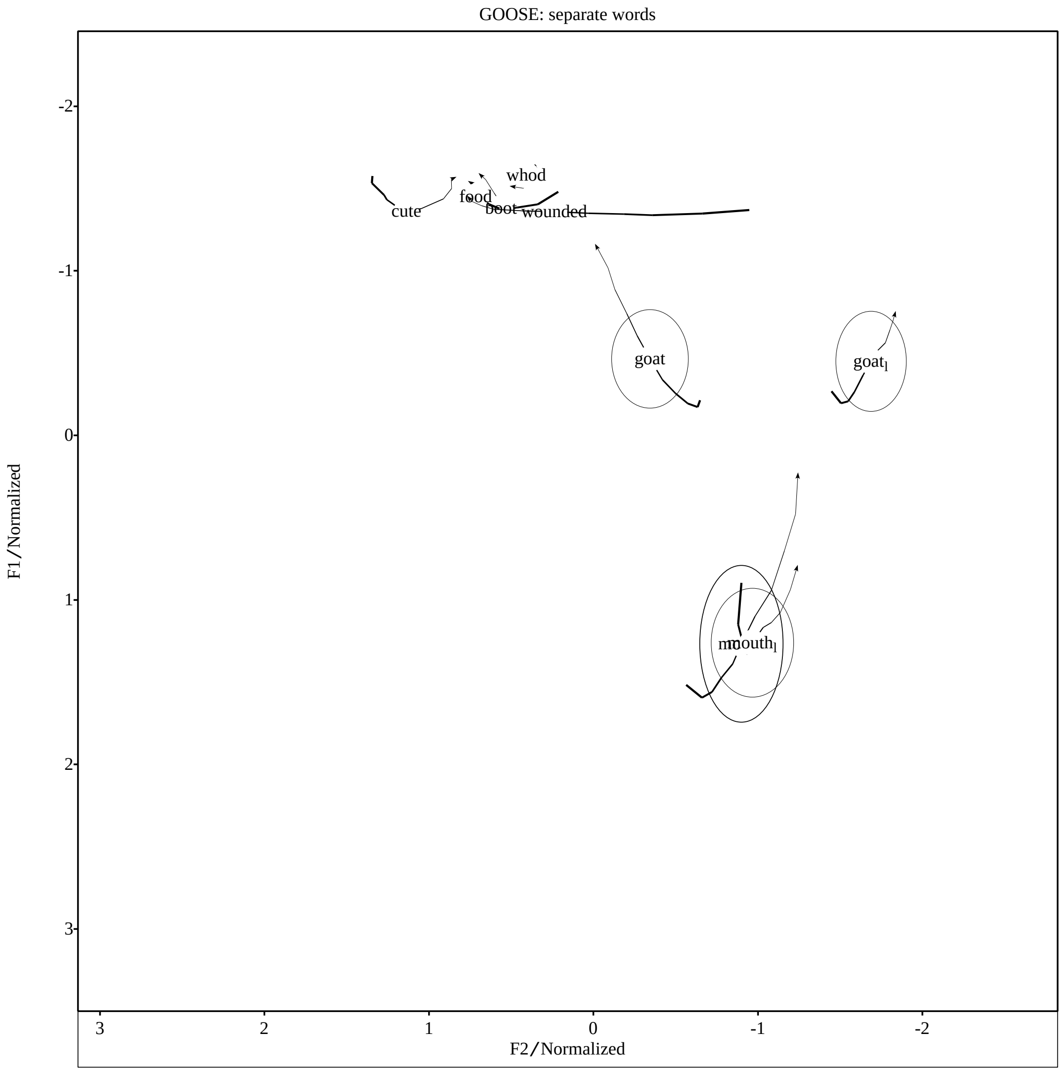


Figure 8.11: GOOSE: Separate words

As depicted in Figure 8.1 on page 295, Webb (1983), in fact, gives four separate allophones of GOOSE: after /j/, after liquids, an unmarked allophone and GOOSE before final /l/. Leaving the later aside for the moment, Figure 8.11 tracks the movement of all 5 words in this study. The effect of an initial 'yod' as well as a preceding /w/ can be clearly seen. While there is a degree of correspondence between this figure and Webb's (1983) figure, the tendency is for all of these vowels to 'target or aim at' a more central unmarked position, defined in this figure by the words *who'd*, *foot* and *boot*. Figure 8.1 on page 295, while correct in its own way, tends to obscure the fact that the various allophones of GOOSE still appear to be targeting the same value. It is unlikely, therefore, at least for the current subjects, and most probably for Webb as well, that the various allophones of GOOSE have been phonologized in the sense of establishing different targets in different positions; except for GOOSE before final /l/, all of the allophonic variation can be explained in terms of coarticulatory effects. Moreover, the fact that there is not a marked difference between the current data and that of Webb (1983), also seems to suggest that, if anything, GOOSE-fronting is an indicator of SAEP i.e. it is largely below the level of consciousness and not subject to much style-shifting. This, in its turn, suggests that the change is an endogenous, system-internal one, and it is, in fact, an expected movement in terms of Labov's (1994) Pattern-3 and the Southern Shift more generally.

Figure 8.12 on the next page compares the SAE GOOSE data with the same from other accents of English. As can be clearly seen from this figure, the various Southern Hemisphere Englishes are far more advanced than RP in the fronting of the GOOSE vowel, with, it appears, AusE and SAE 'leading the way'. It is interesting to note that, for all intents and purposes, the current data for GOOSE is identical with the AusE data, particularly with that of Cox (2006), who focussed on female teenagers, all speakers of Standard (General) AusE. There is some evidence, however, that unlike in SAE, where a fronted GOOSE has become associated (at least unconsciously) with prestige (§8.2.2), that this same feature is more characteristic of Broad accents of AusE (Harrington et al. 1997:163).

With respect to the RP data, it needs to be added, however, that the values displayed in Figure 8.12 on the facing page are based on recordings reported on in 1990 (Deterding 1997) at which time the average age of the female speakers was 36 (range 22 to 49) (Deterding n.d.). By 2005, then, the range would have been in the region of 37 to 64, with an average age of 51 years. While it is not clear when Harrington et al. (2008) did their recordings, it is clear that it was not before 2005. Their younger (RP-speaking) subjects were between 18 and 20 at the time, while there older subjects were all older than 50 "with an average age of 69.2 years". The 1990 subjects whose speech is reported on in Deterding (1997) are thus closer in age to Harrington et al.'s (2008) older group than to their younger group. The rapid change that GOOSE has undergone in RP over the last few decades is clearly indicated in Figure 8.13 on page 310, taken from Figure 2 in Harrington et al. (2008:2829).

On a superficial comparison between Figure 8.13 and earlier figures displaying the current SAE data, it would appears that younger speakers of RP have advanced their GOOSE vowel to a position even more fronted than that found in the Southern Hemisphere. This seem unlikely, however, if we take a number of factors into account. Firstly, Harrington et al.'s (2008) figures are based a large percentage (56%) of GOOSE tokens with a preced-

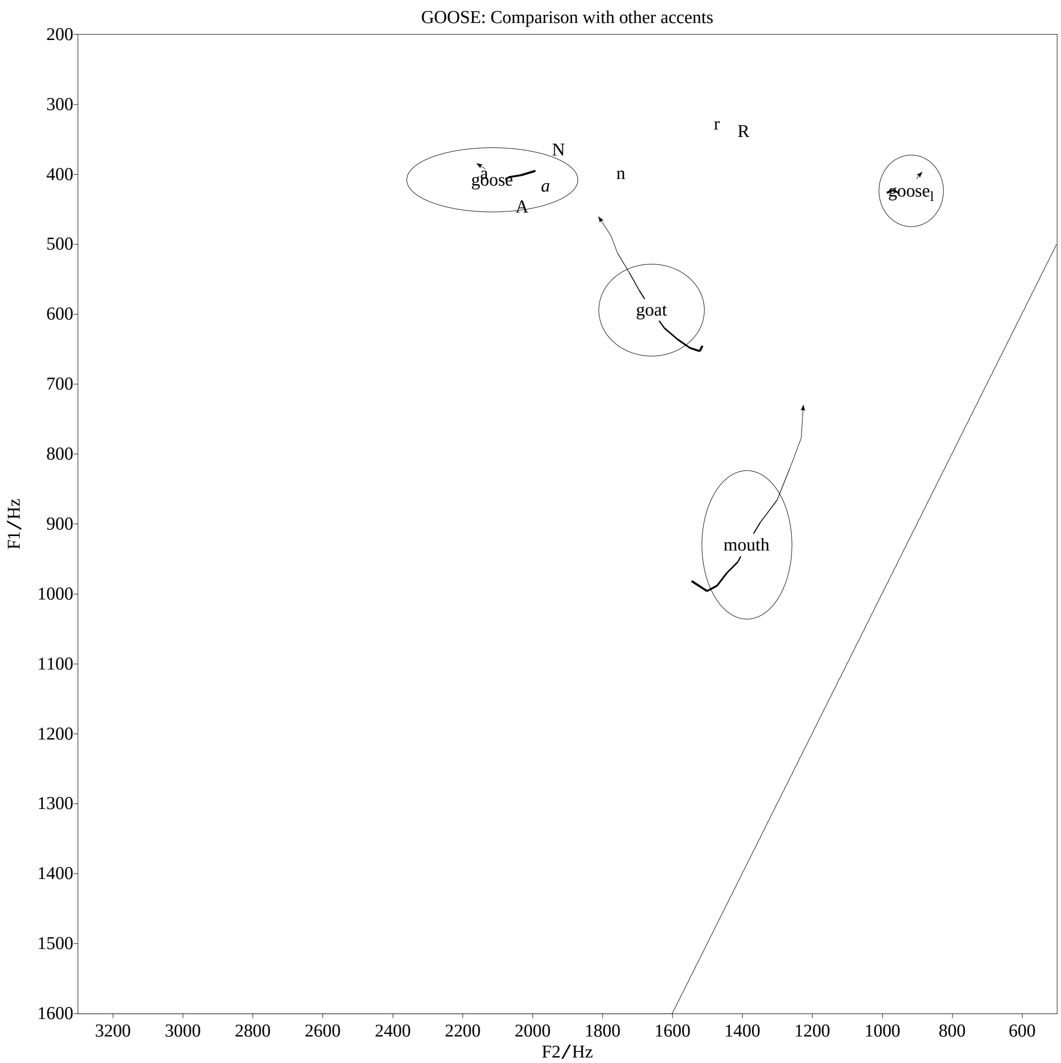


Figure 8.12: SAE GOOSE vs. other English accents

ing yod or coronal context. Secondly, the diagrams in Figure 8.13 have been bark-scaled with the resultant compression of the high F_2 values – see §2.2.3. A comparison of the absolute Hz values for Harrington et al.'s (2008) young female speakers (top axis of the bottom left diagram in Figure 8.13) and the current speakers (see Figure 8.12 on the preceding page) should convince the reader that the two averages are similar. Lastly, Harrington et al. (2008) use 95% confidence ellipses (thus perhaps accounting, along with the prepodarence of post-yod GOOSE tokens, for the overlap between GOOSE and FLEECE) while for the current research standard-deviation ellipses were used (i.e. 50% confidence). Overall, while it is not clear that RP has in any sense overtaken SAE and the other Southern Hemisphere Englishes in terms of GOOSE-fronting, it certainly appears to have caught up over the last 20 years or so.

Another possible difference between the (recently) fronted GOOSE of RP and the front status of GOOSE in SAE (and other Southern Hemisphere accent) relates to the fact that in RP it appears that the fronting is largely the result of "a realignment in production of the phonetically back allophones of /u/ toward a phonetically front position" (Harrington et al. 2008:2830). In the process, the allophony of RP-GOOSE is reduced. In SAE this, as

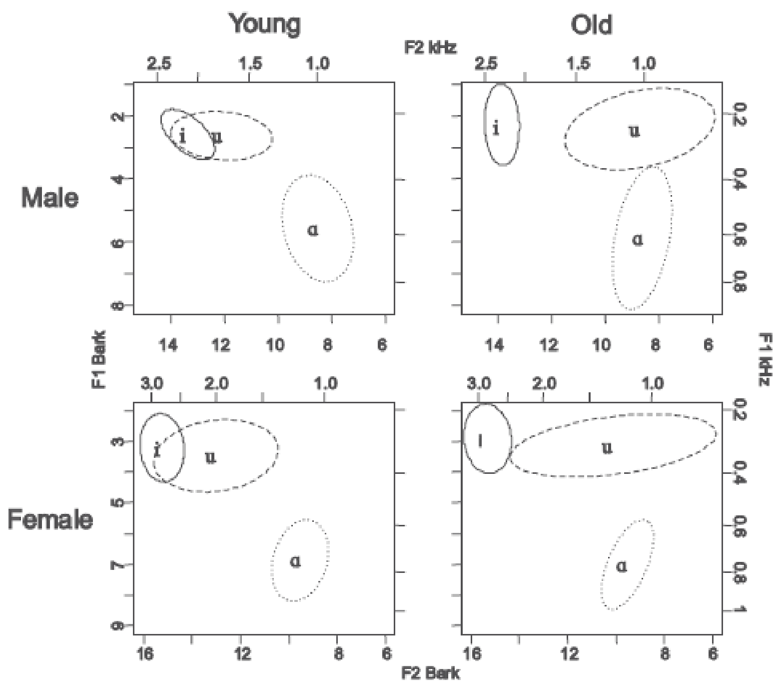


Figure 8.13: GOOSE data from Harrington et al. (2008)

is evident from Figure 6.1 on page 172, does not apply to GOOSE before final /l/. Unfortunately, Harrington et al. (2008) do not provide data for RP GOOSE in this environment so it is unclear whether GOOSE in this environment has remained untouched by the more general fronting in RP or not, although Trudgill (2004:58) makes the impressionistic claim that it has.

8.2.4 BTG: Synopsis

The acoustic analysis of SAE BATH shows the expected retracted value, as clearly indicated by the comparison with NZE and AusE acoustic data, but some unexpected, although tentative, evidence for an in-glide in the case of START. The fact that the current data does not differ substantially from that of Webb (1983) also seems to indicate that this value is near-to-categorical in GenSAE and further Pattern-3-like raising has, in all likelihood, been prevented by the stigmatization of BrSAE values, although one notes that there is some recent indication that this stigmatization is waning.

As far as the diachronic status of BATH is concerned, we have noted in §8.2.2, that some commentators have attempted to attribute the backed SAE BATH to the influence of the southeast of England, presumably through 1820-settler speech. In §3.4.3 we have, though, shown Trudgill (2004) providing relatively convincing evidence that BATH-Backing could not have occurred during the 1820-period, a perspective almost undeniably solidified by the presence of a fronted [aː]-value in AusE (earlier than CE) *and* NZE (later than CE). Trudgill's (2004) conclusion is that the backed variant was either brought in by normative NE-speakers and/or, more likely, and, interestingly-enough, by late 19th-century RP *and* lower-class southeast-of-England speakers to early-Johannesburg[10]. Natal, as well as parts of the *Eastern Cape*, and as *hinted* by Hopwood (1928) in §8.2.2, were, on this analysis, characterized by fronted values. Hopwood's (1928) early obsevations also lead to the conclusion that the source of an even further-backed BATH value in BrSAE was a possible difference between the RP and working-class values brought to Johannesburg i.e. both backed, but the former less so than the other. This would also, of course, account for the indexical difference between these two variants.

More importantly, the development of BATH in SAE provides another example, along with MOUTH (see §6.2.4), of the inadequacy of Trudgill's (2004) model of new-dialect formation in tabula rasa contexts, and one that is implicitly provided by Trudgill (2004) himself. This author in fact concedes that the only way to account for a backed-BATH vowel

[10]Even today, "/ɑː/ ... in START-BATH-PALM ... [is] often much the same in London speech as in RP" (Wells 1982:305; my parenthesis). Still, while, "a fully back variant ... characterizes ... Cockney" (Wells 1982:305), some of the values indicated for RP, in §8.2.1, have a more front value.

in SAE (and its absence in AusE and NZE) is to assume that it was brought in by middle-to-upper class, normatively-minded, north-of-England NE-speakers, as well as through late-19th-century RP and working-class individuals from London and the southeast of England – see §3.4.3[11]. On the other hand, we can assume, following Trudgill (2004) himself, that fronted variants would have been brought in by CE-speakers, the less normatively-minded NE speakers, and working-class individuals from the north-of-England who immigrated to early-Johannesburg (who, as we have seen in §3.3.1, constituted a sizeable presence). Although in the absence of clear historical and demographic data there is no guarantee that one's impressions in this regard are correct[12], it does seem unlikely that speakers evidencing BATH-retraction would have, in any obvious sense, been in a clear majority. Naturally, in cases where there are variants in relatively equal competition one possible outcome, in terms of Trudgill's (2004) model, would be the reallocation of the competing variants to various points on the sociolinguistic spectrum, as was indeed the case with unrounded (backed) BATH (allocated to GenSAE) vs. rounded (backed) BATH (allocated to BrSAE). What still remains to be explained, however, is the complete *absence* of fronted variants in SAE given that these constituted, by all accounts, a sizeable input into the historical process. Trudgill's (2004) account is no doubt historically accurate, but still not convincing as a form of support for his model. On the other hand, if we assume that the adoption of a backed, unrounded variant was driven by prestige factors, then all becomes, I would submit, crystal clear. In defense of this analysis, there are, of course, the added issues that Johannesburg constituted, from early-on, a clearly and highly class-stratified society and the fact that any BATH-retraction brought in by NE speakers was, in its own right, a prestige-drive phenomenon. All-in-all, then, it would appear, to me at least, that the default interpretation would be that the adoption of BATH-retraction was a prestige-driven phenomenon. Fronted variants were disregarded altogether and an extremely-backed variant was rejected in favour of what was an essentially RP-based variant.

With respect to THOUGHT etc. in SAE, the acoustic evidence provides some support for the notion that this vowel has undergone endogenous Pattern-3-like raising in GenSAE, along with many of the other 'Southern' accents, although this possibility certainly requires further research. The 'source' of SAE seems uncontroversially an 1820-input, followed by

[11] It should be noted here again that 'backed' BATH came in two varieties: the backed, but unrounded RP-variant and the backed, rounded and raised Cockney variant. In terms of the latter, it appears that the potential role played by Afrikaans can be discounted. The early Hopwood (1928), for one, claims that the extremely tensed, raised and rounded, [ɔ]-variant, so stereotypical of BrSAE, is derived from Cockney. This author also alludes to [a]-like variants in Natal and at least parts of the Eastern Cape, stressing that the Cockney-derived back BATH is a *late*-19th-century and *not* an early 19th one. Lanham (1978) also claims a 'Home Counties' origin for the BrSAE value.

[12] A possible focus for future research – see §10.3.

later inputs from RP and the southeast of England (via early-Johannesburg) followed, in turn, by an endogenous raising of this vowel.

In the case of GOOSE the acoustic data shows the expected fronted variant alongside a sharply retracted variant before final /l/. The Southern Hemisphere Englishes appear to have been at the leading edge of a Pattern-3 like endogenous fronting of this vowel, a case of Trudgill's (2004) 'drift', although it appears that RP has 'caught up' more recently. The 'source' of SAE GOOSE appears to be similar to that of THOUGHT – input from a variety of southern English accents, during the 1820 settlement as well during the establishment of Johannesburg, followed by endogenous development. There are, however, and as far as I can see, no obvious solutions to the monophthongal status of GOOSE in SAE, in comparison to their diphthongal status in the other two main Southern Hemisphere Englishes. More generally, however, one might wish to link the absence of GOOSE-Diphthongization with the more generally reported resistance of SAE to the Diphthong Shift.

8.3 GOAT and CURE (GC)

Wells (1982:146) defines the lexical set GOAT "as comprising those words whose citation form has the stressed vowel /əʊ/ in RP and /o/ in GenAm". The author gives the phonetic value of the related vowel as being [o ∼ oʊ] in the latter reference accent. According to Labov (1994:208), GOAT is a "result of the 17th-century diphthongization of /iː, eː, uː, oː/ to produce /iy, ey, uw, ow/". As outlined in §2.3.2.1, this author stresses that neither GOAT-Fronting nor GOAT-Lowering participate directly in any of the major chain-shift patterns. In either case the phenomenon represents a generalization of GOOSE-fronting (Pattern 3) or FACE-lowering (Pattern 4) respectively. Interestingly, Palethorpe and Cox (2003) mention that historically, the first target of GOAT was closely related to /ɒ/ i.e. LOT; in many accents it would appear to retain this value in pre-/l/ contexts.

Wells' (1982) definition of CURE is somewhat unusual in the sense that the two reference accents in question are GenAm and "conservative RP" (Wells 1982:162), as opposed to mainstream RP, in which there is a growing tendency to merge CURE with NORTH-FORCE e.g. *sure* [ʃʊə] becomes [ʃɔː]; what the author refers to as "the Second FORCE Merger" (Wells 1982:237).

Leaving this issue aside for the moment, the lexical set can thus be defined in the following fashion:

> "As comprising those words whose citation form contains the stressed vowel /ʊə/ in conservative RP and the sequence /ʊr/ in GenAm . . . Phonetically, RP /ʊə/ is a

centring diphthong with a starting point that is weakly rounded, somewhat close and back, [ʊ], moving towards a mid central [ə] quality. In the varieties of RP where, as is now usual, there is no contrastive /ɔə/, the starting-point may alternatively be rather opener, namely about mid [o̞+] {[o̞˖]}" (Wells 1982:162–3; my parentheses).

By "the contrastive /ɔə/" Wells (1982:162–3) is referring to the FORCE vowel, which until recently had [ɔə] as its quality; the merger of FORCE with NORTH and THOUGHT is referred to by Wells (1982:235) as the "First FORCE Merger", as opposed to the Second FORCE Merger which, as mentioned above, involves the merger of CURE with /ɔː/.

8.3.1 GC: Other Accents of English

Received Pronunciation: As far as RP is concerned, Lanham (1967:39; 41–2) gives [əʊ] for GOAT and implies an older version of RP with a [o] (American-like) quality to the onset. Similarly, Wells (1982:293–4) mentions a degree of variability with respect to this vowel, both chronologically and across different sub-types of RP. Thus an older pre-World War I version has [ö̞ʊ] which has given "way to one with a central and usually unrounded starting-point, [ɜʊ ∼ ɵʊ]" (Wells 1982:293). There is often very weak rounding of the second element of the diphthong and occasionally the vowel is monophthongized. Wells (1982:294) also draws a distinction between the above values and a fronter, typically Refined RP, onset to the diphthong, [ë̞ ∼ ɛ̞̈], which is "now widely considered 'affected', and has ceased to be fashionable among younger speakers". Cruttenden (2001:135) generally confirms the above analysis and adds that a glide-weakened GOAT vowel is particularly common in the context of a following dark-l ([ɫ]) in the process leading to a potential neutralization of the distinction between GOAT and NURSE in this particular context.

For RP CURE , Lanham (1967:43) gives "rounded [ʊ][26]{ʊ}. Glide ends at [ə][44]{ə}" and also mentions the tendency for it to be replaced by /ɔː/. Cruttenden (2001:145) mentions the possibility of a lowered CURE, [ɔə], for those RP speakers who have undergone the First FORCE Merger (i.e. FORCE is /ɔː/ and no longer /ɔə/). "This glide [ɔə] may in turn be levelled with the realization of /ɔː/. So *Shaw*, *sure*, *shore*, still pronounced by some /ʃɔː, ʃʊə, ʃɔə/, are levelled by many others to /ʃɔː/ for all three words" (Cruttenden 2001:145). In general, the change from /ʊə/ to /ɔː/ appears to be a change-in-progress, with certain environments favouring the change more than others and with substantial variety between individuals (and, no doubt, styles) in terms of the realization of particular lexical items.

London (Cockney): Cockney GOAT has, by contrast, a generally opener and fronter quality for the onset, with an unrounded [ɤ̈] being the most common value for the glide. Wells (1982:308–9) gives [æ̈ ∼ ɐ ∼ a] as possible values for the onset. Other possibilities are worth quoting in full:

> "There is also a monophthongal variety, a frontish [ʌː]{[æ̈ː ∼ ɐː]}[13], reflected by eye-dialect spellings such as 'nah' or 'nuh' *no*. And there is also a diphthongal variant gliding to a front rounded endpoint, [ʌø] or [ɶ̈ø]; this has a flavour of 'refined' Cockney . . . used by girls only. In non-prominent environments a monophthongal [ɶ̈] is sometimes encountered" (Wells 1982:309; my parenthesis and footnote).

Of interest is to the so-called 'GOAT split' in Cockney, as described in Wells (1982:312–3). Without entering into the technicalities, this development has essentially entailed the creation of a phonemic split from what was originally an allophonic relationship between the GOAT vowel before [ɫ] and the GOAT vowel elsewhere, the former having a [ɒʊ]-quality, the latter having the values already provided above. L-vocalization, a common feature of Cockney English, involves the vocalization of syllable-final [l] such that, for example, *fill* becomes [fɪo] and *people* becomes [pɪipo] (Wells 1982:313). This process, among other things, has created, with respect to GOAT, minimal "pairs such as *sole–soul* [sɒʊ] vs. *so–sew* [sʌʊ]{[sɐʊ]}" (Wells 1982:313).

The tendency for CURE to merge with THOUGHT also appears to be advanced in Cockney (Wells 1982:305), although in this case the merger is with /oː/ *or* /oə/ as a result of the THOUGHT-split mentioned in §8.2.

Other accents of England: It appears that Newcastle is in the process of adopting "a fronted variant, [ɵː] for GOAT" (Torgersen and Kerswill 2004:25). This is a newer (supralocal) variant in contrast with the more local, traditional form: [ʊə]. From Labov's (1994) summary of Trudgill's work in Norwich it is also evident that (at least part of) GOAT has fronted in this area of England as well. Overall, according to Torgersen and Kerswill (2004:28), "GOAT fronting is . . . a pattern that can be seen in many varieties in England, having a fronted monophthongal . . . or diphthongal . . . realization, regardless of the presence or otherwise of a vowel shift". The somewhat tenuous nature of GOAT fronting (and lowering) in terms of the Pattern-3 chain-shift has already

[13] Earlier on Wells (1982:308) states that "The vowel in Cockney GOAT typically starts in the area of London STRUT, [æ̈ ∼ ɐ], which we write [ʌ] for convenience".

been noted in §2.3.2.1. As with FACE (§6.4), it should also be noted that Trudgill (2004:52–5) provides convincing evidence to suggest that Diphthong-Shifted GOAT (and even GOAT subject to Long Mid Diphthongization[14]) was still a rarity in south-east regional dialects of 19th-century Britain, as well as elsewhere, especially the north of England.

For CURE, Wells (1982:361) claims that the north-of-England "exhibits a complicated patchwork of [uːə, ɪuə, uə, ʊə, oə, ɔːə]".

AusE and NZE: According to Cruttenden (2001:136) AusE has similar values to Cockney for the GOAT vowel, "although the monophthongal variant does not seem to occur". Wells (1982:597) generally confirms this, giving the following values for the three sociolects of AusE:

Cultivated AusE: [ö̜ʊ]{[ö̞ʊ]}

General AusE: [ʌʉ]{ɐ̟ʉ}[15];

Broad AusE: [ʌːʊ ∼ a-ːʉ]{[ɐ̟ːʊ ∼ a̠ːʉ]}

Of further interest is the following comment from Wells (1982:598):

> "Mitchell & Delbridge (1965: 84) report a 'curiously variable' glide in their discussion of GOAT in that subtype of Cultivated Australian, perceived as 'affected', which they label 'Modified Australian'. In South Australia it was found to range from [ɛ̈ʊ] to [ɛ̈ÿ] and from [ɔʊ] to [ɒy]. They claim that a glide starting [ɒ] is very typical of this Modified accent".

Harrington et al. (1997:174) provide acoustic evidence that a fronted first target for GOAT appears to be associated with Broader accents. The position of the second target in AusE GOAT is a little unclear but it is to be found somewhere between a back FOOT and a fronted GOOSE.

NZE appears to have a similar value to both Cockney and AusE, with Wells (1982:609) giving [ʌʊ] as the central value.

In AusE and NZE the Second FORCE Merger appears to be well underway (Wells 1982:600; 609), with the Broader accents showing a greater degree of merger than

[14] In other words the diphthongization of [oː] to [oʊ] which preceeds Diphthong-Shifting to values such as [ɐʊ].

[15] As already alluded to above, Wells (1982:597) says in his description of the Diphthong Shift in AusE that "the symbol [ʌ] is to be interpreted as implying the quality of Australian STRUT, i.e. between half-open and open, just fronter than central, unrounded". Thus my choice of [ɐ̟] for this value.

the more cultivated ones. Harrington et al. (1997:174) seem to agree that AusE seems more advanced in terms of the Second FORCE Merger than Southern British English. In terms of the AusE CURE quality, it appears that the first target, "is slightly opener and fronted relative to the centroid of /ʊ/" (Harrington et al. 1997:177), the quality of the latter being relatively high and back in the data. The offset of the diphthong terminates in the TRAP-START space i.e. low-front to central.

U.S. Accents: As confirmed by Dodsworth (2008:36), "the fronting of the nucleus in the diphthong /o/ as in *home* and *coat* is a component of the Southern Shift . . . but has been found in several geographically- and ethnically-differentiated varieties of U.S. English". Interestingly enough, "central Ohio /o/ is generally fronted except before /l/ (where a back-vowel merger holds) or pre-nasally" (Dodsworth 2008:37). The same tendency for GOAT to front in most varieties of the US is confirmed by Fridland (2008:69), although not as prominently as GOOSE and FOOT fronting – see §8.4 and §8.2. In California, "/ow/ fronting carries an association with "Valley girl" talk" (Fridland 2008:73). According to Labov (1994:80–1), GOAT-fronting is also a prominent sociolinguistic marker in Philadelphia, with a rounded or unrounded first element. The exception is GOAT before final /l/ where, like with Philadelphia GOOSE, the allophone is "in extreme back position, sharply differentiated from the other vowels" (Labov 1994:108).

8.3.2 GC in SAE: The Impressionistic Data

Hopwood (1928): Moving on to SAEP, Hopwood (1928:13) gives **ɑù**{[ɐ̞u̟]} for GOAT. The author continues as follows:

> "The first element varies in the degree to which it is lowered and fronted . . . from half-open **ɔ** through low middle **ɑ**{[ɐ̞]}, and sometimes to low front **a** . . . the change could not have taken place to the extent it has done without the strong pull of the extreme Cockney E. change to **aʊ** . . . Speakers who have strong Cockney tendencies are inclined to weaken the second element . . . or even to omit it; *e.g. no* . . . **nɑˑ**{[nɐ̞ˑ]}" (Hopwood 1928:13).

Interestingly, the last value, i.e. something roughly in the region of [ɐː ∼ aː], but with added nasalization, seems, in the author's experience, to be characteristic (or at least stereotypical) of the so-called 'Kugel' accent, prevalent in the Northern Suburbs of Johannesburg - see §3.3.2.

For SAE CURE, Hopwood (1928:19) gives, for example, "*tour* ˈ**tʊ˞ə** > "too(r)" **tù˞r**{[tu˞r]} (or "tewr" **tÿür**)".

Lanham and Traill (1962), Lanham (1965; 1967): Lanham and Traill (1962:29) also give unrounded values for the onset of GOAT in both SARP 'A' and SARP 'B': [ɜ][35/45]{[ɘ̠˕]} and [ə][45/54]{[ɜ̠˔]} respectively, adding that a growing trend (particularly in non-SARP SAE) is a progressive lowering of this onset: "a vowel position at [ɔ][54]{[ɜ̠]} is common" (Lanham and Traill 1962:29). There is no mention of the possibility of rounding for the onset of GOAT in non-SARP SAE. Lip-rounding is, however, characteristic of the up-glide of SARP and the value given is [ü][25]{[ʉ̠˕]}. Glide-weakening is, however, identified as a growing trend in SARP 'B' (and, presumably, non-SARP SAE) and the authors mention the possibility of confusion between the NURSE and GOAT vowels: "thus *own* almost rhymes with *earn*" (Lanham and Traill 1962:29). Lanham (1965:91) generally confirms this picture:

> "Notice here the change in S.A.E. in the first vowel from a high central somewhat retracted [ɜ˃]{[ɘ̠]} to [ʌˆ]{[ɜ ∼ ə]}, somewhat higher than in *but* or *bun*. The second change is the loss of post-vowel glide with the result that /ɜw/ is replaced by /ʌ˙/, phonetically [ʌˆ˙]{[ɜ˙ ∼ ə˙]}".[16]

Turning to Lanham (1967), there is no mention of lip-rounding either, with the analyses provided above generally confirmed: [ə][44]{[ə]} "or possibly slightly higher and retracted" (Lanham 1967:90) with a glide ending at lip-rounded [ϋ][25]{[ʉ̠˕]} for conservative SAE. For non-conservative SAE, Lanham (1967:86; 90–1) stresses the lowered and retracted nature of the vowel – the value given is [ʌ][54/64]{[ɜ̠]} – as well as its glide-weakened nature: "/ʌ˙/ is the cognate of RP /əu/" (Lanham 1967:86). The author does, however, stress that "many South Africans fluctuate in their speech along the scales: [ʌ] – [ə] (for the vocoid) and "glide – no glide"" (Lanham 1967:91). Lanham (1967:66;91) also adds an observation as to the effect of final /l/ on GOAT, giving it a quality similar to that of LOT i.e. [ɒː][65]{[ɒ̞̈]}, but with added length.

[16]This quote requires a degree of interpretation. Firstly, Lanham (1965) is, no doubt, using [ɜ] to refer to modern-day [ɘ], as he does in Lanham and Traill (1962) – see Figure 5.2 on page 146. Secondly, it is, in fact, rather difficult to assess what value Lanham (1965) is referring to with respect to the [ʌˆ] symbol. It is unlikely that the [ʌ]-symbol is being used here to refer to the cardinal value, given that in Lanham and Traill (1962:23) STRUT is given the value of [ɔ][54]{[ɜ ∼ ɜ̠]} for SARP, while "RP has a lower and more retracted vowel". Given that more advanced values for STRUT tend to be more fronted it is highly unlikely that the cardinal value is thus meant as the value for STRUT in non-SARP SAE. A value approximating [ɐ] is much more likely and thus the values given in the translation.

For SARP 'A' CURE, Lanham and Traill (1962:31) give "slit–rounded [ʋ]^{26}{[ʊ]}, or almost unrounded [o]$^{35/36}${[ɘ̠]}, or some intermediate position" for the onset[17]. On occasion an alternation between these two values is conditioned by a preceding labial (/p, b, m/) – this context selects the higher, more rounded option. This option is also more prevalent in SARP 'B', which reflects the value in non–SARP SAE. "The glide is short, ending at approximately [ɜ]$^{34/44}${[ɘ̞]}" (Lanham and Traill 1962:31). Lanham and Traill (1962:31–2) also mention an even more low value, in many cases resulting in homophony with THOUGHT. Lanham (1965:92) claims a tendency towards monophthongization and a replacement of "/uᵊ/ with either /u˙/ or /ɔ˙/". Thus, according to this author, "*sure*, *poor* are hardly ever /šuᵊ/ and /puᵊ/ in this country; but /šɔ˙/ and /pɔ˙/" (Lanham 1965:92). This is, however, certainly not the case in present-day SAEP - while a monophthongized *sure* is common, the alternative is also used. A monophthongized *poor* is, on the other hand, less common than its alternative and would, in many cases, come across as somewhat affected. Lanham (1967), gives examples such as /kruːl/ and /ʤuːlrɪ/ for *cruel* and *jewelry* for the replacement of CURE by GOOSE. Lanham (1967:94) also mentions the possibility of di-syllabic variants /u˙+ə/ in the most extreme forms of SAE[18]. For conservative varieties of SAE this author claims a value of [ʋ]^{26}{[ʊ]} for the onset (or slightly lower) and [ɜ]^{34}{[ɘ]} for the terminus of the glide. For non-conservative SAE, on the other hand, "the vocoid may be higher [u]^{16}{[u̟]} with more lip-rounding, particularly in idiolects influenced by Afrikaans" (Lanham 1967:94).

Lanham (1978), Lanham and Macdonald (1979), Lanham (1982): Lanham (1978:151–2) adds little to the above analysis. The link between BrSAE and a retracted, open and glide-weakened value for GOAT is mentioned: "extensively corrected and involved in wide stylistic variation" (Lanham 1978:152). The author does not mention the possibility of a rounded onset. The value of GOAT before final tautosyllabic /l/ is given as [ɒ˙].

Lanham and Macdonald (1979:43) give [ɒ̈˙] as the most advanced (GenSAE) value for this vowel before /l/ and unrounded [ʌ˙] for the most advanced variant more generally. According to Lanham and Macdonald (1979:46), the variable appears to be a stable one with no evidence of it receding. There appears to be a correlation, idiolectally, between a backed, retracted and glide-weakened GOAT vowel and a similar

[17] In this regard, I suspect that [o]$^{35/36}$ is a typographical error and should be [ɜ]$^{35/36}$ – thus the translation as {[ɘ̠]}.

[18] A possible source being north-of-England varieties, see §8.3.1, although the general association of NE with prestige renders this more than tentative.

FACE vowel. A backed, lowered and glide-weakened GOAT vowel is identified as a Gen SAE variable. Lanham (1982:339) confirms the above and makes no mention of the possibility of rounding with respect to the onset of this vowel.

Lass (1990; 1995; 2004): According to Lass (1990:280), for his GenSAE Cape Town speakers, the onset of GOAT is *rounded* and centralised i.e. [œ̈], and the second element is unrounded i.e. [ɤ̈ ∼ ɤ]. A typical value would thus be [œ̈ɤ̈]. The author claims prestige value for a front and rounded onset. The onset is generally lower than NURSE, so that in cases of monophthongisation (which the author claims are not uncommon), the qualities do differ. Rounded offglides are, however, possible in GenSAE and [ʉ̞ ∼ ʊ̈] are given as possible values. Lass (1990:280) adds the following comment:

> “Auditorily, the transition between morae is exceedingly gradual: both lip-spreading and tongue retraction appear to begin quite soon after onset . . . phonetics students have more trouble deciding how to transcribe this than any other vowel in the system”.

According to Lass (1990:273), an extreme, BrSAE value for GOAT would be a lowered and retracted [ʌ̈ʊ]. Lass (1995:100) confirms that GOAT in SAE is “an important social marker” and gives common CulSAE values as [ɛʊ ∼ œʊ]. This is contrasted to GenSAE which does not have an unrounded onset, with the typical onset being [œ]. “The second element may be central [ʉ] or unrounded [ɤ̈], and monophthongisation is common, especially in younger speakers” (Lass 1995:100), but does not appear to be gender-specific. Such monophthongisation creates a potential near-merger with NURSE ([øː] vs. [œː]) and, according to Lass (1995:100), non-SAE speakers often have trouble distinguishing between these two vowels. Lass (1995:100), for example, tells the humerous story of hearing *the turtle onslaught* for *the Total Onslaught*. Like CulSAE, BrSAE also has an unrounded onset, but it is lowered and retracted, generally in the area of STRUT, thus overlapping partially with BrSAE FACE i.e. [ʌʊ] vs. [ʌɪ].

Lass (2004:377) generally confirms his earlier analyses, giving [œʊ] as the most common value for GenSAE speakers. For CulSAE and GenSAE speakers the onset is usually more open than NURSE. The possibility of unrounded second element values as well as those in the region of FOOT or GOOSE is also mentioned.

According to Lass (1990:278–9), the quality of CURE in the speech of his subjects is given as [ʊ̈ə]. The author mentions a partial, variable merger with THOUGHT particularly characteristic of younger speakers. The word *sure* is mentioned as being particularly susceptible in this regard.

According to Lass (1995:100–1), the CURE vowel is in the process, particularly among BrSAE and younger GenSAE speakers, of merging with the THOUGHT vowel i.e. [oː], which accounts for the homophony between *your* and *yore* as well as *shore* and *sure* in SAE. It is unclear, however, how extensive this merger is or what its social value is. What is certain is that a /(C)j–/ context, as in words such as *fury*, *pure* and *cure*, appear to be resistant to the merger. In BrSAE, and in cases where THOUGHT has transplanted CURE, there is also the occasional monophthongisation to a GOOSE-like quality, "especially in the contraction *you're*" (Lass 1995:101), as is reflected in Hopwood (1928) (see above).

Lass (2004:378) confirms the distinction between CURE and THOUGHT in SAE although he labels it recessive and makes the claim that with respect to "words with final historical /r/ (e.g. *sure* vs. *shore*) . . . speakers in their twenties rarely make the distinction". CURE is also often used in Afrikaans loan-words (like *kloof* 'ravine') although GOOSE can also be used here. The most common value for GenSAE speakers is given as [ʊə].

Lastly, it is interesting to note that the merger appears to be far more extensive in Rhodesian English than in SAE. Orpen (n.d.) provides some preliminary data confirming this impression.

Wells (1982), Bailey (1984), Branford (1994), Bowerman (2004), Da Silva (2007): Beginning with GOAT, Wells (1982:614) gives [əʊ] for CulSAE and [ʌʊ] or monophthongal [ʌː] for "a popular accent", with the half-long [ʌˑ] being the most advanced variant (Wells 1982:621). The author also mentions substantial style-shifting (and sociolinguistic variation) with respect to this vowel. According to Branford (1994:483) this vowel is subject to a high degree of variability:

> "Alan Paton, broadcasting in later life, had [ə̹ʊ] (perhaps [ʌʊ]?) in *no* and *those*. A student speaker from Johannesburg had [ə̹ʊ] in *Joburg* and in *home*. Another had [ɐ] in *only*. An East Cape student had [ɑː] in *foals* and [oʊ] in *old*, both in formal speech and within the same utterance".

This author confirms the conservative variant as being [əʊ] with a monophthongal "[ʌː] in less conservative speech" (Branford 1994:483). A glide with a rounded and fronted onset but an unrounded terminus is also mentioned: basically Lass' (1990) [œ̈ɤ̈]-value.

For CulSAE, Bowerman (2004:938) claims the same alternation between rounded and unrounded onsets as Lass (1995). Common values for GenSAE are given as [œʉ ∼ œɤ̈ ∼ œː] showing a categorical fronting and rounding of the first element, but a range of values (including monophthongisation) for the off-glide i.e. fronting/centralization and/or unrounding. BrSAE is confirmed as having a more retracted and open onset: [ʌʊ].

It is unsurprising that Da Silva (2007:119) did not detect "a single low back vowel [ɑː]" as a variant of GOAT. Her inclusion of this variant is based on Hopwood's (1928) analysis of this vowel, but as mentioned above Hopwood's (1928) **ɑ** symbol needs to be interpreted as [ɐ̞]. Of more interest is Da Silva's (2007) discovery of an American-like [oʊ]-variant of GOAT, which although used predominantly by 'black'-speakers appears to be, on Da Silva's (2007) analysis, encroaching upon the traditionally 'white' dialect[19].

In line with much else in the literature, Bailey (1984:16) claims that CURE is being eliminated in SAE and being replaced with the vowel in THOUGHT. Branford (1994:482) confirms the partial merger of CURE and THOUGHT in SAE. Bowerman (2004:938–9) confirms [ʊə] for CulSAE and GenSAE as well as the growing tendency of the later to use [oː] when there is no preceding yod. The implication is that this is a Broad value, "perhaps slightly lower than THOUGHT" (Bowerman 2004:939). The mispelling of *you're* and *your* is referred to as evidence of this tendency. Interestingly, Trudgill (2004:145; my emphasis) claims that the Second FORCE Merger "appears . . . to be *absent* . . . from South African English".

8.3.3 GC: A Comparative Acoustic Analysis

The separation of GOAT into two main allophones i.e. an unmarked one and another before final /l/ is relatively clear from Figure 5.5 on page 152, which provides a graphic illustration of Webb's (1983) data. The marked allophone is represented by $goat_l$ and relates to only one token in Webb's (1983) data i.e. *wholly*. The GOAT nucleus also appears substantially lowered, a fact which becomes clearer if we superimpose the GOAT values over the short-vowel data, as provided in Figure 8.14 on the facing page, where we note a relatively clear overlap between GOAT and STRUT. It should also be clear that, in contradiction to certain pronouncements in the impressionistic literature, the is little evidence for an overlap of

[19] Although only minimally: according to Da Silva (2007:159), for her lect 1, the [əʊ]-variant (which subsumes GOAT with a [œ], [ə] or [ʌ] nucleus) has a variant mean score of 95%, while the [oʊ]-variant has a score of 5%.

GOAT before final /l/ and LOT .

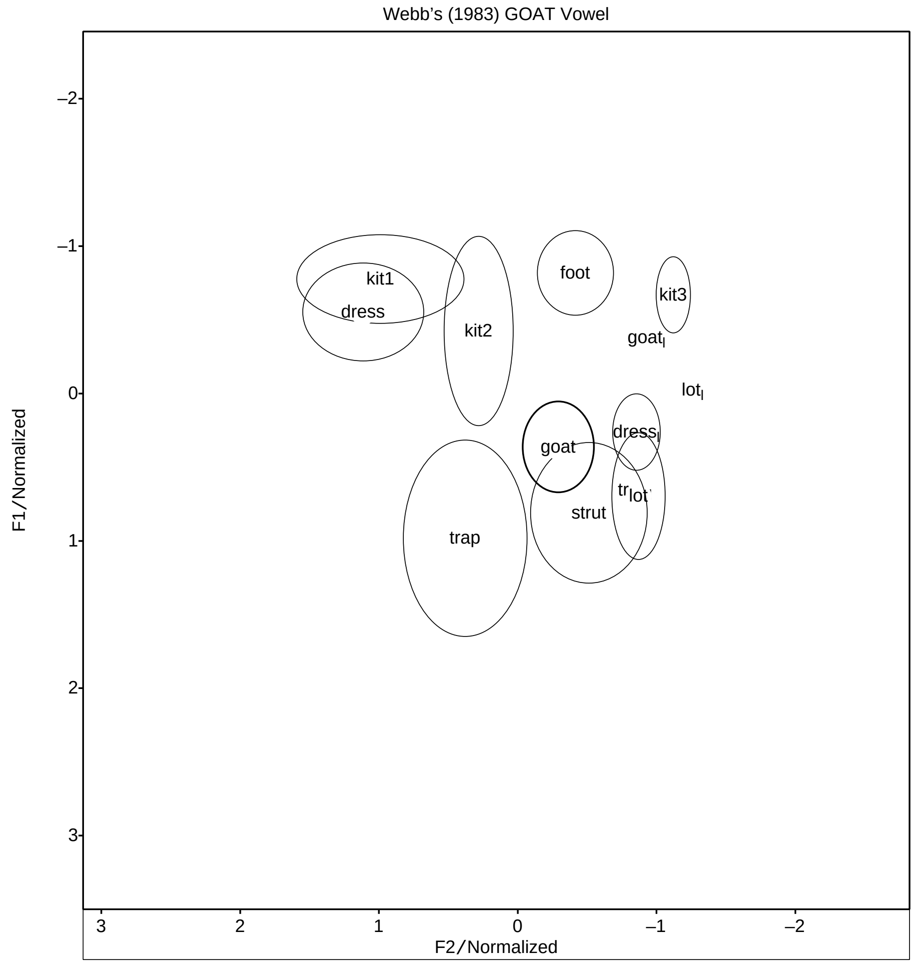


Figure 8.14: Webb's (1983) GOAT Vowel

As far as the glide is concerned, Webb (1983:155) notes a lack of patterning with respect to a short glide which "is sometimes back (by only 86Hz) but more often "forward" (by 155Hz)" (Webb 1983:155).

Data for CURE is not represented in Figure 5.5 on page 152, which gives data on Webb's (1983) long-vowel and diphthong values, since the two putative CURE tokens (both *sure*) were phonetically in the [oː] region; they have thus been included under THOUGHT, as discussed in §8.2.

With respect to the current data, Figure 6.1 on page 172 situates GOAT in the context of the other glides-to-[u]. This graphic representation of GOAT was based on 1826 measurements across 83 tokens across 27 subjects, at an average of 3.07 tokens per subject. The

N-G words were *load*, ***so**fa*, *hoe* and *moat*, while the O-G words were *oat* and *rose*. In the case of GOAT before final /l/, the representation was based on 35 tokens, the N-G words being *hole* and *told*.

If, in the case of the non-pre-final-/l/ variant, we assume some rounding of the first element, as advanced by Lass (1995), then we appear to have a general agreement between the impressionistic literature and the acoustic data i.e. a rounded, but relatively mid onset, [œ], along with a centralized terminus to the off-glide i.e. [ʉ]. Although the glide is comparatively weak there is clear evidence of diphthongization.

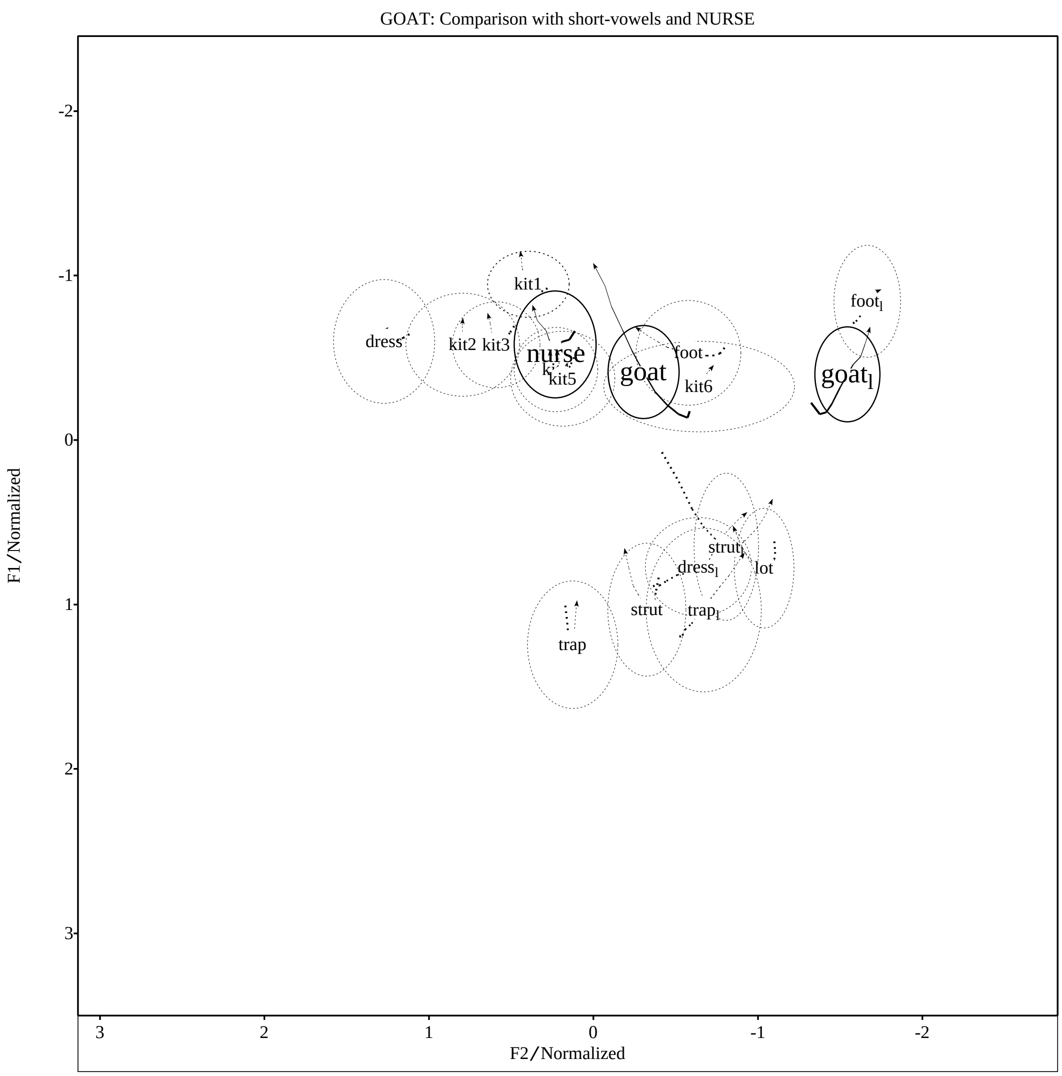


Figure 8.15: GOAT: Comparison with short-vowels and NURSE

If we compare the GOAT data with the other short vowels and, in particular, with NURSE, as has been done in Figure 8.15 on the preceding page, a slightly different picture emerges. Firstly, it should be emphasized that there has been general agreement in the impressionistic literature that NURSE in SAE has a rounded value (see §9.2). If we assume a [ø̈]-value for this vowel then it is clear that while slightly more lowered than NURSE, the onset of GOAT is also more centralized. The onset is therefore more likely to be in the region of [ɵ] or something slightly lower.

We also note that, as in the case with Webb's (1983) data, the pre-/l/ allophone of GOAT does not approximate LOT, but is in fact closer, if anything, to FOOT before final /l/. Given its length it appears to almost have an [ɔː]-like quality.

Figure 8.16 on the following page compares the current GOAT data with that of Webb (1983). As can be clearly see, Webb's (1983) GOAT vowel is substantially further lowered than that of the current data. As shown in Figure 8.14 on page 323, the point of inflection of this vowel practically overlaps with STRUT. What this indicates, in addition, is that like FACE (§6.4), GOAT shows substantial style-shifting and is likely to be sensitive to various sociolinguistic variables such as gender and social-class; a clear sociolinguistic marker in SAEP. This is, of course, also confirmed by the impressionistic literature reviewed above.

Figure 8.17 on page 327 provides the comparison with other accents of English; we note the retracted second target for RP GOAT, as well as the comparably fronted targets for both AusE sets of data. Compared to the RP data, GOAT in the current data is more fronted; while compared to the AusE data, it is somewhat closer (in terms of the onset) and shows a greater degree of glide-weakening. The onset of the AusE data (particularly a1 – data from Cox (2006) on Australian female teenagers) is reminiscent of the values given for Webb (1983) in Figure 8.16 on the next page. It seems that, as was the case with FACE, a lowered onset for GOAT has taken a greater hold in AusE than in SAE. It should be emphasized again that the data from Cox (2006) is from citation-form and from General AusE speakers – thus directly comparable to the present data. What SAE *does* share with AusE is extensive fronting of this vowel.

Figure 8.2 on page 296 also places CURE in the context of the other Part-System D vowels. It is based on 1364 measurements across 62 tokens across 27 subjects, at an average of 2.3 tokens per subject. The N-G tokens were *cure*, *pure*, *moor* while the O-G token was *poor*. From the relevant figure a clear glide is evident, beginning from a high central position and moving to a schwa-like position.

The rather unexpected starting-point for this vowel, as represented in Figure 8.2, is, however, clarified if we view the tokens mentioned above separately, as in Figure 8.18 on page 328. Here we note the movement of *cure* and *pure* (with a preceding yod) from a

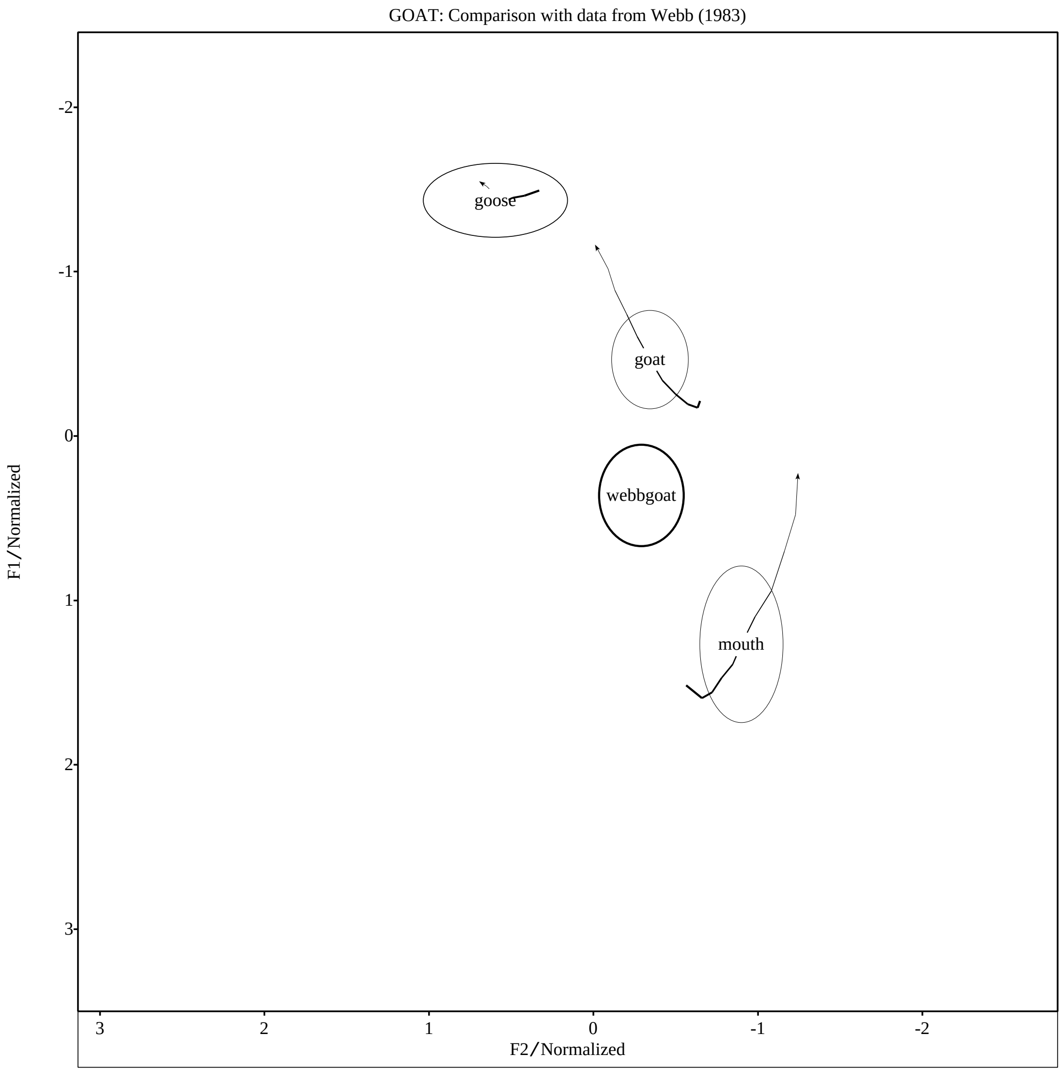


Figure 8.16: GOAT: Comparison with data from Webb (1983)

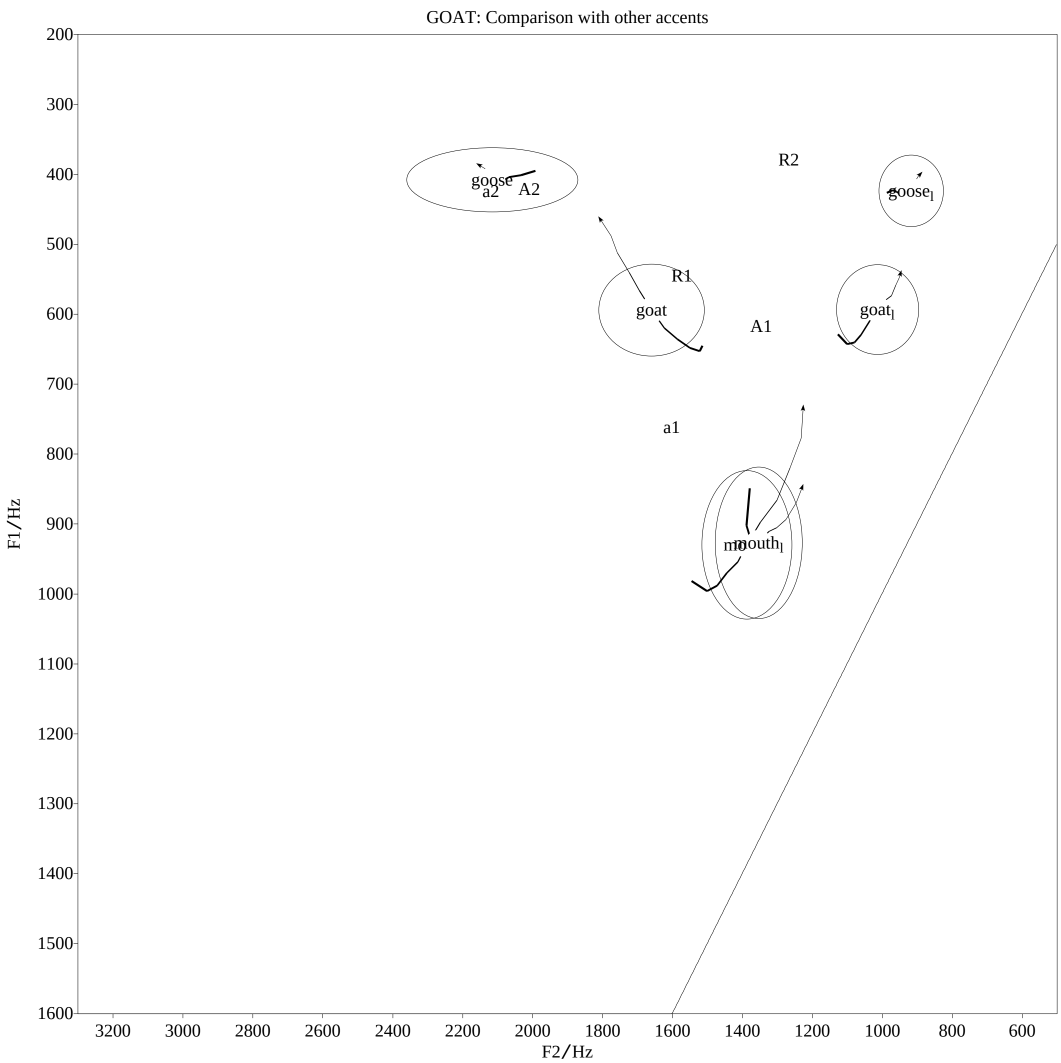


Figure 8.17: SAE GOAT vs. other English accents

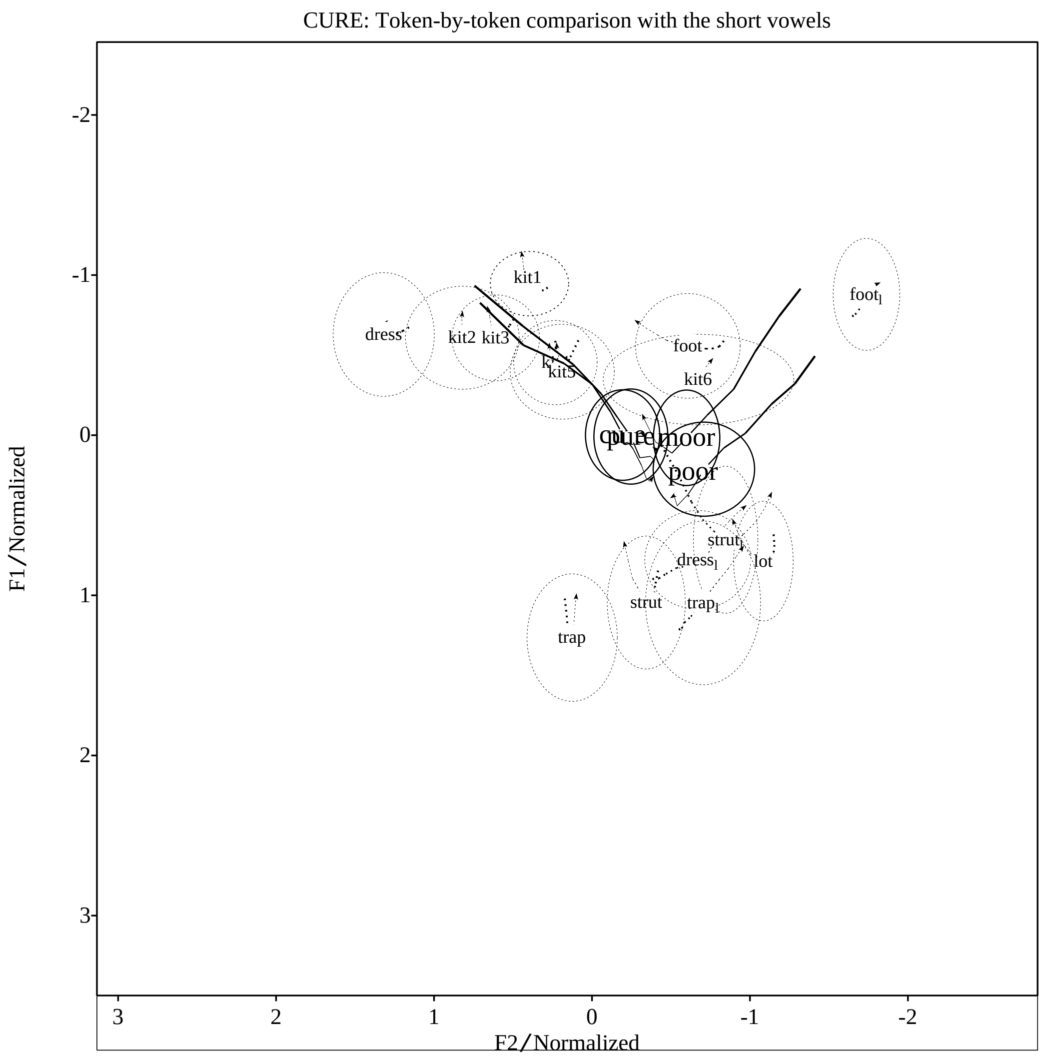


Figure 8.18: CURE: Token-by-token comparison with short-vowels

high-front to a schwa position[20], while the remaining two tokens (*poor* and *moor*) move clearly from a high back position to mid-central. There is no evidence here for a CURE-THOUGHT merger i.e. for the Second FORCE Merger.

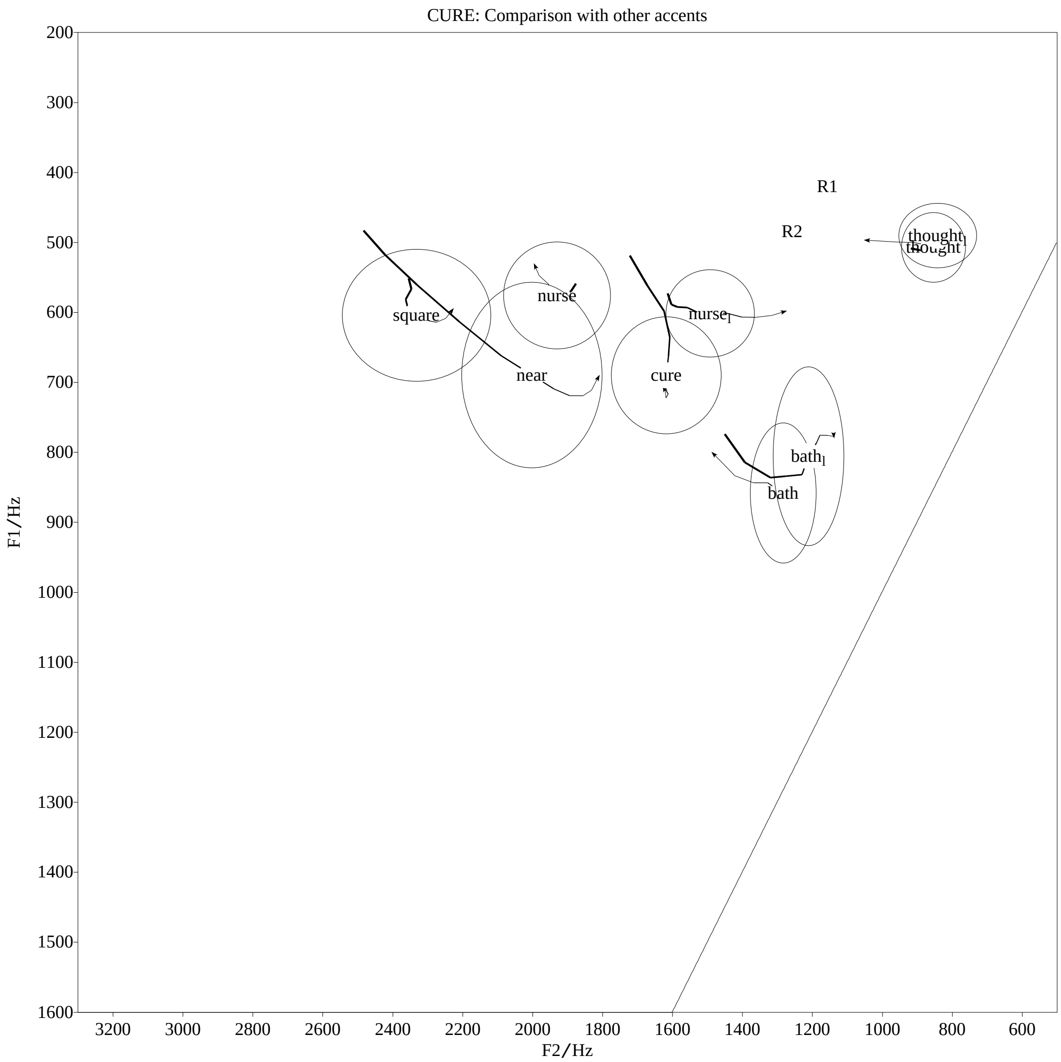


Figure 8.19: SAE CURE vs. other English accents

If Figure 8.19, which provides RP data for CURE, is compared with Figure 8.18 on the preceding page, it is clear that RP has substantially reduced, compared to SAE, the glide of this vowel.

[20]The two labels overlap so the ‘picture’ is a little obscured in this regard.

8.3.4 GC: Synopsis

As in the case of FACE, GOAT in SAE displays substantial stylistic shifting, a fact clearly emphasized in the impressionistic literature as well as in the acoustic analysis provided above. It remains, again like FACE, a clear sociolinguistic marker in SAE. In citation-form at least, the more prestigious variant is, however, not as fronted as implied in some of the literature. GOAT also shows a clear allophonic difference between an unmarked variant and a variant that precedes final, tautosyllabic /l/. The acoustic data also provides evidence for a substantially fronted second target, in line with, for example, Lass' (1995) analysis, and while there is no true monophthongisation, there is substantial glide-weakening; far more than in AusE. The existence of a relatively close version of GOAT in GenSAE again emphasizes SAE's 'resistance' to Diphthong Shift. In GenSAE there is a additional tendency to front[21] rather than fall.

Diachronically, however, the development of GOAT appears to differ from that of FACE. While in both cases there is some evidence for narrowly diphthongal (RP), or even monophthongal (north-of-England) tokens as part of the linguistic mix of early-Johannesburg, this only explains the development of such variants in GenSAE FACE. In the case of GOAT, on the other hand, the eventual development was in the direction of a fronted and rounded GOAT vowel, while north-of-England input would have, if anything, provided monophthongal *back* variants. In addition, while RP would have, indeed, provided a narrower diphthong, and perhaps even a relatively front one, it would have had an *unrounded* onset. As such, the only possible solution is that a fronted, rounded GOAT vowel was an endogenous development, reflecting similar developments in, for example, the USA (§8.3.1). This is also reflected in the fact that GOAT with a fronted, rounded onset has only recently been noticed in the impressionistic literature.

In the case of the more open variant of GOAT, one should note again, as was the case with FACE, that it was, in all likelihood, only the *tendency* towards Diphthong Shift that was inherited in CE. By the time CE arrived in Johannesburg, however, the process was, in all likelihood, well under way, and combined with similar values brought over with working-class individuals from the southeast of England. Until recently, it appears that this Cockney-like GOAT variant has lived side-by-side with an RP-like unrounded variant in stable sociolinguistic variation.

Note that the adoption of a 'narrow' fronted GOAT, e.g. [œɤ], into GenSAE cannnot be used as counter-evidence for Trudgill's (2004) model of new-dialect formation, even though

[21] For Labov (1994), a movement that is in parallel with the fronting of GOOSE, rather than with the falling of FACE.

it was probably unaccounted for in the early-Johannesburg linguistic ‘mix’. Important here is the fact that this development appears to be an endogenously-motivated one and, if the impressionistic literature is anything to go by, a recent development to boot. The same proviso applies, of course, to all variants that constitute the end-products of endogenous shifting or ‘drifting’.

The acoustic analysis of CURE, as given above, provides no surprises. Still, it would appear that the extent of the Second FORCE Merger in SAE has been over-reported and recent studies such as Orpen (n.d.) and the above acoustic analysis seems to support this assessment. Historically, there appear to have been no shortage of inputs to explain the current SAE value, with the Second FORCE Merger in both RP and Cockney, for example, being a relatively recent phenomenon. The north-of-England might also have supplied [ʊə]-like values.

8.4 The ‘Remaining’ SECS Vowels: STRUT, LOT, FOOT (SLF)

As pointed out in §2.14, the SECS-Shift involves a counter-clockwise movement of the traditionally short vowels of English. The three front short vowels have already been dealt with in §6.3 (TRAP), Chapter 7 (KIT) and §6.4 (DRESS, alongside FLEECE and FACE). This sections deals with the three traditionally back short vowels of English i.e. STRUT, LOT-CLOTH and FOOT. Figure 8.20 is an exact replica of the one provided in Chapter 2, repeated here for the reader’s convenience.

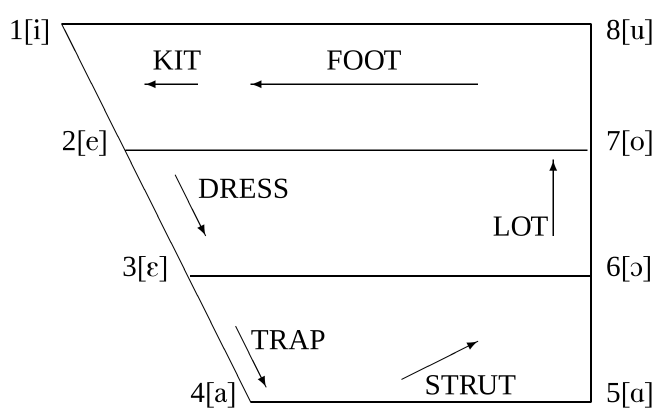


Figure 8.20: A Vowel Chart showing Torgersen and Kerswill’s (2004) SECS

A brief word on LOT-CLOTH:

> “The standard lexical set LOT is defined as comprising those words whose citation

form has the stressed vowel /ɒ/ in RP and /ɑ/ in GenAm. This comprises a large majority of cases of RP /ɒ/, the remainder being allocated to CLOTH" (Wells 1982:130).

The standard lexical set CLOTH, on the other hand, has /ɒ/ in General RP but /ɔ/ in GenAm (Wells 1982:136). In the American notational system LOT is given as /o/ or /a/ (Labov 1991:40; footnote 13). The vowel is often referred to as "Short O".

8.4.1 SLF: Other Accents of English

Received Pronunciation: For RP, Lanham (1967:33) gives two possible qualities for STRUT: from "unrounded, central [ʌ]54{[ɜ̠]}" for the older generation to "a currently more prevalent low central [a]64/73{[ɐ̠]}". According to Cruttenden (2001:113), in RP the STRUT vowel is "a centralized and slightly raised C[a]" i.e. [ä̝], while Cockney STRUT approaches C[a]. From a reading of Fabricius (2007*a*), it is clear that an understanding of STRUT in RP cannot be divorced from an understanding of its relationship with its nearest short-vowel neighbour i.e. TRAP. A brief summary of this relationship as found in RP has already been provided in §6.3.

With respect to the quality of LOT, for RP Lanham (1967:33) gives [ɑ]66{[ɒ̝]}, and Cruttenden (2001:117) confirms that "the realization of /ɒ/ varies little within RP" which he gives as basically C[ɒ]. Diachronically, it appears that "LOT has been remarkably stable within RP over the past century" (Fabricius 2007*b*:1477).

For RP FOOT, Lanham (1967:33) mentions two variants which are often in free variation: a weakly rounded [ʊ]25/26{[ʊ̠̈]} or an unrounded [o]36{[ɤ̈]}. Cruttenden (2001:121) also mentions a growing tendency towards an unrounded quality: "the quality is that of a centralized C[ɤ] or [o] = [ö] or [ɤ̈]". The FOOT vowel has been identified by Hawkins and Midgley (2005:193) as participating in a chain-shift in RP, one part of which involves the unrounding *and fronting* of the GOOSE and FOOT vowels. As pointed out by Fabricius (2007*b*:1477), this is in line with Labov's (1994) Principle III (§2.3.2) by which close back vowels have a tendency to front. It appears that FOOT-fronting in RP began in the 60s and was established by the 70s and early 1980s (Fabricius 2007*b*:1479–80).

London (Cockney): For London, it appears that STRUT has undergone a process of fronting *and* retraction. According to this hypothesis, as, for example, contained in Torgersen and Kerswill (2004), STRUT-fronting had completed itself by the middle of the 20th-century, at which time, under the influence of a lowered TRAP vowel, it began to retract. Thus a retracted STRUT forms part of Torgersen and Kerswill's (2004) SECS, as depicted in Figure 8.20 on page 331.

Torgersen and Kerswill (2004:32;33) appear to indicate that, for London, [ɔ], slightly more raised above the RP value, appears to be a common value for LOT; this is confirmed in Wells (1982:305). LOT-raising also appears to be a part of Torgersen and Kerswill's (2004) SECS-Shift. It should be added, in this regard, that LOT-raising constitutes a possible violation of Labov's (1994) Principle II[22]; unless one can find evidence for the peripheral status of this vowel.

According to Torgersen and Kerswill (2004:28), a fronted FOOT vowel has "been reported across several varieties in England", and appears to part of their SECS; again, see Figure 8.20 on page 331.

Other Accents of England: Beginning with STRUT, we can, of course, ignore north-of England varieties given that, in general, these accents have not undergone the FOOT-STRUT Split. In the case of LOT, however, it is interesting to note that, according to Wells (1982:356), "the LOT vowel ... is generally fully open ... perceptibly opener than mainstream RP". There is no indication, however, that it is unrounded, unlike in the southwest of England in which there is, indeed, some evidence for unrounding to [ɑ] (Cruttenden 2001:117). Trudgill (1999*a*:127) also reports an unrounded LOT for traditional Norwich.

Moving on to FOOT, it is interesting to note in Wells (1982:351–3) that many northern near-RP speakers[23] tend to a hypercorrective use of /ʊ/ in FOOT words: "pronunciations such as *sugar* [ˈʃəgə ∼ ˈʃʌgə]", which generally involve FOOT lowering and unrounding.

AusE, NZE and Falkland Island English: As already mentioned in §3.4.3, a lowered and fronted STRUT is seen to be typical of Southern Hemisphere Englishes. Thus, with respect to AusE, for example, Wells (1982:599) refers to the "fronting of the /ʌ/ of STRUT towards the cardinal 4 area, [a--]{[a̠]}". For NZE, Gordon et al. (2004:26) give either [a̟] or [ɐ̞].

For NZE, Gordon et al. (2004:26) give a raised and central quality for LOT: [ɒ̝̈]. Interestingly, Harrington et al. (1997:178) imply [ɔ] as being a central value in AusE. Evans and Watson (2004:197) claim that LOT (the authors use /ɔ/) shows little variation across NZE, AusE and "British English". Intriguingly, Trudgill (2004:48) gives

[22] Only if this vowel is participating in a chain-shift, however.

[23] A label that one might have easily applied to many NE-speakers.

an *unrounded* LOT as common in Falkland Islands English, a phenomenon which he attributes to the fact that "unrounded LOT was formerly much more common in Britain than it is today" including the south of England.

U.S. Accents: A backed and raised STRUT appears to be an incipient change in Philadelphia, connected with a more general chain-shift involving the backing and falling of KIT and DRESS (Labov 1994:82). A fronting of LOT (short **o**) appears to be an integral part of Labov's (1994) Northern Cities Shift; in essence it 'fills' the space left by the tensed and raised reflex of short **a** i.e. /æh/ which is a combination of BATH and certain (or all) tokens of TRAP. FOOT-fronting appears, according to Fridland (2008:69), to be a feature common to most US dialects, regardless of region.

GenAm is characterized by Wells (1982:136) as a broad-CLOTH accent (i.e. CLOTH has the same vowel as THOUGHT). General RP is, on the other hand, a flat-CLOTH accent (i.e. CLOTH uses the same vowel quality as in LOT). Generally broad-CLOTH accents are conservative and reflect an older merger of CLOTH and THOUGHT. On the other hand flat-CLOTH accents reflect a more modern reversal of this merger (Trudgill 2004:66) thus explaining why an older form of RP *does* use the THOUGHT vowel in a sub-set of CLOTH words e.g. in *off*, *cough*, *soft* (Wells 1982:136–7), a practice it shares, interestingly enough, with Cockney (Cruttenden 2001:117) and with many of the Southern Hemisphere Englishes. With respect to these later varieties, Trudgill (2004:67) confirms that "/ɔː/ is still today a good deal more common in New Zealand, Australia and the Falklands than it is in England," thus emphasizing the conservative nature of these accents. The distinction is, of course, moot in those accents of American English which display the THOUGHT-LOT merger (Wells 1982:473–6) i.e. *cot* and *caught* are homophones.

8.4.2 SLF in SAE: The Impressionistic Data

Hopwood (1928), Breckwoldt (1961): The STRUT vowel appears to have a relatively large range of variation in SAEP. Beginning with Hopwood (1928:27), we find that "**ʌ** > **œ** in stressed syllables, and sometimes **ë** in weakly stressed syllables"[24]. For the 'standard' value, Hopwood (1928:89) states that "E. **ʌ** is nearer the middle of the mouth, not Half-open back". From his vowel chart[25] it would appear that the author has something like the modern [ɐ̠] in mind. In this regard note the value provided

[24]Da Silva (2007:90; my parenthesis) states that "this seems like an unlikely variant, as it [has] not been documented as a possible option by any other researcher on the topic". I agree that this seems unlikely, but we will note that Breckwoldt (1961) provides a similar value.

[25]see Figure 5.1 on page 144.

by Trudgill (2004) and discussed in §3.4.3 i.e. [ʌ̟]. Hopwood (1928) also mentions the possibility of "the Cockney EP. **a** for **ʌ**{[ɐ̠]} . . . *e.g.* E. enough **ɪnʌf**{[ɪnɐ̠f]} > Cockney **ɪnaf**". Of particular interest is Hopwood's (1928) following claim:

> "A minor variation of this Cockney change of **ʌ** to **a**, and based on it, may be heard on the Rand [Johannesburg and environs] and sometimes in Natal and the Eastern Province; viz. **ʌ** (> **a**) > **ɛ̃**, noticeable in the word "chuck"; *e.g.* E. **čʌk** (> **čak**) > **čɛ̃k**" (Hopwood 1928:27; my parenthesis).

The nasalization of the vowel aside, we will see below that one of the prestige values in GenSAE is a STRUT vowel with a value in the region of C[ɛ].

According to Hopwood (1928:14), LOT is realised as [ɔ ∼ ɔˑ ∼ oˑ] in SAEP, although the values he provides seem extremely Broad. Hopwood (1928) seems to imply a combined effect of substitution from Afrikaans and Cockney origin for these particular values. For FOOT, and as far as BrSAEP is concerned, Hopwood (1928:21) implies that RP [ʊ] becomes [u̟(ˑ)], and even fully-lengthened "when the spelling is *oo*" (Hopwood 1928:21). Thus RP [hʊk] becomes [hu̟ːk].

Breckwoldt (1961:8) also gives a fronted variant of the STRUT vowel; thus "bʌt > bœt". Admittedly, though, the other values provided by this author seem extremely unlikely, thus casting some doubt on the accuracy of this author's transcriptions. Thus Breckwoldt (1961:8) claims "dʒʌst > dʒɩst . . . kʌvə > kɔvə(r)". For LOT-CLOTH Breckwoldt (1961:7–8) provides Broad values, with, for example "mɒdən > moːdə(r)n" (Breckwoldt 1961:8). For FOOT, Breckwoldt (1961:8) also gives examples such as "bɷk > buːk".

Lanham and Traill (1962), Lanham (1967), Lanham and Macdonald (1979), Lanham (1965):

For STRUT, a less extreme form of fronting is claimed by Lanham and Traill (1962:23) for SARP i.e. [ɔ]54{[ɜ ∼ ɜ̠]}, while "RP has a lower and more retracted vowel". For non-conservative SAE, Lanham (1967:82) gives an "unrounded, central [ʌ]54{[ɜ̠]}" and claims that its value before /l/ is in the region of [ɒ]65{[ɒ̤̈]} "indistinguishable from the normal allophone of /ɒ/"[26]. In general, "SAE short /ʌ/ . . . does not follow the lowering trend towards [a] as in RP and its allophones remain at the mid level of tongue height" (Lanham 1967:84).

For LOT-CLOTH, Lanham and Traill (1962:22–3) give [*a*]65 for SARP, claiming that the vowel is unrounded; and thus approximately [ä̝] or [ɑ̤̈]. SARP 'B' shows

[26] Also see Lanham and Macdonald (1979:43; footnote) in this regard.

a tendency, like non–SARP SAE, to highten, tense and round this vowel. Lanham (1967:82) also confirms the rounding of this vowel as a feature of non-conservative SAE, with a value somewhat higher than conservative unrounded [ɒ]^65^{[ɑ̞̈]}.

The earlier assessments of BrSAE FOOT are supported by Lanham and Traill (1962:23) who give [u]^17/26^{[u̟˕]} for FOOT in this sociolect. For SARP, Lanham and Traill (1962:23) give "[ʊ]^26^ (with weak slit rounding) to lower more central and virtually unrounded [o]^35/36^," with the latter being more typical of formal registers – thus, [ʊ̜ ∼ ɘ̠][27]. Lanham and Traill (1962:34) also mention the influence of a following /l/ on FOOT: "the influence is mainly a lowering one producing [θ]^46/55^{[ɔ̞̈]}". Lanham (1965:93) confirms this, claiming that in *wolf* and *wool* FOOT has almost the same value as LOT. The author does not, however, predict the neutralization of these two vowels: "there is no immediate evidence of the collapse of the . . . contrast before /l/ in present-day speech" (Lanham 1965:93). A similar assessment is provided in Lanham (1967:66) who, for non-conservative SAE, provides [ʊ] for the unconditioned allophone, but [ɵ]^46/47^{[ö̞]} for FOOT before final, tautosyllabic /l/; also see Lanham and Macdonald (1979:43; footnote) in this regard.

Lass (1990; 1995; 2004): For STRUT, Lass (1990:277) views the rather raised [ɜ] as "one common and perhaps pivotal value," while stressing the impressive phonetic range of this vowel. The author characterizes this range as extending from as far back as [ʌ̈] (usually only in word-list style) to a front (and open) [ä] more in line with values common in other Southern Hemisphere Englishes such as AusE and NZE.

Lass (1995:98) mentions STRUT as being a weak social variable, the values fluctuating "from low central to centralized front half-close", with the norm as central [ɐ̶]{[ɐ̞]} to centralised [ä]. According to this author, "the backer and opener values are associated with Conservative and older Respectable speakers, and the fronter and higher ones, going as far as even [ɛ̈], with younger, especially female, Respectable ones" (Lass 1995:98).

Lass (2004:376) refers to this vowel as a 'floater', given that, according to him, "there is nothing particularly close to it in the vowel space". The frontest value is given as [ä], with the highest at [ɛ̈]. The higher the value the younger the speaker. The most common values given for GenSAE are [ɐ] or [ʌ].

According to Lass (1990:277), the LOT vowel is, in the dialect of GenSAE he investigates, generally weakly rounded and centralised: [ɒ̈]. This author also indicates

[27] In this regard, [o]^35/36^ is almost certainly an error and should be [ɜ]^35/36^ – thus its translation into modern terms as [ɘ̠].

that it is normally raised above the cardinal position (i.e. [ɒ̣̈]{[ɒ̝̈]}) and that sometimes the lip-rounding is weak enough to suggest [ɑ̣̈]{[ɑ̝̈]} or even something [ʌ]-like if the vowel is high enough. “It seems impressionistically that the lower the prominence, the higher and less rounded the realisation” (Lass 1990:277). Likewise, Lass (1995:98) claims that “younger Cape Town and Natal Respectable speakers seem to be manifesting raising and further unrounding, giving a quality in the vicinity of a rather central [ʌ]”. The prototypical quality is, however, a weakly-rounded, quite centralized [ɒ̈].

While [ɒ] still appears to be the most typical value for GenSAE speakers, Lass (2004:376) confirms the common tendency for this vowel to have only weak-rounding. It is not always fully-open. The author gives [ɑ̝] as one common value. “Some regions, especially Natal, may have a raised version that is virtually an advanced [ʌ] ... This is never subject to confusion with the STRUT vowel, which is much fronter” (Lass 2004:376).

According to (Lass 1990:277), in GenSAE the FOOT vowel “bears roughly the same relation to RP [ʊ] as SAE [ɪ̈] does to RP [ɪ]” i.e. it is more central than [ʊ] i.e. [ʊ̈], sometimes even as front as [ʉ̜]{[ʉ̞]}. The effect of a final /l/ is to render the variant a backer and, often, rounder one. The prototypical value for FOOT in SAE is given by Lass (1995:98) as [ʊ] or something slightly more fronted. An even more front [ʉ̞] is not uncommon among “younger (especially female) Respectable speakers” (Lass 1995:98) however. In Lass (2004:376) a raised and fronted [ʉ̞] is, however, given for young *CulSAE* speakers. The most typical value for GenSAE speakers is given as [ʊ].

Wells (1982), Branford (1994), Bowerman (2004), Da Silva (2007): Wells (1982:615) claims that STRUT in SAE is “central to back in South African English (as against central to front in Australian English)”. Branford (1994:475) stresses the wide phonetic range of this vowel, giving [ɐ] or [ʌ̈] for CulSAE, with a range in the region of [ɐ ∼ ɜ] as being more typical of GenSAE. According to Bowerman (2004:937), STRUT in SAE is “typically a low to mid, centralised vowel ... [ä] to [ɐ]”. Da Silva (2007:188) seems to imply [ʌ] and a low-front [a] as constituting two main variants, although her use of [ʌ] to encode “close-mid or mid”(Da Silva 2007:188) is somewhat uninterpretable[28]. In terms of her results it appears that the [ʌ]-variant is used categorically

[28]The author possibly meant open-mid or mid in which case [ɐ] would have been a more accurate, less broad transcription. On the other hand, in her Table 8 (Da Silva 2007:114), the author gives [ʉ] as one possible variant of STRUT; the latter seems extremely unlikely.

in her primarily ‘white’ lect 1 (Da Silva 2007:159)[29].

With respect to LOT-CLOTH, Branford (1994:476) stresses the weak lip-rounding and centralization of this vowel in SAEP. An [ʌ̈]-value is linked to Natal speakers in particular. Branford (1994:476) also claims that [o] is found in the Eastern Cape. Bowerman (2004:937) gives [ɒ̈] to [ɔ] as the range of LOT in SAE. He also mentions the GenSAE [ʌ̈]-value prevalent among “younger Cape Town and Natal speakers” (Bowerman 2004:937).

For FOOT, Bailey (1984:16; my parenthesis) confirms the possible confusion between KIT and FOOT in the following rather awkward passage:

> “The short stressed vowel /u/ seems to be realized as /i/ with its central allophone, before /i/ [meant no doubt to be /l/]. Young people in Durban don’t all distinguish between ‘pill’ and ‘pull’ and have some difficulty spelling ‘fulfil”’.

As mentioned in §7.3, Bailey (1984) fails to distinguish between the quality of the unmarked schwa-like KIT allophone and KIT before final /l/. That there is in general an overlap between KIT and FOOT is, however, given further confirmation. Branford (1994:476) also mentions variation between the ‘standard’ [ʊ] and a more centralized [ʊ̈]. According to Bowerman (2004:937) this vowel has little variation and has [ʊ] as its canonical value. He adds, however, that “there is very little lip rounding relative to other L1 varieties of English world-wide. The pronunciation [ʊ̜] (with added lip-rounding)[30] is associated with Broad” (Bowerman 2004:937).

With respect to the distribution of CLOTH words in SAE, Bailey (1984:19) mentions the possibility of a THOUGHT vowel being used in CLOTH words such as *cross*, *loss*, *wrath* and *salt* but implies that it is restricted to highly-conservative accents of SAE. This author confirms the variable use of THOUGHT and LOT vowels in words such as *Austrian*, *Australia* and *auction*, and adds the following:

> “the /oo/ is regularly replaced by /o/ in ‘always, also’ in the speech of some people with British backgrounds that I have heard in South Africa. This indicated to me that the sound change has been more effective in RP than in SAE”.

[29] It is suspected that what the author was ‘picking up’ on here was the distinction between a central-to-front less-than-fully open lax vowel (on Labov’s (1994) non-peripheral track), Da Silva’s (2007) lect 1 [ʌ]-variant, as opposed to a central, fully-open tense vowel (on Labov’s (1994) peripheral track), Da Silva’s (2007) lect 2, predominantly ‘black’, [a]-variant.

[30] Bowerman (2004) thus, no doubt, meant [ʊ̹].

Of more interest is this author's observation that particularly before voiceless consonants and in unstressed position, a shortened form of THOUGHT (i.e. [ɔ] or [o]) is used (variably with LOT) in words such as *forty*, *off* and *'cause* (from *because*), as well as sometimes in *Austria*, *Australia* and *auction*. Thus the author mentions that "some young speakers in Durban have a contrast between 'horse' [hɔs] and 'hoarse' [hɔːs]" (Bailey 1984:19). The author tentatively attributes this feature to the influence of Afrikaans which has a short [ɔ]-sound in words such as *hok* ('cage'), *moffie* ('homosexual'), *pomp* ('pump', but used colloquially to refer to sex) etc., many of which are loan-words in SAE, often used in slang. While SAE is, like RP, generally a flat-CLOTH accent, Lass (1990:278) does claim the occasional use of an [oː]-like value in some lexical items in this set, particularly before /f/. He, for example, claims that "virtually all speakers have /oː/ in *off*, many in *cough*" (Lass 1990:278). The tendency though is that the younger the speaker the more likely it is that CLOTH words will have the same vowel as LOT.

8.4.3 SLF: A Comparative Acoustic Analysis

STRUT in Figure 5.4 on page 151, which reflects Webb's (1983) acoustic data, appears to have a central position between TRAP and LOT and a fair degree of internal variation. Webb (1983:145) makes the following claim:

> "/ʌ/ has 2 allophones, viz. a retracted [ʌ] before nasals, and the [ʌ] elsewhere. The /ʌ/ in young is due to the preceding front glide".

Figure 8.21 on the next page provides a visual representation of the various STRUT 'allophones', as analyzed by Webb (1983). The position of *young* is clear, while the STRUT vowels before a nasal and the unconditioned STRUT tokens are respectively represented in Figure 8.21 on the following page by STUT$_{nas}$ and STRUT$_{un}$. Frankly, the data is not convincing enough to warrant the conclusion of there being two allophones. While the mean of STUT$_{nas}$ is certainly more retracted than STRUT$_{un}$ this might well be a matter of chance or measurement error, particularly since STRUT$_{un}$ was only based on the measurement of four tokens (*rugby*, *tough*, *us*, *mother*) with STUT$_{nas}$ based on the measurement of nine tokens[31].

[31] There seems, furthermore, to be some confusion in this regard in Webb (1983:142–3). In this author's three tables for STRUT the author places *young* under table 1, two tokens of *one* under table 2 and the rest of the tokens including *scrum*, *function*, *money*, *run*, *some*, *son* and *some* under table 3. The different tables are, presumably, as they are elsewhere, meant to indicate different allophones and, indeed, the F_2 of table 2 is far lower than the F_2 of table 3 (990Hz vs 1216Hz) – and it is perhaps this difference that Webb (1983) is referring to. Yet, the majority of tokens in table 3 contain a STRUT vowel before a nasal.

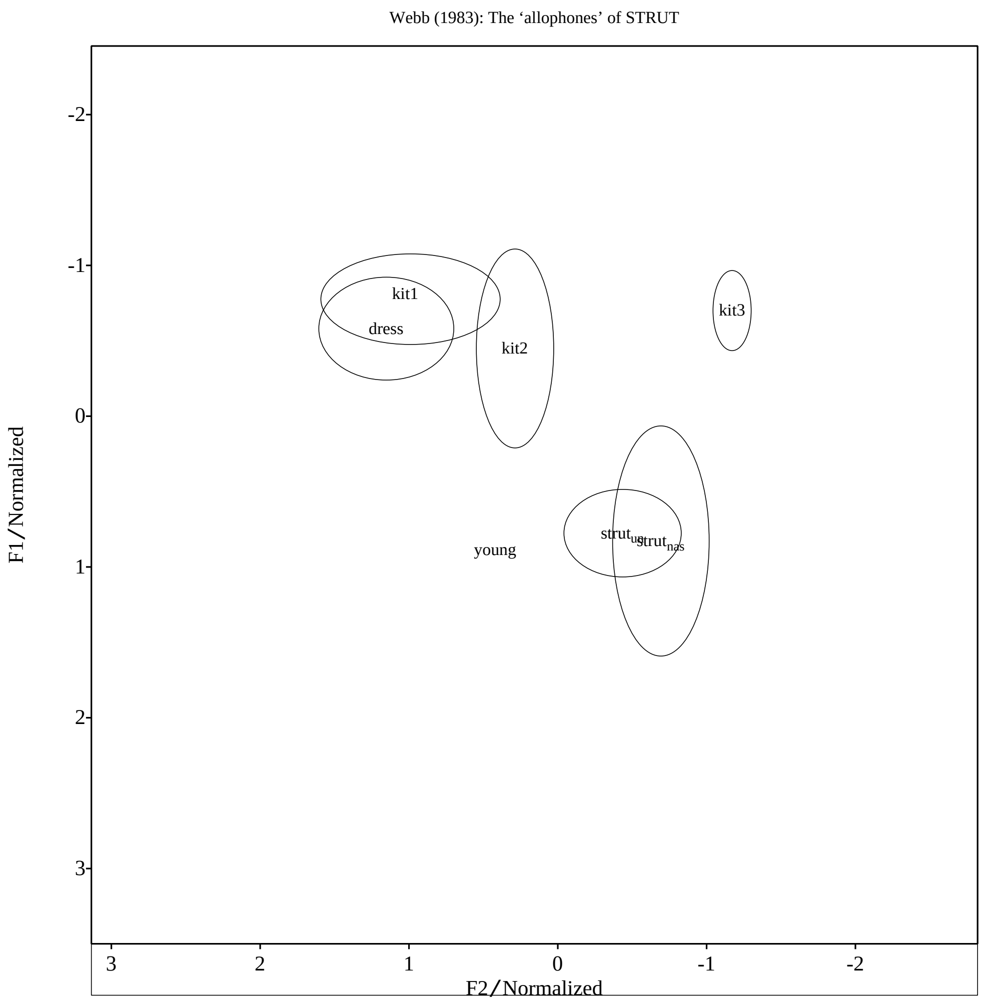


Figure 8.21: Webb’s (1983) STRUT allophones

Figure 5.4 on page 151 also provides a graphic illustration of Webb's (1983) acoustic data for LOT. According to (Webb 1983:149), "there is a considerable overlap between the long vowel /ɑ/ ... and /ɒ/". This state-of-affairs is represented in Figure 8.22, which is essentially Figure 5.4 with the BATH vowel added. Given Lass' (1990) claim for derounding of the LOT vowel in SAEP there is a distinct possibility, therefore, that LOT and BATH are, in many idiolects of SAE, only distinguishable in terms of length.

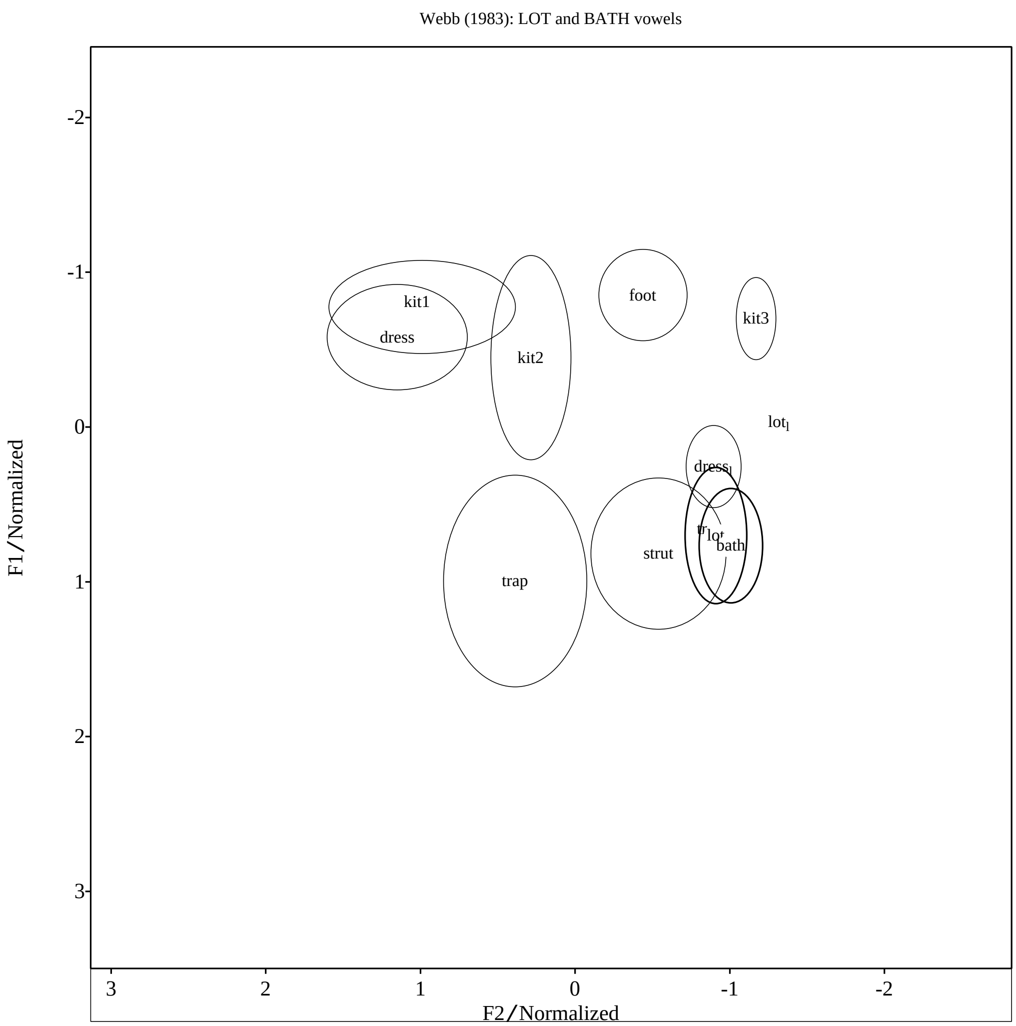


Figure 8.22: Webb's (1983) BATH and LOT Vowels

With respect to Webb's (1983) acoustic analysis of FOOT – see Figure 5.4 – it is surprising that he does not mention the rather fronted quality of this vowel. One notes for example that the KIT vowel before /l/, generally often said to be confusable with the FOOT vowel

Lexical Set	**No. of Tokens**	**No. of Measurements**	**N-G Tokens**	**O-G Tokens**
STRUT	203	4466	*hud*, *bus*, *buzz*, *buck*, *but*, *bud*, *bun*, *snuff*	*but*, ***mo**ther*, *tuck*, *suck*, *duck*, *stuck*, *cut*, *pup*
STRUT before /l/	23	506	*skull*	*skulk*
LOT-CLOTH	94	2068	*hod*, *hot*, ***co**py*	*pot*, ***co**mma*, *cot*, *got*, *Scott*
FOOT	63	1386	*hood*, *foot*	*foot*, *look*, ***pu**lley*
FOOT before /l/	28	616	*full*	*pull*

Table 8.1: Summary of STRUT, LOT-CLOTH and FOOT data

(e.g. *bill* heard as *bull* by non-SAE speakers), is far more retracted than FOOT.

Moving on to the current data, Figure 6.13 on page 197 provides a clear picture of the positions of STRUT, LOT and FOOT vs. the other SAE short vowels. The various positions are based on the data provided in Table 8.1.

Beginning with STRUT, it is clear from the representation provided in Figure 6.13, that the slight spectral movement evident in the data is in all likelihood due to coarticulatory processes. More important is the relatively fronted position of STRUT *vis-à-vis* LOT and its relative closeness to TRAP.

Figure 8.23 on the next page provides a comparison between the STRUT data collected as part of the current project and the data collected as part of Webb (1983), as originally given in Figure 5.4 on page 151 and Figure 8.21 on page 340. While there is a substantial degree of overlap there is also some indication that STRUT has moved a little more down and forward since Webb (1983). As was mentioned above, Webb (1983) postulates slight retraction of STRUT before nasals (‘ws$_n$’ in Figure 8.23)[32]. Figure 8.24 on page 344, on the other hand, illustrates the relationship between STRUT before nasals and unmarked STRUT in the current data. There was only one token of STRUT before a nasal (i.e. *bun*) giving a total of 18 tokens for this putative allophone; results are thus somewhat tentative. Still, from Figure 8.24 it is clear that, for the current data at least, there is little evidence for STRUT retraction before a nasal.

[32]Unmarked STRUT in Figure 8.23 is labeled ‘ws$_u$’ and the token *young* is ‘ws$_y$’.

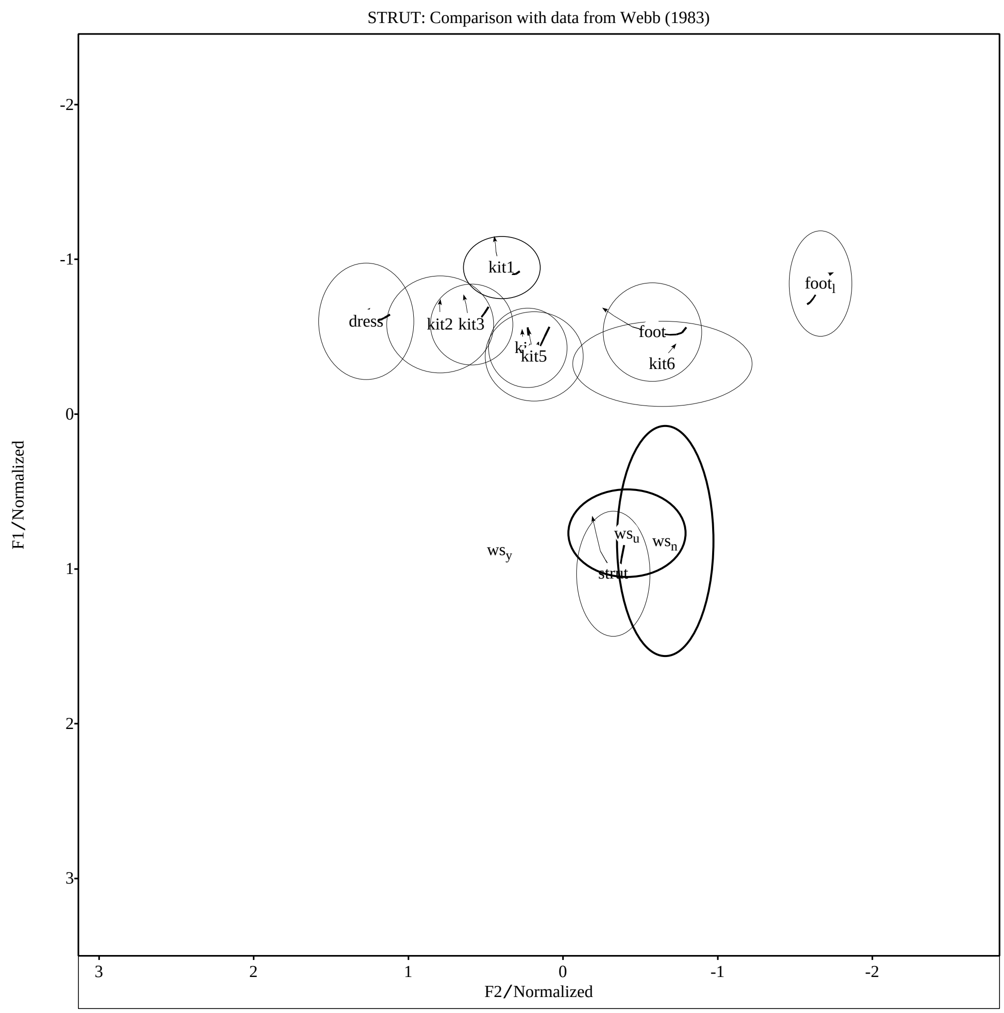


Figure 8.23: STRUT: Comparison with Webb (1983)

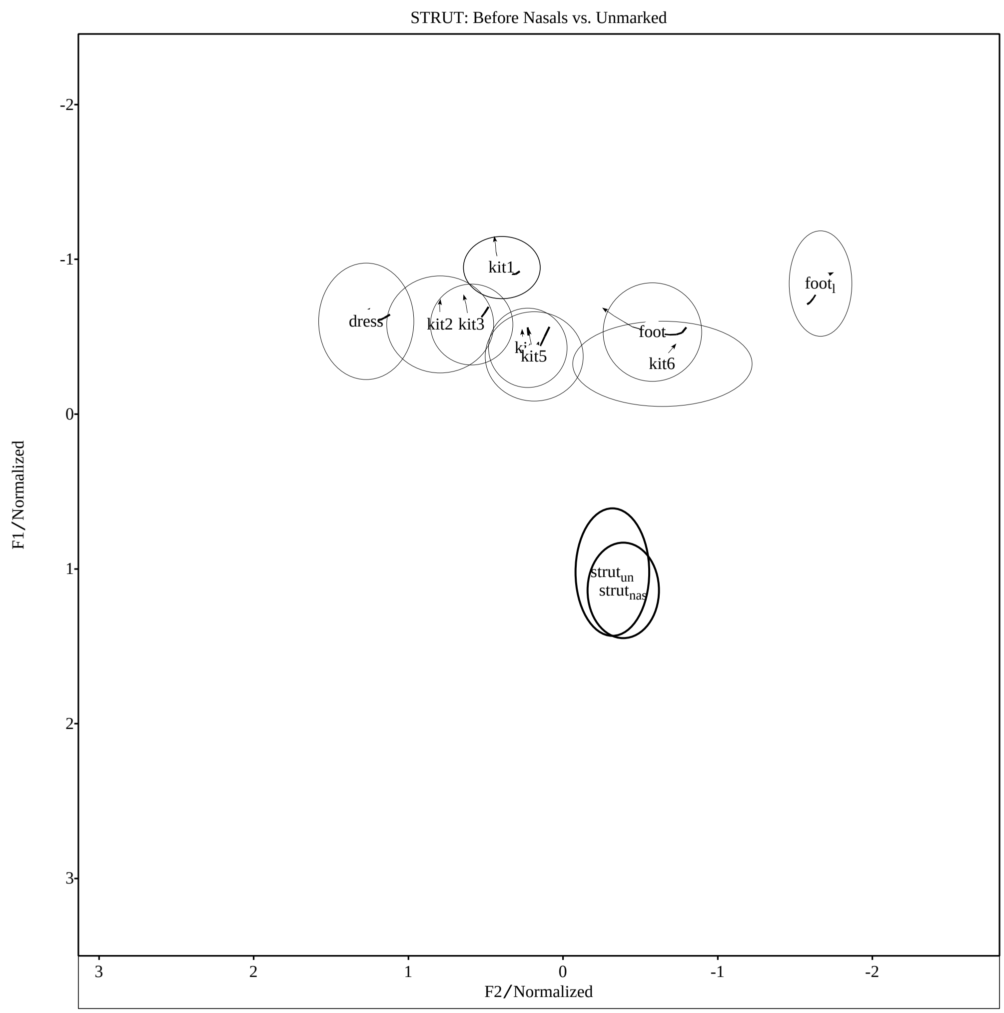


Figure 8.24: STRUT: Before Nasals vs. Unmarked

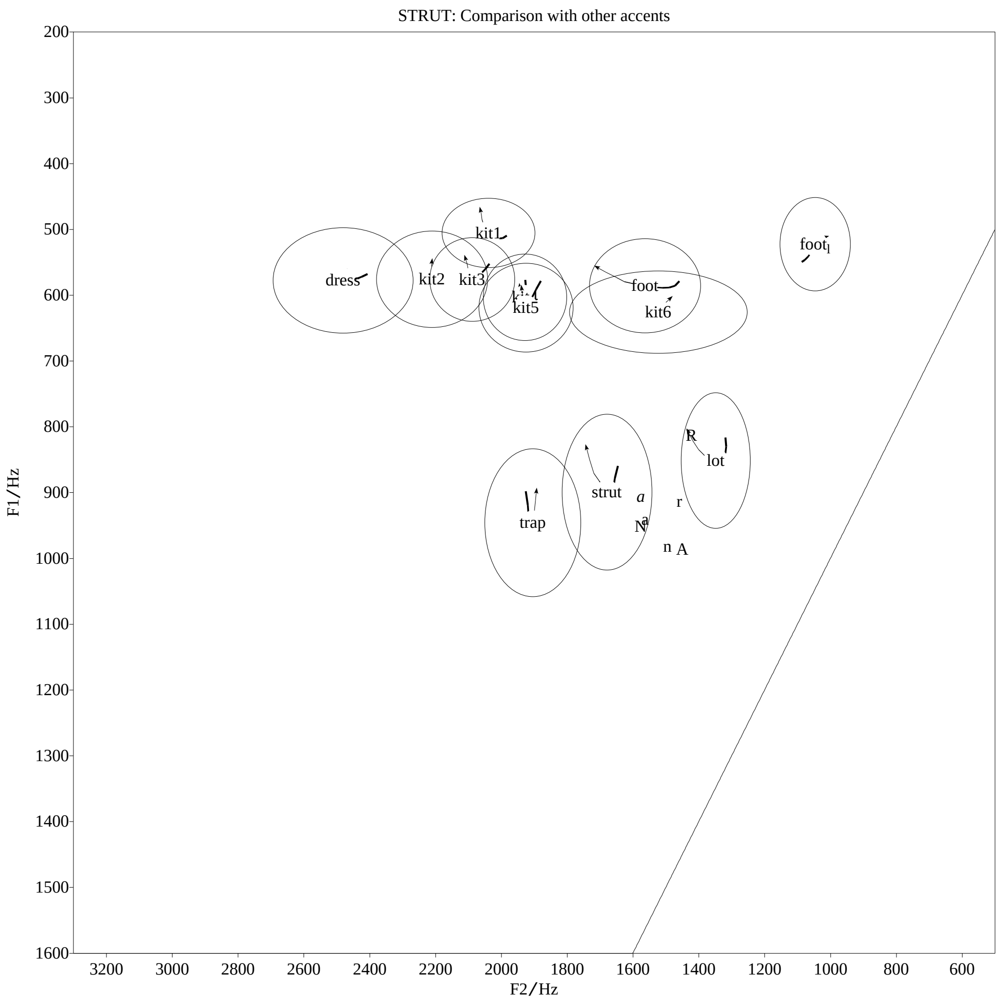


Figure 8.25: SAE STRUT vs. other English accents

Returning to the issue of possible fronting of STRUT in SAEP, Figure 8.25 on the previous page compares SAE STRUT with STRUT found in RP and the other Southern Hemisphere Englishes. The results are particularly illuminating for two reasons. Firstly, we find an overlap between SAE LOT and RP STRUT (both in citation and connected-speech form, but particularly in the former). This, of course, lends credence to the notion that SAE LOT has undergone a substantial process of unrounding, at least for the lect under investigation. Secondly, we note the substantial forwarding (and lowering) of AusE and NZE STRUT in comparison with the RP-values, as already mentioned in the impressionistic literature and reviewed in §8.4.1. More importantly, we note the relative frontedness of *SAE* STRUT. Although perhaps not as lowered as the NZE and AusE data, STRUT in SAE is even more fronted than in the other two Southern Hemisphere Englishes. The provides substantive support for the various observations reported in §8.4.2 of a front and *relatively* high value for STRUT being common among young SAE speakers. Furthermore, the fact that STRUT is found to be so fronted in citation-form style means that it is at least possible that SAE STRUT has undergone even further fronting in less formal styles. The fronted (although not lowered) nature of STRUT is also evident from the slight overlap between STRUT and TRAP, as in evident in Figure 6.13 on page 197. In a previous publication, Bekker and Eley (2007), this overlap was attributed to possible TRAP lowering, but it appears that it is, in all likelihood, due to TRAP-lowering *and* STRUT-fronting. Lastly, it will be recalled that, in §3.4.3, Trudgill (2004:135–5) is quoted as claiming that “STRUT fronting is not a prominent feature of South African English”. It appears that Trudgill (2004) was incorrect in this regard; with respect to citation-form data at least, SAE appears to show *more* fronting than AusE or NZE, although not as much lowering.

Figure 8.26 on the next page compares the SAE STRUT data with that from Torgersen and Kerswill (2004). The bullets show mean values for each older-generation female, while the diamonds show data for the younger females. As was mentioned in §2.14, STRUT retraction is seen as part of SECS, and it is clear from this figure that SAE is far behind the southeast of England in the implementation of this retraction. It would appear that SAE is in fact the most conservative dialect with respect to this phenomemon in the sense that it retains the traditionally fronted Cockney value. While in the southeast of England, “STRUT fronting had been reversed by the middle of the 20th century” (Torgersen and Kerswill 2004:32), and while RP, AusE and NZE all show evidence of similar retraction, SAE retains a fully-front variant. This is, however, perhaps a very recent development, if the impressionistic literature is anything to go by (see §8.4.2) and, perhaps, explains why Trudgill (2004) does not mention STRUT-fronting as a feature of SAEP; thus older SAE generations still show signs of retracted STRUT – evident in Figure 8.23 on page 343 where the current

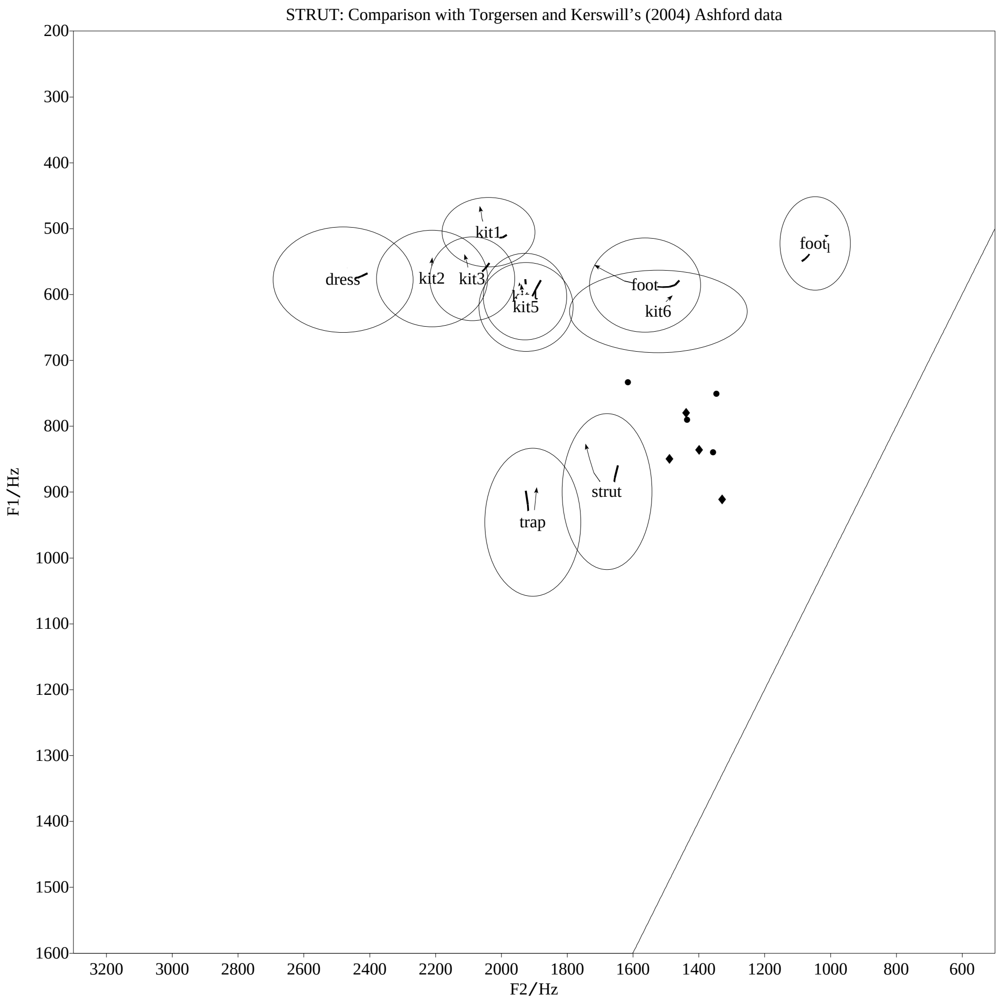


Figure 8.26: SAE STRUT: Comparison with Torgersen and Kerswill's (2004) Ashford data

data is compared with that of Webb (1983). Thus, interestingly enough, older SAE speakers would have values more similar to younger speakers from the southeast of England; only that in SAE the retracted STRUT values are, as it were, first-generation values, while the Ashford speakers, for example, are showing 2nd-generation STRUT retraction.

Figure 6.13 on page 197 shows the LOT-CLOTH vowel in its standard low back position. The citation-form data only includes LOT tokens, but as per the discussion in §8.4.2, the majority of CLOTH words in SAE can be assumed to have an identical value. While LOT-CLOTH in Figure 6.13 shows some evidence of spectral movement this is, in all likelihood, a simple coarticulatory effect to and away from the vowel target, as is common particularly in open vowels.

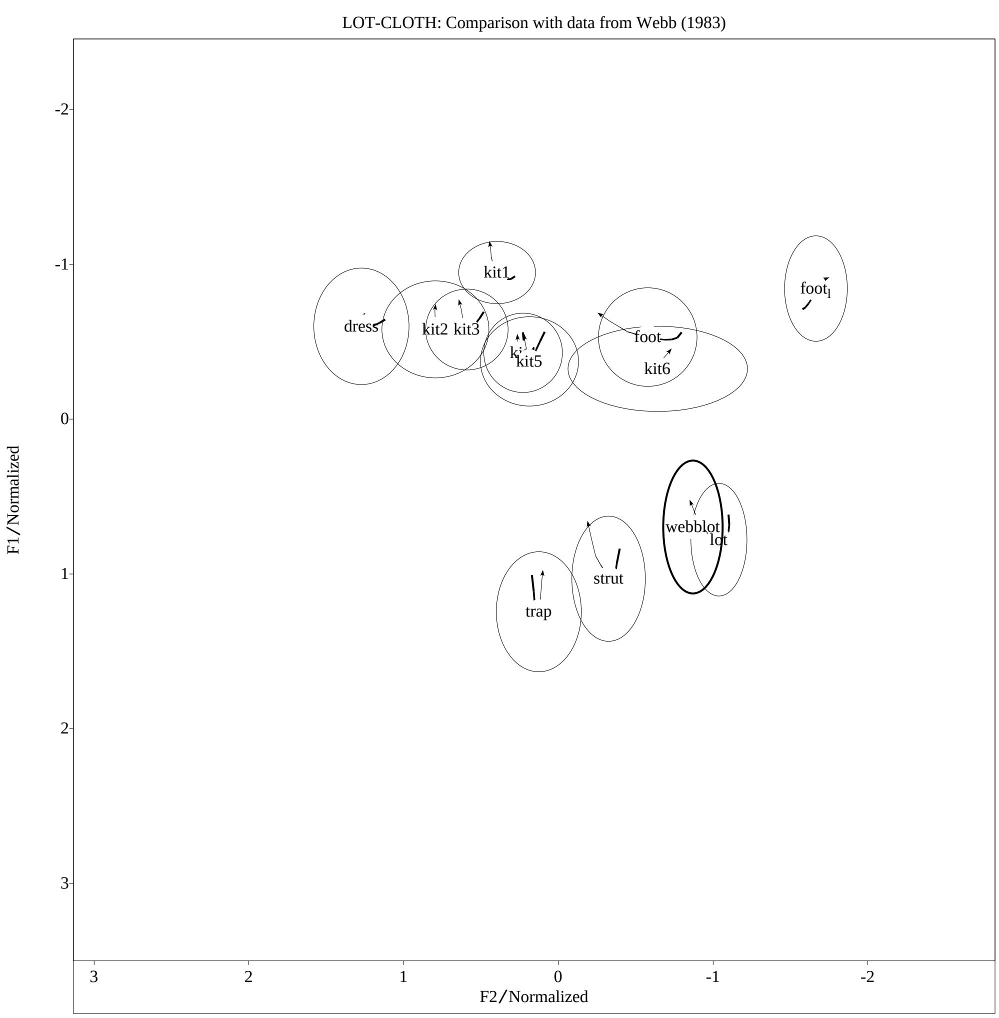


Figure 8.27: LOT-CLOTH: Comparison with Webb (1983)

Figure 8.27 on the facing page provides a direct comparison between the LOT-CLOTH data collected for this research project and that of Webb (1983) as originally provided in Figure 8.22 on page 341. In this figure, Webb's (1983) data (mean and standard deviation) is labeled as 'webblot'. As can be seen, there is a substantial degree of overlap between the two sets of data with respect to the LOT-CLOTH vowel. This is somewhat surprising, given that the current data is based on citation-form tokens derived from young, female subjects, while Webb's (1983) data was garnered from a conversation held between the author and a colleague. The author (i.e. Webb) is male and the recording was done over twenty years ago. What this seems to suggest is that LOT-CLOTH in SAE is a substantially stable phoneme i.e. it does not undergo much style-shifting and has not changed its basic phonetic value during the course of the last few decades. It would appear to also be insensitive to gender-differences.

In addition, as shown in Figure 8.22 on page 341, Webb's (1983) data shows substantial overlap between the BATH and LOT sets, the phonemic difference being 'carried', no doubt, by a distinction in length. Figure 8.28 on the next page provides a similar comparison between LOT and BATH for the current data. The relationship between these two vowels is remarkably similar across the two data sets and confirms some of the impressionistic data provided in §8.4.2 i.e. with respect to the unrounding of LOT (and perhaps slight rounding of BATH) and its similar value (in terms of quality) to BATH. The phonemic distinction between these two phonemes is maintained, of course, by a difference in length. The average duration for BATH in the current data is 227ms while the average duration for LOT is 116ms, a difference of 111ms; more than enough duration to maintain a phonemic distinction.

Turning now to a comparison between the SAE data and acoustic data from other accents of English, Figure 8.29 on page 351 provides rather surprising data. What is particularly intriguing is that, while SAE LOT is reported on in the impressionistic literature (§8.4.2) as somewhat 'raised' above the standard RP [ɒ]-level, Figure 8.29 hardly provides evidence for such raising; in fact the contrary appears to be the case. What it does, however, provide is evidence of some degree of centralization of this vowel in SAE, as has been confirmed in the impressionistic literature as well. Similar centralization has been noted for other accents by Easton and Bauer (2000:107), who, in their comparison of acoustic data-sets across RP, AusE and NZE, have the following to say about the LOT-CLOTH vowel:

> "If F2 figures alone are considered . . . there seems to be a slight general shift in LOT to a fronter (probably better described as a more centralized) position . . . This trend is visible in RP, with the pronunciation in context . . . showing a greater shift than isolated-

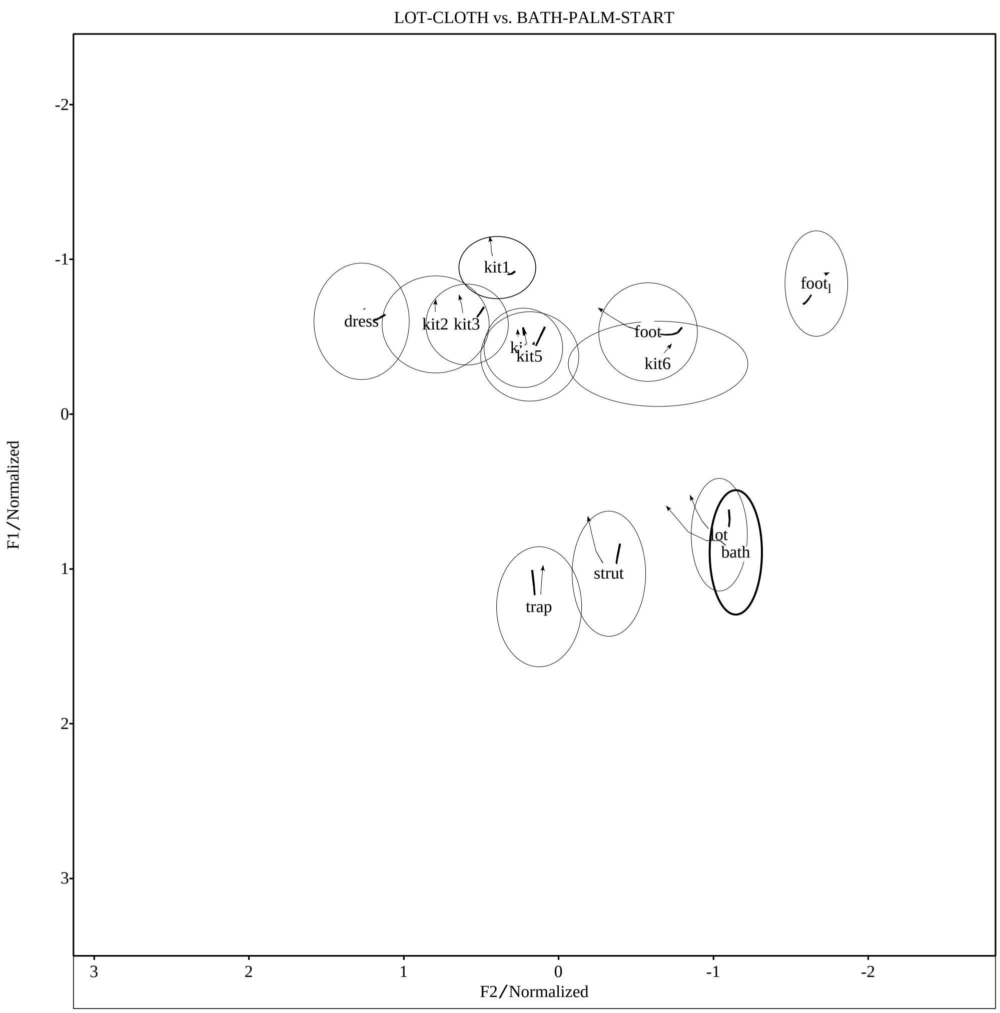


Figure 8.28: LOT-CLOTH vs. BATH-PALM-START

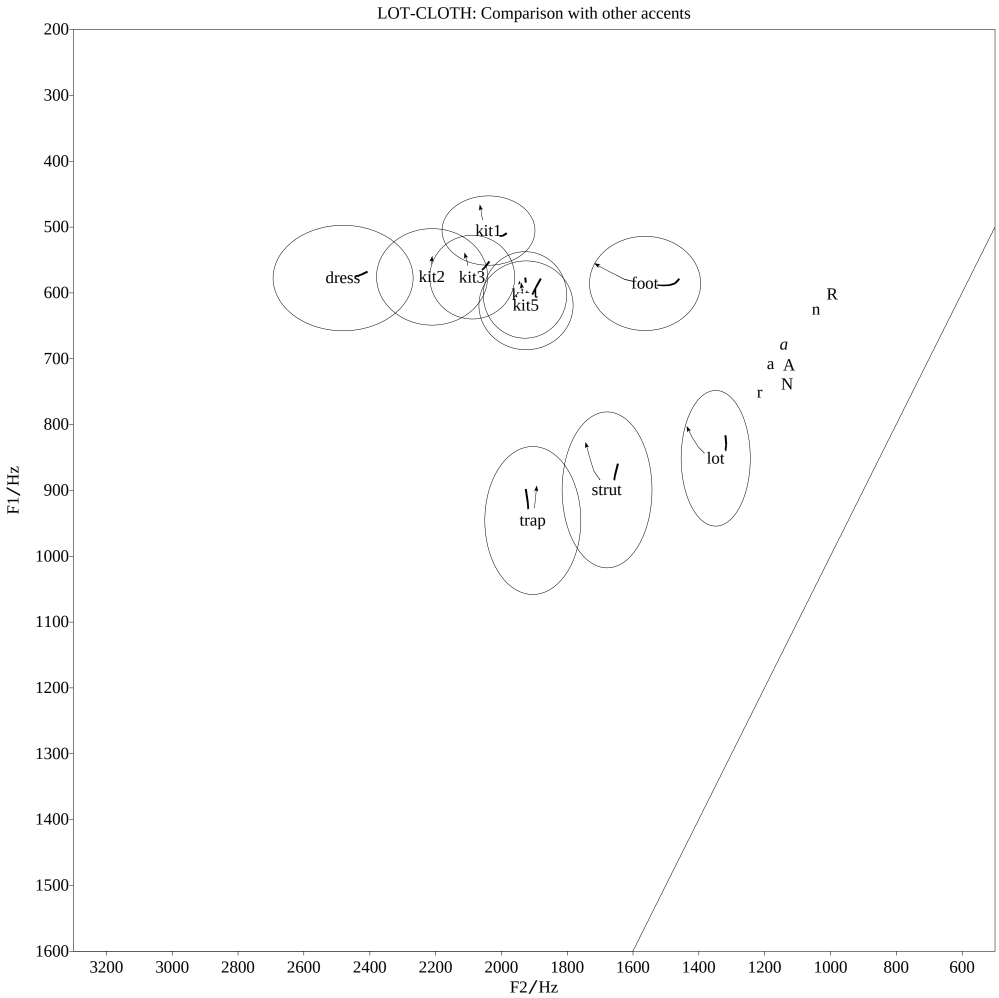


Figure 8.29: SAE LOT-CLOTH vs. other English accents

word pronunciation . . . However if (F2-F1) values are considered this trend is no longer visible".

The difference between the connected-speech (r) and citation-form (R) RP data, as illustrated in Figure 8.29, clearly confirms this observation.

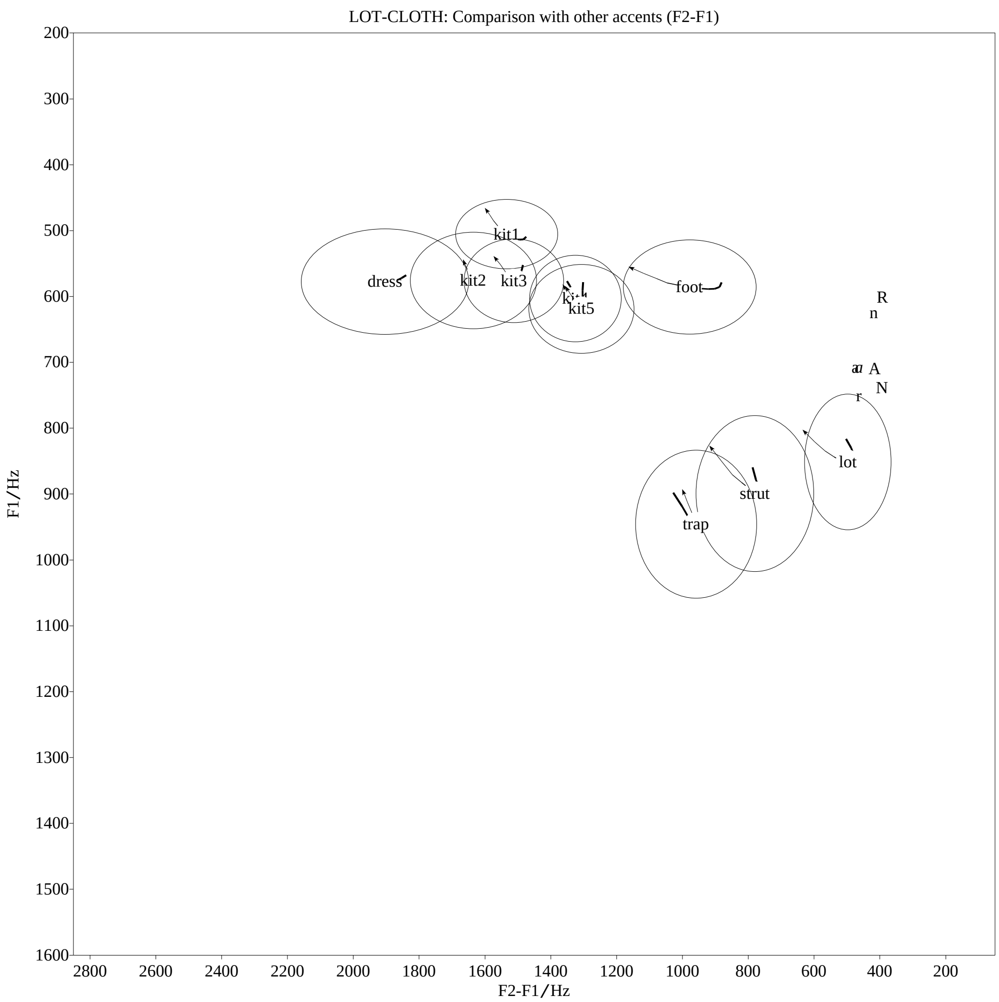


Figure 8.30: SAE LOT-CLOTH vs. other English accents (F2-F1)

To test Easton and Bauer's (2000) claim regarding the use of F2-F1 values as opposed to F2 values, the various data sets were subject to the relevant transformation and plotted in Figure 8.30. Note that in this figure the Australian data from Cox (2006) (a) and Butcher (n.d.) (*a*) overlap. While the data in Figure 8.30 certainly places the various realizations of LOT-CLOTH on an 'equal footing' in terms of backness, it certainly does not explain the relative positions of these vowels in terms of height. Easton and Bauer (2000:107) ‘fudge’ this issue by concluding that “female speakers show considerable variation in F1 values, but this does not appear to correlate with the national variety spoken”. What is required is a principled explanation for the relative position of these accents in terms of the height of the LOT-CLOTH vowel, particularly in light of comments such as the following from Watson, Harrington and Evans (1998:203) in their acoustic comparison of AusE and NZE: “HOD is transcribed with [ɔ] to reflect that this vowel is somewhat raised relative to RP; since our experimental data suggest very few differences between the NZE and AE qualities for this vowel, we have proposed the same symbol for NZE”. While their NZE (N) and AusE (A) data certainly do match, their description of the NZE and AusE qualities *vis-à-vis* RP, particular given the data collected previously by Deterding (1990), does not convince. Either they are incorrect, did not bother to consult Deterding (1997) in which the citation data from the earlier Deterding (1990) is provided, or Deterding’s (1990) data is corrupt on some serious level. The last option is, however, unlikely given that a principled explanation lies in the phenomenon of lip-rounding. As discussed in §2.2.2.4, an increase in lip-rounding generally leads to a lowering of F_1 i.e. to greater (acoustic) vowel height. If we thus view the F_2-F_1 chart in Figure 8.30 on the preceding page as giving a relatively accurate representation of the (articulatory) front-back dimension then it appears probable that the attested differences between the LOT-CLOTH values for these various accents of English is mainly the result of degree of lip-rounding, with citation-form RP having the highest degree of this feature. It appears, then, that SAE has kept pace with, or led, a change towards unrounding of LOT-CLOTH affecting all the Southern Hemisphere Englishes. RP, given its by-definition conservative character, has lagged behind, although the tendency is prevalent in the unrounding of LOT-CLOTH in connected-speech. The unrounding of LOT-CLOTH in SAE has been explicitly mentioned in §8.4.2.

Unfortunately, Torgersen and Kerswill (2004) do not provide acoustic data for LOT for their female Ashford speakers.

Figure 6.13 on page 197 provides a clear picture of the position of FOOT vs. the other SAE short vowels. What is particularly striking about FOOT in this figure is its fronted position. Recent impressionistic analyses, as reviewed in §8.4.2, have, however, pointed out a fronted FOOT as being typical of younger (particularly female) General and Cultivated SAE speakers.

This suggests that FOOT in the current data should be more fronted than that provided in Webb (1983). Figure 8.31 on the next page thus comes as somewhat of a surprise, with Webb’s (1983) data in fact showing, if anything, slightly more fronting of this vowel. There is enough overlap though to question any significant difference here, a view augmented

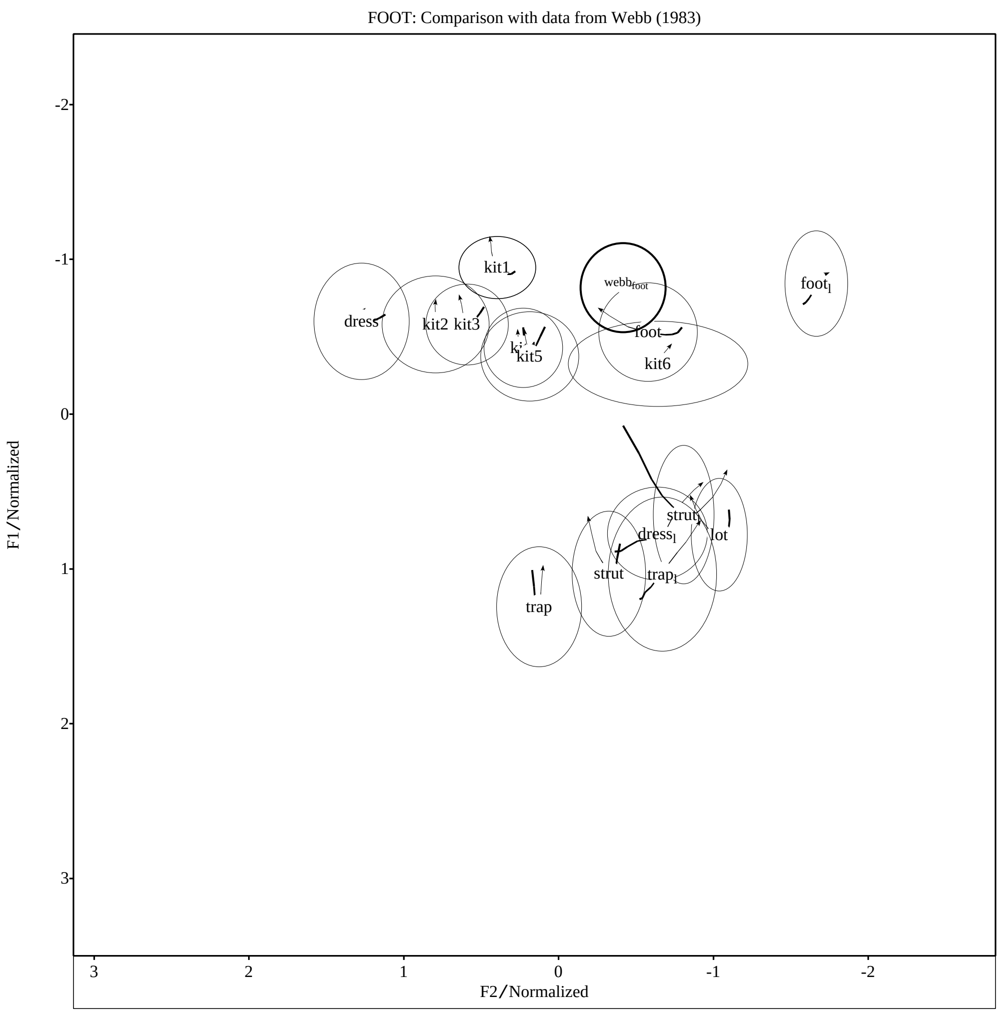


Figure 8.31: FOOT: Comparison with Webb (1983)

by the fact that Webb (1983) bases his analysis on only 9 tokens. What this overlap does suggest, however, is that a fronted FOOT has become an indicator of SAE – with very little style shifting and substantially below the level of consciousness. There is nothing in the literature suggesting that there is any overt awareness of this feature of SAE; given the data from Webb (1983) it is also quite possible that impressionistic analyses had failed to pick this movement up until recently; and even then that its full extent has been underestimated.

A further feature evident in Figure 6.13 on page 197 is that FOOT has a definite diphthongal quality to it i.e. there is substantial movement towards a more front position. All in all, therefore, there is substantial evidence here to show that FOOT has undergone and is perhaps still undergoing FOOT-fronting.

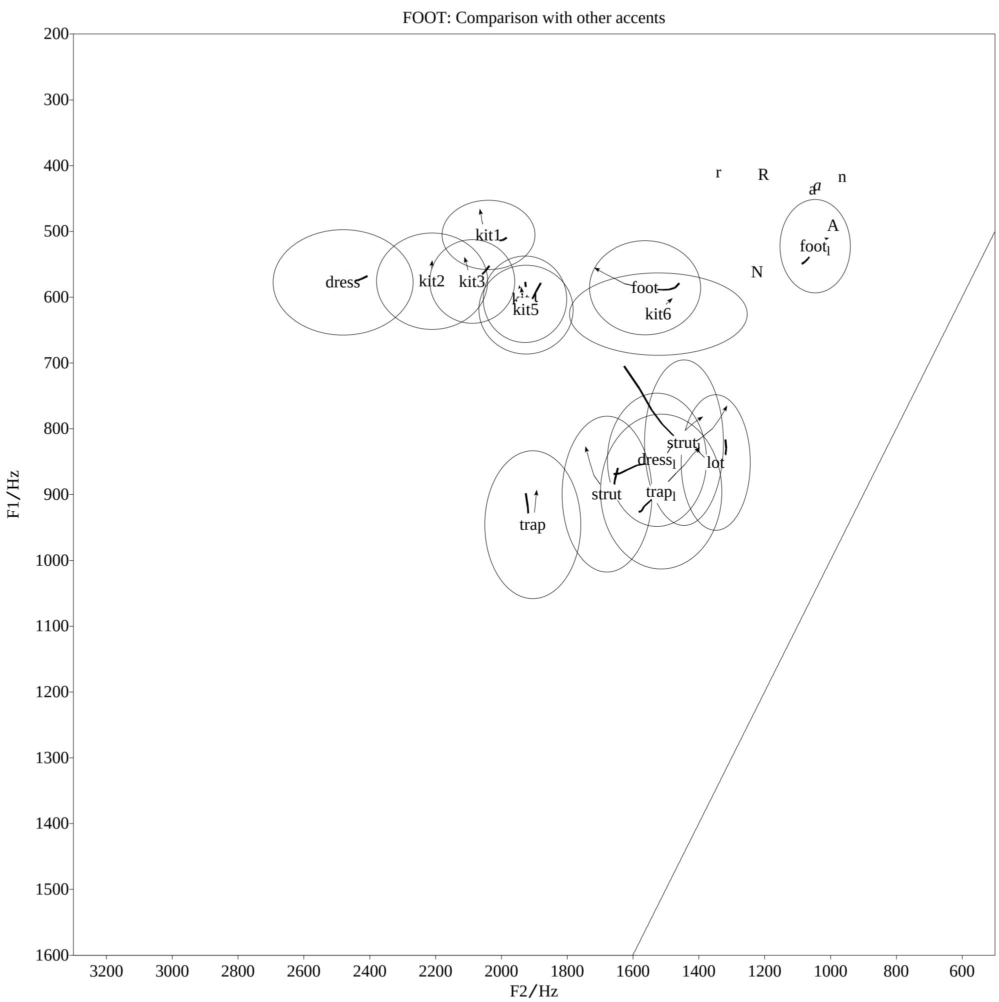


Figure 8.32: SAE FOOT vs. other English accents

The degree of fronting becomes even more apparent when we compare the SAE FOOT data with data from other accents of English. Figure 8.32 on the previous page provides the relevant comparison. From this figure it appears that SAE is far in advance in terms of FOOT-fronting in comparison to RP or even the other Southern Hemisphere Englishes.

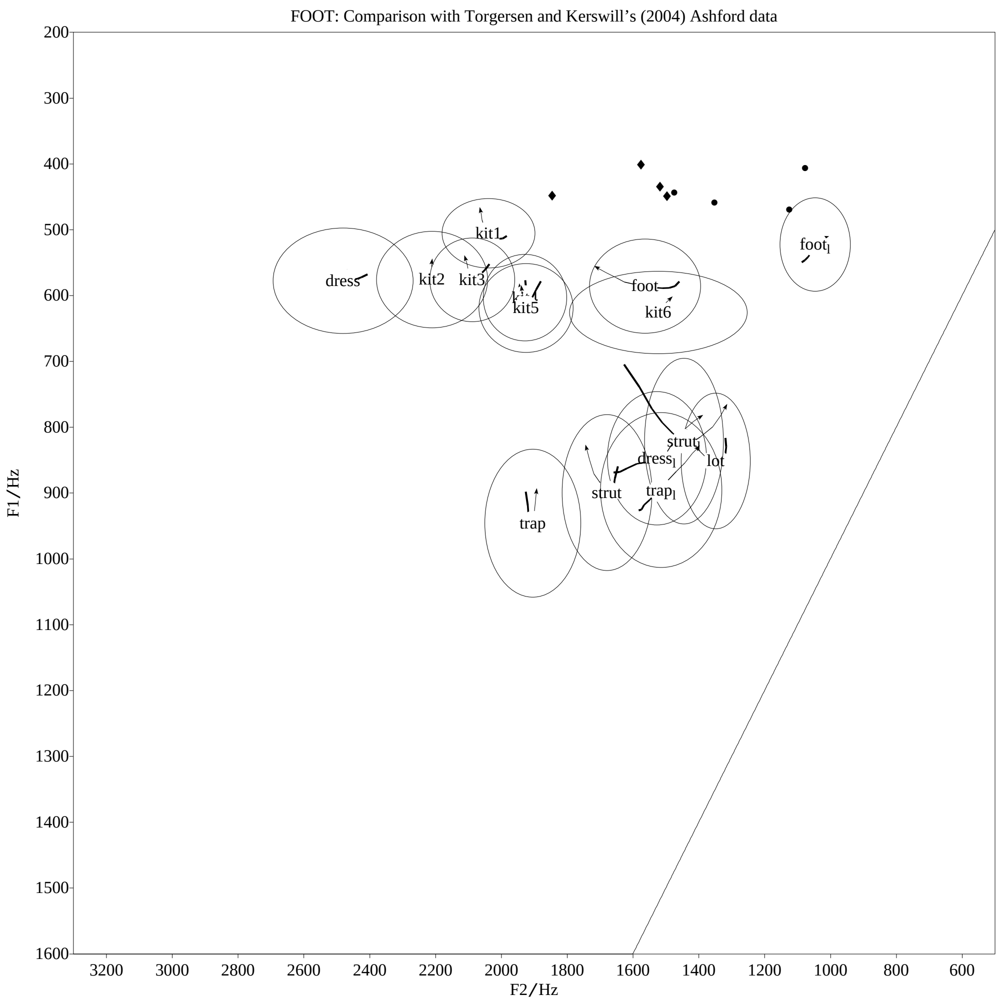


Figure 8.33: SAE FOOT: Comparison with Torgersen and Kerswill's (2004) Ashford data

Figure 8.33 provides a comparison between the SAE data and Torgersen and Kerswill's (2004) Ashford data, with the diamonds showing the FOOT values for the younger females. As can be seen, the degree of FOOT-fronting in SAE appears to be on par with that of contemporary values in the southeast of England. What is markedly different, however, about the SAE data, is the degree to which FOOT has been subject to simultaneous lowering, almost in direct analogy with KIT.

8.4.4 SLF: Synopsis

Beginning with STRUT, the acoustic analysis given above, provides compelling evidence, *contra* Trudgill (2004), that STRUT, at least in modern GenSAE, is substantially fronted. More generally, there is also *some* evidence to suggest that this is a change-in-progress, given the differences between the current data and that provided by Webb (1983).

As stated in various sections preceding this one, STRUT is viewed as an integral component of Torgersen and Kerswill's (2004) SECS-Shift. An overall comparison of the current data with these authors' Ashford values, as well as further acoustic data derived from RP, AusE and NZE, all seems to support the hypothesis that SAE is behind these other accents in the implementation of this shift. In particular, STRUT has not undergone the secondary STRUT-retraction that seems to result, in these other accents, from the lowering of TRAP, the initiating move of the SECS-Shift.

The hypothesis that a fronted STRUT is a recent phenomenon in SAE is bolstered by the observation that STRUT-fronting in London (Cockney) was only complete by the middle of the 20th-century. As such, SAE, along with the other Southern Hemisphere Englishes, was likely to have received, as input, a relatively retracted variant. Its subsequent fronting, in all these varieties, was thus an endongenous development, a fact confirmed by Trudgill (2004), and as outlined in §3.4.3. Furthermore, SAE seems to have been the most conservative in terms of the implementation of this process, perhaps as a result of the greater influence of RP and the resultant stigmatization of Cockney-like forms such as fronted STRUT. However, since the post-WWII 'waning' of CulSAE (and the more general influence of RP), STRUT-fronting has, on this reading, been 'set free' for further endogenous development.

Moving on to LOT, we note the relative stability of this vowel with respect to the two SAE data sets, suggesting that this is an indicator of GenSAE. We also note the *qualitative* overlap between LOT and BATH, suggesting either or both the weakening of lip-rounding in the former and slight lip-rounding in the latter. The hypothesis that LOT in GenSAE has a quasi-American-like quality is supported by the comparison of the SAE data with that derived from other accents. In particular, and as noted in the discussion that accompanies Figure 8.29 on page 351, a reduction of lip-rounding is generally linked to a reduction of vowel height. It is, of course, also linked to a greater degree of *acoustic* fronting. This provides a principled explanation of the difference between SAE and the other English accents.

While Torgersen and Kerswill (2004) do not provide relevant acoustic data for LOT, it is clear that a retracted LOT forms part of the SECS-Shift as well[33]. As such, a relatively

[33] See Figure 8.20 on page 331.

fronted (through unrounding) LOT vowel would make sense in the SAE context, given the relatively conservative nature of SAE in this regard, and, in particular, given the fronted nature of STRUT and the lack of secondary STRUT-retraction. We have also noted, in the impressionistic literature reported on in §8.4.2, that there have been a number of recent observations, particularly by Lass, of an unrounded (and thus at least *acoustically* fronted) LOT vowel,[34] particularly among younger GenSAE speakers. Any related heightening of tongue position in GenSAE, as claimed in the impressionistic literature, has no doubt been obscured by the acoustic effect of unrounding.

On the other hand, the relatively *categorial* nature of the LOT value across the two SAE data sets raises the suspicion that this is not a change-in-progress, but has been inherited from the past. In this regard, it is worth mentioning the unrounded variants reported for CulSAE and, even more speculatively, that the southwest of England (including Cornwall) and Norwich have unrounded variants[35]. Furthermore, as reported on in §8.4.1, Trudgill (2004:48) claims that "unrounded LOT was formerly much more common in Britain than it is today". Completing this picture is the fact that raised, heightened and tensed Cockney-like varieties have been reported for Broad varieties of SAE, a feature shared, *perhaps*, by at least AusE. While the etiology is most unclear here, there at least enough evidence to speculate that unrounded LOT variants entered into the Johannesburg feature pool from a variety of sources, that in the establishment of GenSAE there was a prestige-driven aversion towards *extreme*, perhaps Afrikaans-influence, tensing and rounding[36] and that these 'trends' were augmented by a structurally-motivated tendency in the same direction.

In the case of FOOT, there is evidence for a similar, possible, confluence of exogenous and endogenous 'sources'. The acoustic data provides evidence for a substantially fronted FOOT vowel (overlapping with KIT before final /l/), with a sharply retracted allophone before final /l/. The fronting of FOOT has, in fact, been mentioned in the impressionistic literature with high (peripheral) values being, in general, associated with BrSAE and *somewhat* opener, fronted values being associated particularly with (modern) GenSAE

[34] It is important here to stress, again, the distinction between the acoustic vowel-chart and the more traditional IPA chart. While unrounding of a vowel will be reflected by a fronter value on an acoustic level, the traditional vowel chart is, on the contrary, focussed more on tongue position, with rounded and unrounded variants of the same tongue position being placed at the same point. As illustrated in Figure 8.30 on page 352, one mechanism for aligning these two representations is to use F_2-F_1 for the x-axis of the acoustic plot, rather than simply F_2. And in Figure 8.30 we do, in fact, see that SAE LOT is not considerably fronted in comparison with the other accents. The use of the Labovian 'vowel-triangle' (i.e. F_2 alone) makes more sense, however, if we do not wish to obscure potential 'pathways' of linguistic change.

[35] The reader will perhaps recall that Cornish miners appear to have constituted a relatively sizeable input into early-Johannesburg.

[36] We have to remember here that the RP of the time had, by all accounts, a rounded [ɒ]-like value.

speakers[37]. And in the acoustic data we do, indeed, find dramatic evidence for such FOOT fronting *and* lowering, very much in parallel with the central allophones of KIT. A comparison of SAE FOOT with acoustic data from other English accents confirms the dramatic fronting and lowering of FOOT in SAE, and, in the case of Torgersen and Kerswill's (2004) Ashford data, SAE FOOT is on a par with the most advanced variants.

The report of FOOT-fronting in the early impressionistic literature for *Br*SAE as well as the preponderance of reportage for FOOT-fronting in *Gen*SAE in the more recent impressionistic literature, seems to carry the hallmarks of endogenous change i.e. one which has spread, effectively, 'from below' (in all senses of this expression). Furthermore, similar changes have been reported widely (and independently) across the English-speaking world (e.g. RP and the USA), a change in line with Labov's (1994) Principle III. It has, of course, also been linked with Torgersen and Kerswill's (2004) SECS-shift. In the latter regard, however, the fact that SAE shows FOOT-fronting of an equivalent or even more advanced level than that presenting in the southeast of England, coupled with SAE's otherwise conservative nature *vis-à-vis* the SECS-shift[38], seems to indicate that FOOT-fronting is not linked to any chain-shift but is operating under a more general principle and, in fact, Torgersen and Kerswill (2004:45) hint at this very possibility when they confirm the following:

> "…the fronting of FOOT (and that of GOOSE and GOAT) is widely observed in present-day English throughout the world …we can safely say that there is something 'natural' about these changes, neatly encapsulated in Principle III, even though the presence of a chain shift seems not to be a precondition".

While the application of Principle III as a general principle of endogenous development helps to explain the dramatic *fronting* of FOOT in SAE, it does not, however, explain SAE FOOT's dramatic *lowering*, in comparison with, for example, Torgersen and Kerswill's (2004) data. One option is the notion of a parallel shift, mentioned briefly in §2.3.2.1 and seen to be used to explain movements that operate to restore parallelism to the vowel system; used by Labov (1994), for example, to explain both GOAT fronting and lowering, as touched on briefly in both §2.3.2.1 and §8.3. A further, admittedly speculative, possibility lies, however, in the history of SAE. While FOOT-fronting in the southeast of England and RP appears to be relatively recent phenomena, and are thus excluded as possible inputs, we have noted in §8.4.1 that normatively-minded north-of-England individuals, as a result of having

[37] As an aside, the author has quite recently been the 'butt' of a number of jokes concerning his use of [pət] for [pʊt] i.e. *put*. I have managed to console myself with the observation that my research seems place me, in this regard, in the young and modern 'bracket', although Lass' (1995) description, "younger (especially female) Respectable speakers" remains a little troubling.

[38] As reported on above for STRUT and LOT.

no FOOT-STRUT Split in their traditional accents, often hypercorrect the use of FOOT to a more open variant; thus the use of, for example, [ʃəgə] *sugar* or [bətə] *butter*. Assuming the enthusiastic adoption of such variants by other, particularly L1-Yiddish, speakers of early-Johannesburg, we have a possible exogenous explanation as well.

8.5 Conclusion

As expected, the acoustic data confirms a retracted BATH vowel in GenSAE. While there is some evidence in the impressionistic literature that BATH-rounding and tensing is loosing its 'negative' indexical value, this is not supported in the acoustic data in terms of any endogenous, Principle II-based (acoustic) raising. Instead the acoustic data reveals a relatively categorical value for BATH in GenSAE, with some tentative evidence of an in-glide. The historical reconstruction of SAE BATH seems to indicate, in accordance with Trudgill's (2004) observations in this regard, a mid-to-late 19th-century prestige RP-based influence, with the BrSAE value being based on a late-19th-century Cockney-input, with, perhaps, some role being played by Afrikaans.

Still, while Trudgill's (2004) observations are broadly valid, the facts of BATH-raising in SAE still provide relatively convincing evidence against his indexicality-'free' model of new-dialect formation in tabula rasa contexts.

THOUGHT, on the other hand, presents a rather simple 'picture', both on a synchronic and diachronic level. Diachronically, the input to the development of a raised variant has been diverse (although southeast-of-England), no doubt supported by endogenous forces. The acoustic analysis supports this general analysis.

The development of a monophthongal, dramatically-fronted GOOSE in GenSAE presents a virtually identical picture to THOUGHT, the only real 'difference' between the two vowels (apart from the obvious) being a dramatically severed-off pre-/l/ allophone in the case of GOOSE. In GenSAE, and SAE more generally, THOUGHT and GOOSE have thus followed a standard Pattern 3 movement and the resultant and respective raised and fronted values can be viewed as general indicators of this variety.

In contrast, GOAT, as with FACE, is almost certainly a social marker, and the acoustic analysis shows widely divergent values across the two SAE data sets. For the more prestigious, citation-form GenSAE variant, the acoustic data shows a value a little more centralized than generally assumed in the impressionistic literature, as well as substantial glide-weakening. The comparison of the SAE data with RP and AusE values confirms, again, SAE's relative resistance to Diphthong-Shift movements. GOAT differs from FACE in its probable historical development given that the less-open variant of GOAT has less precursors in the linguistic mix that characterized SAE's development; all of which, along with its only-recent mention in the impressionistic literature, seems to indicate endogenous development; in Labov's (1994) terms a parallel-shift in sympathy with GOOSE-fronting, and in accordance with his Principle III. The more open variant is explained via a tendency toward Diphthong-Shift 'drift' based on the original input to the Cape, bolstered by late-19th-century working-class input into early Johannesburg. In Labov's (1994) terms the endogenous 'force' behind the lowering of the GOAT-nucleus would be a tendency to restore parallelism with the equally lowered FACE vowel.

CURE, on the other hand, appears to retain its [ʊə]-like status in SAE, displaying regular and easily accounted-for allophonic distribution, and still, but contrary to some reports in the literature, displaying relatively little evidence for the Second FORCE Merger, at least in citation-form style. Recent research provides corroborating evidence in this regard and the historical reconstruction of the SAE CURE value provides no specific challenges.

Moving to the short back vowels of SAE, STRUT in SAE currently displays, *contra* Trudgill (2004), substantial fronting, and can be viewed, in this regard, as an imminent participant in Torgersen and Kerswill's (2004) SECS-Shift; in addition, a classic case of drift *and* 'colonial lag'. Said drift also probably reflects a current change-in-progress, although this needs to be confirmed by further research.

A relatively unrounded, centralized (and perhaps tongue-raised) LOT appears to be relatively categorical in modern GenSAE. It has the same acoustic quality as BATH, although the distinction is maintained via a difference in length. An unrounded (and thus acoustically lowered and fronted) LOT vowel is in accordance with SAE's conservative nature *vis-à-vis* the SECS-Shift. The etiology of this vowel remains unclear although it is certain that unrounded (e.g. Cornwall, Norwich), rounded (e.g. RP) and 'extra'-rounded (e.g. Cockney and Afrikaans) variants were all present during the early koinéization process.

For Gen(SAE) FOOT there is evidence for advanced fronting *and* lowering. The first is, in all likelihood, endogenously-motivated and in accordance with Labov's (1994) Principle III although, I would argue, unrelated to Torgersen and Kerswill's (2004) SECS-Shift. In fact, the presence of conservative SECS values for SAE STRUT and LOT coupled with advanced values for SAE FOOT seems to provide evidence *against* the notion of FOOT-fronting participating in a broader anti-clockwise chain-shift of the short vowels of English (i.e. SECS). In other words, while the broad *movement* encapsulated under SECS is undeniable, the fronting of FOOT appears to be structurally unrelated to, for example, the raising of LOT.

The lowering of FOOT is also, perhaps, best explained in terms of an endogenous parallel-shift development (FOOT 'in sympathy' with KIT), although one possible exogenous source for a lowered FOOT is the hypercorrective efforts of north-of-England immigrants.

The next chapter completes the vowel-by-vowel analysis of SAE, focussing on NEAR, NURSE, SQUARE, the various ‘weak’ vowels of SAE, as well as the nature and effect of final, tautosyllabic /l/ on preceding vowel quality.

CHAPTER 9

The Remaining Vowels and the Nature and Effect of Final-/l/

> "The close examination of the present shows that much of the past is still with us" (Labov 1994:27).

9.1 Introduction

This chapter deals with the remaining vowels of SAE (including the various weak vowels of English), as well as the nature of final-/l/ and its effect on preceding vowel quality. Given that many of the relevant vowels have an historical 'r' associated with them, still pronounced in all rhotic accents, it is interesting to note that, as early as Bailey (1984:21), there has been some (mostly informal) debate about the re-introduction of rhoticity into SAE. This is not to be confused with the semi-rhoticity that is a feature of BrSAE; neo-rhoticity in SAE appears, if anything, to be a prestige variant. Hartmann and Zerbian (n.d.) report rhoticity in Black South African English, and associate its presence with gender (female), affluence and the media. The effect that this adoption of rhoticity, as well as other American-like variants, is having on the traditionally 'white' accent is still unclear[1].

9.2 NURSE, NEAR and SQUARE (NNS)

The following definition of NURSE is provided by Wells (1982:137):

[1] In the author's personal experience, rhoticity in GenSAE *does* occur, particularly, although not exclusively, among young women and even in unmonitored speech. It is still highly sporadic and socially marked.

> “The standard lexical set NURSE is defined as comprising those words whose citation form contains the stressed vowel /ɜː/ in RP and /ɜr/ = [ɝ] in GenAm . . . phonetically, it is a relatively long unrounded mid central vocoid, [əː]; in GenAm, though not in RP, it is r-coloured” (Wells 1982:137).

The issue of how to phonemicize NURSE in rhotic American English accents does arise; partly because there appear to be two alternate pronunciations e.g. [nɜɹs] or [nɝs] for *nurse*; partly because of rhyming pairs such as *furry* and *hurry*. As summed up by Wells (1982:138), NURSE “is not necessarily distinct from the sequence consisting of the STRUT vowel plus /r/”: unlike RP, which has [fɜːɹɪ] and [hʌɹɪ], American English often has [fɝi](or [fɜɹi]) and [hɝi] (or [hɜɹi]). One option, therefore, as suggested by Wells (1982:480), is for words such as *nurse* to be phonemicized in rhotic American English as /nʌrs/ i.e. the short-vowel /ʌ/ plus /r/, with [ɜ] as a positional allophone of /ʌ/ before /r/ (in the case of [nɜɹs]) or with [ɝ] as a realization of /ʌr/ in the case of [nɝs].

In Labov (1994), NURSE is transcribed as /ɜh/. If there is, in fact, no separate NURSE vowel in the rhotic accents of the USA, as suggested above, then /ɜh/ is, of course, only applicable to the non-rhotic varieties.

With respect to NEAR, Wells (1982:153; my emphasis) provides the following definition for this standard lexical set:

> “comprising those words whose citation form contains the stressed vowel /ɪə/ in RP . . . and the *sequence* /ɪr/ in GenAm . . . phonetically, RP /ɪə/ is a centring diphthong with a starting-point that is unrounded and fairly close and front, [ɪ], moving towards a mid central [ə] quality”.

Examples of words with the NEAR vowel include *beer*, *deer*, *fear* and *appear* (Wells 1982:155). It is important to note that words such as *idea*, *real* and *museum* (i.e. those without a following historical ‘r’) are analyzed by Wells (1982:153) as having, in non-rhotic accents, the *sequence* of phonemes /ɪ/ and /ə/ i.e. as opposed to a single diphthongal phoneme. Lass (1990:279), however, lumps words with final orthographic ‘r’ and the “few non-Germanic loans like *idea* and the suffix *-rrhea*” together.

The vowel used in words in this lexical set is often transcribed as /ih/ for non-rhotic accents of American English, such as that used in New York City. In the case of rhotic accents, such as GenAm, Wells (1982:153) would mostly claim an analysis of /ɪ/ plus /r/. According to the American tradition, /ih/ does, however, occur (marginally) in rhotic accents, in particular in words such as *idea* and *theatre* (Labov 1991:40; footnote) i.e. in the non-Germanic loans mentioned above.

In Wells' (1982) terms, the SQUARE vowel is defined as follows:

> "as comprising those words whose citation form contains the stressed vowel /ɛə/ in RP ... and the sequence /ɛr/ ... in GenAm ... phonetically, RP /ɛə/ is a centring diphthong with a starting point which is front, unrounded and approximately half-open ... moving towards a mid-central quality" (Wells 1982:155–6).

GenAm, being a rhotic accent, does not have a separate SQUARE phoneme. One exception mentioned by Labov (1991:40) is *yeah*, which he subsumes under /eh/[2]. SQUARE (or /eh/) is, of course, found in non-rhotic accents of the USA.

9.2.1 NNS in Other Accents

Received Pronunciation: For General RP NURSE, Lanham (1967:42) gives [ɜ]^{34} {[ɘ]} for the nucleus while a slightly lower [ə]^{44}{[ə]} is characteristic of Refined RP. If there is a glide at all, it is toward a fractionally lower position. Cruttenden (2001:126) claims "considerable individual variation in its realization, with variations from close-mid to open-mid ... a pronunciation somewhat below open-mid is characteristic of Refined RP".

For RP NEAR, Lanham (1967:42) gives [ɪ]^{22}{[ɪ]} or something slightly more mid-central for the nucleus and [ə] as the target of the glide. This author also mentions the replacement of NEAR with NURSE in certain cases e.g. *here* and *beard*. Cruttenden (2001:143) mentions the increasing use of a monophthongal [ɪː] and also confirms the use of NURSE (or even BATH) in the place of NEAR in Refined RP.

For RP SQUARE, Lanham (1967:42) gives [ɛ]^{51} for the nucleus and [ə]^{44} for the in-glide, thus [ɛə] in modern IPA terms. Wells (1982:157) mentions the possibility of glide-weakening in RP SQUARE and for Refined RP mentions "a particularly open starting-point for /ɛə/, giving [æə]" (Wells 1982:293). Bailey (1984:16) briefly confirms that the traditional in-glide in RP SQUARE is absent or is in the process of being eliminated in this accent. According to Cruttenden (2001:144), a [ɜ]-like quality is usual for the off-glide when the vowel is in final-position, while in a closed position the off-glide of SQUARE is more likely to have a completely mid-central quality i.e. [ə]. According to this author "nowadays a long monophthong [ɛː] is a completely acceptable alternative in General RP" (Cruttenden 2001:144).

[2] Wells (1982) would, no doubt, analyze *yeah* as ending with the *sequence* /e/ plus /ə/.

Other Accents of England: According to Wells (1982:305), there is often not much difference between London and RP NURSE, although "somewhat front-ed and/or lightly rounded . . . Cockney variants such as [ɜ+ː]{ɜ̟ː}, [œ̈ː]" are found. Cruttenden (2001:126) mentions a rounded version, [œ̈], for Liverpool. Watt and Milroy (1999:33) give a rounded vowel, [øː], as one of the possible values for NURSE in Newcastle, although this appears to be a relatively recent development. Trudgill (2004:143) adds that "a number of Lancashire . . . and southeast Welsh accents . . . have a rounded [œː] vowel". Generally, however, north-of-England accents appear to have an unrounded NURSE vowel (Wells 1982:360–1), often leading to a NURSE-SQUARE Merger.

In the case of NEAR, both Wells (1982:305) and Cruttenden (2001:143) mention a closer starting point for Cockney English, [iə], although monophthongal variants do occur, [ɪː], particularly word-internally. "Triphthongal variants also occur, and are perceived as very strongly Cockney; they . . . are restricted to sentence-final position, as in . . . [ˈœʊvrˈɪjɐ] *over here*" (Wells 1982:306). Interestingly enough, Wells (1982:361) notes the following regarding the north-of-England:

> "the more conservative northern accents . . . use disyllabic or varisyllabic pronunciations of the type [fiːə] *fear*, [biːə] *beer*, which is best regarded as a phonemic sequence /iːə/. Word-internally before /r/, in words such as *serious*, *period*, these accents usually have simple /iː/ with no following /ə/".

For Cockney SQUARE, there is also, generally, a closer-than-RP starting point to the diphthong. Wells (1982:304–5) uses [eə] to represent this vowel, although also gives the monophthongal [ɛː] as a possible variant, and Cruttenden (2001:144) gives [eə̞] as the precise quality, while Branford (1994:482–3) mentions that "this vowel has glideless manifestations in British dialects, for example for some working-class speakers in Norwich". SQUARE in the north of England tends to be monophthongal with values provided by Wells (1982:361) including [ɜː ∼ ɛː ∼ ɛə]. "In Liverpool the most characteristic qualities are reported to be the centralized front monophthongs [ëː] and [ɛ̈ː]; but [ɛː] and [ɵː] also occur".

AusE and NZE: Turning to the Southern Hemisphere Englishes, Wells (1982:599) mentions the closeness of the NURSE vowel in AusE as being the main difference between it and its equivalent in RP – "qualitatively it is thus a kind of lowered [ɨː]", – although it is also often front of central. The closeness of AusE NURSE is also mentioned in Cruttenden (2001:126). Harrington et al. (1997:163) provide acoustic evidence for an association between Broad (female) speakers and fronted NURSE in

AusE. The general implication appears to be that AusE does not have rounded variants of this vowel; thus Harrington et al.'s (1997) transcription-system proposal for this vowel is [ɜː]. Trudgill (2004:143), however, makes specific mention of raised, *rounded* and fronted variants of NURSE in AusE (as well as Falkland Islands English).

Since NZE and SAEP are often similar with regard to the pronunciation of NURSE, Wells' (1982) description of NURSE in the former accent is worth quoting in full:

> "The long monophthong /ɜː/ (NURSE) is rounded in New Zealand English; it is realized as a centralized front mid vowel, [œ̣̈ː]{[œ̞̈ː]}. In some broader accents it is as close as cardinal 2, thus [ø̈ː]. Hearing the New Zealand linguist Turner pronounce his surname, I was very much reminded of German *Töne*" (Wells 1982:607–8).

For NZE NURSE, Trudgill (2004:143) gives [øː ∼ œː ∼ ɞː].

For NEAR, AusE shows a tendency toward monophthongization, "particularly before /r/" (Wells 1982:599), thus [wɪːɹəi] for /wɪəriː/ *weary*. Furthermore, according to Wells (1982:600), the broader variants of AusE tend to insert a /iː/ plus /ə/ sequence in final position instead of /ɪə/ which would be the norm in the Cultivated sociolect: "Broad Australian *beer* is [bəɪə], which we write phonemically as /biːə/". A similar analysis is provided in Harrington et al. (1997) who, essentially, claim a continuum from monophthongal NEAR (particularly in a closed syllable) to full bisyllabicity. As in the case of SQUARE (see below), the authors found a predominance (in their citation-form data) of diphthongal variants as well as a (minor) tendency toward rhoticization. In the case of the diphthongal variants the first target was in the region of KIT, but somewhat raised or fronted in the case of Broad accents. The glide of the diphthongal variants terminated in the TRAP-START space i.e. low-front to central.

In NZE, the starting point of the NEAR diphthong is often very close, [i], and, according to Wells (1982:608), "may not be distinct from /iː/ plus /ə/". In other cases, and as mentioned in the discussion of SQUARE (see below), NEAR merges with the latter vowel, resulting in a [ɪə] quality for both. Trudgill (2004:145) gives [e̝ː(ᵊ) ∼ ɪː(ᵊ)] for this merged quality.

In AusE, the SQUARE vowel is close, as in Cockney, but also very often monophthongized, as confirmed in Branford (1994:483). The difference between *shed* and *shared* is, thus, often signalled by length alone (Wells 1982:599-600). As in the case of AusE NEAR, Harrington et al. (1997) suggest two main versions of SQUARE i.e. a monophthongal version (particularly in closed syllables) with a bisyllabic version as the other option. In their acoustic study (citation-form data), these authors found a preference for a full diphthongal version (whether bisyllabic or not) of this vowel. These authors also discovered some evidence of rhoticization in SQUARE (as with NEAR), which they put down, along with the preference for diphthongal versions, to hyperarticulation under citation-form conditions. In the case of the diphthongal variants the first target was usually in the region of DRESS, although “Broad female /ɛə/ has a clearly fronted and slightly raised first target compared with that of the other accents” (Harrington et al. 1997:177). The offset of the diphthong terminates, as in the case of NEAR, in the TRAP-START space i.e. low-front to central.

NZE is characterized by the NEAR–SQUARE Merger in which the two relevant vowels have an identical quality i.e. [ɪə]. Thus, for example, *bear* and *beer* are often homophonous (Wells 1982:608).

9.2.2 NNS in SAE: The Impressionistic Data

Hopwood (1928), Breckwoldt (1961): Hopwood (1928) claims, for SAE NURSE, that “*sir* **sɜː** > “suh” **sœˑ(r)** or **sœˑɹ**”. The rhoticity implies BrSAEP. The alternatives implied are thus [sœˑr], [sœˑɹ] and [sœˑ]. With respect to the last of these, we have the possibility of a half–long monophthong. Breckwoldt (1961:8) also notes a rounded variant of this vowel in SAE i.e, [œ].

For SAEP NEAR, Hopwood (1928:18) claims that “E **ɪˑə** > **ìˑ(r)**{[iˑ(r)]}: *e.g.* E. really ˈ**ɹɪˑəlɪ** > “reelee” ˈ**rì(ˑ)lì(j)** {[ˈri(ˑ)li(ː)]} or ˈ**r˕eˑlì(j)** {[ˈre̞ˑli(ː)]} . . . E. tear ˈ**tɪˑə** > **tìˑr**{[ˈtiˑr]}”. Again, the rolled /r/ in the SAE transcriptions indicates that we are, as in most of Hopwood’s (1928) examples, dealing with BrSAEP. It should be mentioned that in modern-day GenSAE [ˈɹe̞ˑliː] seems, to me, to be the norm i.e. the /r/ becomes an RP-like approximant, but the vowel remains monopthongised as in Hopwood’s (1928) transcription. [ˈtiˑr], though, is definitely a BrSAE value. GenSAE would have [tɪə] or, at most, [tiə].

In Hopwood (1928:19), we find that RP SQUARE “*there* ˈ**ðɛˑə** > “thay(r)” **ðɛˑr** or **ðeˑ(ɹ)**”. The rhoticity implies BrSAE, as does the half-long non-rhotic alternative [ðeˑ] implied by the optional /r/ in **ðeˑ(ɹ)**. More important is the indication that the quality of the nucleus is close, as well as the absence of an in-glide.

Lanham and Traill (1962), Lanham (1964?; 1965): For SARP NURSE, Lanham and Traill

(1962) give the unrounded [ɜ]$^{34/44}${[ə̩]} (with a short glide to the center), except in the context of a preceding /ʃ, w, y/, in which case the onset is raised to [ɜ]^{34}{[ɘ]} or even [ï]$^{24/34}${[ə̝]}. Non–SARP SAE, however, shows rounding, raising, fronting and monopthongisation: Lanham and Traill (1962:30) give [ɜ]^{33}{[ɵ̝]} as the value. The authors also claim that the latter quality was, at the time, "a tendency in SARP 'B"' (Lanham and Traill 1962:30). Lanham (1964?:27) mentions a shortening of this vowel with no implication that such shortening is solely a BrSAE value.

For NEAR, Hopwood's (1928) obviously BrSAE value, i.e. **ì·(r)**, stands in contrast with Lanham and Traill's (1962) SARP 'A' value which is given as "[ɪ]$^{22/32}${[ë˔]} or [e]^{32}{[ë]} [with] the glide reach[ing] about [ɜ]$^{33/43}${[ə̝]}" (Lanham and Traill 1962:30; my parenthesis), the difference here being a more open onset as well as evidence of a glide.

For non-SARP SAE, Lanham and Traill (1962:30) report three trends for NEAR, the first involving a more close onset than that reported for SARP[3] and the second a more open onset. The values given are [i]^{12}{[i̠]} for the first and roughly between [ɛ]$^{51/61}${[ɛ̞]} and [E]^{41}{[ɛ̝]} for the second, a fact which leads to NEAR and SQUARE becoming almost identical, although at a different quality to that used in the NEAR-SQUARE Merger in NZE. The values given here for this second option are, in fact, based on Lanham and Traill's (1962) description of the *SQUARE* vowel (see below) since they do not provide an exact value for this variant of NEAR but simply claim that "the vowel is lowered – so far in some cases that *pier* . . . and *pear* . . . become almost identical" (Lanham and Traill 1962:30). Unlike Hopwood (1928), there is no *explicit* mention of the potential monophthongisation of this vowel, but the quote above suggests as much.

The third and more recent trend (at the time) in non-SARP SAE, and one also strongly evident in Lanham and Traill's (1962) SARP 'B' variety, is for the onset to be centralised to a [ï]^{23}{[ï]} position. This is usually conditioned by a preceding /r/, but the authors evoke a minimal pair to claim (marginal) phonemic status for this value:

> "A limited number of SARP speakers seem to have this nucleus in a small number of words that usually include *reared*, *weir*. Some have *Ria* (a girl's name) as /riH/ and *rear* /rïH/ as a minimal pair. The vowel is approximately [ï]^{23}{[ï]} and the glide to about [ɜ]$^{33/43}${[ə̝]}" (Lanham and Traill 1962:32).

As mentioned above, the potential monophthongization of this vowel is not explic-

[3] i.e. similar to Hopwood's (1928) value for NEAR.

itly mentioned in Lanham and Traill (1962:92). Lanham (1965), however, claims that there is a growing tendency in SAEP “to replace /iᵊ/ with either /iˑ/ or /eˑ/”. The author gives /reˑli/ or /riˑli/ for *really* and /peˑs/ for *Pearce*. Thus Lanham (1965:92) predicts the eventual homonymity of *reel* and *real*. While the given values for *really* are common in modern-day SAEP, the value given for *Pearce* seems less likely, although certainly conceivable. In general it appears that while some polysyllabic words seem to have been effected by this process (so, for example, *nearly* also often has a monophthongized variant) it doesn’t often occur in monosyllables. Thus *near* and *fear* seem, to me, to almost always have a diphthongal quality[4]. With respect to *reel* and *real*, while these are certainly homonyms in SAEP, the nature (and thus source) of this homonymity is different to that implied by Lanham (1965). In Jones (1924) the vowel of *reel* is identified as belonging to the FLEECE lexical set i.e. as [riːl] in broad notation, while in Wells (1982:155) the vowel in *real* is given as belonging to NEAR. If Lanham’s (1965) predictions were correct then both vowels should be pronounced [riːl] in modern-day SAEP. The opposite has in fact occurred: due, no doubt, to the influence of the following /l/, *reel* is pronounced [ɹɪəl] or [ɹiəl], as is *real*.

With respect to SQUARE, and as noted by Lanham and Traill (1962:30), “in extreme SAE there is a general shortening of the whole nucleus,” a feature which is coupled with monophthongisation and a comparatively close vowel with a value around [E]^{41}\{[ɛ̝]\}. SARP ‘B’, on the other hand, has a [ɛ]^{51}\{[ɛ]\} value, while SARP ‘A’ finds itself at [ɛ]$^{51/61}$\{[ɛ̞]\} with a glide to [ɜ]$^{33/43}$\{[ɘ̟˕]\} or even as high as [ï]$^{23/33}$\{[ï̞]\}. In Lanham (1964?:27) shortening of the nucleas is mentioned although there is no implication that this is a feature solely found in BrSAE lects.

Lanham (1967): For non-conservative SAEP NURSE, Lanham (1967:64) gives rounded [ɜː]^{33}\{[ɵ̟]\} as the value and confirms the lack of any centring glide. While conservative SAE generally has unrounded [ɜ]^{34}\{[ɘ]\}, a rounded equivalent is also attested to (Lanham 1967:90).

For NEAR, Lanham (1967:95) mentions the marginal contrast between /iH/ and /ïH/ which “has been recorded in Johannesburg and the Eastern Province in idiolects intermediate between extreme and conservative and mainly in the under-30 age group”. For conservative SAE NEAR, Lanham (1967:93) gives [ɪ]^{22}\{[ɪ]\} as the value of the onset and [ə]^{43}\{[ə̟]\} as the value of the terminus of the glide, mainly in CV syllables,

[4]See, however, Bowerman’s (2004) latter comments in this regard. It has been mentioned to me, in addition, that *beer* is often monophthongized to [beː] in the accent used by Eastern Cape farmers (Todd n.d.).

the implication being that in non-CV syllables the glide will be weaker. The author notes, however, that this value may alternate with yod+NURSE giving, for example, /njɜː/ instead of /nɪə/ for *near*. In many non-conservative SAE idiolects similar values are commonly found, though the following two trends are also attested to:

1. In the most extreme (Afrikaans-influenced) varieties the onset can be as high as [i]^{12}{[i̲]}; or /ɪə/ becomes the AusE-like di-syllabic /iˑ + ə/.
2. The vowel can monophthongize to /eː/ or /iː/

Lanham (1967:63) confirms the participation of SQUARE in the general tendency toward glide-weakening of diphthongs in SAE, as well as the presence of a shortened nucleus in the most extreme versions of the accent. This author gives "[eː]^{31}{[eː]} or [E:]^{41} {[ɛ̝ː]} in, respectively, more extreme, less extreme SAE" (Lanham 1967:63). It is interesting in this regard to note the closer values reported on in Lanham (1967) as opposed to those provided in Lanham and Traill (1962) above, a hint that the movement of SQUARE to a C[2] position has been an endogenous development in SAEP. For conservative SAE, Lanham (1967:90), gives the onset as [ɛ]^{51}{[ɛ]} with the glide ending at [ɜ]$^{33/43}${[ɘ̝˕]}, generally confirming the earlier Lanham and Traill (1962) in this regard.

Lanham (1978), Lanham and Macdonald (1979): Lanham (1978:149) identifies a raised NURSE vowel as a (possible) Natal variable and as an identifying feature of Gen-SAE. More specifically, the most advanced variant is a "fronted, slightly rounded high schwa ... corrected to low schwa, hypercorrected to [ʌː] by those sensitive to R St" (Lanham 1978:154). For most, however, the value of NURSE in SAE is seen to be below the level of social consciousness. According to this author, it has a possible Northern England origin, and, as reported on above, rounded values have been reported for both Liverpool ([œ̈ː]) and Newcastle ([øː])[5].

[5]These do not appear, however, to be the traditional or most common values for these respective areas. According to Wells (1982:372;374–5), [ëː ∼ ɛ̈ː] are the two most common Scouse values and, in Tyneside, the broadest value is [ɔː], while "in a less broad Newcastle accent, NURSE words have [ɜː] or something similar, e.g. rounded, centralized-front [øː]". From Trudgill (2004:143), and as already mentioned above in §9.2.1, it appears that a rounded NURSE is traditionally associated with the Lancashire area and accents of the southeast of Wales. According to this author, however, "the accents of these areas would have been far too much in the minority in the Southern Hemisphere mixtures to have had any influence in the new dialects" (Trudgill 2004:144). Trudgill (2004) thus resorts to the notion of 'drift' to explain the (according to him) independent development of a fronted, raised and rounded NURSE vowel across all the Southern Hemisphere Englishes. The presence of "fronted and/or lightly rounded" (Wells 1982:305) variants of NURSE in Cockney is suggestive, however.

Interestingly, Lanham and Macdonald (1979:43) mention the possibility of a retracted NURSE vowel before /l/ as a feature of GenSAE. The value given for the most advanced variant by these authors is [ɜ˃ː]{[ɘ̠ː]}. The value seems highly restricted lexically with the authors only giving the pronunciation of *girl* as an example. More generally, these authors confirm a "raised, fronted, rounded əə" (Lanham and Macdonald 1979:44) as a Natal English variable and thus a defining feature of GenSAE. The most advanced variant is given as a weakly rounded [ɜ˂ː] i.e. [ɘ̟ː], "in Natal most favoured after /y/{/j/} (e.g. *year*), but contextually unconditioned in SAE generally . . . quantities of Fronted əə are regionally highest in the Cape (East and West)" (Lanham and Macdonald 1979:44). Hypercorrection to a low [ɐ] is only prevalent in Natal and there appears to be little evidence of style-shifting. The authors also mention the possible north-east England origin of this value i.e. Newcastle and environs.

Lanham (1978:152) claims that in certain idiolects the tongue height for SQUARE overlaps with that of DRESS. The glideless nature of the vowel is confirmed, though a glide (together with a lowered tongue position) can occur in more formal styles. Lanham and Macdonald (1979:46) concur with this assessment, identifying a raised and glideless SQUARE vowel as a GenSAE variable and implying an implicational link between it and a close DRESS vowel i.e. if a close DRESS vowel exists in a particular idiolect it is likely that a close and glideless SQUARE vowel will exist in that same idiolect.

Lass (1990; 1995; 2004): Lass (1990:273; 278) confirms that the NURSE vowel distinguishes CulSAE which, like RP, has unrounded [ɜː], from GenSAE and BrSAE which always have rounded and centralised [ø̈ː ∼ œ̈ː]. As an aside, and according to Lass (1990:278), there is also in SAE a characteristic shortening of this and other long nuclei in disyllabic words e.g. *first* [fø̈ːst], but *firstly* [fø̈stli]. According to Lass (1995:98), the GenSAE and BrSAE value for NURSE is "similar in quality to the vowel in German *schön*, French *peu*, or the second element in the Afrikaans diphthong in *seun*".

Lass (2004:376) essentially repeats his earlier analyses. He implies, however, that the rounding of this vowel is not of particular interest given that the same feature is found "in most of the English-speaking world where ME /ir, ur, er/ have merged" (Lass 2004:376)[6]; the typical [ø̈ː] value is given as slightly more retracted that the similar

[6]This is, of course, an interesting observation *vis-à-vis* Trudgill's (2004) claim that this feature could not have had its origins in the accents brought over from Britain to South Africa and the other Southern Hemisphere colonies. See §9.2.4 for more on this.

vowel in "standard NorthGerman *schön*" (Lass 2004:376–7). Unrounded variants are common in CulSAE and, with respect to younger speakers, among those coming from prestigious and private girl's schools, an observation that needed to be taken into account when analyzing the acoustic data collected for the purposes of this study. The unrounded variant is "judged positively or negatively according to the politics of the judger" (Lass 2004:377).

According to Lass (1990:279), [ɪə] would be the principle value for SAEP NEAR. The onset of NEAR is equivalent in quality to the fronter allophones of SAE KIT, while the terminus of the glide varies from [ɘ] "to a quite open [ɐ]; the latter is most common in *here* [çɪ̯ɐː]" (Lass 1990:279)[7]. With respect to the *year* lexical item, the author mentions a rather open NURSE quality, [œ̈], as being one characteristic pronunciation, but claims that this is more likely a matter of "dissimilation after /j/ rather than a class-transfer" (Lass 1990:279).

Lass (1995:100) claims that NEAR in SAE, "is normally [ɪə] for all speakers, though in Extreme it may be monophthongised to [ɪː]". According to Lass (2004:378), "there is some monophthongisation in younger speakers, but this is not nearly as salient as it is in New Zealand". The most common value for GenSAE speakers is given as [ɪə].

In relation to the potential overlap of SQUARE and DRESS, Lass (1990:274) points out the possibility of a minimal length contrast in SAE i.e. as opposed to most other dialects of English where a difference in length is accompanied by a difference in quality. Thus Lass (1990:277) makes the following claim about what he refers to as "Respectable (=(Upper)Middle Class) White Cape Town English":

> "Except in word-list style and very formal speech, where a slight lowering and diphthongisation may occur giving values like [e̜ːᵊ] {[e̞ːᵊ]}, this is normally a long monophthong, qualitatively identical to the short vowel of DRESS".

The difference between Lass' (1990) values and those of the earlier Lanham and Traill (1962) and Lanham (1967) provide further evidence for the possible raising of this vowel as an endogenous development. Interestingly enough, Lass (1990:277) also hints at possible confusion between SQUARE and FACE, given the extremely short (weak) glide that Lass (1990) reports for the latter for his subjects – see §6.4.

Lass (1995:100) provides a similar picture to that already painted above: CulSAE has invariant [ɛə], GenSAE has [ɛə], [ɛː] or [eː], with the latter more common among

[7]With respect to the first value provided for the terminus, the original text gives [ə], but the description of the value as half-close leads me to suspect that the correct value should have been [ɘ].

younger speakers, while [eː] or an even closer variant is found in BrSAE. According to Lass (1995:100), "this is a highly salient variable, and many Respectable speakers (even those who monophthongise) stigmatise non-diphthongal variants".

Lass (2004:376) gives a monophthongised variant as the most typical for young GenSAE speakers, with [eə] or [ɛə] only found among Cultivated speakers "and the posh end of" (Lass 2004:376) GenSAE. The length contrast between SQUARE and DRESS (*bared* vs. *bed*) is also mentioned.

Wells (1982), Bailey (1984), Bowerman (2004), Da Silva (2007): In the case of NURSE, Wells (1982:615) confirms the general picture presented so far, giving a weakly-rounded [ø̈ː] as the main trend, with [ɜː] as the Cultivated variant, with hypercorrection to [ɐː]. Bowerman (2004:937) confirms the CulSAE RP-like [ɜː]-value, while non-CulSAE has a rounder and fronter variant: [øː ∼ ø̈ː]. Of interest is Da Silva's (2007) comment to the effect that for NURSE "[ɵ] was given, as it was not perceived to be as tense and high as German *ö* as in *schön* or as front as *Götter*" (Da Silva 2007:120)[8]. This author also mentions the sporadic use of a rhotic NURSE in her data, but its use was not common among her 'white' speakers.

For Bailey (1984:16), NEAR "may be replaced by /ee/ in some words such as 'serious' /seeri.s/ but is more often replaced by /yəə/ as in 'near' /nyəə/". Branford (1994:482) mentions the possibility of a "raised point of onset, resulting in an [iə]". This is, by all accounts, a BrSAE value, "rather similar to the diphthong used by many mother-tongue speakers of Afrikaans in words like *geel* ('yellow')" (Branford 1994:482). Bowerman (2004:938) adds that monophthongisation in BrSAE is most common after [j], as in *near* [njɪː].

Bailey (1984:16) confirms the lack of an in-glide in non-RP-influenced SAEP SQUARE. Branford (1994:481;482) implies a long and glideless "vowel in the region of cardinal 2", as the Broad (stereotyped) variant. A lower but still monophthongized [ɛː] is also heard in non-CulSAE speakers. It is only in CulSAE, however, that a diphthongal quality is retained. Bowerman (2004:938) gives the most common values for CulSAE, GenSAE and BrSAE as [ɛə], [ɛː] and [eː] respectively.

9.2.3 NNS: A Comparative Acoustic Analysis

With respect to Webb's (1983) acoustic data, as presented in Figure 5.5 on page 152, we note the raised NURSE vowel and the retracted variant before final /l/ (the token was *girl*). If

[8] Of interest particularly since the author is an L1-speaker of German.

we assume lip-rounding then the vowel is relatively fronted as well, given that lip-rounding lowers F_2. We note too that the *year* token (near$_{year}$) is relatively close to NURSE. As hinted at above, NURSE often replaces NEAR as the vowel in this word and is, as confirmed by Webb (1983:155), "a shibboleth of Natal English, which is the variety of English acquired by the author". The value represented in Figure 5.5 is based on the formant values of four tokens, 2x*year* and 2x*years*.

The author's 'take' on his NURSE data is of some interest. He claims that "/ɜ/ is fronted after a segment which is either [+cor] or [+ant]" (Webb 1983:149). In short, the tokens *were*, *works*, *worked*, *girl*(x2) have an average F_2 of 1281Hz, while the tokens *jerking*, *her*, *first*, *church* and *attorney* have an average F_2 of 1582Hz[9]. While it certainly true that there is a relatively obvious difference between these two groups, it should be pointed out that the two tokens with the lowest F_2 are the two *girl* tokens i.e. NURSE before final /l/. If these are separated out from the relevant group the average increases to 1320Hz. It is quite likely that the initial /w/ has a retracting influence here, but the number of tokens (i.e. three) makes the validity of this analysis difficult to assess.

Webb's (1983) data on the NEAR vowel is also represented graphically in Figure 5.5 on page 152. This figure seems to indicate a relatively lax onset/nucleus, a fact that can be ascertained from an inspection of Figure 9.1 on the following page, which superimposes, in bold, the NEAR data (minus *year*) over Webb's (1983) short-vowel data, as represented originally in Figure 5.4 on page 151. As can be seen there is substantial overlap between NEAR and both KIT and DRESS[10]. It is possible, though, that the lax position of this vowel is an artefact of Webb's (1983) measurement procedure. The point of measurement was generally-speaking the point of inflection (see §2.2.4.2) or, in the case of a steady-state vowel, the temporal mid-point (Webb 1983:137). The author also took care "not to select spectra at the beginning of the vowel because of the influence of prevocalic segments" (Webb 1983:137). Since the formants of NEAR generally present a smooth glide running from a high-F_2 and low-F_1 starting point to a lower-F_2 and higher-F_1 end-point with no real point of inflection (or steady-state for that matter), the representation of, for example, the onset, depends crucially on how 'close' the measured segment was to the beginning of the vowel. This is unclear from Webb (1983), but given his comments above it seems likely that the vowels concerned 'began' at a somewhat higher and fronter position.

Figure 5.5 on page 152 also provides a representation of Webb's (1983) SQUARE vowel. The quality of SQUARE in this particular idiolect becomes more apparent, how-

[9] It is unclear how the /h/ in *her* should be classified as either [+cor] or [+ant].

[10] In Figure 9.1 the label for the fronter allophone of KIT is 'hiding' behind the label for Webb's (1983) NEAR.

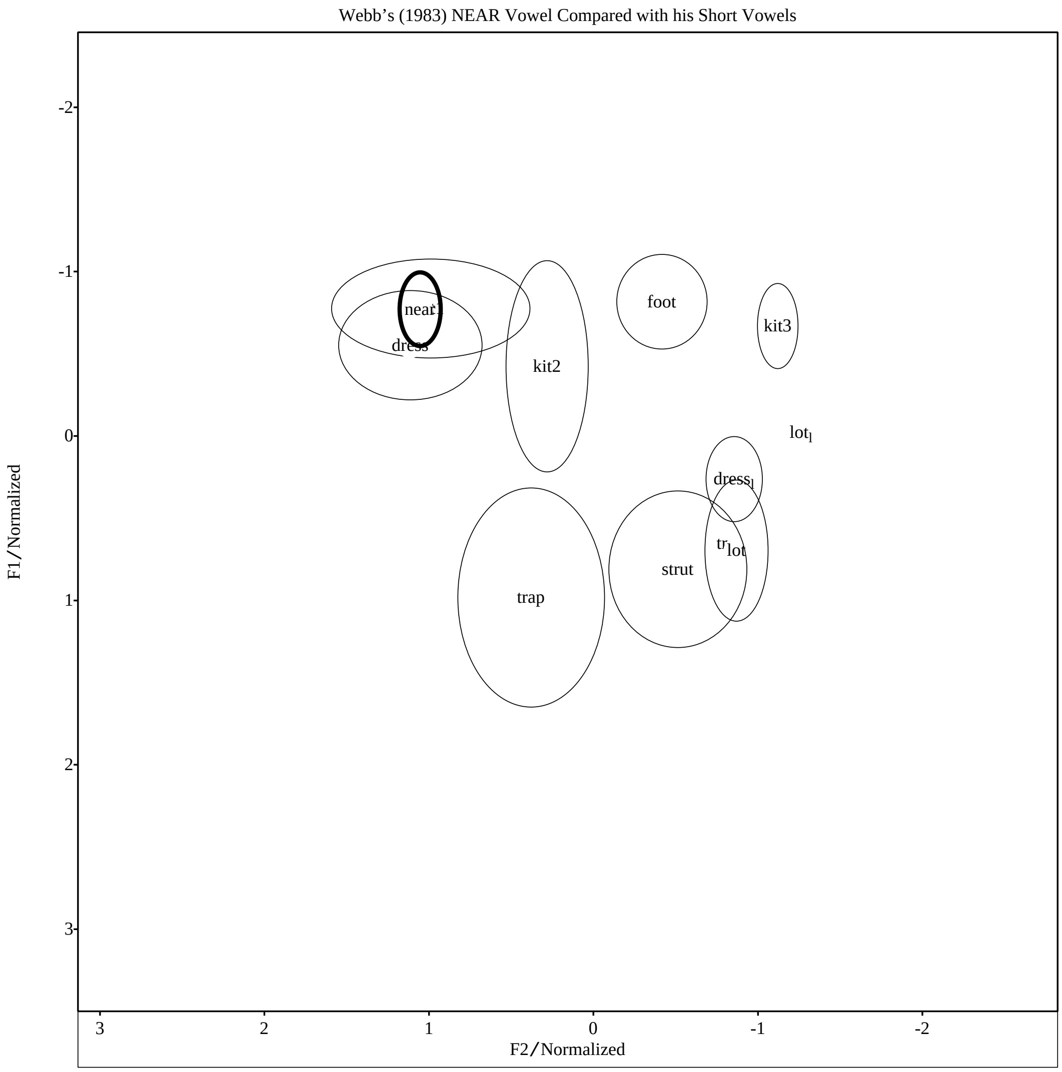


Figure 9.1: Webb's (1983) NEAR Vowel Compared with his Short Vowels

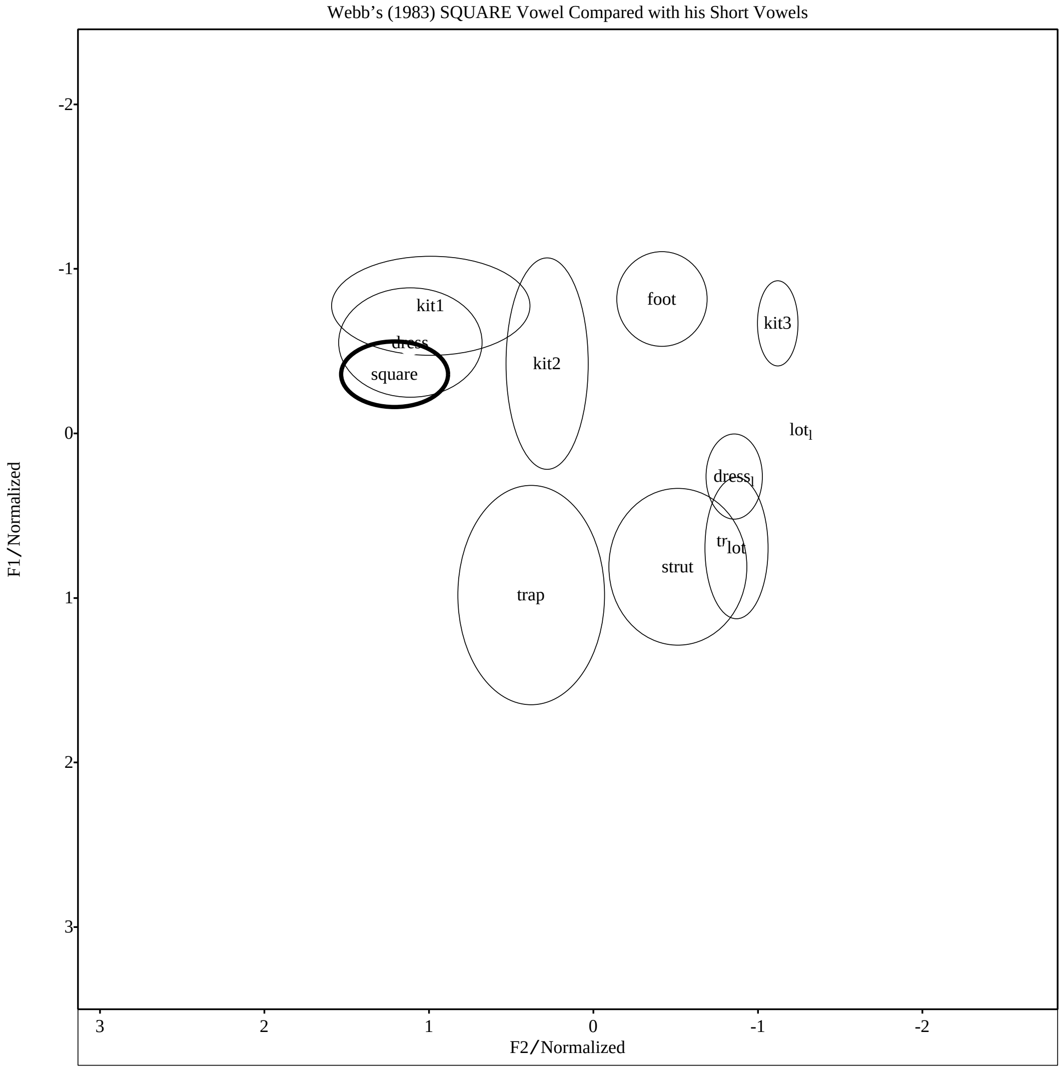


Figure 9.2: Webb's (1983) SQUARE Vowel Compared with his Short Vowels

Lexical Set	No. of Tokens	No. of Measurements	N-G Tokens	O-G Tokens
NURSE	44	968	*herd*, *hurt*	*hurt*
NURSE before /l/	18	396	*hurl*	N/A
NEAR	60	1320	*beard*, *here*, *beer*	*here*
SQUARE	61	1342	*hair*, *scare*, *share*	*fair*

Table 9.1: Summary of NURSE, NEAR and SQUARE data

ever, if we superimpose the data for this vowel over Webb's (1983) short-vowel data, as provided in Figure 9.2 on the previous page. The overlap between DRESS and SQUARE is apparent from this figure. The fact that SQUARE is somewhat lower than DRESS is suggestive but remains just that given that only five tokens (*square*(x2), *scared*, *affair* and *there*) were included in Webb's (1983) data. As far as the putative glide of SQUARE is concerned, Webb (1983:155) makes the following largely uninterpretable (and unconvincing) comment:

> "In the case of the inglide /eə/ there is very little information. The available information, however, points to a strong glide which is 500Hz more "central" than the nucleus".

Turning to the current data, Figure 8.2 on page 296 situates the NURSE, NEAR and SQUARE vowels in the context of the other Part-System D vowels. The representations are based on the data provided in Table 9.1. As can be clearly seen from Figure 8.2, as well as Figure 8.15 on page 324, where NURSE is compared with GOAT, the NURSE vowel is relatively raised (it is, for example, on a par with both DRESS and SQUARE) and assuming rounding, is therefore likely to be representative of a [ø̈]-like value i.e. close and front(ish) as expected from a reading of the impressionistic literature. The retracting influence of a final, 'dark'-/l/ is also apparent and also provides evidence for such retraction outside of the lexically-restricted context of *girl*. As can be seen from Table 9.1 the relevant token in the current data was *hurl*.

Figure 9.3 on the facing page, which superimposes, in bold, Webb's (1983) NURSE data over the current Part-System D data, provides striking evidence for the invariability of this vowel's target across the two data-sets: as can be seen, after Lobanov normalization, virtually equivalent representations are produced. This provides some support for the conclusion that the *current* subjects have a rounded NURSE vowel, important given Lass'

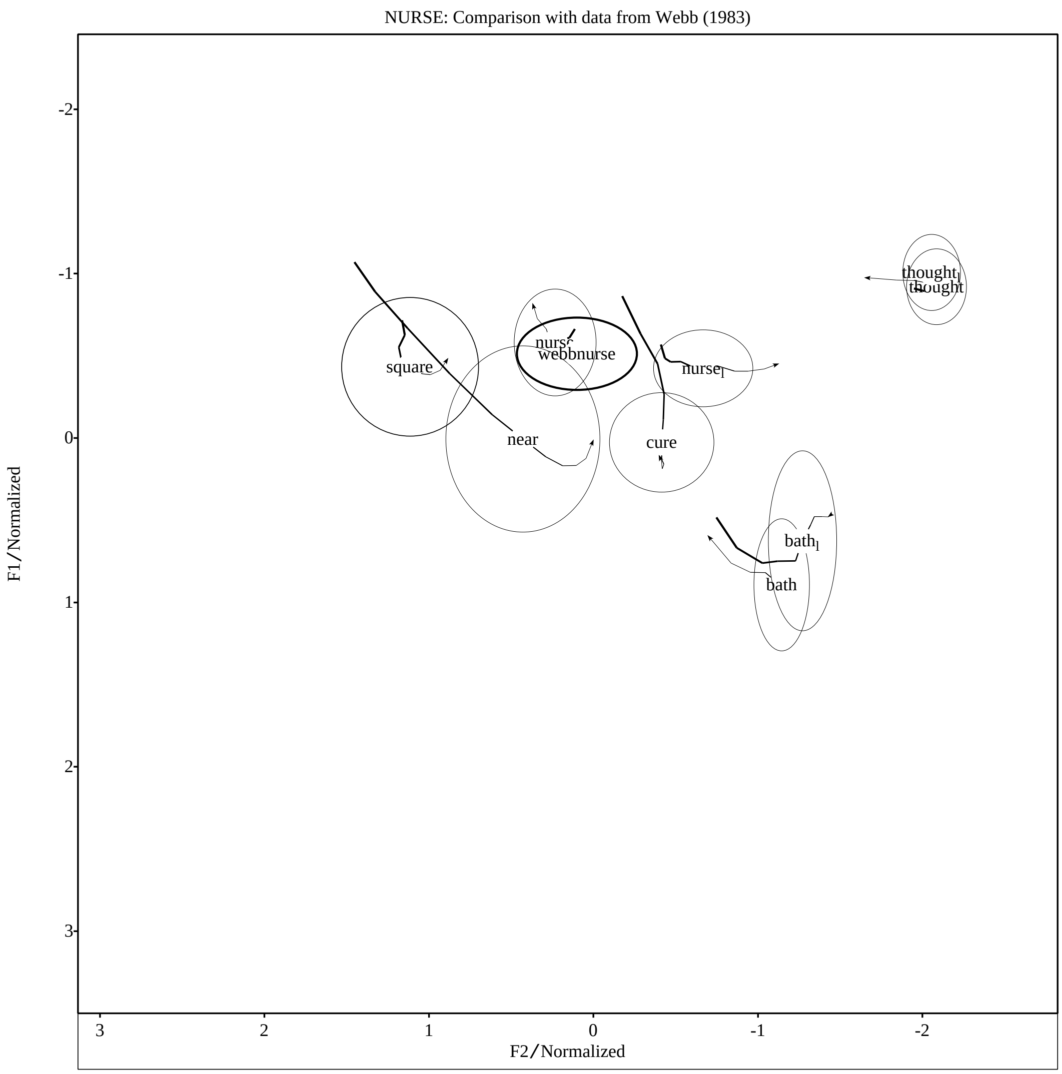


Figure 9.3: NURSE: Comparison with data from Webb (1983)

(2004) earlier comments regarding unrounded variants still being found in the speech of young females, particularly those from private schools[11]. The slightly higher standard deviation of Webb's (1983) NURSE on the front-back dimension is perhaps due to the fact that 3 of the 8 NURSE tokens used were preceded by /w/, which would have had an obvious retracting influence[12].

Figure 9.4 on the next page is a wonderful representation of the gradual heightening and fronting of this vowel in the Southern Hemisphere Englishes. One notes, however, that, surprisingly, NZE appears to be in the forefront of this movement, while AusE and SAE appear to have similar values. RP, as predicted, has a lower value, more evident in citation-form than in connected-speech. The fact that the Southern Hemisphere Englishes retain fronter *acoustic* values despite the influence of lip-rounding (which generally lowers F_2) speaks to the fronted tongue articulation used in the pronunciation of this vowel.

Turning to NEAR, in Figure 8.2 on page 296 a clearly diphthongal version of this vowel is found in the context of the other Part-System D vowels. As can be seen in Figure 9.5 on page 382, NEAR in the current data begins from a relatively front and close position (above DRESS and certainly somewhat removed from the fronter allophones of KIT) and ends in a central schwa-like position. Note that if we assume [ə] as being the value of the terminus of the glide, this provides evidence for the relatively high positions of the central values of KIT – see Chapter 7. Note that the average value for the terminus is far higher than STRUT, precluding a possible [ɐ] value as the main one. The standard-deviation of NEAR is, however, rather large, implying a substantial degree of variation.

This variation is very clearly illustrated in Figure 9.6 on page 383 where the bolded NEAR tokens are superimposed over the other short-vowel Part-System A vowels. The first difference we notice is that between *beard* (in a closed syllable) and *here* and *beer* (in open syllables). The first, *beard*, begins in a position similar to the group as a whole, but ends between the fronter and more centralized allophones of KIT. It is likely that the full glide-'target' has been truncated due to the coarticulatory effect of a following obstruent.

Although terminating in a similar enough position, there are some interesting differences between the remaining two tokens: *beer* and *here*. Firstly, we note that the latter

[11] It is, however, possible that these putatively unrounded variants are produced with a comparatively retracted tongue position, which would render their *acoustic* value equivalent, at least on an F_1 and F_2 level, to a rounded variant with a somewhat more fronted tongue position. This possibility requires further investigation and it is tempting to speculate on the effect that this potential mismatch between articulation and acoustics has had on impressionistic transcriptions of this vowel across different accents of English. Of course lip-rounding also has a discernable effect on F_3 and listeners and expert phoneticians no doubt use this dimension in identifying the vowel in question. Still, the overall point remains, given that two sounds with differences on all three formant levels would, presumably, be easier to distinguish than two sounds which differ solely in terms of F_3.

[12] Webb's (1983) *girl* tokens were excluded from this part of the analysis.

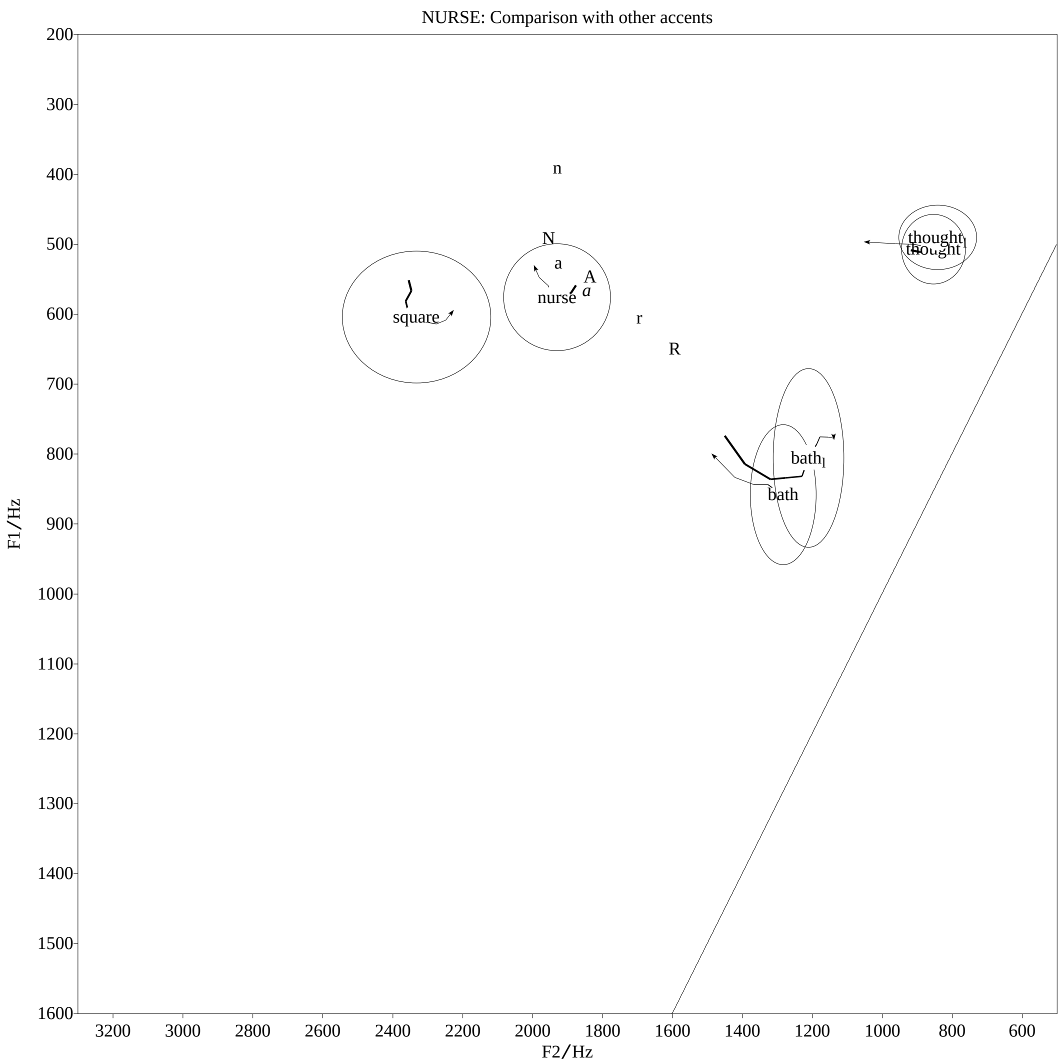


Figure 9.4: SAE NURSE vs. other English accents

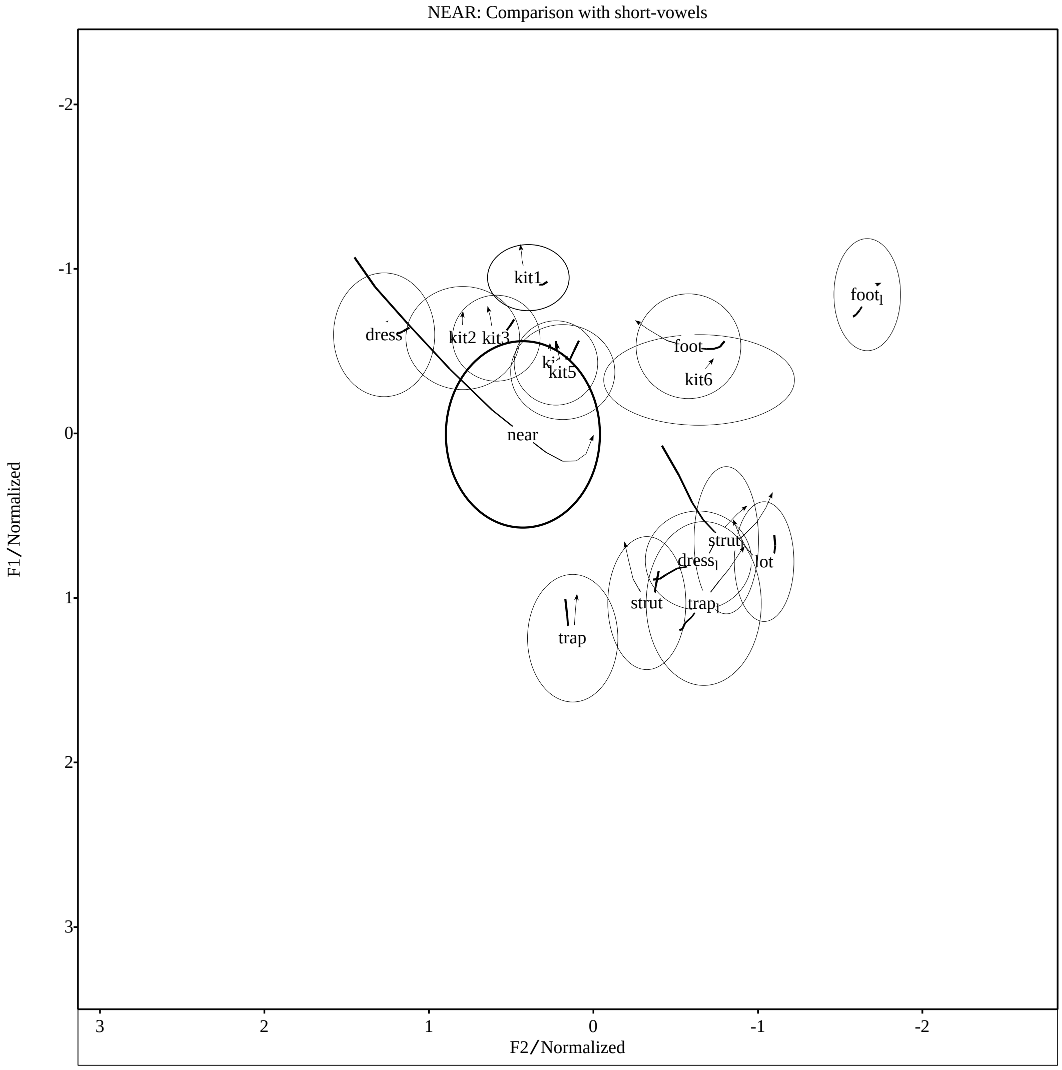


Figure 9.5: NEAR: Comparison with the short-vowels

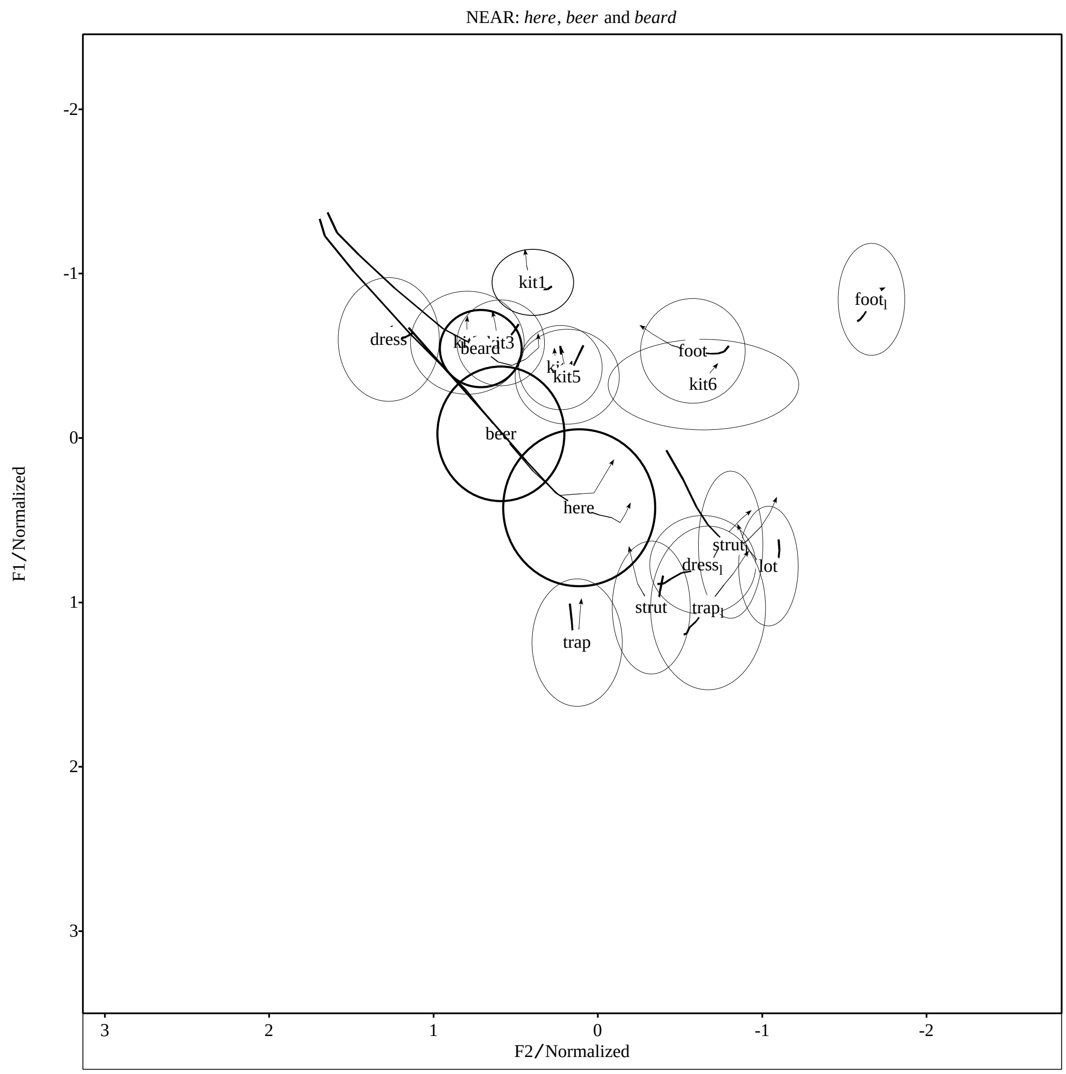


Figure 9.6: NEAR: *here*, *beer* and *beard*

begins at a far more open position (overlapping primarily with DRESS) than *beer*. Secondly, the terminus of the glide appears to be reached far sooner for *here* than for *beer*. This is evident from the fact that the label (placed on the 50th percentile) is far lower for the former than the latter. In *here*, therefore, the second target is likely to receive a fair degree of prominence. We notice, too, given the standard deviations of both these tokens that there are certainly cases of both *beer* and, especially *here*, having second targets approaching an [ɐ]-like value. The use of an [ɐ]-like value has already been mentioned in the review of the impressionistic literature above.

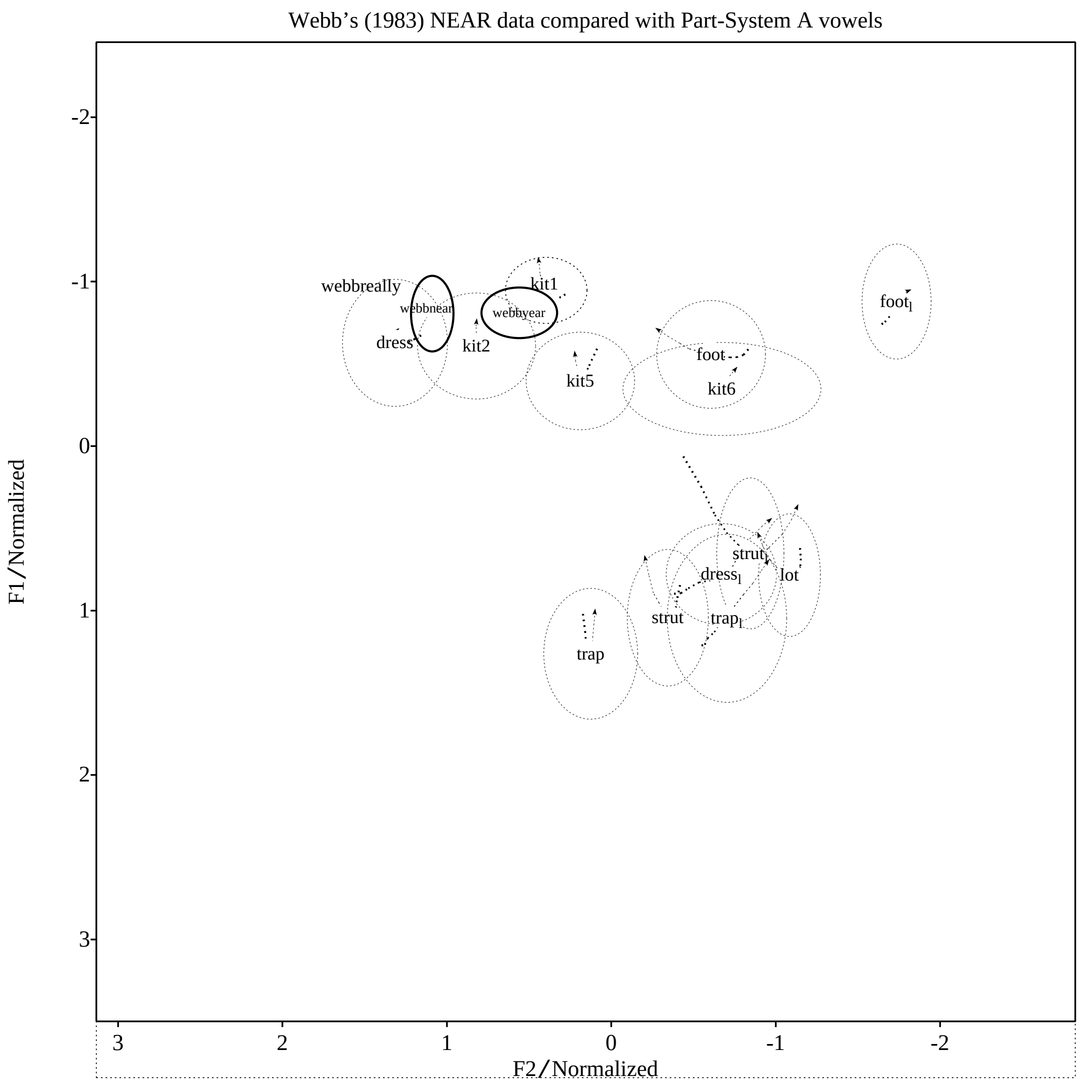


Figure 9.7: Webb's (1983) NEAR data compared with Part-System A vowels

Figure 9.7 on the preceding page superimposes, in bold, Webb's (1983) NEAR data over the current Part-System A data; and should be compared with Figure 9.6 on page 383. On the surface the value for NEAR (i.e. excluding *year* and *really* for the moment) appears to have a value close to the fronter allophones of KIT. However, as emphasized above, it is possible that the relatively lax position of this vowel in Webb's (1983) data is an artefact of his measurement procedure. As far as *really* is concerned, if we assume monophthongisation of this vowel, then a DRESS-like value seems well within expectations. Lastly, with respect to *year*, it was mentioned above that this vowel could, effectively, be analyzed under NURSE. This is evident in Figure 5.5 on page 152 as well as in Figure 9.8 on the following page, which superimposes Webb's (1983) *year* over the current-data NURSE vowel.

Figure 9.9 on page 387 provides a comparison between the current NEAR data and that from other accents of English: RP data from Cruttenden (2001:99) and AusE data from Cox (2006:175); for the latter only one data point is provided since, according to the author, in AusE, NEAR in an hVd context is typically realized as monophthongal (Cox 2006:150). We note the relatively close position of NEAR for AusE in this regard. Perhaps more interestingly, we note the completely different glide for RP. From a relatively high-front position to a high-central one; this is in itself somewhat surprising given the usual transcriptions for this vowel in RP.

Turning to SQUARE, Figure 8.2 on page 296 shows very little evidence of in-gliding in the data. There is, however, a small amount of spectral movement, perhaps due to coarticulatory pressures. The relatively large standard deviation does indicate a fair degree of variability with respect to this vowel. Plotting the various words separately, as provided in Figure 9.10 on page 388, explains the variation, and the slight degree of spectral movement, but doesn't suggest a principled explanation: *scare* and *hair* both have discernable glides, *share* less so, while *fair* seems to be stable and lower than the rest.

The SQUARE vowel is, as expected, also in a relatively close position, a fact confirmed clearly in Figure 9.11 on page 389, where there is evidence for a substantial overlap between DRESS and SQUARE. It appears that SAE has a genuine case of two phonemes being distinguishable solely in terms of their length: the average length of DRESS in the current data being 0.150 seconds (N=121, sd=0.036) while SQUARE is, on average, 0.287 seconds long (N=61, sd=0.053). The FACE vowel has also been included in this figure, because it clearly shows that, in the current data at least, FACE and SQUARE have similar values. It is clear from this diagram that the first half of the FACE vowel overlaps with SQUARE and that all that distinguishes the two vowels is, in the case of FACE, a very weak-glide to a slightly more raised position. The position of FACE *vis-à-vis* SQUARE provides evidence that in at least citation-form style GenSAE returns FACE to the peripheral track i.e. it

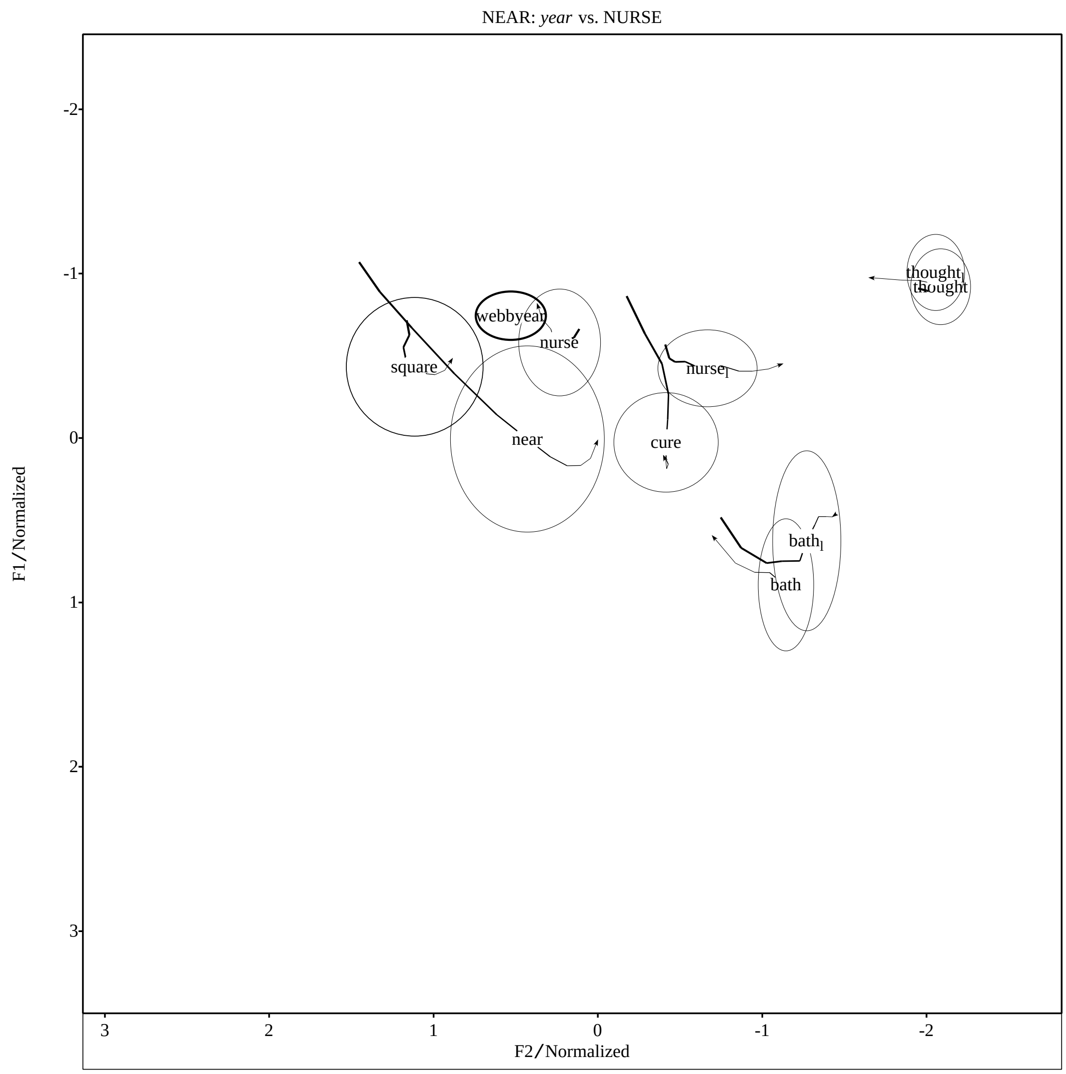


Figure 9.8: NEAR: Webb's (1983) *year* vs. NURSE

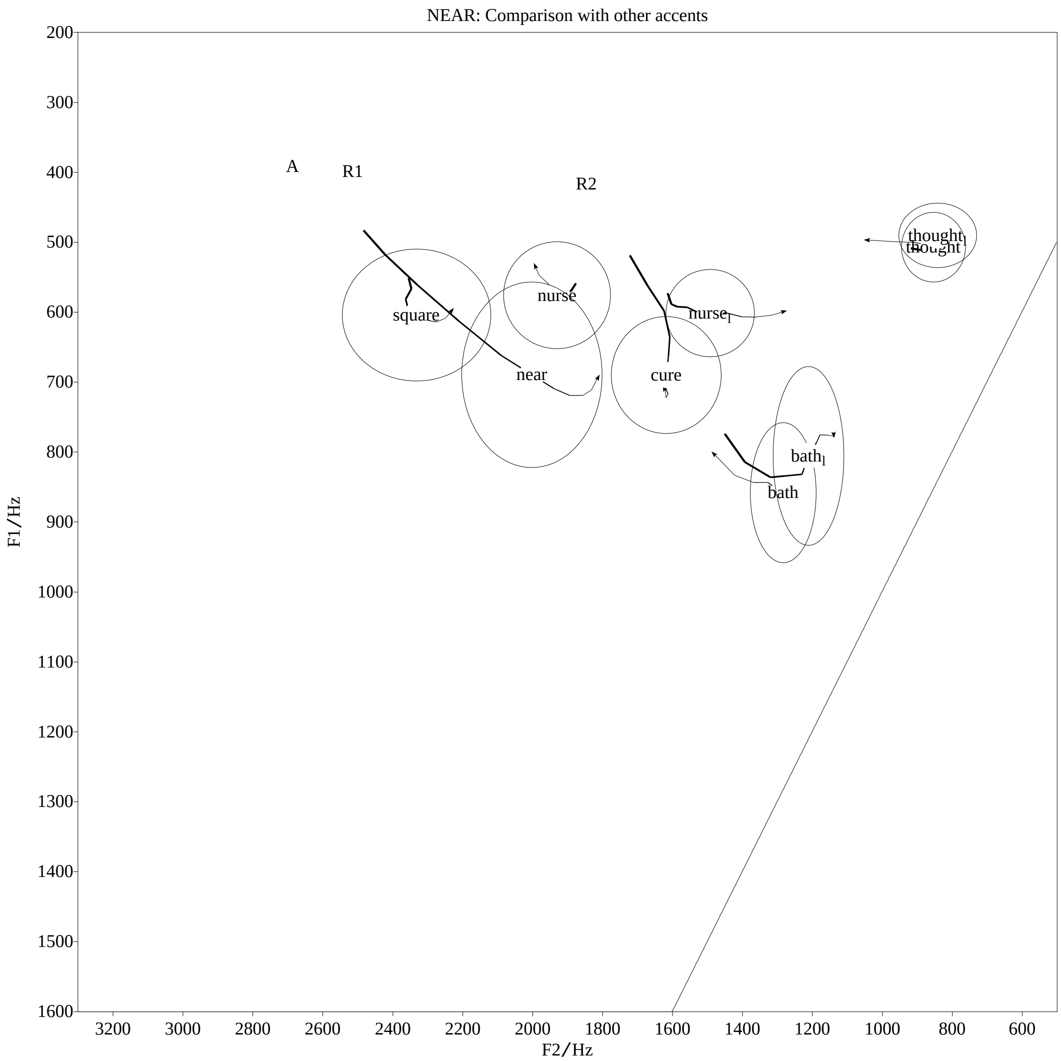


Figure 9.9: SAE NEAR vs. other English accents

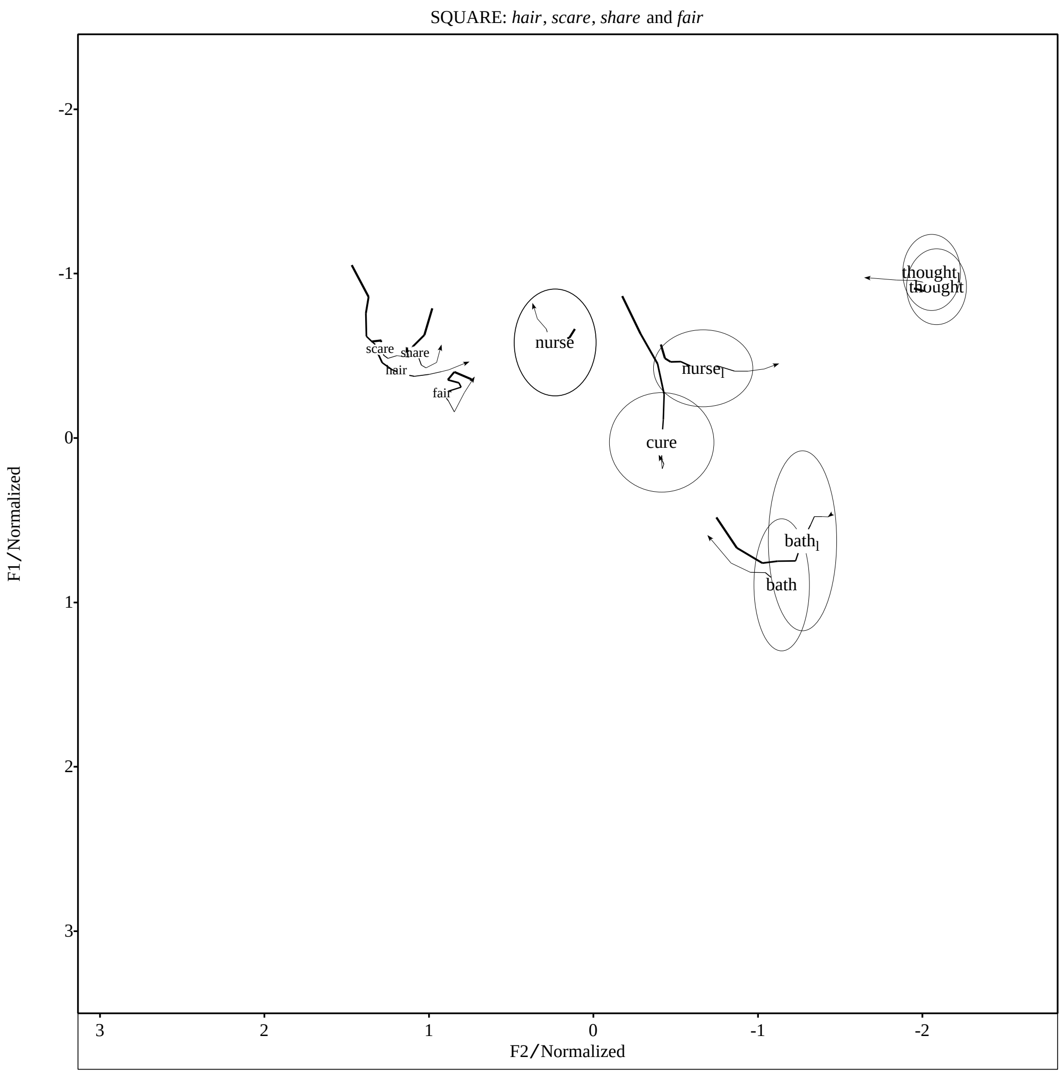


Figure 9.10: SQUARE: *hair*, *scare*, *share* and *fair*

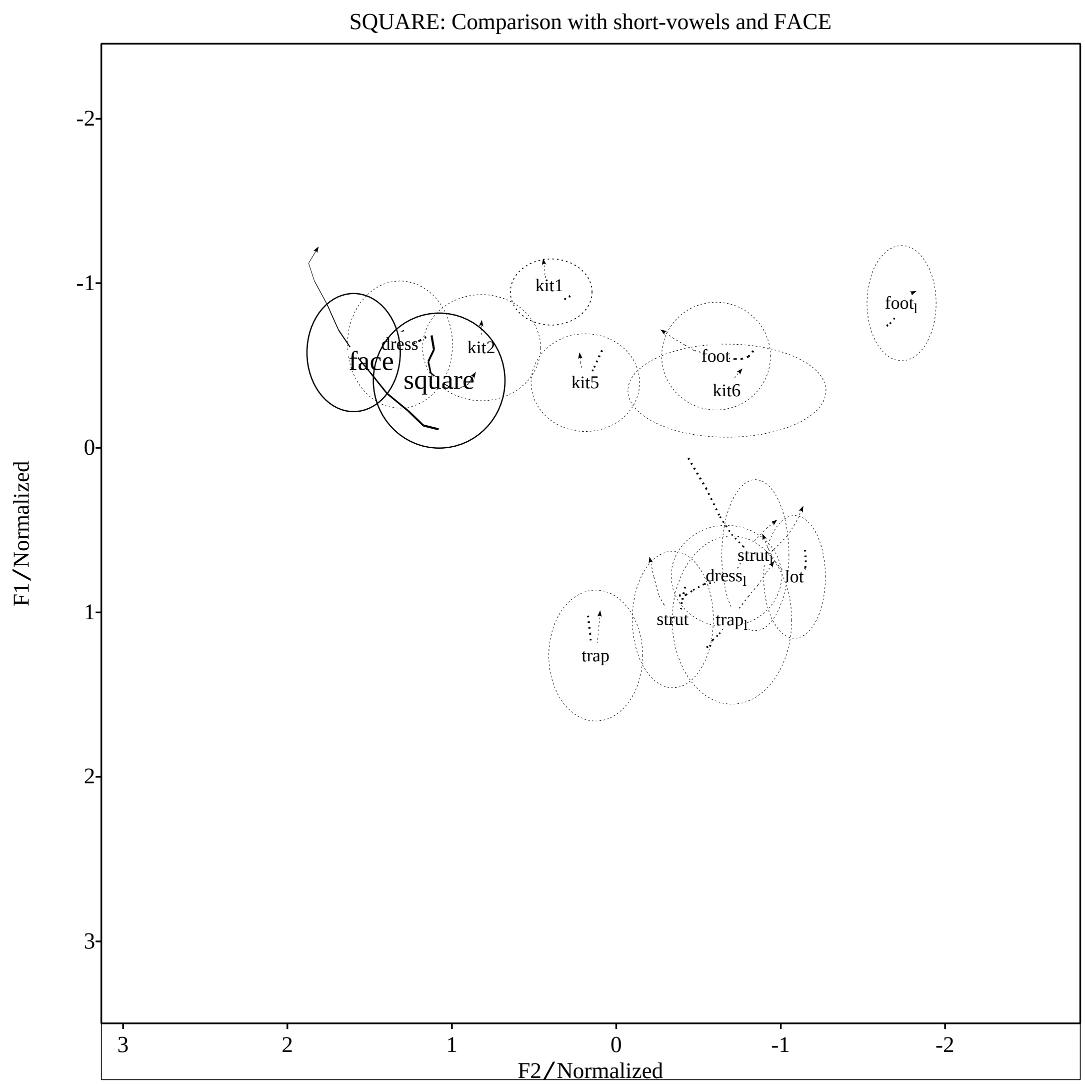


Figure 9.11: SQUARE: Comparison with Short Vowels and FACE

effectively reverses the effects of a Pattern-4 chain-shift movement.

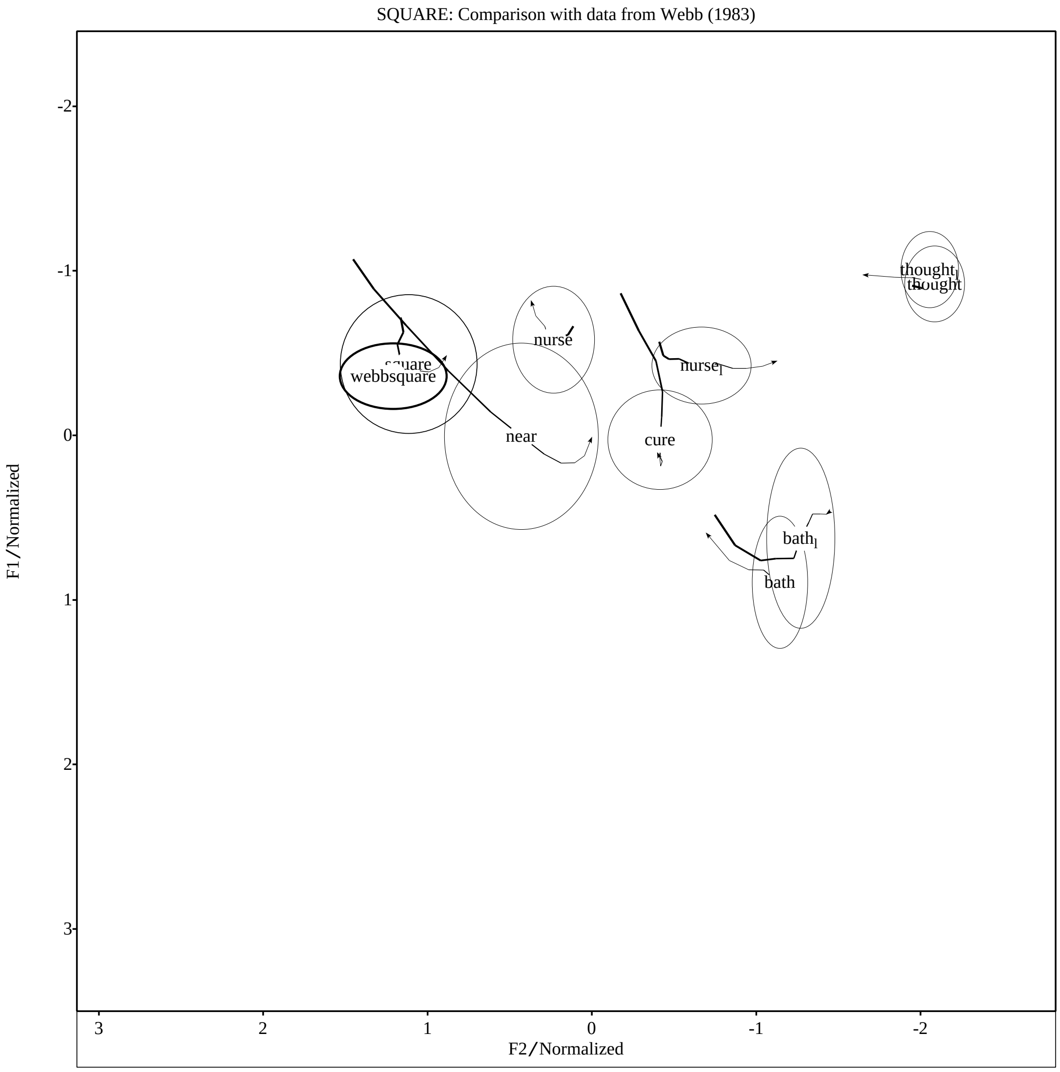


Figure 9.12: SQUARE: Comparison with data from Webb (1983)

Figure 9.12 compares SQUARE across the two SAE data-sets, while Figure 9.13 on the facing page compares the current data with that of other accents of English, in particular RP and General AusE. In the first figure it is clear that SQUARE in SAE is pretty-much categorical across the two data sets, with little evidence of variability, while in the second figure one notes that SAE and AusE have similar values for this vowel, with the later appearing to be slightly more close. The diphthongal character of RP SQUARE is apparent.

Lastly, it should be mentioned that in many NEAR and SQUARE tokens (but less so in

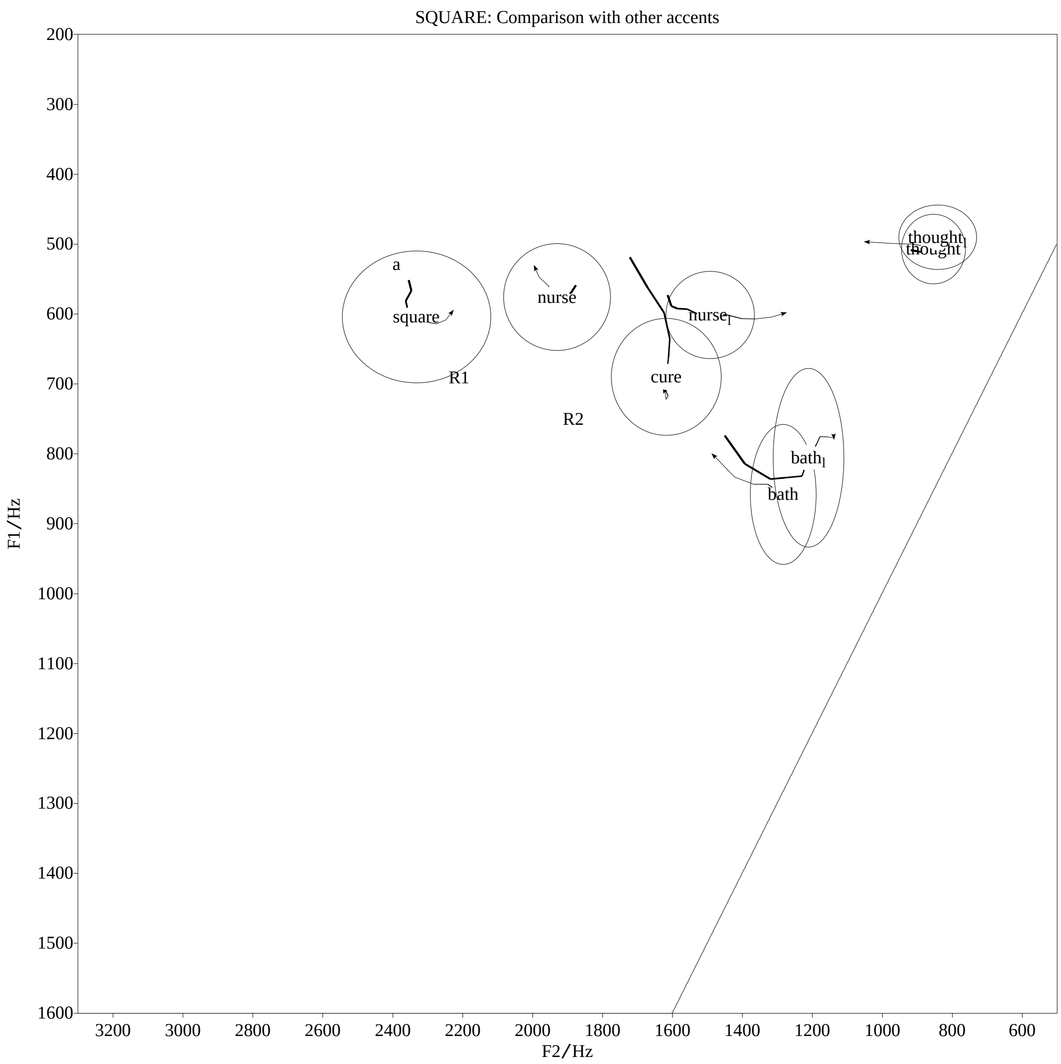


Figure 9.13: SAE SQUARE vs. other English accents

NURSE) there were clear indications of post-vocalic /r/ i.e. the neo-rhoticity mentioned in §9.1 and also found in Harrington et al.'s (1997) AusE citation-form data – as reported on in §9.2.1.

9.2.4 NNS: Synopsis

Beginning with NURSE, it is interesting to note that, as reported on in 9.2.2, Lanham (1967) gives the rather close, but unrounded, [ɘː] as the value for General RP and Cruttenden (2001) also gives 'close-mid' as one possible realization of NURSE in this reference accent. Lower values are characteristic of Refined RP as well as CulSAE. As reported on in §3.4.3, Trudgill (2004) views the *raising*, rounding and fronting of NURSE across the Southern Hemisphere Englishes as an example of language drift i.e. it cannot be accounted for in terms of the original linguistic input that went into the formation of these Extraterritorial Englishes. The somewhat close values reported on by Lanham (1967) and Cruttenden (2001) for General as opposed to Refined RP are perhaps indicative therefore of a similar (but retarded) process of language drift taking place in RP. Indeed the values for RP shown in Figure 9.4 on page 381 seem to indicate a *relatively* close vowel but one still behind the Southern Hemisphere Englishes in terms of both raising and, more particularly, fronting. The data represented in this Figure, assuming rounding of NURSE in the Southern Hemisphere varieties, also provides strong evidence for the extensive fronting of this vowel. An anomaly is the somewhat more raised position of the NZE NURSE data *vis-à-vis* the AusE and SAE values.

While rounding of this vowel has, as far as I can ascertain, not been reported for RP, rounded versions of this vowel are found in various accents of England, including Cockney, Liverpool, Newcastle and Lancashire. There appears to be some disagreement in the literature regarding the unrounded or rounded status of NURSE in AusE, while its rounded status in NZE and non-CulSAE sociolects of SAE appears to be uncontroversial. Lass' (2004) remarks concerning the appearance of a rounded NURSE vowel in "most of the English-speaking world where ME /ir, ur, er/ have merged," (Lass 2004:376) are potentially troubling for Trudgill's (2004) above-mentioned position since, although the general appearance of a rounded NURSE vowel could, hypothetically, be used to defend the 'drift' thesis, it, more likely, points to an underestimation of the incidence of NURSE rounding, both synchronically and diachronically. This in turn links to the fact that, on an F_1 and F_2 level at least, a rounded and unrounded NURSE (with different compensatory tongue positions) can have virtually the same value; a phenomenon which could, no doubt, lead to incorrect transcription.

If we add to these observations the position taken by Lanham in his various publications on the influence of Natal English on the development of GenSAE, there appears to be enough evidence against Trudgill's (2004) position to warrant a more thorough (particularly diachronic) investigation of NURSE. In particular, the fact that North-of-England varieties were substantial influences in the development of Natal English (and thus GenSAE), renders Trudgill's (2004) claims to the contrary suspect.

Historical arguments aside, the SAE acoustic data analyzed above provides evidence for a relatively invariant, monophthongized, comparatively raised, rounded and fronted value for NURSE in GenSAE. Given that there is little reason to assume that Webb's (1983) NURSE vowel was unrounded, the acoustic equivalence of the current data with that of Webb (1983) provides relatively strong evidence for the conclusion that the rounding of this vowel in SAE has filtered into the speech of the upper-echelons, and even into their citation-form speech[13] The relatively fronted position of the SAE data *vis-à-vis* the (presumably) unrounded RP value, despite the effects of lip-rounding, also speaks to the extensive articulatory fronting of this vowel, as supported in the impressionistic literature.

Moving on to the NEAR vowel, the acoustic data seems to indicate a fair degree of variation, both of a phonetically-conditioned and lexical nature. This, in general, supports the extant literature, as does the [ɐ]-like value for the *here*-tokens. There is no evidence for monophthongisation, but the acoustic data does suggest a relatively higher onset than might have been expected. More indirectly, the [ə]-like value of the terminus of the NEAR glide provides evidence for the non-schwa value of the centralized allophones of KIT.

As far as the historical origins of SAE NEAR is concerned, there is ample evidence of input from various sources:

- RP [ɪə], coupled with the use of NURSE in *here* etc, as well as the monophthongisation of this vowel to [ɪː].
- Monophthongal variants in Cockney, as well as a closer starting point i.e. [iə][14]. Also the use of triphthongal variants, [ɪjɐ], sentence-finally.
- In north-of-England varieties, word-internal monophthongisations in words such as *serious*, as well as disyllabic variants[15].

[13] As emphasized already, it is technically possible that an unrounded vowel with a more retracted tongue position could have produced the same acoustic effect. The resolution of this issue demands additional research.

[14] Also viewed as a possible influence from Afrikaans.

[15] A reading of the impressionistic literature, particularly Lanham (1967), would seem to imply that such disyllabic variants are only found in BrSAE and are Afrikaans-influenced. I suspect that the use of disyllabic variants in SAE has been underestimated. North-of-England varieties might have a role to play [hiːə].

While the monophthongal realization of a number of NEAR-tokens in SAE does lead to a *partial* merger of NEAR and SQUARE, this merged value is different to the more well-known and more extensive NZE one, [ɪə].

In the case of SQUARE, there is some evidence, mainly from a critical reading of the available impressionistic literature[16], for the post-settlement endogenous raising of this vowel, in according with Labov's (1994) Principle 1, as originally given in §2.3.2, although the fact that the two SAE data sets show little difference in this regard suggests that, if endogenous, this change has almost reached completion. A close [eː]-value appears to be a categorical value in modern non-Cultivated SAEP, even in citation-form GenSAE, and thus a broad indicator of SAEP. The distinction between DRESS and SQUARE is based solely on duration and there is, in at least some styles of GenSAE, very little to distinguish FACE and SQUARE either.

While endogenous raising, from a lower and glided RP-like version, appears to be one possible source for the development of a [eː]-like quality in SAE, there were also possibly some historical inputs into this development. Cockney, for example, currently shows both close (but diphthongal) and monophthongal (but open) values and monophthongal variants appear to be common in many varieties of Britain, including the north-of-England; all, presumably, reflecting possible qualities at the time of SAE's conception.

That there would not have been a structural pressure on SAE to select a *glideless* variant is evident from the fact that, as reported on briefly in §2.3.2, one of the common characteristic of tense vowels rising along the peripheral track is their acquisition of a distinct in-glide, as, for example, has been the case with short **a** raising in certain accents of the USA: [mɪən], or such-like, for /mæn/ *man*.

All-in-all it appears, thus, that the modern SAE value is, in all likelihood, a result of both a large variety of values in the original pre-koinéization feature pools, including open, close, monophthongal and diphthongal variants, coupled with a endogenous pressure to select raised variants. Monophthongal variants were, presumably, in the majority.

9.3 The Weak Vowels: WV

This section deals with the vowels in so-called 'weak' (i.e. unstressed) position. These include lettER, commA, happY and weak-vowels in non-final position.

According to Wells (1982:166), the lettER set relates to word-final orthographic *r* and *re* and is "regularly pronounced with /ər/ (usually realized as [ɚ]) in rhotic accents … and

[16] In particular, the impressionistic literature seems to *progressively* report more close (and monophthongal) GenSAE values.

with plain /ə/ in non-rhotic accents". Examples are *paper*, *offer*, *usher*, *calendar*, *martyr*, *figure* and *tenure*. The vowels lettER and commA are usually pronounced in the same way in non-rhotic accents i.e. as [ə]. The latter set has orthographic *-a* word-finally and includes words such as *sofa*, *opera*, *saga* and *phobia* (Wells 1982:167).

According to Wells (1982:165), the happY set includes "words with orthographic *-y*, *-ie* and *-i* ... and also words with orthographic *-ee*, *ey*, *ea*". Examples include *copy*, *fluffy*, *khaki*, *birdie*, *committee*, *money* and *Chelsea*.

9.3.1 WV in Other Accents

Received Pronunciation: RP commA and lettER generally have [ə]. happY is realized as [ɪ], although a lower DRESS-like quality is often heard in the speech of speakers of Refined RP (Wells 1982:165;281). Lanham (1967:31) gives [ɪ][22]{[ɪ]} as the quality of happY in RP. Wells (1982:294) does, however, mention a growing tendency for speakers of RP to use a closer quality, much like that of FLEECE. A recent acoustic study by Fabricius (2002) focuses on, among other things, the happY vowel in modern RP. This author comes to the conclusion that "younger [RP] speakers [tend] to settle upon an intermediate phonetic value for happY words" (Fabricius 2002:232; my parenthesis) i.e. intermediate between the two extremes of [i ∼ iː] and [ɪ]. Figure 9.14, taken from Figure 5 in Fabricius (2002:225), provides an example of such intermediate happY values. As can be seen, the circled areas mostly encompass the KIT([ɪ]) and FLEECE([iː]) vowels, while the happY tokens lie mostly in-between these two extremes[17].

Traditionally, RP has not undergone the Weak Vowel Merger i.e. it retains a contrast between /ə/ and /ɪ/ in weak vowels in non-final position. An example provided by Wells (1982:167) is *rabbit* vs. *abbot* which do not rhyme for some speakers of RP i.e. [ræbɪt] vs. [æbət] in broad phonetic transcription. Bailey (1984:15), however, claims that, with respect to RP [ɪ] and [ə] in unstressed position, "the former is giving way to the latter in many contexts".

Lanham's (1967) characterization of weak vowels in non-final position in RP de-

[17] Figure 9.14 on the following page is based on happY tokens from "adverbs derived by the addition of the suffix *-ly*" (Fabricius 2002:217) i.e. unlike the tokens in this study which are mostly monomorphemic – see §9.3.3. Fabricius (2002), however, found the same pattern across '*-ly*' and '*-y*'-type happY tokens; thus, with respect to the latter, she remarks that "in the case for [sic] happY-tensing in *-y* forms, all speakers showed a general tendency towards an intermediate or fronted value for the final vowel in *-y* words, either midway between the KIT and FLEECE ellipses or with the majority of tokens within the FLEECE area" (Fabricius 2002:228; my parenthesis). Figure 9.14 was chosen as a relatively representative (and easy to 'read') representation of Fabricius's (2002) overall results.

pends, in particular, on his notion that this variety has a true (although marginal) phonemic KIT split – see Chapter 7. For this author, the split between /i/ (realized as [ɪ]^{22} {[ɪ]}) and /ɜ/ (realized as [ɪ̈]^{24}{[ɨ̞]}) is based on words such as *pin* /pin/ which render the complementation between these two qualities non-categorical and thus indicative of a marginal contrast. According to Lanham (1967:45), in unstressed syllables /ɜ/ is contrasted with /ə/ accounting for the difference in RP between, for example, *allusion* and *illusion*, with the former beginning with /ə/ and the latter beginning with /ɜ/. Phonetically, /ə/ is realized as [ə]^{44}{[ə]} and /ɜ/ as [ɜ]^{34}{[ɘ]}. [ɨ̞] and [ɘ] are thus two allophones of /ɜ/, the first occurring in stressed syllables, the second in unstressed syllables. A word such as *visibility*, as pronounced with an RP accent, would therefore be transcribable, in Lanham's (1967) terms, though using modern IPA symbols, as [vɨ̞səbɨ̞lətɪ], given that the second and fourth syllables are unstressed. It is important to emphasize that for this author the contrast between /ɜ/ and /ə/ is the only one operating under conditions of weak stress. The use of /ɪ/ is thus excluded and only occurs under conditions of tertiary stress or higher. Lanham's (1967) conception of the weak-vowel contrast in RP is thus different to that of Wells (1982), at least in terms of the phonetic values assigned to the two poles of the

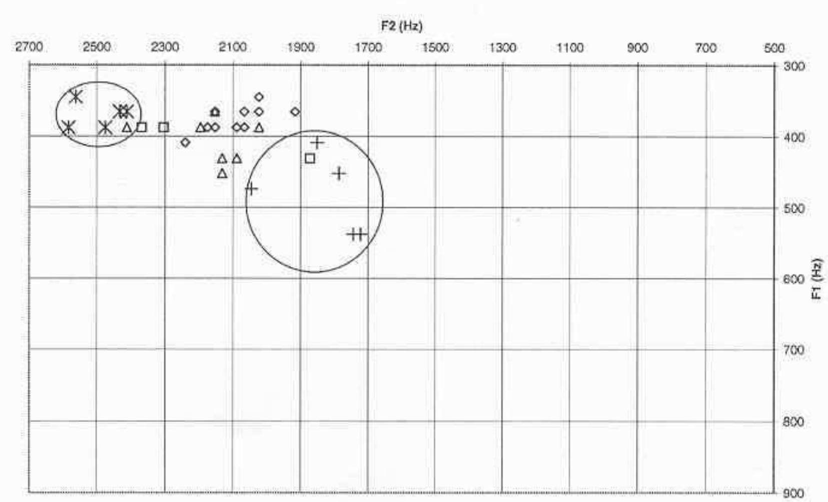


FIGURE 5. happY, *-ly*, Catherine. ◇ *-ly*#C, □ *-ly*#P, △ *-ly*#V, ⋇ FLEECE, + KIT.

Figure 9.14: happY in modern RP, taken from Fabricius (2002:225)

contrast.

Other Accents of England: As in RP, Cockney lettER and commA appear to be realized as [ə].

A closer quality for happY is, however, according to Wells (1982:166), characteristic of the south of England. Wells (1982:319) gives [əi ∼ ɪi ∼ i] for London. This is in contrast with a typical northern accent which mainly has an RP-like [ɪ] although more open and more close FLEECE-like variants are also found, depending on the specific area under consideration (Wells 1982:362).

As far as the Weak Vowel Merger is concerned, a number of accents are variable in terms of retaining the distinction "along with many accents of England including some which could perhaps be considered to fall within RP" (Wells 1982:167). Perhaps more importantly, Trudgill (2004:117–20) shows convincingly, as already alluded to in §3.4.1, that the Weak Vowel Merger was not a characteristic of London and the southeast of England during the period which SAE and the other Southern Hemisphere Englishes were established. In fact, it appears that the Weak Vowel Merger was only characteristic of, at the time, East Anglia[18], the West Country, the far north of England and Ireland. Scotland and most of England retained the relevant distinction.

AusE and NZE: As in RP and Cockney, AusE and NZE lettER and commA are realized as [ə]. A close quality for happY is the most common quality for these two accents as has already been mentioned in §3.4.3, as part of the description of drift in Southern Hemisphere Englishes, as advanced by Trudgill (2004). Wells (1982:602) gives [iː] for AusE and the same quality for NZE (Wells 1982:607). The Weak Vowel Merger occurs in the Southern Hemisphere varieties i.e. there is no distinction between the final vowels of *Lenin* and *Lennon*, which in RP would be pronounced as [ˈlenɪn] and [ˈlenən] respectively (Wells 1982:167). Wells' (1982) description of the Weak Vowel Merger in AusE is pertinent because it highlights the fact that the merger need not lead to the use of only one of the relevant qualities. The issue here is one of phonemic contrast vs. complementary distribution. Thus, except in the case of speakers of Cultivated AusE, there is a tendency for [ə] and [ɪ] to occur in complementary distribution, with the latter finding itself preceding velar consonants e.g. *paddock* as [pædɪk] (Wells 1982:601).

[18] In Trudgill (1999*a*:126–7), the author claims that "this feature of East Anglian English may have had some input into the formation of southern hemisphere Englishes such as Australian English".

U.S. Accents: With respect to commA and lettER, in the case of the Southern USA "there is fluctuation ... between /ɪ/ and /ə/ in commA, although /ɪ/ here is definitely regarded as non-standard" (Wells 1982:552).

According to Wells (1982:167), it appears that southern American English accents also retain the contrast between /ɪ/ and /ə/ in non-final position i.e. they have not undergone the Weak Vowel Merger. GenAm tends to be variable in this regard.

Flemming and Johnson (2007) provide an acoustic analysis of the commA vowel as well as weak vowels in non-final position in non-southern American English. As confirmed by these authors, transcriptional practice has usually given [ə] as the value for non-final unstressed vowels (as in *begin* [bəgin]), with a barred-i ([ɨ]) being used to transcribe the sound in certain affixes e.g. the vowel in the plural suffix in *roses* [roʊzɨz] and *hinges* [hinʤɨz]. Using acoustic evidence the authors show convincingly that, while [ə] should be retained for commA, all other non-final unstressed vowels have a high central quality and should, consequently, be transcribed as [ɨ]. While Flemming and Johnson (2007) are focussed on American English they provide some interesting speculations about weak vowels and the Weak Vowel Merger in RP. In essence they argue (tentatively) for [ə] for commA and lettER, [ɪ] for words such as *Lenin* and [ɨ] for *Lennon*.

A close happY vowel appears to be the most common quality for GenAm (Wells 1982:166). The American South is, however, characterized by the use of [ɪ], as well as more conservative varieties of GenAm (Wells 1982:551; 165).

9.3.2 WV in SAE: The Impressionistic Data

Hopwood (1928): With respect to happY in SAE, Hopwood (1928:18) claims that '**ɪ** > ... **ì(j)**{[i(ː)]} in final syllables e.g. E. mealy ˈ**mì·jlɪ**{[ˈmiːlɪ]}, SAE **mì·lì(j)**{[miːli(ː)]}'. Hopwood (1928) believes that this is both the result of sound substitution with Afrikaans[19] "and in many instances also an A. spelling substitution, as in English lad*ie*s" (Hopwood 1928:18); the letters *ie* are usually pronounced as [i] in Afrikaans.

Turning to weak vowels in non-final position, for SAEP and according to Hopwood (1928:17), words such as *spirit, Phyllis, profit, wicked, sounded* etc. all have [ë] or [ə] (as opposed to [ɪ]) in their second-syllable i.e. in *closed* as opposed to open unstressed final syllables. Interestingly, Hopwood (1928:19) seems to suggest that the source of this merger might be Scottish English: "The pronunciation of **ë** or its

[19] Thus Afrikaans *mielie* "corn" is transcribed as ˈ**mìlì**{[ˈmili]} in Hopwood (1928:18).

omission in weakly stressed syllables of SAE . . . *e.g.* . . . E. freshness ˈ**fɹɛʃnɪs**, SAE ˈ**frɛʃnës**, ScE **frèʃnès** or ˈ**frèʃnəs** . . . is paralleled by the alternative pronunciation of the neutral vowel or a stressed vowel in ScE secondary syllables".

Lanham and Traill (1962), Lanham (1967): Lanham and Traill (1962:19) confirm that lettER and commA have a [ə]-value in both SARP 'A' and SARP 'B'.

For SARP happY, Lanham and Traill (1962:19–20) claim a short (relatively untensed) vowel, although a lengthened, tensed version "is an occassional alternative" (Lanham and Traill 1962:19; footnote). The actual values are "in the region of [ɪ]^{22}{[ɪ]} although a more retracted [ɪ]$^{22/23}${[ɪ̠]} is possible" (Lanham and Traill 1962:20) i.e. the exact same values Lanham and Traill (1962) provide for KIT in word-initial position or following /h/ – see Chapter 7.

For non-SARP happY, however, Lanham and Traill (1962:21) provide the values [ɪ]^{21}{[e̞]} to [i]$^{11/21}${[i̞]}, although no mention is made of extra length. In terms of the vowel's quality, Lanham (1967:68; 81) goes a little further, providing [i]^{11}{[i]} and [ɪ]^{21}{[e̞]} as indicative of the possible range of values.

With respect to weak vowels in non-final position, Lanham and Traill (1962:19) confirm that unlike SARP 'A', which very infrequently has a [ə]-value in weak closed syllables (or closed syllables under tertiary stress), SARP 'B' often displays this phenomenon, choosing [ə]^{44}{[ə]} above [ɜ]$^{33/43}${[ɘ̟˕]} "in the syllables with weak or tertiary stress in the following words: *óffìce*, *báskèt*, *fócùs*, *èxqúisite* (final syllable), *chíldrèn*, *wíckèd*, *mìstáke*" (Lanham and Traill 1962:19)[20].

It should be emphasised here that Lanham and Traill (1962) are not equating SARP 'A' with RP. The use of [ɜ]$^{33/43}${[ɘ̟˕]} "is a lower representative of /ɜ/ than is its counterpart in RP" (Lanham and Traill 1962:19): Lanham and Traill (1962:22) give [ɜ]^{33}{[ɘ̟]} to [ɪ]$^{22/23}${[ɪ̠]} as the RP value [21]. RP also often has /i/{/ɪ/} where SARP 'A' has /ɜ/. Thus, a word such as *wicked* would have been pronounced in the RP of the time as [wɪkɪd], SARP 'A' would have had [wɪkɘ̟˕d], while SARP 'B' would have had [wɪkəd].

The position is, however, far from categorical in that not *all* SARP 'A' weak syllables have in the region of [ɘ]. Thus Lanham and Traill (1962:25) claim, for SARP 'A', as follows:

[20] To orientate the reader, in the authors' transcription system "primary, secondary and tertiary stress . . . are marked by /í î ì/ and weak stress is usually left unmarked" (Lanham and Traill 1962:14; footnote).

[21] While this is in minor contrast to the value given in Lanham (1967), which is [ɜ]^{34}{[ɘ]} (see below), the point remains the same.

> "Both [ɜ]{[ɘ]} and [ə] are heard in weakly stressed syllables, some SARP 'A' speakers contrasting them in the identical final syllables of the pair: *hággis* /hægɜs/: *bógus* /bɜwgəs/".

The Weak Vowel Merger has, thus, not completed itself in SARP 'A', although the two values in question appear to be different to that found in RP. This seems to be confirmed by Lanham and Traill (1962:25; footnote) when they claim that in "a more formal style . . . /ɜ/ . . . changes to /i/ and /ɜ/ substitutes for /ə/". Of importance is the authors' conclusion that "because [ə] is not part of the vowel pattern in strongly stressed syllables in SARP 'A', /ə, ɜ/ must be treated as two vowels in a separate subsystem occurring only with weak stress" (Lanham and Traill 1962:25).

The trend seems to be a gradual replacement of /ɜ/ with /ə/ as the accent gains distance from SARP 'A'. Thus SARP 'B' generally has [ə] as opposed to [ɘ̟˕], while with respect to non–SARP SAE "it is possible that dialects exist in which [ɜ] is completely replaced" (Lanham and Traill 1962:19).

Lanham (1967) adds little to the above analysis, except to confirm that in conservative SAE /ɜ/ and /ə/ are distinguishable, as they are in RP, while in non-conservative SAE the trend is towards the dissolution of this contrast.

Lanham (1978), Lanham and Macdonald (1979), Lanham (1982): Of relevance to happY-tensing in SAE, is Lanham's (1978) description of "stress-raised word-final weak syllables" (Lanham 1978:154), whereby weak word-final syllables (such as the final syllable of *mo<u>ney</u>*) receive a higher degree of stress (with rising pitch) than is usual in RP. It is seen "as a feature of the Respectable variety" (Lanham 1982:340). This is ultimately an intonational phenomenon with the implication that it only occurs "before pause without finality, i.e. signalling temporary suspension of utterance" (Lanham 1978:154). Lanham (1978:154) claims that stress-raised word-final weak syllables are a defining feature of GenSAE, particularly prevalent among "young upper and middle-class females of Eastern European descent group". This intonational feature is applicable to other weak vowels i.e. those in a final syllable but not *word*-final themselves, as well as to (presumably) lettER and commA.

Lanham and Macdonald (1979:47) generally confirm the above, identifying happY-tensing as a GenSAE variable. They provide the following description of happY in SAE:

> ". . . cardinal i, with additional length and stress-raising in the most advanced variants. This phonetic trend is widespread, particularly in the under-45 age group,

> and is entirely below the level of social consciousness. It is never corrected" (Lanham and Macdonald 1979:47).

Of interest is the claim of Lanham (1982:338) that "stressed-raised word-final weak-syllables" are also found in CulSAE. It is unclear whether this reflects a reassessment of CulSAE or is indicative of the growing influence of the local norms. As in the case of GenSAE, this feature applies not only to happY but to weak-vowels in non-final position (i.e. in the final syllable) and (probably) to commA and lettER as well.

Lanham (1978:152) uses the term "low schwa" to indicate the value of weak vowels (and centralized values of KIT) in advanced varieties of SAE. As already mentioned in Chapter 7, which focusses on the KIT vowel, the prevalent use of "low schwa" in SAE is, according to this author, above the level of social consciousness and is subject to (hyper)correction. "Low schwa" gains prominence in weak syllables generally in SAE i.e. in commA, lettER and in weak-vowels in non-final position.

Lanham and Macdonald (1979:46) add little, confirming that "[ɪ]-quality, unstressed vowels (as in the final syllables of *wanted*, *chicken*) of RP are low schwa" in SAE. The prevalence of "low schwa" in SAE is identified as a GenSAE variable.

Wells (1982), Bailey (1984): Wells' (1982) assessment of happY-Tensing in SAE is of interest because he clarifies the two sides of the happY-Tensing 'coin': a closer, fronter value on the one hand, added duration on the other. Wells (1982:613) explains as follows:

> "The final vowel in *happY* words may be short but close and front, ranging from [ɪ̣]{[ɪ̝]} to [i] . . . But another possibility, apparently increasing, is a long [iː], particularly in pre-pause position. This is a consequence of a more general trend towards strengthening final unstressed syllables at the end of an intonation group in which the nuclear tone is rising"

From this perspective happY-Tensing refers, strictly speaking, to the use of a more tense vowel quality (without an increase in duration), while the increase in duration is rather a suprasegmental phenomenon. These two options, in fact, receive, in Wells' (1982) analysis, two differing phonemicizations i.e. /ɪ/ and /iː/. The first analysis is based on the fact that a short, fronted [i] is found as a realization of KIT, particularly in broader accents of SAE.

For Bailey (1984:18), happY is not equivalent to FLEECE. While acknowledging that it is tensed in SAE, "it is distinct from the tense long vowel /iy/ with which it is in contrast". Thus *refugee* /refyuwjiy/ but *city* /siti/ in which the last syllable of the former is stressed, but not the last syllable of the later[22]. According to this author, tensing of happY (as well as GOAT in *follow* etc.) occurs *after* the assignment of stress in the derivation of the word. In general "English word-final unstressed vowels can . . . be described as exploiting (in a way) maximal phonetic difference corresponding to the well-known [i, u, a] triangle. English realizes these as [i, əu̯, ə], the only unstressed word-final vowels" (Bailey 1984:5). This amounts to claiming that the tensing of this vowel has not yet been subject to phonologization, but is rather a phonetic reflex of added constituent-final duration[23].

Wells (1982:612) confirms the Weak Vowel Merger in non-CulSAE and uses it as a basis for arguing for the phonemic status of /ə/ in SAE. As discussed in Chapter 7, however, the 'split' seems tenuous at best.

Bailey (1984:15), although not confirming a complete merger in SAE, confirms the progressive replacement of [ɪ] with [ə] in unstressed positions. He also points out that [ɪ] still occurs "before velars, hushing fricatives and hushing affricates" (Bailey 1984:18), i.e. /k, ɡ, ŋ, ʃ, ʧ, ʤ, nʤ/, in final syllables. Examples provided are *relic*, *brillig*, *building*, *rubbish*, *Greenwich*, *cabbage*, *orange*. Word-initially, after a glottal stop, is also seen as conditioning [ɪ], as in *explain* or *event*. It should be pointed out, however, that in my own idiolect, words such as *cabbage*, *orange* and *event*, particularly in informal speech, often carry [ə]. Bailey (1984:18) also claims that in a few minor cases a contrast between [ɪ] and [ə] is possible in unstressed position. Thus he claims that "shwa [sic] may occur before velars as in 'recórd' (verb), stómach' which may be compared with 'rélic"'; thus /rəkord/ and /stamək/, but /relik/[24].

Lass (1990), Branford (1994), Lass (1995), Bowerman (2004): According to Branford (1994:477), "*letter* and *comma* words are typically pronounced with final [ə], though this might be strengthened to [a] or [ɐ] by certain speakers". Lass (1995:101) states that [ə] is the main value for BrSAE and GenSAE lettER and commA, "though in Conservative it may be opener, even in some older speakers as open as [ɐ]{[ɐ̞]}". According to Bowerman (2004:939), lettER is schwa in all varieties and "very often omitted before another consonant: [kɪtn̩] *kitten*". The example here is, however, infelicitous since, firstly, *kitten* does not belong to the lettER set and, secondly, rather than being characterized by the omission of unstressed vowels in words such as *kitten*, (Broad) SAE

[22]It should be added that in the author's scheme-of-things there is no secondary stress in English.

[23]See Barnes (2006) for more on this and related matters.

[24]For me *stomach* often *does* have [ɪ] in the second syllable. Secondly, Bailey (1984) seems to miss the point that the velar sound in *record* is not in the same syllable as the [ə].

often uses [ən] or [əl] *instead* of syllabifying the consonant (Wells 1982:613).

According to Bowerman (2004), commA is most often schwa but can have a more open [ɐ] value "in Cultivated SAE; and also in Broad varieties close to Afrikaans English" (Bowerman 2004:939).

According to (Lass 1990:273), like RP, CulSAE has [ɪ] in happY. With respect to GenSAE it is worth quoting Lass (1990:284; my parenthesis) fully:

> "The *happY* class has /iː/ (FLEECE) in the post-tonic syllable, most often with marked secondary stress; e.g. *city* has the same prominence-contour as *sit-in*, that is in essence the contour of a disyllabic compound. Words [ending] in *–y*, *ie* thus have two independent feet, like compounds or simplex words like *rabbi*, *motto*, *feline*".

As discussed in Chapter 7 above, Branford (1994) uses the [ɩ]-symbol to, in all likelihood, represent a [i], and *not* a [ɪ], quality. His comment that "the final vowel of *happy* is usually not conservative [ɪ] but a vowel of [ɩ] quality, of varying lengths" (Branford 1994:477) thus needs to be interpreted in this light and is, in addition, therefore in line with what has been provided above.

Lass (1995:101) confirms that happY-Tensing to [i(ː)], as found in BrSAE and GenSAE, is not a salient sociolinguistic variable. It is only in CulSAE that it has an RP-like and untensed [ɪ]. It often attracts secondary stress: "hence the same vowel in the weak syllable of *city* as in the second element of *ice-cream*" (Lass 1995:101).

According to Bowerman (2004:939), with respect to happY, "the unstressed . . . or secondarily stressed . . . vowel is usually /iː/, but half-long [iˑ]".

Branford (1994:475) observes that "following a velar consonant [sic] in unstressed syllables, such as the second syllables of *ending*, *frantic* and *plastic*, many speakers of SAEP use . . . a relatively high [ɪ]". The author no doubt meant in the context of a *following* velar consonant.

9.3.3 WV: A Comparative Acoustic Analysis

Figure 9.15 on page 405 places the short weak vowels (lettER, commA and weak vowels in non-final position) in the context of the stressed short vowels (i.e. Part-System A). These representations are based on the data provided in Table 9.2 on the following page[25]. As

[25]While the second vowel in *taxis* is, strictly speaking, not word-final, there is no doubt that it has a happY vowel. The weak vowels in non-final position are labeled 'nf' in Figure 9.15.

Lexical Set	No. of Tokens	No. of Measurements	N-G Tokens	O-G Tokens
lettER	41	902	*to**wer**, pa**per***	*mo**ther***
commA	42	924	*are**na**, so**fa***	*co**mma***
happY	85	1870	*co**py**, ci**ty***	*sa**lly**, si**lly**, te**lly**, ta**xis**, ci**ty**, pu**lley***
non-final	48	1056	***a**bbot, ra**bbit***	*woun**ded**, ta**xes***

Table 9.2: Summary of Weak Vowel data

can be seen from Figure 9.15, the values for commA and lettER fall, as expected, in a mid-central position and, for all intents and purposes, overlap in quality. The quality is, in fact, less of interest in its own right than in terms of confirming the non-[ə] position of the centralized allophones of KIT. While commA and lettER do, minimally, overlap with STRUT, they are, in general, at a higher position than STRUT, thus precluding the possibility of a [ɐ]-like value for most of these tokens.

Figure 9.16 on page 406 provides a comparison between the commA and lettER values collected as part of the current project and that provided by Flemming and Johnson (2007:86), for "a variety of US dialects, mainly Western". The latter data is, given the rhotic nature of most US dialects, based solely on commA tokens (like *sofa*) and does not include lettER tokens, which of course, receive a final [ɚ] in rhotic accents. We note from this figure the large overlap between Flemming and Johnson's (2007) data and both commA/lettER on the one hand, and, on the other hand, the more centralized allophones of KIT. Given that Flemming and Johnson (2007:94) transcribe this value as [ə] one is tempted to reanalyze the relevant SAE values as high-schwa ([ɘ]) for central KIT and low-schwa for commA and lettER ([ɜ]).

Figure 9.15 on the facing page also places the weak-vowels in non-final position, labeled 'nf', in their relevant context. We note that, as a whole, this vowel has a much higher position than that found in commA and lettER tokens. This is, in fact, completely in line with the research of Flemming and Johnson (2007) who show "a fundamental distinction between word-final schwa and other reduced vowels".

Figure 9.17 on page 407 confirms this higher position for weak vowels in non-final position, separating out, as it does, the various tokens. Note that *rabbit* (R) is 'hiding behind'

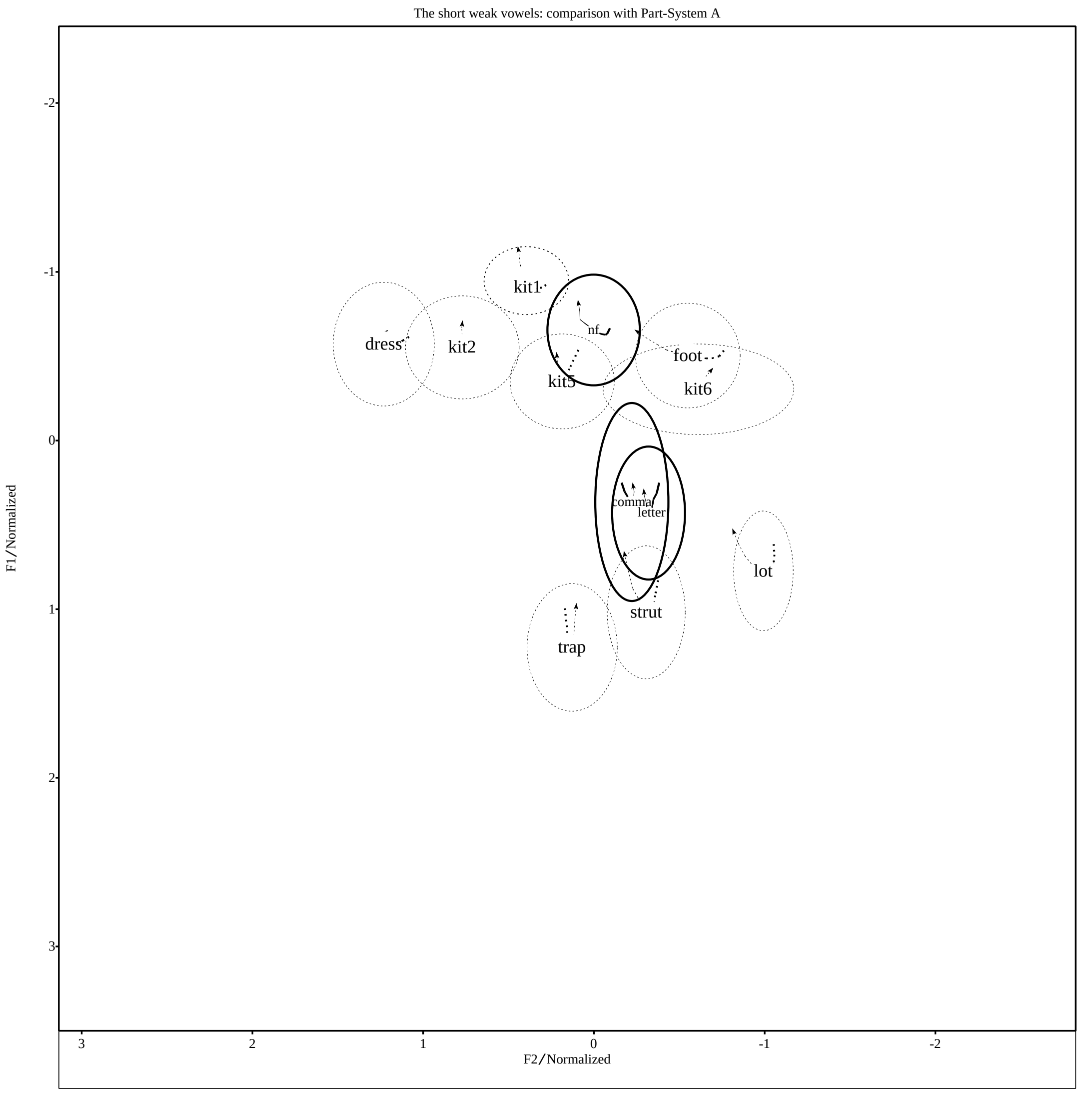


Figure 9.15: The Short Weak Vowels: Comparison with Part-System A

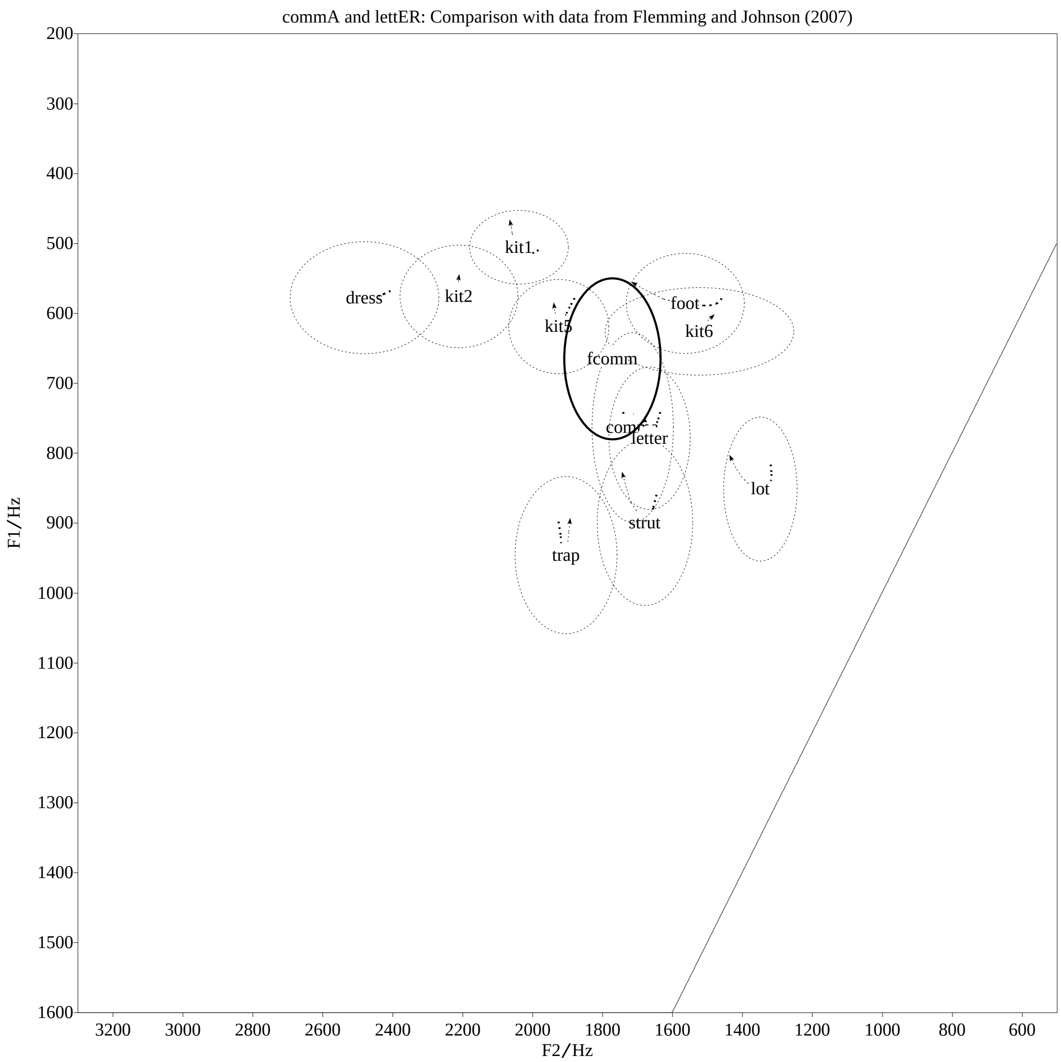


Figure 9.16: commA and lettER: Comparison with Data from Flemming and Johnson (2007)

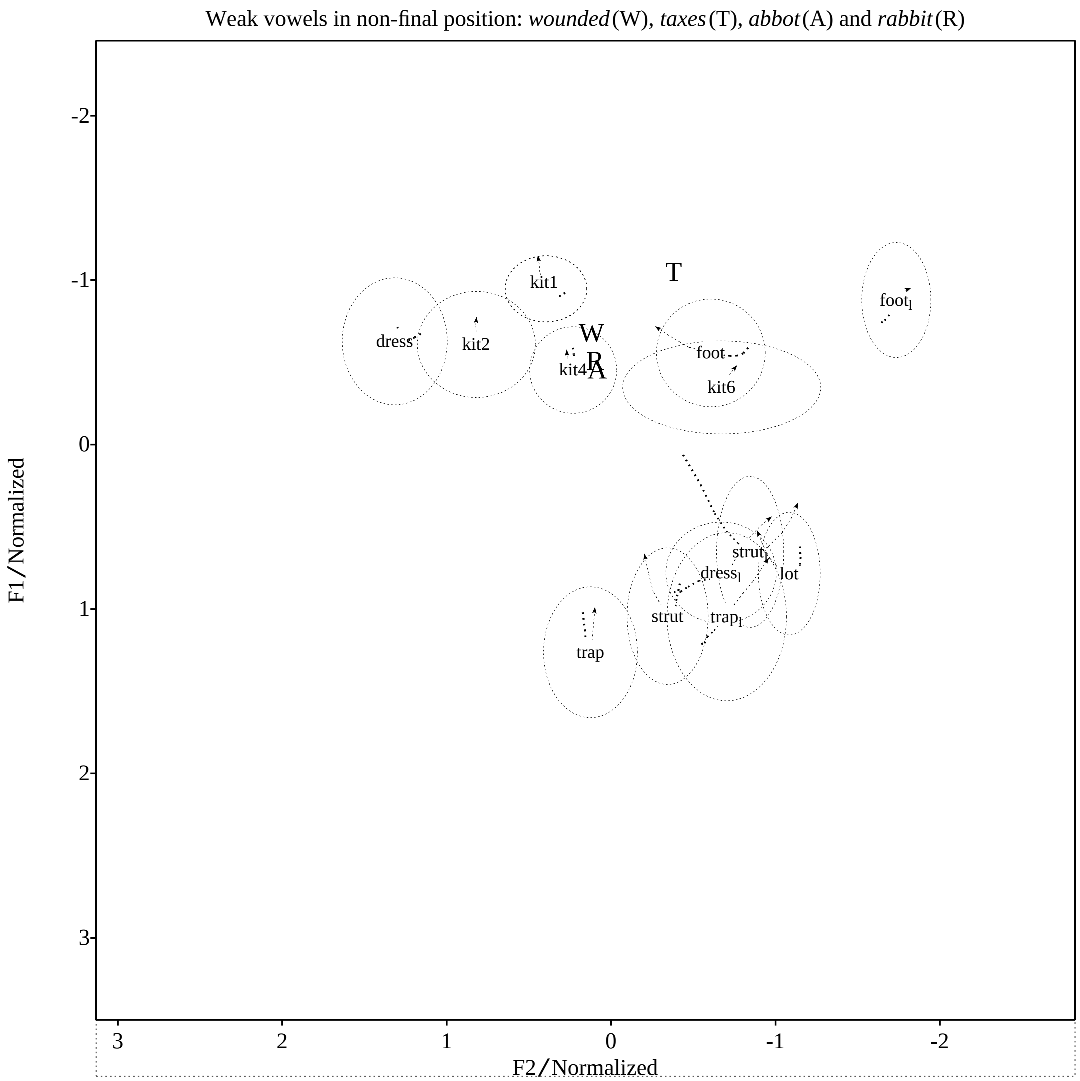


Figure 9.17: Weak vowels in non-final position: *wounded*(W), *taxes*(T), *abbot*(A) and *rabbit*(R)

abbot (A)[26]. We note the slightly removed (and closer) position for *taxes*. We note too that the values for the remaining three tokens (*abbot*, *rabbit* and *wounded*) overlap with the centralized values of KIT. From this representation it is also clear that SAE has undergone the Weak Vowel Merger, with *abbot* and *rabbit* having virtually identical values (traditionally transcribed as [ə] for *abbot* and [ɪ] for *rabbit* for non-merged dialects of English).

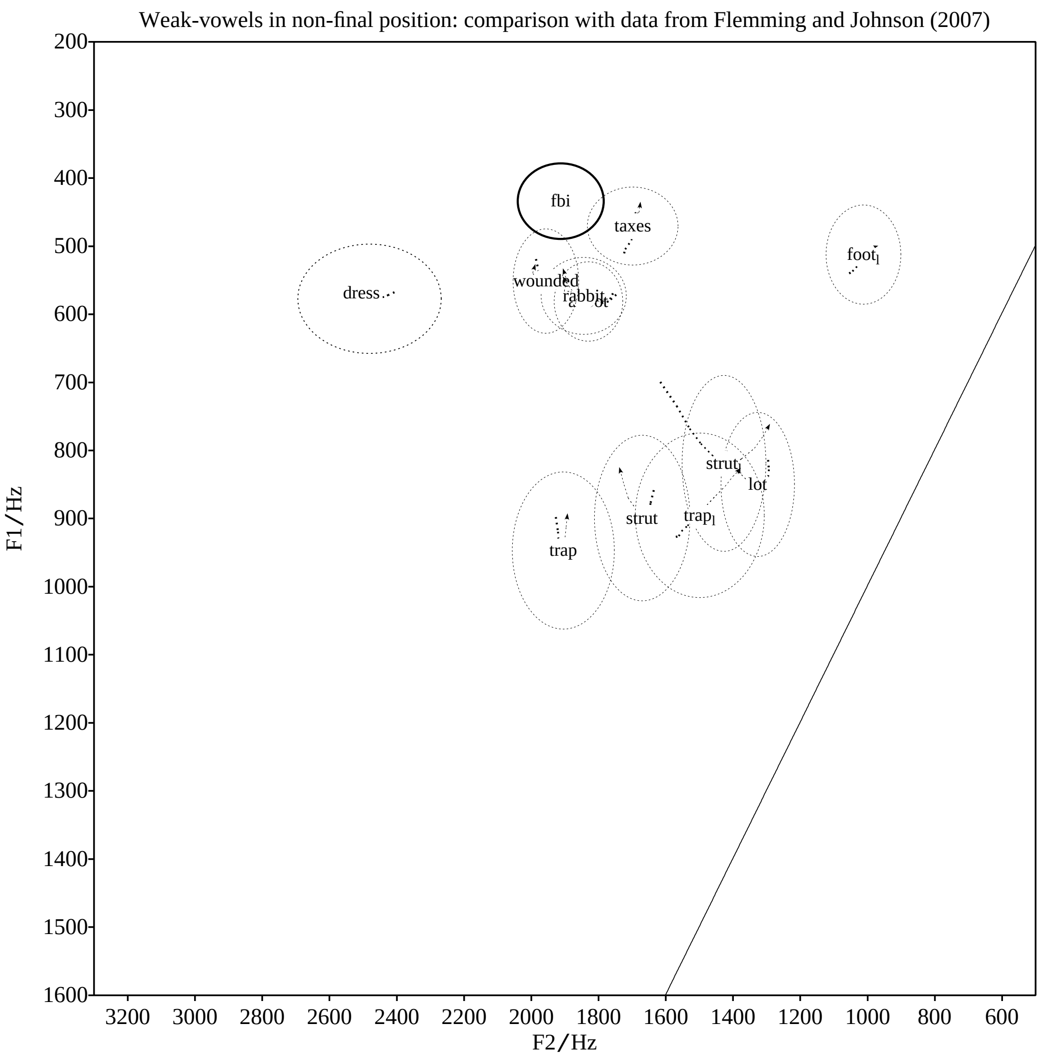


Figure 9.18: Weak vowels in non-final position: Comparison with data from Flemming and Johnson (2007)

[26]Neither the standard deviations nor the formant tracks are shown in Figure 9.17 simply because their inclusion makes the graph difficult to read. Figure 9.18, however, displays this information and we note the relatively low standard deviations for each token.

Figure 9.18 on the preceding page provides a comparison between the current data on the weak vowels in non-final position and data derived from Flemming and Johnson (2007). The data extracted from Flemming and Johnson (2007) relates specifically to those weak-vowels found in plural suffixes after words which do *not* have schwa finally i.e. *roses* (singular /roʊz/) as opposed to *sofas* (singular /soʊfə̠/). We note that *taxes*, in particular, has a similar height (although it is somewhat more retracted) and, the singular of *taxes*, does not, of course, end in schwa. Flemming and Johnson (2007) transcribe their value as [ɨ], which tempts one to do the same for *taxes* as well. In fact, the differences between the various tokens appear slight enough to posit a similar distinction as that advanced by these authors i.e. [ə] for final schwa (including weak vowels in plural contexts where the singular form *does* have final schwa e.g. *sofa*[27]) and [ɨ] for all (reduced) weak vowels in non-final context i.e. including not only words such as *taxes*, but also *wounded*, *abbot* and *rabbit*. Overall though (i.e. for both lettER/commA and weak vowels in non-final position) there appears to be a tendency to render the value with a more open quality than in US English: as already hinted, [ɜ] appears to be a more suitable transcription for lettER and commA in SAE, while [ɘ] appears appropriate for most weak vowels in non-final position (and for the more central allophones of KIT).

Figure 9.19 on the following page provides a representation of the happY vowel for the current data in the context of the other Part-System B vowels. As can be seen from this figure, there is substantive overlap between FLEECE and happY and thus evidence for happY-tensing in the data, confirming the impressionistic literature. There is also no evidence of spectral movement in the happY data. The average length of the FLEECE tokens was 0.183 seconds (N=58, sd=0.0618) while the average length for the happY tokens was 0.181 seconds (N=85; sd=0.0481); no obvious difference between these two vowels exists therefore, at least in citation-form. Given the citation-form nature of the acoustic data it is impossible, though, to separate out segmental from suprasegmental effects i.e. whether added duration is an intonational or segmental phenomenon.

9.3.4 WV: Synopsis

The various analyses provided above provide reasonable-enough evidence to conclude that FLEECE and happY have essentially the same vowel quality in SAEP. As intimated in the previous section, the precise role played by added duration still requires further investigation. As pointed out in §3.4.3, Trudgill (2004:137), for one, claims that happy-Tensing was

[27] Although this particular detail would need to be subject to further testing since there were no equivalent tokens included in the current study.

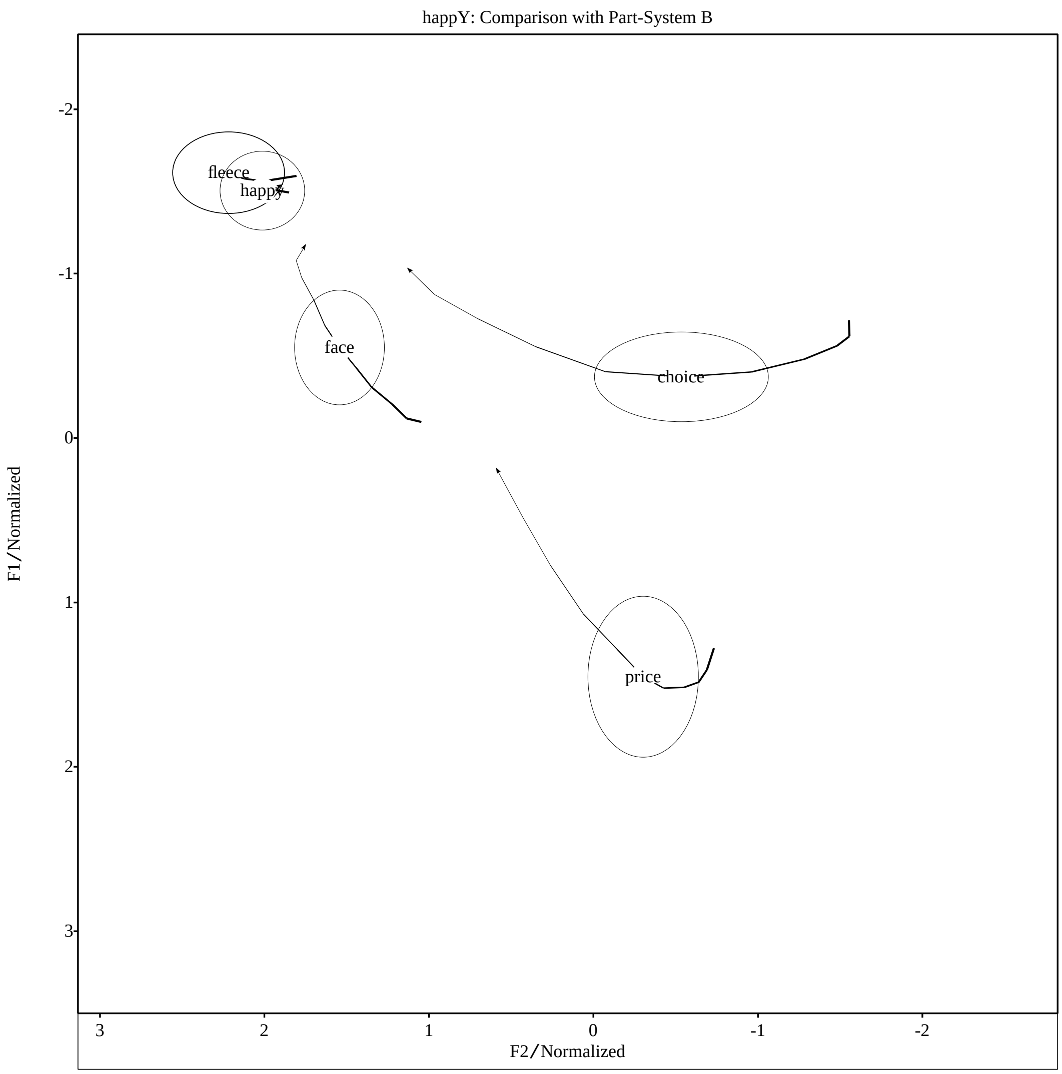


Figure 9.19: happY: Comparison with Part-System B

"very rare in mid–nineteenth century Britain" and, from this perspective, therefore, an unlikely input into the SAE koinéization process, involving instead 'drift' and an endogenous change in line with similar developments in Cockney and RP. Still the fact that front, high (and potentially) lengthened variants were observed by the early Hopwood (1928) raises suspicions in this regard. A reading of both Jones (2006:294–302) as well as Fabricius (2002) seems to, in fact, suggest that happY-Tensing was far more common than suggested by Trudgill (2004) and that this issue requires, as as whole, further investigation. The influence of Afrikaans input into the development of happY-Tensing cannot be discounted either.

The acoustic data also provides general overall support for the notion that the Weak Vowel Merger has taken place in SAEP and that, in general, there is a comparatively open quality to the short weak vowels (i.e. excluding happY). In particular, there is some suggestion that the commA and lettER vowels might be best described as [ɜ] (i.e. low schwa), at least with respect to the more formal end of the sociolinguistic and stylistic scales. Furthermore, it would appear that while the Weak Vowel Merger has indeed taken place, that weak vowels in non-final position have not settled on a [ə]-like quality, but something closer i.e. [ɘ] or [ɨ̞][28]. The distinct quality of commA and lettER vs. weak vowels in non-final position, as well as the overlap of the latter with the more centralized allophones of SAEP KIT, provides compelling evidence for the latter having a [ɘ] rather than a [ə] like quality as well.

Still, there are two possible arguments against the [ɘ]-rather-than-[ə] analysis. Firstly, one might argue that such close qualities are indicative of hypercorrection as would seem plausible from a reading of, for example, Lanham (1978). While this is, indeed, a possibility, the relatively low standard deviation of the weak vowels in non-final position seems to point in the opposite direction. Hypercorrection by its very nature implies variability as opposed to categoriality. The highly 'focussed' articulation that the acoustic data reflects thus implies the latter and not the former. A second possible, and more convincing, argument against a [ɘ]-rather-than-[ə] analysis, is that weak-vowels in non-final position are, by definition, very often found in contexts in which the vowel is both preceded and followed by another segment. This fact, as well as the observation that a common reflex of their unstressed nature would be limited duration, means that, unlike in the word-final context characteristic of commA and lettER, a putative [ə]-like target would in many cases be undershot. The issue, therefore, dissolves ultimately into the question whether the [ɘ]-

[28]This would still, however, appear to be more open that the quality in American English, as reported on in Flemming and Johnson (2007). As intimated above, *taxes* would appear to have a [ɨ̞] value, while the somewhat more open [ɘ] seems to be a more faithful transcription for *wounded*, *rabbit* and *abbot*.

like acoustic quality is simply a coarticulatory phenomenon that listeners compensate for or whether this quality has been phonologized; in other words, do listeners *perceive* the relevant vowels in non-final position to be [ə] or [ɘ] or, from an articulatory perspective, is the target [ə] or [ɘ]? Again, the low standard deviation of this vowel seems to indicate that [ɘ] has become phonologised as the vowel target, given that in the case of formant undershoot one would expect a higher degree of variability given the inevitable differences between speakers in terms of tempo. As a whole, however, the issue requires further research.

As far as the historical development of SAE commA, lettER and the Weak Vowel Merger is concerned, the first two provide little 'trouble', as a brief reading of §9.3.1 should make clear. Weak Vowel Merger, on the other hand, has been shown by Trudgill (2004:117–120) to have been conspicuously absent from the southeast of England during the 19th-century. It was, however, present in the accents of East Anglia, Ireland, the southwest of England and the far north, all of which, in varying degrees and at varying times, could have 'provided' for the appearance of this feature in SAE. More importantly, perhaps, Trudgill (2004:119) argues that the appearance of this 'minority' form in NZE was the result of the unmarked nature of [ə], as opposed to [ɪ], an argument that *might* have relevance to the SAE context as well.

9.4 The Nature and Effect of Final-/l/

While the focus of this study is on vowel quality, some attention needs to be paid to the phonetic characteristics of /l/ given that it is the quality of this sonorant that is often taken as the basis for its effect, when in post-vocalic position, on preceding vowel quality. Section 9.4.1 begins, therefore, with a brief synopsis of the quality and behaviour of /l/ in various accents of English, including, of course, SAEP. Section 9.4.2 then focusses on the impressionistic literature in terms of the reported effects of so called 'dark'-/l/ on preceding vowel quality, while §9.4.3 is focussed on the acoustic data.

9.4.1 'Clear' vs. 'Dark' /l/

In some varieties of English, in England mainly in the south and in RP, /l/ has two main allophones: a clear, non-velarised variety in syllable-initial position (as in *leaf*) and a dark, velarised variety in coda position (as in *feel* or *milk*). It should be stressed that this distinction between 'clear'- and 'dark'-/l/ in fact subsumes "a continuum of productions" (Palethorpe and Cox 2003). In the north-of-England, on the other hand, there is a tendency for /l/ to be neutral in all environments (Wells 1982:370–1).

With respect to so-called 'dark'-/l/, there appear to be two main, not necessarily mutually exclusive, developments across various accents of English :

- The emergence of L-vocalization, whereby the /l/ becomes vocalized 'into' a [ʊ]-like quality e.g. *silk* is pronounced as [sɪʊk]. This is a characteristic feature of Cockney English, but is also prevalent in at least one Southern Hemisphere variety i.e. NZE (Gordon et al. 2004).
- The retraction (and occasional lowering) of vowels preceding dark-/l/ – this issue is dealt with further in §9.4.2.

It is important to stress, however, that this allophonic distribution appears to be a relatively recent one in English and Trudgill (2004:79) provides evidence to suggest that the pattern was unlikely to be the only input into the development of the Southern Hemisphere Englishes. This variation is implicit in the following quote:

> "In Tristanian English ... a clear /l/ is usual after front vowels. In the modern Southern Hemisphere varieties, the pronunciation of /l/ in most or all environments in Australasian English tends to be rather dark, possibly pharyngealised, and the distribution of 'clear' and 'dark' allophonic variants is certainly not as prominent as in many English accents ... L Vocalisation is now under way in prepausal and preconsonantal position in New Zealand, but it is obviously a twentieth-century innovation ... Falkland Islands English has the English English type of allophony with dark /l/ being rather markedly velarised" (Trudgill 2004:79).

It is thus interesting to note that while more recent authors such as Bowerman (2004:940) have taken the classic clear-dark allophony in SAE for granted, the relatively early Lanham (1967:65;98) concludes that /l/ in SAE is clear in all positions, the retraction of the various vowels being explained by the 'migration' of the /ʊ/-like quality from the /l/ to the nucleus. Wells (1982:617) concurs with this assessment by asserting that "South African /l/ is not really dark in any environment". According to this author, /l/ is always clear or neutral. Trudgill (2004:79) relies on Wells' (1982) description of SAE /l/ as part of his broader analysis of /l/ in the Southern Hemisphere Englishes.

It is in Lass (1990:284;footnote 8) that we have the first dissenting voice. The author claims the following:

> "According to Wells (1982: §8.3.6), SAE /l/ 'is not really dark in any environment'; but this is not true of the variety discussed here, nor any type that I am familiar with.

> What are sometimes perhaps taken as 'clear' /l/ are simply ones with coarticulation of a less than fully back quality ... In Cape Town standard, /l/ is typically very clear (palatalised) syllable-initially and in interludes (*silly*, *follow*), and with some backer co-articulation in codas: normally palato-velarised or velarised ([ɨ] to [ɤ]-coloured) after high vowels and uvularised ([ɔ] or [ʌ]-coloured) after lower ones".

Lass (1995:102) adds some further comments on this matter: for BrSAE and GenSAE speakers (and most CulSAE speakers) initial and intersyllabic /l/ is "neutral or slightly palatalized", with the latter being more common "among younger female Respectable speakers". Final /l/ is 'dark'; either velarized or uvularised. According to Lass (2004:278–9), "final /l/ may be very dark (often uvularised or pharyngealised, even after front vowels) ... /l/ is clear syllable-initially, rather palatalised in posher varieties, regardless of the following vowel". Bowerman (2004:940) confirms the above adding that "when /l/ occurs at the end of a word, but before another word beginning with a vowel, it tends to be realised as clear in Cultivated WSAfE".

Interestingly enough, Bailey (1984:21) claims L-vocalisation in his idiolect of SAE, particularly in casual speech. He implies, furthermore, that for most South Africans the vocalized segment is unrounded, making it particularly difficult to detect (on an acoustic or auditory level at least). Thus while unrounded variants do exist in Cockney (e.g. [pɪɤ] for *pill*), for example, the most common variant is a rounded one i.e. [pɪʊ] (Wells 1982:313). In SAE, on the other hand, [pəɤ] would be more common than [pəʊ] on Bailey's (1984) analysis. Lass (2004:278–9), however, claims that final /l/ only vocalises in "extremely fast speech, and even there rarely".

A more recent acoustic study, Giudici (n.d.), conducted on the same data as the current one, has provided relatively convincing evidence that the acoustic quality of final 'dark'-/l/ in SAEP is categorically velar i.e. no uvular or pharyngealized variants were found, while prevocalic /l/ is neutral or, more precisely, [ɨ]-coloured. The main conclusion to be drawn is one of two:

1. That Lanham (1967) and, consequently, Wells (1982) and Trudgill (2004), were incorrect in their assessment of SAE /l/; or

2. more tantalizingly, that there has been a change from a 'clear' to a 'dark' value over the last 50 years or so.

9.4.2 The Effects of Final-/l/: The Impressionistic Data

NZE is a variety in which vowel retraction is a characteristic feature leading to a number of neutralizations (Wells 1982, Gordon et al. 2004). Similarly, but with respect to pre-dark-/l/ DRESS, from Labov's (1994) summary of Trudgill's work in Norwich it is evident that DRESS before /l/ has undergone a major change in this area, such that by the 1980s *hell* and *hull* were homonyms.

As stressed by Lanham (1967:67), however, the issue in these and other accents is likely to be one of *degree* of retraction, some degree of retraction being characteristic of many British dialects.

In general, the influence of dark-/l/ seems to be particularly evident with respect to the short front vowels of English, given "the physical incompatibility of the required gestures and the propensity of short vowels to embrace co-articulatory effects" (Palethorpe and Cox 2003). Thus, as confirmed by Schmitt (2007:322), in the context of a discussion on RP, "there are two major allophonic varieties of all front checked vowels in English . . . a raised variety before velars . . . and a lowered one before dark /l/"[29].

According to Palethorpe and Cox (2003), the effect of a following /l/ on a preceding long high front vowel (either monophthongal or with a glide to [i]) is very often one of creating a bisyllabic effect. Perhaps more strikingly, modification of vowel quality before final /l/ often leads to phonemic neutralization; thus, by way of example, Palethorpe and Cox (2003) mention the common neutralization of LOT and GOAT in AusE.

It seems quite clear that a final /l/ conditions the retraction (and sometimes lowering) of many SAE vowels, particularly those of DRESS, KIT and GOAT. Lanham (1967:64-5) summarizes this effect in the following manner:

> The effect is strongest on front and central vowels which have retracted allophones before /l/; in the case of /e/ the allophone is both retracted and lowered. When the back vowel /u/ is affected . . . the allophone is lowered.

It appears that for the /l/ to have the relevant effect it needs to be tautosyllabic. As stressed by Lanham (1967:116; note 3), for example, there is no evidence of vowel retraction in words such as *hamlet* or *silly*[30].

[29] With respect to velars, this might, on the surface, appear to be a contradiction in terms, but one needs to bear in mind that the velars that Schmitt (2007) is referring to are, in fact, phonetically palatal. Dark 'l' is, on the other hand, 'truly' velar.

[30] One exception that I've noticed is *yellow*, which very often has a retracted DRESS vowel in SAEP. In constrast, the first vowel in *telly* is never, as far as I can tell, retracted.

The effect of final-/l/ on the values of various particular vowels has already been dealt with, on a piecemeal basis, in previous chapters. It is, however, worthwhile isolating and summarizing the relevant literature at this point.

- Lanham (1978:149) identifies vowel retraction before /l/ as a Natal variable with its origin in relevant British dialects[31] and as a defining feature of GenSAE. Furthermore, while the feature is "largely insensitive to age, sex and class variables . . . this variable has diffused widely [and] is prominent in speech of upper-class under-30s particularly in Johannesburg" (Lanham 1978:152–3; my parenthesis). There is little correction of this feature in formal styles with the exception of KIT before /l/. Lanham and Macdonald (1979:43) also provide evidence that "the trend is advancing through age grades".

- Lanham and Macdonald (1979:43; my parenthesis) mention a certain degree of regionality with respect to vowel retraction:

  > "Young males and females on the Witwatersrand favour /el/, young males in Natal /əul/, with the latter sample coming closest to having the variants of the three most affected nuclei [DRESS, KIT and GOAT] more or less equally advanced".

  These authors emphasize that vowel retraction appears to be below the level of social consciousness (except for KIT) and that little variation is found when style-shifting takes place.

- Wells (1982:617) pretty much agrees with these assessments, claiming that the tendency towards retraction and lowering is particularly relevant to DRESS, KIT and GOAT "in the environments _lC and _l#", with [æ̈], [ɤ̈] and [ɒː] being given as the relevant values. Only KIT before final /l/ is subject to correction.

It is worthwhile mentioning that Lanham (1967:78) claims that, even for conservative SAE, final, tautosyllabic /l/ has a retracting effect. He explains as follows:

> "Some South African speech teachers reveal their limited command of RP which they try to teach, by allowing the most severely /l/-influenced allophones of vowels to pass in a context approximating to "advanced" RP. It is not uncommon to hear <u>shell</u>, <u>bell</u>, etc. as [šæl], [bæl], in school presentations of poetry and drama. This particular hallmark of SAE passes unnoticed by many dialect-sensitive South Africans".

[31] A curious observation given the standard association of Natal with north-of-England varieties coupled with the relatively neutral values for /l/ given by Wells (1982:370–1) for these varieties.

Lanham (1982:338) provides the same conclusion, mentioning vowel retraction for DRESS, KIT and GOAT before /l/ in CulSAE.

9.4.3 The Effects of Final-/l/: The Acoustic Analysis

With respect to Webb's (1983) acoustic data as represented, in particular, in Figures 5.4 on page 151 and 5.5 on page 152, the retraction of a number of vowels before final /l/ is apparent and, in fact, surpasses the predictions made by the earlier impressionistic literature. Thus, from Figure 5.4 on page 151, it is clear that DRESS before /l/ has a quality close to LOT (as opposed to a somewhat centralized TRAP), while KIT before final /l/ is even more further retracted than FOOT. While the relative positions of these allophones might have as much to do with an unrounded LOT and a fronted FOOT as with retraction before /l/, it is clear that, overall, the degree of retraction is greater than has generally been reported, a fact emphasized by the data presented graphically in Figure 5.5 on page 152, where we have evidence for vowel retraction before /l/ in the case of GOOSE, NURSE, CHOICE and GOAT.

With respect to the current data, the reader is referred back to the following figures in order to get an overall sense of the effect of final tautosyllabic /l/ on preceding vowel quality:

- Figure 6.13 on page 197, which provides the short vowels of SAEP along with their pre-/l/ allophones [32].
- Figure 6.2 on page 173, which shows the Part-System B vowels and their respective pre-/l/ allophones i.e. FLEECE, FACE, PRICE and CHOICE.
- Figure 6.1 on page 172, which shows the Part-System C vowels and their respective pre-/l/ allophones i.e. GOOSE, GOAT and MOUTH.
- Figure 8.2 on page 296, which shows the Part-System D vowels. The relevant vowels are NURSE, THOUGHT and BATH.

With respect to Figure 6.13 on page 197, and the short vowels of SAE, we note, firstly, the confluence of values for DRESS, TRAP and STRUT before final /l/. It would appear that these vowels are subject to neutralization to a position between STRUT and LOT in this particular context. Apart from this fact of neutralization, which is of interest in itself, the degree of retraction of these allophones, particularly in the case of DRESS, appears

[32] Note that in the case of KIT, the pre-/l/ allophone is labeled 'kit6'. For the other lexical sets, the relevant allophone has a subscript 'l' following it. There were no pre-/l/ LOT tokens in the study.

to have been underestimated in the impressionistic literature. Thus, by way of example, Lass (1995:98) gives [æ] as the most retracted value for DRESS before final /l/. While the impressionistic data certainly confirms a value as open as TRAP, the retracted (i.e. in terms of backness proper) value of this vowel has not been given its full due.

With respect to KIT before final /l/, Figure 6.13 on page 197 needs to be read in conjunction with Figure 7.3 on page 270, where *with* is separated out from KIT before /l/. It is interesting to note in this regard that FOOT and KIT before /l/ have close to the same value. This is in contrast to some reports in the literature which have tended to claim near-homophony between the two pre-/l/ allophones of these two lexical sets. Note that FOOT has a decided spectral glide toward the front, while KIT before final /l/ keeps gliding towards the back area of the acoustic chart. FOOT before /l/ is, on the other hand, even more fully retracted thus preventing possible confusion between, for example, *will* and *wool* – see Figure 6.13 on page 197 again. This difference is most likely due to slightly more lip-rounding in the case of FOOT before /l/ i.e. explaining the lowered F_2 value for this vowel *vis-à-vis* KIT before /l/.

Figure 9.20 on the next page provides an illustration of the relationship between GOAT and LOT in the current data, compared with data from Webb (1983). This figure is provided mainly because GOAT before /l/ has been mentioned in the impressionistic literature as having a similar value to LOT. In Figure 9.20 it is clear that while GOAT before /l/ is subject to substantial retraction there appears to be little evidence for its overlapping with LOT. The same applies to the data provided in Webb (1983), as illustrated in Figure 9.20 as well. Unfortunately, no data for LOT before final /l/ was collected as part of the current research, but it is clear from Figure 9.20 that, at least in terms of Webb's (1983) data, GOAT before final /l/ is far closer to LOT before final /l/ than unmarked LOT. Note, however, that Webb's (1983) data for this allophone of LOT is based on only one token i.e. *involved*. It is interesting to note, however, that there is acoustic data showing that these two vowels in pre-/l/ context have virtually the same quality in AusE, although they do differ in terms of length (Palethorpe and Cox 2003). This issue requires further investigation.

In the case of the various short vowels before /l/ there is a case to be made for the phonologization of these various allophones i.e. the target has not remained the same as that of the unmarked allophone(s), but rather a new target has been adopted. The allophony has thus becomes fully grammaticalized and cannot be explained purely in coarticulatory terms. This appears to be the case particularly for DRESS, KIT and FOOT (i.e. the three high vowels) where there is ample evidence to suggest that the speakers are not targeting the canonical values. This can be fruitfully compared with the data for Part-System B vowels before final /l/ (given in Figure 6.2 on page 173). It is clearer in this case that the original

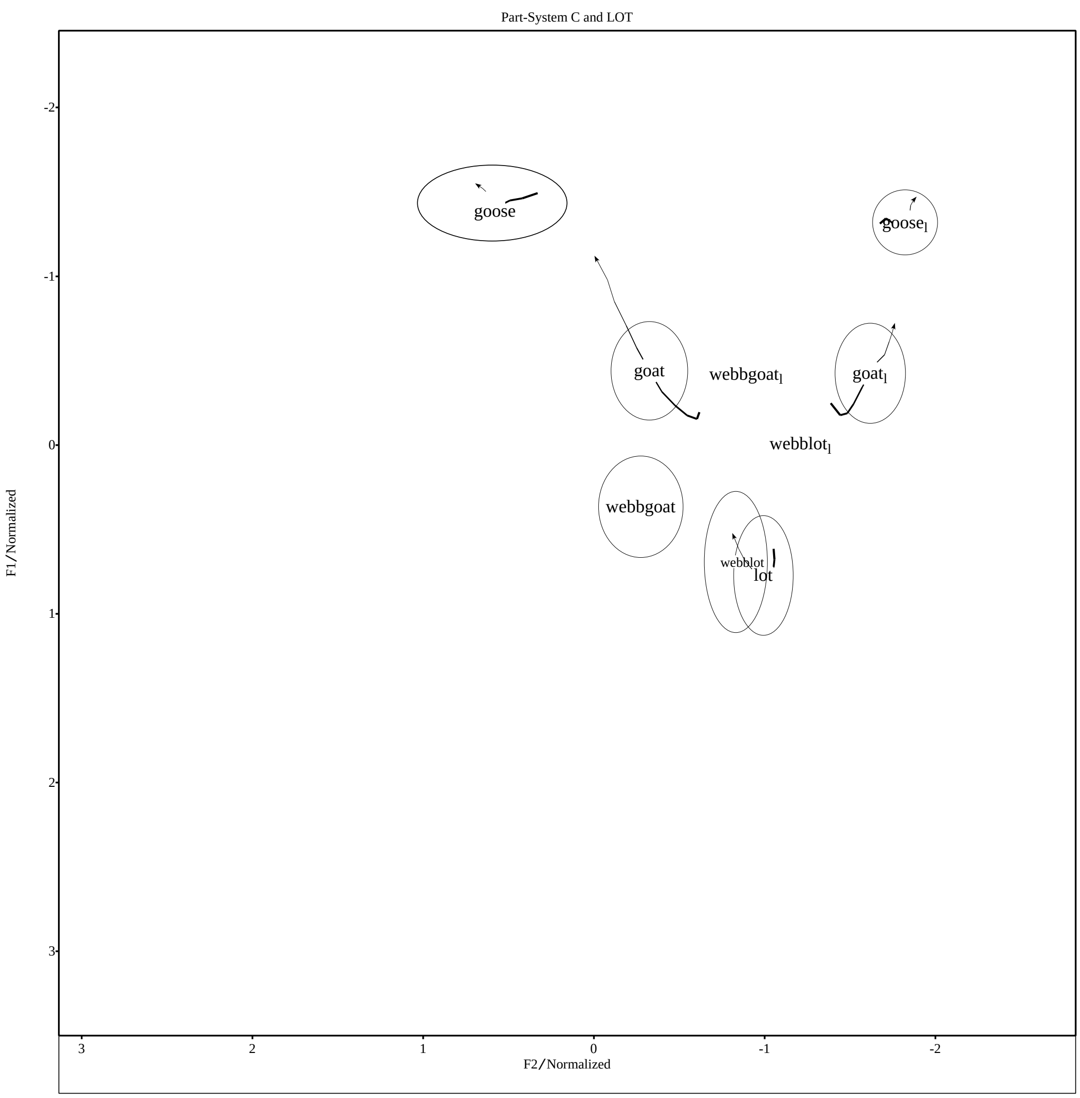


Figure 9.20: GOAT and LOT before /l/: comparison with data from Webb (1983)

targets for the various before–/l/ allophones are the canonical ones and that the difference in vowel quality is due solely to the coarticulatory force of the following dark /l/. Thus in the case of FLEECE before /l/ the beginning of the glide comes close to the unmarked FLEECE value, while in the case of PRICE the F_1 maximum (i.e. point of inflection) of both the marked and unmarked allophones are virtually identical. While FACE before final /l/ is somewhat more dissociated from its canonical value than the other two vowels, it is equally clear that there is at least a gesture in the 'right direction'.

From this perspective, the Part-System C vowels are a mixed-bunch – see Figure 6.1 on page 172. While final /l/ has virtually no effect on MOUTH, the dissociation (in terms of target) becomes clearly larger the closer the vowel. Thus the two GOAT allophones are clearly separated in acoustic space (as predicted by the literature) and the effect is even more dramatic for GOOSE. Part-System D provides only one other example where one might wish to argue for an alternative target (and thus the phonologization) of the pre-/l/ allophone i.e. NURSE.

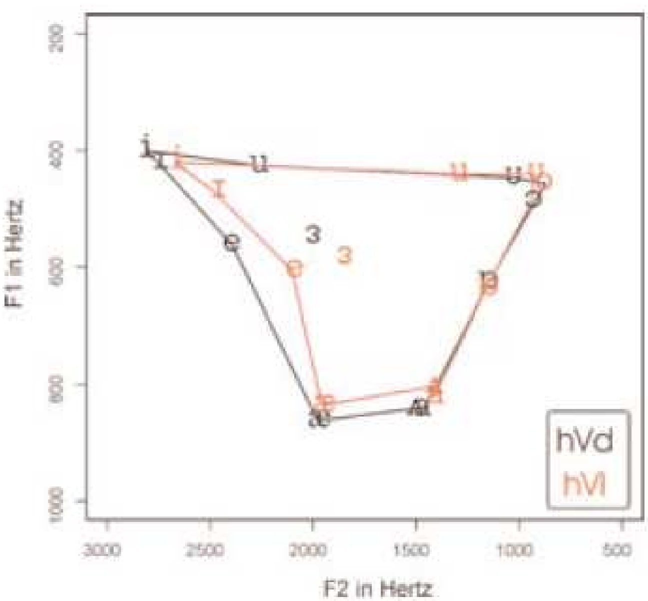


Figure 9.21: Australian Monophthongs before /l/: from Palethorpe and Cox (2003)

Figure 9.21 provides an illustration of the effect of final /l/ on AusE monophthongs – see §4.6.2 for more on this study. As can be seen from this figure, the most substantial effect on pre-/l/ vowel quality for monophthongal AusE vowels is on GOOSE. From Figure 6.1 on page 172 it is clear that a similar dramatic effect is evident in SAE. There is also much similarity between the AusE and SAE data with respect to FLEECE, NURSE, BATH and

THOUGHT. With respect to the short-vowels, however, it is clear from a comparison of Figures 9.21 on the facing page and 6.13 on page 197 that SAE 'dark'-/l/ has a far greater effect on the preceding vowel than does its AusE counterpart. In the case of FOOT, however, this might have as much to do with the fronted nature of the main allophone of this vowel as with vowel retraction *per se*.

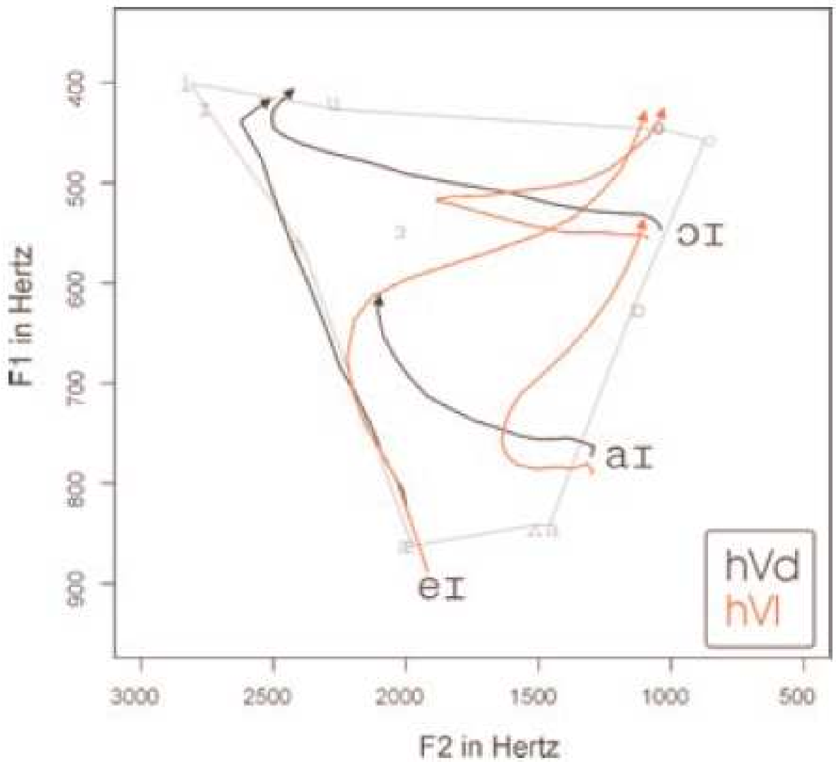


Figure 9.22: Australian Front-Gliding Diphthongs before /l/: from Palethorpe and Cox (2003)

Figure 9.22 provides a graphic illustration from the same study on the effect of final-/l/ on the front upgliding diphthongs of AusE. It appears that what has been tracked in this diagram is the vowel *plus* final /l/. A comparison between this figure and Figure 6.2 on page 173 makes it reasonably clear that the effect of final /l/ on the relevant vowels across the two accents is relatively similar – the main differences being attributable to vowel quality itself. As in SAE, it appears that the effect of final /l/ on Part-System B vowels in AusE is purely co-articulatory i.e. unlike in the case of GOOSE, the targets for the unmarked and marked allophones appear to be the same.

9.4.4 The Nature and Effect of Final-/l/: Synopsis

SAEP would appear to have a velarized 'dark'-/l/ in syllable-final position. The secondary articulation that leads to this velarization has an impact on the quality of most preceding vowels. In some cases this impact is purely coarticulatory, while in other cases it would appear that the retracted quality has provided a new 'target' for the phonology. The latter seems to be the most obvious analysis particularly with respect to the DRESS and GOOSE vowels. With respect to the former there has also been a tendency in the impressionistic literature to underestimate the degree of retraction of this allophone. Furthermore, SAE 'dark'-/l/ appears to have a comparatively radical effect on preceding DRESS, KIT and FOOT, at least in comparison with AusE.

As far as the etiology of dark, final-/l/ as well as preceding vowel retraction in SAE is concerned, one has to keep in mind that the impressionistic literature seems to be divided on at least the first point. If one assumes, however, a final, velarized, dark-/l/ for SAE[33] then it seems clear that this quality, along with the associated clear-dark allophony, was inherited from the southeast of England.

Vowel retraction before final /l/ is a common and only natural feature of any variety of English (or language for that matter) that contains a 'dark' version of said segment. What has to be accounted for, historically, assuming, as has been done above, that SAE *has* a dark-/l/ finally, is the *radical* retraction of vowel quality, as illustrated in the impressionistic literature on SAE, but even more dramatically in the acoustic data provided above. One possibility relates to an endogenous development, linked in its turn to the notion of phonologization discussed in §1.4. This has already been alluded to above, and would have, diachronically, involved the misanalysis, by children, of 'natural' coarticulatory effects as elements of the grammar (broadly conceived).

Another intriguing possibility relates to the on-the-surface anomalous observations by Lanham (1978) and Lanham and Macdonald (1979), reviewed in the previous section, regarding the supposed Natal-based origin of such radical pre-/l/ retraction of vowel quality in SAE. As already pointed out in a related footnote, these observations are anomalous given the attribution of a neutral quality, by Wells (1982), to /l/ in most north-of-England accents, as well as given the preponderance of exactly these accents in the original input into NE. One resolution of this seeming paradox lies in Trudgill's (2004) notion of 'hyperadaptation', first mentioned in §3.3.1, and explained as a process by "which speakers attempt to use forms from higher status accents, but employ an incorrect analysis and extend changes to items where they are inappropriate" (Trudgill 2004:87). This is, unlike phonologiza-

[33] As mentioned above, Giudici (n.d.) provides acoustic evidence in this regard

tion, a process mainly involving adult-speakers, but its adoption by normatively-minded NE speakers (i.e. as part of a process of hypercorrecting to RP-values) *could* have had a similar (although no doubt variable) outcome to that of phonologization; one which further generations could have rendered categorical. On the other hand, Lanham (1978) and Lanham and Macdonald (1979) might have been incorrect.

9.5 Conclusion

The results of the analyses contained in this chapter seem to indicate that, with respect to the NURSE vowel, SAEP has a fronted, raised and rounded variant; basically [øː] or somewhat more centralized. This value appears to be a categorical value in SAE and thus constitutes an indicator of this variety. NZE and AusE appear to have similar values, although the former's NURSE vowel appears to be somewhat more raised, contrary to expectations. Of more interest is that there is enough evidence to *suggest*, *contra* Trudgill (2004), that the value for this vowel has its roots not in an endogenous development, but in the original values brought in via Natal English from the north of England. In contrast there appears to be evidence for the post-settlement raising and further tensing of SQUARE in SAEP, in accordance with Labov's (1994) Principle I. If the change is endogenous, it appears to have almost reached completion, the value appearing to be categorically [eː], with duration being the sole basis for the distinction between SQUARE and DRESS. While structural pressures might have, indeed, been a primary force in the raising of this vowel, it would appear that its glide-weakening was mainly a result of a preponderance of monophthongal variants in the original feature pools. The analysis of NEAR reveals nothing surprising; the basic value is [ɪə], or with a somewhat closer onset, with little evidence of monophthongisation. There is a high degree of variability however, both of a phonetic and lexical nature. Historically, there is evidence for input from multiple British-based sources.

Turning to the weak-vowels, the analysis shows happY-Tensing in SAEP, although the exact role played by added duration awaits further investigation. The existing evidence does not provide clear support either for or against the notion that happY-Tensing was an endogenous development within SAEP. With respect to commA and lettER, there is some evidence to suggest a low-schwa value for this vowel i.e. [ɜ]. Of more interest is the fact that weak-vowels in non-final position appear to have a closer value than expected. While hypercorrection can not be ruled out, on the surface it would appear that an [ɘ]-like *acoustic* value needs to be posited for this vowel. Whether this value has been phonologized or is simply a mechanical reflex of a decrease in vowel-duration remains to be seen. Overall, there is convincing evidence for the Weak Vowel Merger in SAEP. Historically, while the input into SAE had, in all likelihood, a minority of variants showing Weak-Vowel Merger, the unmarked nature of [ə] might, indeed, have led to its adoption into SAE.

Recent research indicates that SAE displays a classic 'southern' clear-dark /l/-allophony. The investigation of vowel quality before final-/l/ in SAEP reveals that the overall degree of retraction has been underestimated in the impressionistic literature. In support of this general observation it would appear that, in comparison with at least one other Southern Hemisphere English (i.e. AusE), SAEP appears to license a fairly radical degree of said retraction. The development of this retraction can be linked to one of two processes (or perhaps even both): firstly, vowel retraction could be related to an endogenous process of phonologization or, alternatively, it could be the outcome of the hyperadaptation of normative-minded speakers of accents without final dark-/l/, particularly those from the north-of-England. The last analysis would account for reports of a Natal-bias in this regard.

This chapter concludes the critical and comparative analysis of vowel quality in SAEP. The next chapter is focussed on providing a summary of the results contained in previous chapters as well as recommendations for future research.

CHAPTER 10

Summary of Findings and Recommendations

10.1 Introduction

This thesis has attempted to generate new insights into the development and current nature of the SAE vowel system. On a synchronic level, it provides the first *comprehensive* application[1] of modern acoustic techniques to the study of this system. On a diachronic level, it views the development of SAE, and GenSAE in particular, through the 'lens' of a three-stage koinéization model. By embedding these two approaches within a broader comparative analysis, which includes the use of additional acoustic data from SAE (Webb 1983) and other relevant accents of English as well as a thorough and critical review of the SAE impressionistic literature, it generates a number of new findings and new perspectives. This chapter, therefore provides, in §10.2, a final, integrated account of the findings generated as a result of this analysis, and as contained in previous chapters. Furthermore, and as contained in §10.3, a set of recommendations for future research is also included. The chapter ends, in §10.4, with a final, general conclusion.

[1] Webb's (1983) acoustic analysis of this own idiolect, while useful in its own way, can hardly be considered comprehensive in any substantive sense.

10.2 Summary of Findings

This section provides a summary of the findings of more specific and more general interest respectively. Given the level of detail provided in the comparative analysis contained in this thesis, this section is of particular importance in terms of its role in distilling the essence of said analysis.

As such, §10.2.1 to §10.2.4 are focussed on providing a vowel-by-vowel summary of the relevant findings, as well as on highlighting the few consonantal phenomena of interest which were discovered during the analysis of the data. The foci of the various subsections are as follows:

- Section 10.2.1 is focussed on the vowels analysed in Chapters 6 and 7 i.e. those vowels linked directly to Labov's (1994) Pattern 1 extension and Pattern 4 component of the Southern Shift. The relevant vowels are MOUTH, PRICE, CHOICE, TRAP, FLEECE, DRESS, FACE and KIT.

- Section 10.2.2 focusses on Chapter 8 i.e. those vowels involved in Labov's (1994) Pattern 3 component of the Southern Shift, as well as those short vowels not dealt with in Chapters 6 and 7. The relevant vowels are BATH, THOUGHT, GOOSE, GOAT, CURE, STRUT, LOT and FOOT.

- Section 10.2.3 relates to Chapter 9 i.e. the remaining vowels as well as the nature and effect of final, tautosyllabic /l/ on preceding vowel quality. The relevant vowels are NURSE, NEAR and SQUARE as well as the various weak-vowels: happY, commA, lettER and weak-vowels in non-final position.

- Section 10.2.4 deals with the few consonantal features of interest that were discovered as a side-effect of the comparative analysis of vowel quality. In all but one case, these have already been reported on in the thesis, albeit briefly. Although not central to this thesis, it was felt that their inclusion might constitute a rough guide to possible future research in this area.

In all cases, attention has been paid to both synchronic and diachronic findings of interest; and both on a descriptive and theoretical level. Given that the focus here is on 'distillation', the style is somewhat telegraphic and some detail has inevitably been lost. As a consequence, some of the tentativeness inherent in a number of the arguments contained in previous chapters has also been lost in the service of brevity. The reader is referred back to the relevant sections in order to garner a more complete reading. The most relevant section number, as well as the page on which each section begins, has been provided in each case in order to assist the reader in this respect.

The last subsection, §10.2.5, is, on the other hand, focussed on providing the more general findings of the thesis; in particular the extent to which the SAE data can be explained by the various frameworks highlighted during the course of this thesis. The relevant frameworks are those provided in Labov (1994), Torgersen and Kerswill (2004), Trudgill (2004) and Schneider (2003; 2007).

10.2.1 Summary of Findings in Chapters 6 and 7

This section provides a vowel-by-vowel summary of the findings presented in Chapters 6 and 7 i.e. MOUTH, PRICE, CHOICE, TRAP, FACE, FLEECE, DRESS and KIT.

MOUTH; § 6.2.4 on page 186: A peripheral, tense vowel but no sign of PRICE-MOUTH Crossover in GenSAE; more retracted than PRICE, even (arguably) in the connected speech of Webb (1983). Evidence of glide-weakening. Comparison with AusE data provides support for Wells' (1982) contention that SAE comparatively unaffected by Diphthong Shift. An indicator of GenSAE. According to Lanham and Macdonald (1979), BrSAE fronted MOUTH appears to be in retreat: evidence for an arrestment of Labov's (1994) Pattern 1 extension and Diphthong Shift and subsequent diffusion of the backed variety. Historically, backed MOUTH appears to have originally constituted a hypercorrection i.e. a response to the fronted CE variant. This, in turn, provides evidence against Trudgill's (2004) model of new-dialect formation, particularly since it was unattested in the input to koinéization.

PRICE; § 6.2.4 on page 186: Fronted value in comparison to other 'Southern' accents, except the US South; as with MOUTH, support for Wells' (1982) view on the lack of effect of Diphthong Shift on SAE. Although vowel is glide-weakened, monophthongisation not evident in citation-form data; provides support for Lass' (1990) analysis *contra* Lanham and Macdonald (1979) who viewed fronted-PRICE *monophthongisation* as a prestige variant. Thus support for a recent shift in indexicality of PRICE-monophthongisation. Lanham and Macdonald (1979) claim a retreat from the back BrSAE value; all points to an arrestment of Labov's (1994) Pattern 4 and subsequent diffusion of GenSAE fronted variant. Fronted, monophthongised PRICE possible vestige of north-of-England influence; likely the result, however, of both exogenous and endogenous forces.

CHOICE; § 6.2.4 on page 186: Some evidence for an endogenously-motivated non-peripheral, lowered nucleus; Such lowering would be consonant with a 'reversal' of the PRICE-MOUTH Crossover, as mentioned above.

TRAP; § 6.3.4 on page 204: Relatively strong evidence for (recent) lowering of the TRAP vowel in SAE, perhaps as the initiating move of a SECS-like (Torgersen and Kerswill 2004) anti-clockwise rotation of the short-vowel system; evidence for Trudgill's (2004) 'colonial lag' principle i.e. SAE conservative with respect to this shift. Historical sources of traditionally raised TRAP mainly southeast of England, including London and RP.

FACE; § 6.4.4 on page 231: A (probably stable) social marker of GenSAE. Lowering not as endemic as in AusE and in citation-form regains peripheral status. Evidence of Afrikaans-influence coupled with 'drift' for lowered variant. RP and north-of-England influence for 'narrow' variant. Peripheral citation-form variant provides evidence for non-peripheral DRESS in SAE.

DRESS; § 6.4.4 on page 231: Evidence for an overlap between DRESS and the fronted allophones of KIT. Appears to be nearly-invariant across styles etc. i.e. indicator of SAE. Evidence though for the non-peripheral status of this vowel; this makes sense in terms of SECS-like (Torgersen and Kerswill 2004) rotation; contrary to, or at least development from, Labov's (1994) Southern Shift. No lowering yet: underlines SAE's conservative status with respect to SECS; 'colonial lag'. Derived, historically, from southeast England (via CE) as well as from earlier version of RP.

FLEECE; § 6.4.4 on page 231: Evidence points to an invariant monophthongal vowel, both synchronically and historically; *contra* Labov's (1994) Pattern 4 component of the Southern Shift. Afrikaans influence perhaps played role in preventing Diphthong Shift, which *did* take place in AusE and NZE. Some north-of-England inputs also had 'pure' vowel.

KIT; § 7.5 on page 275: Shows unexpected lowering in comparison to the RP data, particularly with respect to fronted allophones. Little evidence of style-shifting etc. across the two SAE data sets; indicator of GenSAE. The fine-grained phonetic differences between the non-final-/l/ allophones provide support for the position outlined in Lass and Wright (1985) i.e. lack of polarization, at least in GenSAE. Acoustic evidence on commA, lettER and weak-vowel in non-final position (see below) provide evidence for the non-[ə] value of the centralized allophones of KIT i.e. a [ɘ]-like value. Modern *Gen*SAE KIT perhaps closer to NZE KIT than often supposed. Historically, high-front [i]-variants brought to Johannesburg via CE (with possible Afrikaans influence) or late-19th-century working-class accents. Development of GenSAE characterised by avoidance of CE/working class [i], coupled (perhaps) with endogenous pressure to avoid merger with FLEECE. Centralized variants (perhaps) via CE (again with possible Afrikaans influence) as well as accents of the north-of-England and Scotland via NE. A ‘shading’ of values also provided by an earlier version of RP. Centralization also motivated by structural pressure to avoid merger with raised DRESS.

10.2.2 Summary of Findings in Chapter 8

This section provides a vowel-by-vowel summary of the findings contained in Chapter 8 i.e. BATH, THOUGHT, GOOSE, GOAT, CURE, STRUT, LOT and FOOT.

BATH-START-PALM; § **8.2.4 on page 311:** Overlap (in terms of quality, but not quantity) of BATH with LOT. Some evidence of in-gliding in the case of START. As expected, backed in comparison with other Southern Hemisphere Englishes. Near-to-categorical in GenSAE; an indicator of this variety. Historically, backed varieties sourced from normatively-minded NE-speakers and speakers from the late-19th-century southeast of England (including RP). Difference between GenSAE and (raised and rounded) BrSAE BATH linked to relatively extreme ‘backing’ by working-class (London etc.) speakers (with possible influence from Afrikaans). Trudgill’s (2004) analysis of BATH-Backing not completely convincing; unable to account for the non-adoption of fronted [aː]-like variants.

THOUGHT-NORTH-FORCE; § **8.2.4 on page 311:** As expected, substantially raised. Some evidence for endogenous, post-settlement (Pattern-3-like) raising of this vowel. Historical sources include southeast of England (CE and early-Johannesburg), including late-19th-century RP, followed by endogenous raising.

GOOSE; § **8.2.4 on page 311:** As expected, evidence of Pattern-3 fronting; as advanced as AusE and NZE, with RP having recently ‘caught up’. Evidence for Trudgill’s (2004) notion of ‘drift’. Historical sources identical to THOUGHT (see above), with subsequent endogenous *fronting*.

GOAT; § **8.3.4 on page 330:** A clear sociolinguistic marker of GenSAE with two variants sensitive to style-shifting etc. With respect to ‘narrow’ variant, evidence for relatively high and *mid* onset with a centralized to fronted terminus. Glide-weakening, but no monophthongization in citation-form. Has a more front terminus than RP and is closer (in citation form) than AusE. Evidence again for SAE’s resistance to Diphthong

Shift. The development of the ‘narrow’ variant appears to be an endogenous (and recent) development. More open variant the result of Diphthong Shift and ‘drift’ (Trudgill 2004), bolstered by southeast input into early-Johannesburg.

CURE; § **8.3.4 on page 330:** The expected [ʊə] values. As a whole, the extent of the Second FORCE Merger appears to have been overreported in some of the literature. Historically, no shortage of inputs.

STRUT; § **8.4.4 on page 357:** Evidence for a fronted value, *contra* Trudgill (2004) – possibly a recent endogenous (and in-progress) development. Fronted value provides evidence for ‘colonial lag’ and SAE’s conservative nature *vis-à-vis* Torgersen and Kerswill’s (2004) SECS shift. Historical input was likely to have been a relatively retracted variant – southeast of England and RP.

LOT-CLOTH; § **8.4.4 on page 357:** Overlap between SAE LOT and RP STRUT as well as SAE LOT and BATH (distinguished by length). Evidence of an unrounded American-like quality leading to (acoustic) fronting (perhaps ‘pulled’ by fronted STRUT). Fits in with picture of SAE being conservative i.t.o SECS-Shift. Evidence, though, for both exogenous and endogenous ‘forces’. Some unrounded variants in early-Johannesburg ‘mix’, coupled with aversion toward extreme (working-class and Afrikaans-influenced) rounding as well as structural forces operating in the same direction.

FOOT; § **8.4.4 on page 357:** Evidence of substantial fronting in line with SECS; more advanced than most other Southern accents. Also shows substantial degree of lowering. Advanced degree of fronting (in an otherwise conservative SECS dialect) provides evidence, however, for FOOT not being part of SECS, but operating under the more general Principle III. FOOT *lowering* explained either by parallel-shift (in ‘sympathy’ with KIT) or by NE-speaker hyperadaptation.

10.2.3 Summary of Findings in Chapter 9

This section provides a vowel-by-vowel summary of the findings contained in Chapter 9, including the nature and effect of final-/l/.

NURSE; § **9.2.4 on page 392:** Evidence for an invariant close, front and rounded value, as per the impressionistic literature. An indicator of SAE. Similar value to AusE and NZE values. The three Southern Hemisphere Englishes appear to have shown ‘drift’ in this regard (Trudgill 2004), with RP ‘trailing’. The SAE process appears to be near-to-complete. More research, however, is needed to ascertain whether rounded variants of NURSE constituted an effective historical (particularly north-of-England) input.

NEAR; § 9.2.4 on page 392: Value as expected, but difference between SAE and RP values somewhat surprising. Review of literature shows fair degree of phonetic and lexical variability. Evidence for a higher-than-expected onset in SAE. Higher-onset, monophthongal, triphthongal and disyllabic variants (as well as the substitution of NURSE for NEAR) can all be linked to likely historical inputs.

SQUARE; § 9.2.4 on page 392: Substantial qualitative, but not quantitative, overlap between monophthongal SQUARE and DRESS; SAE and AusE values similar. Evidence for endogenous (Principle I-based) raising (although evidence of exogenous input too), a process that is near-to-completion. Categorical in non-Cultivated SAE: an indictor of these varieties. In citation-form GenSAE there is little difference between SQUARE and FACE. Monophthongisation of SQUARE is, as opposed to raising, more likely to be rooted in historical input.

happY; § 9.3.4 on page 409: Evidence for happY-Tensing and qualitative overlap of happY and FLEECE. Quantitative aspects require further investigation, as does its etiology, which shows possible endogenous *and* exogenous sources. Possible Afrikaans influence.

commA and lettER; § 9.3.4 on page 409: Evidence for a low-schwa [ɜ]-value for these sets as well as a non-[ə] value for the centralized allophones of KIT (see above). Historical input from various sources.

Weak-vowels in non-final position; § 9.3.4 on page 409: Substantial evidence for the Weak Vowel Merger, but at a non-[ə] value. Proposed phonetic value is [ɘ]. While unlikely the result of hypercorrection this relatively close value might still be the outcome of (articulatorily-based) formant undershoot i.e. [ə] and not [ɘ] is still the 'target'. Requires further investigation. While the Weak Vowel Merger was a comparatively 'minor' historical input, the unmarked status of [ə ∼ ɘ]-like variants could have led to its 'uptake' (Trudgill 2004).

Nature and Effect of /l/; § 9.4.4 on page 422: /l/ in SAE appears to have a 'dark', velarized quality in syllable-final position. This, as in other 'southern' accents, conditions vowel retraction. DRESS, TRAP and STRUT appear to neutralize before final /l/ in SAE. Appears that *degree* of retraction in SAE before final /l/ underestimated in the impressionistic literature. Evidence for the extensive phonologization of vowel quality in this environment, particularly in cases such as GOOSE, GOAT, DRESS, KIT and FOOT. Historically, this radical retraction of vowel quality the result of endogenous, phonologization processes or, perhaps, the result of hyperadaptation, particularly by normative-minded NE-speakers.

10.2.4 Summary of Consonantal Features of Interest

Neo-rhoticity; §9.1: Some evidence in the acoustic data of post-vocalic /r/, particularly in word-final SQUARE and NEAR. Links to reports of neo-rhoticity in Da Silva (2007) and Hartmann and Zerbian (n.d.).

Pre-aspiration; §4.5.2: Aspiration found at the end of vowels prior to a final voiceless oral stop e.g. in *bat*. Suggestion that this might be a prestige feature.

Creak; §4.5.2: Creak used as a general feature on vowels. As with pre-aspiration, there is a suggestion that this might constitute a prestige feature.

L-vocalisation; §9.4.1: Some evidence for L-vocalisation in the acoustic data, but highly variable and not extensive.

Dentalisation: Dentalisation of /s/ reported in Bekker (2007), but appears in a wider range of environments in the acoustic data, particularly /t/; prevalent in the more 'Kugel'-accents.

10.2.5 Findings of General Interest

Apart from the specific findings reported on in §10.2.1 to §10.2.4, which taken together provide, I believe, a complete and comprehensive 'description' of the synchronic and diachronic status of the SAE vowel system, there are a number of more general findings, linked explicitly to the various frameworks into which the broad comparative analysis contained in this thesis has been embedded. These are as follows:

1. Section 3.3.3 provides a critique of the Dynamic Model of Schneider (2003; 2007) in terms of its application, in Schneider (2007:173–88), to the SAE context. This critique argues for an extension of a principle of flexibility and fluidity that already lies at the heart of the model. In essence, it argues for an application of the model that takes into account successive 'waves' of nativization. In the case of SAE, this implies that while 'white' SAE (viewed *vis-à-vis* Afrikaans in particular) has perhaps

reached far into Phase 4, SAE conceived more broadly (i.e. viewed *vis-à-vis* L1-Bantu speakers) has only begun the process of Phase 3 nativization. In short, the proposed 'adjustment' renders the SAE data more tractable, thus allowing it to fit the model better.

2. The sociohistorical background provided in Chapter 3, more specifically in §3.3.1 to §3.3.2, coupled with the various historical 'reconstructions' provided on a vowel-by-vowel basis in later chapters, provides evidence against Trudgill's (2004) notion that new-dialect formation proceeds on a developmental course that is insensitive to indexical factors. In particular, this model, as described briefly in §1.3.3, fails, I believe, to convincingly capture the development of at least two SAE vowels, MOUTH and BATH. The relevant arguments have been provided in §6.2.4 and §8.2.4. This, moreover, provides support for Schneider's (2007) model, which accepts that face-to-face "accommodation is one of the mechanisms of expressing one's identity choices" (Schneider 2008:264).

3. The comparative analysis provided in the various chapters, assisted particularly, in this case, by the acoustic analysis, has provided relatively convincing evidence that SAE is (beginning to) participate in a shift similar, if not identical, to Torgersen and Kerswill's (2004) SECS shift, as reported on for the southeast of England. In particular, the acoustic evidence collected as part of the current project shows evidence of TRAP-lowering, the initiating movement of the SECS-shift. Evidence of STRUT- and LOT-fronting also supports the notion that SAE is a highly conservative variety in terms of its participation in this shift, providing, in turn, support for Trudgill's (2004) notion of 'colonial lag'. This conservative status is, however, in contrast with the advanced level of FOOT-fronting shown in SAE, a contradiction which supports the notion that Torgersen and Kerswill's (2004) SECS-Shift should not be seen as inclusive of this particular development.

4. In accordance with Wells' (1982) observation that SAE has remained *comparatively* immune to Diphthong Shift, the comparative analysis contained in this thesis provides evidence to support the notion that SAE has participated in an arrestment of the Diphthong Shift and, by implication, an arrestment of a number of Labov's (1994) patterns, particularly his Pattern 1 extension and the Pattern 4 component of his Southern Shift. While Pattern-3 developments (except in the case of BATH) have been allowed to develop in a relatively unobstructed fashion, the spread of GenSAE at the expense of BrSAE, has seen the resultant diffusion of variants of MOUTH and PRICE that are,

in a sense, reversals of these shifts. A backed MOUTH is spreading at the expense of a fronted, Diphthong-Shifted variant, and a backed, Diphthong-Shifted PRICE is, by all accounts, receding. The presence of 'narrow' unshifted FACE and GOAT variants in citation-form GenSAE, in contrast with, for example, citation-form AusE, which retains rather open variants, speaks more generally to the conservative nature of SAE in this regard.

With regard to those short vowels which have been seen to participate in Labov's (1994) Southern Shift (KIT, DRESS and TRAP), the presence of a SECS-like shift in any 'Southern' accent of English creates problems for Labov's (1994) account, as Torgersen and Kerswill (2004) have noted. In particular, if the SECS-Shift is to be viewed as a development from the Southern Shift, some principled mechanism is required to account for the reallocation of the short vowels onto the non-peripheral track in order to account for their subsequent lowering. Perhaps more generally, the SECS-Shift can be viewed as an effective reversal of at least one component of Labov's (1994) Southern Shift i.e. the assignment of the short vowels onto the peripheral track; a 'movement' intimately connected with their tensed and raised status. In the SECS-Shift the vowels, of course, become non-peripheral, lax and prone towards lowering. SAE's imminent participation in the SECS-Shift provides, therefore, potential added 'ammunition' to this broader debate.

Related to this is the fact that the *anti*-Diphthong-Shift developments in SAE, which in many cases (e.g. PRICE) appear to be linked to a north-of-England influence on the koinéization process that led to SAE's genesis, are, in general, linked to prestige norms i.e. changes-from-above. The same applies to the SECS-Shift, or, more specifically, to the lowering of TRAP and the fronting of STRUT and LOT in SAE. We have seen, in §2.3.2.1, an example of a reverse chain-shift (i.e. the Reverse Parisian Chain-Shift) which, as Labov (1994:139–40) stresses, is linked to specific social factors which, in turn, lead to a situation in which "conservative factors in the speech community . . . operate in a systematic manner to reverse chain shifts as a whole". While the arrestment of the Diphthong-Shift and the *diffusion* of GenSAE variants at the expense of BrSAE ones, does not, in essence, quality as a reverse *shift*, the definition might indeed be an appropriate one for possible SECS-Shift developments in SAE. A 'beady' eye therefore needs to be kept on SAE in this regard; in particular, evidence for in-progress, secondary STRUT-retraction, as a result of TRAP-lowering, would provide evidence that SAE was a full, although 'lagging', participant in this regard.

We have seen, however, in §3.3.1, that GenSAE can hardly be considered a 'conservative' variety, and was, at least in earlier years, a distinctly 'local' variant, in opposition to CulSAE. GenSAE is, by all accounts, a *vernacular*, with its own complex 'load' of indexical meaning. As such, and in line with more modern theories of indexicality, as outlined briefly in §1.3.3, I would suggest that post-koinéization developments in SAE can only be explained if a more 'situated' approach to issues of identity is adopted; one that does not simply link endogenous, structurally-motivated, changes-from-below with the interior classes and exogenously-motivated changes-from-above with borrowing from other dialects. An explanation of the *systematic* arrestment, in SAE, of endogenously-motivated changes, the diffusion of prestige variants and, most particularly, the potential systematic reversal from '*above*' of a chain shift pattern, all require a theory of indexicality that takes into account the specific sociohistorical factors that have characterised SAE and its speakers. While an explanation, in all likelihood, lies in the unique class structure of 'white' English-speaking South Africa (§4.2.3.2), as well as the complex indexicality of its various sociolects (§3.3.1), such an explanation will have to await further research; and, furthermore, research focussed on applying a more 'grounded' approach to indexicality, as explained in §1.3.3.

5. With regard to the koinéization processes that led to the development of modern SAE, this thesis has proposed a three-stage koinéization model, which although substantively based on previous work, focusses attention in a useful way, I believe, on the role played by dialect and language-contact on the development of SAE as we know it today. In particular, it hopefully re-energises research into the potential influence of north-of-England varities of English, Afrikaans, and Yiddish on the development of SAE. More generally, the useful results derived from this model emphasize the utility of adopting a 'catholic' approach to the sources and forces involved in new-dialect formation and language change in general; an approach adopted by, for example, Mufwene (2008), who emphasizes that the 'ecology' of a variety involves both structural and exogenous factors.

10.3 Recommendations for Future Research

In line with various comments provided above, I would like to recommend two broad avenues of future research:

Apparent-time investigations: There are a number of possible developments that could be confirmed and clarified by the application of standard variationist techniques for the study of change-in-progress, in particular those of an apparent-time nature. Areas of particular interest include the ongoing participation of SAE in a SECS-like shift as well as the diffusion of GenSAE variants at the expense of BrSAE ones, with a particular focus on those vowels participating in the Diphthong Shift and the related Pattern 1 and Pattern 4 chain-shifts. Studies in both of these directions would be interesting in terms of a broader theory of indexicality, given that these changes appear to be driven by prestige-factors while at the same time retaining a broadly systematic aspect.

Real-time investigations: Possibilities here are of two kinds. Firstly, real-time investigations into SAE would complement the apparent-time forms of research mentioned above. Some initial steps have, in fact, already been taken in this respect, in the form of the collection of real-time SAE data as part of the broader South African Phonetic Corpus Project being conducted in the English Language and Linguistics Department at Rhodes University, and as reported on in De Klerk, Adendorff, de Vos, Hunt, Niesler, Simango and Todd (2006). Secondly, sociohistorical and demographic research on the various proposed stages of SAE koinéization (Eastern Cape, Natal and Johannesburg) would assist in a more thorough linguistic reconstruction of the historical development of SAE, both in terms of its linguistic particulars and its more broadly-conceived social history. Such research would be particularly useful in ‘teasing apart’ the role played by language and dialect contact and the role played by endogenous, structurally-motivated forces. As mentioned above, varieties that deserve particular attention are those of the north-of-England, Afrikaans and the main variety of Yiddish brought over to South Africa as part of the diaspora; where possible with a focus, of course, on 19th-century variants.

A number of other, more minor, details require investigation as well, and have been mentioned, in passing, in previous chapters. Some, tentative, although as yet unpublished, progress has already been made in this regard, with Orpen (n.d.) focussed on a comparison between SAE and Rhodesian English in terms of the Second FORCE Merger and Giudici (n.d.) using acoustic techniques to resolve the debate in the literature concerning the qualitative status of /l/ in syllable-final position. Much more work, however, needs to be done.

10.4 Final Conclusion

This thesis has provided a comparative analysis of vowel quality in SAE, focussed on both its synchronic and diachronic status. It has employed modern acoustic techniques and a focussed approach, in the form of a three-stage koinéization model, toward the reconstruction of SAE's linguistic history and development. It has, at all times, tried to embed this comparative analysis within a number of theoretical and descriptive frameworks. With respect to these frameworks, and in accordance with the analyses provided at various points in the thesis, it provides the following:

- a proposed adjustment to Schneider's (2007) Dynamic Model which takes into account various 'waves' of nativization.
- evidence, with respect to the BATH and MOUTH vowels, showing Trudgill's (2004) model of new-dialect formation to be inadequate in accounting for the SAE data; and, consequently, support for Schneider's (2008) emphasis on the importance of indexical factors in new-dialect formation.
- evidence to suggest that SAE is an imminent and 'conservative' member of Torgersen and Kerswill's (2004) SECS-shift; it, in addition, uses SAE data to question the applicability of the SECS-shift to FOOT-Fronting.
- evidence and analysis to suggest that SAE has undergone, in certain respects, a prestige-driven arrestment of the Diphthong and Southern Shifts and a subsequent diffusion of GenSAE variants at the expense of BrSAE ones. With respect to SAE's possible inclusion in the SECS-Shift, there is also tentative evidence for a chain-shift reversal 'from above'. It proposes that, in order to understand these phenomena, recourse needs to be made to a theory of indexicality that takes into account the unique sociohistorical development of SAE and its speakers.

The future development of SAE remains to be charted, and it is hoped that the present work has provided a solid-enough basis from which to proceed.

Appendix A: The International Phonetic Alphabet

THE INTERNATIONAL PHONETIC ALPHABET (2005)

CONSONANTS (PULMONIC)

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Epi-glottal	Glottal
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ			
Plosive	p b	ȹ ȸ		t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʡ	ʔ
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	ʜ ʢ	h ɦ
Approximant		ʋ		ɹ		ɻ	j	ɰ				
Trill	ʙ			r					ʀ		ᴙ	
Tap, Flap		ⱱ		ɾ		ɽ						
Lateral fricative				ɬ ɮ		ꞎ	𝼆	𝼄				
Lateral approximant				l		ɭ	ʎ	ʟ				
Lateral flap				ɺ		𝼈						

Where symbols appear in pairs, the one to the right represents a modally voiced consonant, except for murmured *ɦ*.
Shaded areas denote articulations judged to be impossible. Light grey letters are unofficial extensions of the IPA.

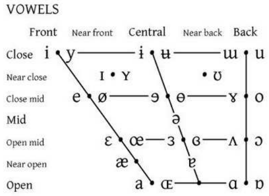
CONSONANTS (NON-PULMONIC)

Anterior click releases (require posterior stops)	Voiced implosives	Ejectives
ʘ Bilabial fricated	ɓ Bilabial	ʼ *Examples:*
ǀ Laminal alveolar fricated ("dental")	ɗ Dental or alveolar	pʼ Bilabial
ǃ Apical (post)alveolar abrupt ("retroflex")	ʄ Palatal	tʼ Dental or alveolar
ǂ Laminal postalveolar abrupt ("palatal")	ɠ Velar	kʼ Velar
ǁ Lateral alveolar fricated ("lateral")	ʛ Uvular	sʼ Alveolar fricative

CONSONANTS (CO-ARTICULATED)

ʍ	Voiceless labialized velar approximant
w	Voiced labialized velar approximant
ɥ	Voiced labialized palatal approximant
ɕ	Voiceless palatalized postalveolar (alveolo-palatal) fricative
ʑ	Voiced palatalized postalveolar (alveolo-palatal) fricative
ɧ	Simultaneous *x* and *ʃ* (disputed)
k͡p t͡s	Affricates and double articulations may be joined by a tie bar

VOWELS



Vowels at right & left of bullets are rounded & unrounded.

SUPRASEGMENTALS

ˈ	Primary stress	ˈˈ	Extra stress
ˌ	Secondary stress	*[ˌfoʊnəˈtɪʃən]*	
eː	Long	eˑ	Half-long
e	Short	ĕ	Extra-short
.	Syllable break	‿	Linking (no break)

INTONATION

|	Minor (foot) break
‖	Major (intonation) break
↗	Global rise
↘	Global fall

TONE

Level tones			*Contour-tone examples:*		
e̋	˥	Top	ě	˩˥	Rising
é	˦	High	ê	˥˩	Falling
ē	˧	Mid	e᷄	˦˥	High rising
è	˨	Low	e᷅	˩˨	Low rising
ȅ	˩	Bottom	e᷇	˥˦	High falling
Tone terracing			e᷆	˨˩	Low falling
ꜛ		Upstep	e᷈	˧˦˧	Peaking
ꜜ		Downstep	e᷉	˧˨˧	Dipping

DIACRITICS

Diacritics may be placed above a symbol with a descender, as *ŋ̊*. Other IPA symbols may appear as diacritics to represent phonetic detail: *tˢ* (fricative release), *bʱ* (breathy voice), *ˀa* (glottal onset), *ᵊ* (epenthetic schwa), *oᵘ* (diphthongization).

SYLLABICITY & RELEASES		PHONATION		PRIMARY ARTICULATION		SECONDARY ARTICULATION			
n̩ ɹ̩	Syllabic	n̥ d̥	Voiceless *or* Slack voice	t̪ b̪	Dental	tʷ dʷ	Labialized	ɔ̹ x̹	More rounded
e̯ ʊ̯	Non-syllabic	s̬ d̬	Modal voice *or* Stiff voice	t̺ d̺	Apical	tʲ dʲ	Palatalized	ɔ̜ x̜ʷ	Less rounded
tʰ ʰt	(Pre)aspirated	n̤ a̤	Breathy voice	t̻ d̻	Laminal	tˠ dˠ	Velarized	ẽ z̃	Nasalized
dⁿ	Nasal release	n̰ a̰	Creaky voice	u̟ t̟	Advanced	tˤ dˤ	Pharyngealized	ɚ ɝ	Rhoticity
dˡ	Lateral release	n͈ a͈	Strident	i̠ t̠	Retracted	ɫ z̴	Velarized or pharyngealized	e̘ o̘	Advanced tongue root
t̚	No audible release	n̼ d̼	Linguolabial	ä j̈	Centralized	ɯ̽	Mid-centralized	e̙ o̙	Retracted tongue root
e̞ β̞	Lowered (*β̞* is a bilabial approximant)			e̝ ɹ̝	Raised (*ɹ̝* is a voiced alveolar non-sibilant fricative)				

Appendix B: The O–G Word–List

The word-list below is derived from that provided by Lass (1990:285).

1. **Stressed and unstressed vowels**: BIT BET BAT FOOT BUT POT BEAT BOOT BOUGHT PALM HURT BITE MATE BOY OUT OAT CUTE HERE FAIR POUR POOR PAW CITY COMMA TAXES WOUNDED MOTHER TAXIS.
2. **Short vowel allophonics**: HIT IT SIT KIT SING LIMP RIM FILL FISH; BAT BAD VAN BACK BAG BANG HAT HAD MAD HAM; PECK BECK SPECK; TUCK DUCK STUCK COT GOT SCOTT; LIP PILL SILLY LAP PAL SALLY LET TELL TELLY LOOK PULL PULLEY.
3. **Loan words (major categories)**: AG KWAGGA BAKKIE KAK SCHALK; STOEP POEP DOEK BOEREWORS; KREEF STEEN STEENBERG VEREENIGING; STOMPIE GOGGA WELKOM BOSMAN
4. **STRUT vs. PAP**: BUT PAP SUCK TANNIE SAK SKULK CUT SCHALK PUP BANTU BLATJANG

- Although recorded in some cases, the Afrikaans loan–words (including those under the '**STRUT vs. PAP**'–category) have not been included in the analysis.
- In individual cases some or all of the following words were added to the list: RED, ROSE, EIGHT.

Appendix C: The N-G Word-List

Short Monophthongs

KIT(14)	DRESS(3)	TRAP(9)	STRUT(9)	LOT(3)	FOOT(3)
IT(initial)	BET	HAD	HUD	HOD	HOOD
HID(after /h/)	TELL	LASS	BUS	HOT	FOOT
SING(before velars)	HEAD	HAS	BUZZ		FULL
KIT(after velars)		BACK	BUCK		
FISH(before palato-alveolars)		BAT	BUT		
BIT(context of labials)		BAD	BUD		
LIT(after /l/)		BAN	BUN		
RID (after /r/)		PAL	SKULL		
TILL(final /l/)		CAP	SNUFF		
WITH(after /w/)					
DISH					
BITCH					
CHIN					
SIT					

Long Monophthongs

FLEECE(3)	GOOSE(4)	NURSE(3)	THOUGHT(3)	BATH(3)	SQUARE(3)
HEED	WHO'D	HERD	HOARD	HARD	HAIR
BEAT	BOOT	HURT	BOUGHT	HEART	SCARE
FIELD	FOOD	HURL	HAUL	SNARL	SHARE
	FOOL				

Diphthongs

FACE(4)	GOAT(4)	NEAR(3)	CURE(3)	PRICE(11)	CHOICE(3)	MOUTH(4)
HAY	HOE	HERE	CURE	HEIGHT	BOIL	HOW
MAID	MOAT	BEER	MOOR	HIDE	EXPLOIT	LOUD
MATE	LOAD	BEARD	PURE	HYPE	VOID	LOUT
FAIL	HOLE			HIKE		FOUL
	TOLD			IRON		
				PRICE		
				PRIZE		
				TRY		
				MINE		
				MILE		
				MICE		

Weak Vowels

happY(2)	lettER(2)	commA(2)	Non-Final(2)
COPY	PAPER	SOFA	RABBIT
CITY	TOWER	ARENA	ABBOT

Appendix D: Formant Means and Standard Deviations

Each table in this appendix links directly to one of the main sections provided in Chapters 6 to 9 and provides means ($\bar{x}$) and standard deviations (σ) for the South African English data recorded and analyzed for the purposes of this research and used to plot the various figures in the above-mentioned chapters. The means and standard deviations are based on 'raw' Hz values[2]. For each vowel (or vowel allophone) the number of tokens analyzed is also provided (e.g. *N*=79). F_1 and F_2 mean values and standard deviations across the duration of the segmented portion of the vowel are also provided, at increments of 10% of duration. $P_{10}\{F_1\}$ thus relates to the F_1 mean and standard deviation values at the 10th percentile of duration and $P_{60}\{F_2\}$ to the F_2 of the 60th percentile. Since most readers would be interested in using the mean (and standard deviation) values closest to the vowel target, the temporal midpoint values (i.e. $P_{50}\{F_1\}$ and $P_{50}\{F_2\}$) have been bolded. Each table also provides, at the end, means and standard deviations for the relevant vowel's duration in seconds[3].

In the case of vowels before syllable-final /l/ the relevant allophone is identifiable, in the various tables, by the use of a subscript $_L$ e.g. PRICE$_L$. In the case of KIT, the reader is referred back to Table 7.1 on page 267 for a full elaboration of the various symbols used (e.g. KIT1, KIT2 etc.).

[2]The individual token values (Hz and Lobanov–normalized) as well as the means and standard deviations of the Lobanov–normalized values are available on request from the author.

[3]As implied in §4.5, there was a difference between the procedure used to calculate the length of any particular vowel and that used to measure VISC. In the former case, the original segmentation boundaries, determined on the basis of Peterson and Lehiste's (1960) heuristics for vowel-length segmentation, were, in general, used to calculate vowel length. In the later case, the middle half or middle 70% of the segmented vowel (depending on whether the vowel was traditionally short or long) was isolated for measurement of VISC. This middle half or middle 70% was then measured at 10% intervals for F_1 and F_2 values.

MOUTH, PRACTICE and CHOICE – See § 6.2 on page 155

	MOUTH N=79		MOUTH$_L$ N=18		PRICE N=117		PRICE$_L$ N=18		CHOICE N=44		CHOICE$_L$ N=18	
	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ
$P_0\{F_1\}$	981.3	121.3	848.8	75.05	930.7	146.1	898.2	137.1	546.6	61.32	508.8	40.04
$P_0\{F_2\}$	1546	140.3	1379	117.9	1465	123.9	1384	120.4	1092	189.2	922.2	92.31
$P_{10}\{F_1\}$	996.0	115.6	902.0	92.90	955.8	133.6	922.5	164.9	565.4	67.35	528.2	33.60
$P_{10}\{F_2\}$	1501	135.9	1389	103.0	1486	130.6	1397	114.3	1091	178.0	937.2	70.96
$P_{20}\{F_1\}$	988.1	106.7	922.8	99.19	970.5	132.0	931.6	165.3	577.1	62.89	549.9	38.12
$P_{20}\{F_2\}$	1472	125.3	1376	109.0	1509	138.3	1414	120.5	1127	171.7	990.7	80.46
$P_{30}\{F_1\}$	970.0	104.1	930.7	93.79	975.7	124.2	953.8	154.8	594.6	64.68	573.2	43.70
$P_{30}\{F_2\}$	1445	123.1	1381	86.39	1548	151.1	1427	124.9	1220	182.0	1091	98.93
$P_{40}\{F_1\}$	954.2	105.8	939.0	101.5	977.3	131.1	974.7	153.0	611.0	62.91	609.8	59.73
$P_{40}\{F_2\}$	1414	124.9	1367	112.3	1607	155.4	1458	128.0	1372	219.2	1225	144.5
$\boldsymbol{P_{50}\{F_1\}}$	**929.7**	**106.1**	**926.9**	**108.2**	**962.8**	**121.7**	**982.8**	**161.4**	**618.2**	**72.09**	**632.3**	**49.56**
$\boldsymbol{P_{50}\{F_2\}}$	**1388**	**128.1**	**1354**	**124.3**	**1665**	**165.8**	**1506**	**147.9**	**1577**	**273.0**	**1399**	**218.8**
$P_{60}\{F_1\}$	897.6	117.5	910.6	111.8	929.2	111.4	967.1	161.8	611.8	77.07	632.7	42.84
$P_{60}\{F_2\}$	1349	131.2	1325	134.7	1740	189.7	1551	169.3	1796	339.3	1588	274.2
$P_{70}\{F_1\}$	866.1	113.7	905.1	114.2	886.0	113.3	933.3	148.6	581.2	81.35	618.8	43.57
$P_{70}\{F_2\}$	1302	137.2	1302	136.3	1836	218.5	1588	184.7	1992	333.1	1746	325.3
$P_{80}\{F_1\}$	819.5	127.8	893.6	118.2	827.8	119.9	900.9	139.7	547.7	75.85	614.4	48.49
$P_{80}\{F_2\}$	1264	144.5	1276	137.2	1929	244.7	1618	202.9	2153	303.7	1835	330.1
$P_{90}\{F_1\}$	777.3	133.2	869.6	124.4	772.4	128.9	863.2	132.3	519.5	71.87	615.0	51.49
$P_{90}\{F_2\}$	1233	153.3	1246	150.0	2008	262.4	1604	195.0	2274	271.0	1838	318.9
$P_{100}\{F_1\}$	729.0	143.8	842.6	136.2	714.5	142.7	823.1	121.3	487.9	75.41	619.6	43.26
$P_{100}\{F_2\}$	1226	156.6	1226	134.3	2085	283.2	1608	198.5	2351	236.1	1795	297.8
sec.	0.223	0.078	0.196	0.059	0.234	0.086	0.226	0.063	0.267	0.082	0.232	0.044

TRAP – See § 6.3 on page 190

	TRAP N=301		TRAP$_L$ N=27	
	$\bar{x}$	σ	$\bar{x}$	σ
$P_0\{F_1\}$	898.0	120.1	926.2	106.7
$P_0\{F_2\}$	1926	159.0	1583	181.5
$P_{10}\{F_1\}$	919.0	116.7	925.2	109.7
$P_{10}\{F_2\}$	1920	153.8	1575	183.0
$P_{20}\{F_1\}$	935.5	115.0	918.0	108.7
$P_{20}\{F_2\}$	1917	154.4	1567	178.5
$P_{30}\{F_1\}$	944.6	113.2	910.2	105.9
$P_{30}\{F_2\}$	1915	148.5	1550	176.2
$P_{40}\{F_1\}$	947.9	111.4	900.8	112.0
$P_{40}\{F_2\}$	1909	149.4	1532	176.6
$\boldsymbol{P_{50}\{F_1\}}$	**945.7**	**112.3**	**895.3**	**117.5**
$\boldsymbol{P_{50}\{F_2\}}$	**1904**	**145.9**	**1515**	**184.3**
$P_{60}\{F_1\}$	943.1	116.0	880.8	110.6
$P_{60}\{F_2\}$	1901	146.1	1494	175.5
$P_{70}\{F_1\}$	937.1	122.7	867.0	108.1
$P_{70}\{F_2\}$	1901	147.7	1468	176.2
$P_{80}\{F_1\}$	926.6	129.8	853.9	105.8
$P_{80}\{F_2\}$	1900	150.5	1442	175.1
$P_{90}\{F_1\}$	911.7	137.9	839.7	105.1
$P_{90}\{F_2\}$	1897	152.7	1420	169.7
$P_{100}\{F_1\}$	892.3	149.8	827.6	106.9
$P_{100}\{F_2\}$	1893	165.8	1397	172.2
sec.	0.198	0.089	0.150	0.046

FACE, DRESS and FLEECE – See § 6.4 on page 206

	FACE N=82		FACE$_L$ N=18		DRESS N=94		DRESS$_L$ N=27		FLEECE N=58		FLEECE$_L$ N=18	
	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ
$P_0\{F_1\}$	667.9	76.18	687.0	64.04	567.6	66.18	868.7	127.9	377.5	56.26	408.2	70.45
$P_0\{F_2\}$	2317	208.5	1970	162.9	2407	243.0	1659	128.1	2656	303.9	2608	155.8
$P_{10}\{F_1\}$	663.2	72.27	701.8	66.49	571.8	68.97	867.7	122.9	381.7	55.65	413.0	69.83
$P_{10}\{F_2\}$	2358	204.7	2026	144.2	2425	233.1	1640	139.9	2729	264.6	2615	176.0
$P_{20}\{F_1\}$	646.7	70.89	693.7	59.72	574.5	72.19	862.7	117.6	381.2	53.51	430.1	61.98
$P_{20}\{F_2\}$	2398	214.1	2090	158.5	2437	228.9	1619	155.8	2776	237.4	2606	149.2
$P_{30}\{F_1\}$	625.8	71.99	684.3	57.76	576.5	76.08	855.4	107.9	379.4	52.48	456.9	71.33
$P_{30}\{F_2\}$	2456	191.9	2128	176.0	2454	226.2	1588	143.3	2805	223.9	2557	165.9
$P_{40}\{F_1\}$	602.7	76.19	664.0	49.42	578.0	78.74	852.5	101.4	375.1	53.30	483.7	84.99
$P_{40}\{F_2\}$	2501	184.7	2162	193.3	2473	213.0	1552	136.8	2836	210.2	2416	239.7
$\boldsymbol{P_{50}\{F_1\}}$	**579.4**	**77.75**	**653.4**	**59.16**	**577.4**	**80.24**	**846.9**	**101.3**	**371.8**	**50.57**	**514.4**	**86.01**
$\boldsymbol{P_{50}\{F_2\}}$	**2548**	**185.1**	**2187**	**192.2**	**2481**	**213.0**	**1527**	**150.3**	**2849**	**196.6**	**2251**	**273.8**
$P_{60}\{F_1\}$	553.6	78.01	641.4	64.68	575.8	81.75	838.2	95.83	371.2	48.49	543.7	84.32
$P_{60}\{F_2\}$	2591	193.4	2193	187.6	2481	209.6	1497	142.6	2868	193.3	2070	291.1
$P_{70}\{F_1\}$	524.8	78.71	642.2	62.24	573.1	78.91	821.9	91.60	373.3	48.52	573.5	91.15
$P_{70}\{F_2\}$	2623	199.1	2132	144.6	2480	209.8	1477	162.8	2883	196.0	1906	305.0
$P_{80}\{F_1\}$	498.8	76.65	641.0	60.10	569.5	76.18	805.5	89.78	376.5	53.63	584.4	77.22
$P_{80}\{F_2\}$	2658	204.4	2041	148.8	2482	203.5	1448	150.5	2890	195.0	1751	298.8
$P_{90}\{F_1\}$	478.5	73.35	647.8	62.69	564.9	76.73	790.9	88.12	377.0	58.70	580.5	78.13
$P_{90}\{F_2\}$	2674	211.2	1948	165.3	2482	201.7	1416	135.7	2874	190.9	1620	291.8
$P_{100}\{F_1\}$	458.6	72.37	649.0	60.08	559.4	76.47	781.4	93.49	375.5	64.99	580.0	74.47
$P_{100}\{F_2\}$	2645	225.5	1807	235.9	2469	198.3	1387	126.8	2842	190.6	1510	271.3
sec.	0.211	0.085	0.211	0.056	0.152	0.040	0.141	0.036	0.183	0.062	0.198	0.059

KIT – See Chapter 7 on page 239

	KIT1		KIT2		KIT3		KIT4		KIT5		KIT6	
	N=36		N=108		N=63		N=45		N=89		N=54	
	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ
$P_0\{F_1\}$	509.5	42.31	566.1	68.14	552.1	63.39	576.5	56.33	578.2	63.02	610.9	85.24
$P_0\{F_2\}$	1984	129.6	2228	197.2	2037	134.0	1927	126.9	1880	143.6	1527	209.0
$P_{10}\{F_1\}$	513.1	45.78	571.0	69.23	560.3	63.83	584.3	59.73	590.6	64.48	616.8	80.73
$P_{10}\{F_2\}$	1996	130.6	2224	186.6	2049	132.6	1926	122.9	1894	143.3	1535	210.8
$P_{20}\{F_1\}$	513.5	49.14	574.6	70.71	566.6	64.06	591.1	60.23	601.6	64.78	620.8	75.02
$P_{20}\{F_2\}$	2009	133.2	2220	180.0	2061	131.3	1924	124.6	1905	144.0	1535	217.5
$P_{30}\{F_1\}$	512.3	51.12	576.2	70.95	571.5	64.17	596.1	64.35	609.9	65.32	623.5	69.44
$P_{30}\{F_2\}$	2021	136.5	2218	175.0	2067	126.0	1924	130.3	1913	144.1	1538	232.2
$P_{40}\{F_1\}$	509.6	52.09	576.7	71.79	575.5	64.14	599.2	65.77	614.9	66.39	625.1	65.43
$P_{40}\{F_2\}$	2031	138.7	2214	167.8	2075	125.4	1923	131.5	1919	143.3	1530	250.2
$\boldsymbol{P_{50}\{F_1\}}$	**505.3**	**52.62**	**575.7**	**73.27**	**576.1**	**63.45**	**603.1**	**65.71**	**618.7**	**67.48**	**625.7**	**62.51**
$\boldsymbol{P_{50}\{F_2\}}$	**2039**	**141.4**	**2211**	**169.1**	**2088**	**129.1**	**1928**	**126.7**	**1925**	**143.1**	**1523**	**270.1**
$P_{60}\{F_1\}$	500.0	53.48	571.4	72.57	573.1	62.61	603.6	66.05	618.4	67.88	622.6	60.94
$P_{60}\{F_2\}$	2047	145.6	2208	174.1	2095	131.6	1935	126.0	1930	145.7	1518	289.1
$P_{70}\{F_1\}$	493.4	54.17	567.3	72.19	569.1	59.85	600.3	66.36	613.5	68.46	618.2	58.95
$P_{70}\{F_2\}$	2054	146.4	2210	181.8	2098	137.1	1939	125.1	1931	151.4	1513	301.0
$P_{80}\{F_1\}$	485.0	55.43	560.7	71.73	562.3	60.23	596.8	68.20	606.1	70.71	614.8	58.31
$P_{80}\{F_2\}$	2059	149.2	2212	185.4	2100	139.2	1942	121.6	1932	160.7	1508	319.0
$P_{90}\{F_1\}$	475.9	57.23	552.7	71.41	551.0	63.99	589.9	68.88	596.3	74.68	608.3	59.42
$P_{90}\{F_2\}$	2061	155.9	2213	190.9	2105	137.6	1944	119.0	1937	164.4	1495	330.6
$P_{100}\{F_1\}$	465.5	58.97	543.6	71.26	538.6	67.02	581.5	72.65	583.9	76.31	601.3	60.56
$P_{100}\{F_2\}$	2066	164.1	2210	212.4	2113	135.1	1945	119.3	1940	170.3	1480	338.9
sec.	0.061	0.014	0.102	0.030	0.126	0.033	0.099	0.022	0.111	0.028	0.129	0.036

BATH, THOUGHT and GOOSE – See § 8.2 on page 279

	BATH N=43		BATH$_L$ N=18		THOUGHT N=66		THOUGHT$_L$ N=18		GOOSE N=76		GOOSE$_L$ N=17	
	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ
$P_0\{F_1\}$	864.9	116.5	773.6	124.9	508.6	41.68	504.2	46.18	395.0	53.45	426.4	43.41
$P_0\{F_2\}$	1252	96.65	1451	124.0	918.7	89.06	863.5	108.9	1992	330.0	987.2	106.5
$P_{10}\{F_1\}$	860.0	110.4	814.3	117.4	510.8	44.47	501.6	47.60	401.4	52.98	422.5	47.65
$P_{10}\{F_2\}$	1264	92.07	1393	132.3	893.1	92.24	856.7	102.1	2034	308.8	970.7	108.0
$P_{20}\{F_1\}$	865.4	106.6	836.1	144.5	512.6	42.44	497.1	47.07	404.2	50.69	424.9	51.84
$P_{20}\{F_2\}$	1263	92.34	1321	129.8	877.3	93.59	852.9	100.4	2068	290.9	959.3	100.6
$P_{30}\{F_1\}$	864.0	102.6	834.1	149.7	511.1	48.55	489.6	45.67	407.8	48.68	430.9	52.88
$P_{30}\{F_2\}$	1267	98.34	1278	118.9	866.0	88.28	856.8	97.29	2082	276.9	941.2	99.77
$P_{40}\{F_1\}$	858.4	102.6	831.9	151.9	509.4	51.08	489.0	45.01	407.9	47.86	432.3	52.06
$P_{40}\{F_2\}$	1275	93.85	1230	103.9	858.7	88.20	853.8	108.3	2101	254.7	926.8	91.71
$P_{50}\{F_1\}$	**858.1**	**100.2**	**805.4**	**127.7**	**507.1**	**49.71**	**490.3**	**46.07**	**407.9**	**45.86**	**423.6**	**51.10**
$P_{50}\{F_2\}$	**1284**	**94.34**	**1211**	**100.9**	**853.9**	**91.72**	**842.0**	**111.3**	**2116**	**244.6**	**917.6**	**92.07**
$P_{60}\{F_1\}$	852.0	101.6	785.8	116.1	508.4	49.71	495.7	46.71	406.3	43.88	419.3	47.14
$P_{60}\{F_2\}$	1303	101.6	1188	102.9	858.1	98.11	841.8	112.6	2122	242.3	915.6	89.34
$P_{70}\{F_1\}$	843.5	107.1	775.2	124.9	503.3	45.20	492.2	48.29	402.0	42.72	414.7	41.92
$P_{70}\{F_2\}$	1328	112.0	1179	104.2	876.0	117.0	847.1	105.0	2117	241.3	904.5	89.00
$P_{80}\{F_1\}$	843.7	97.68	775.9	107.0	500.1	39.99	489.2	42.90	396.7	42.44	407.4	42.04
$P_{80}\{F_2\}$	1368	147.1	1156	91.81	911.6	133.0	846.0	105.3	2123	233.1	902.9	105.7
$P_{90}\{F_1\}$	833.1	95.98	777.5	114.1	499.3	42.21	490.2	43.46	390.6	42.28	401.3	41.38
$P_{90}\{F_2\}$	1422	175.1	1138	91.45	963.2	161.2	846.2	112.8	2140	227.4	895.1	100.3
$P_{100}\{F_1\}$	799.1	113.2	779.9	115.2	496.5	54.51	486.9	50.18	384.3	42.68	396.0	41.16
$P_{100}\{F_2\}$	1488	216.8	1138	85.75	1053	226.3	839.0	111.7	2162	208.1	885.4	115.9
sec.	0.241	0.062	0.193	0.051	0.240	0.067	0.207	0.066	0.186	0.053	0.186	0.050

GOAT and CURE – See § 8.3 on page 313

	GOAT N=83		GOAT$_L$ N=35		CURE N=62	
	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ
$P_0\{F_1\}$	645.1	67.59	628.7	76.22	518.7	80.41
$P_0\{F_2\}$	1515	143.1	1128	116.2	1722	502.4
$P_{10}\{F_1\}$	652.9	60.90	642.7	63.57	561.8	78.69
$P_{10}\{F_2\}$	1523	131.4	1101	120.9	1672	410.1
$P_{20}\{F_1\}$	648.1	61.19	640.5	61.66	598.2	78.24
$P_{20}\{F_2\}$	1551	130.1	1080	113.1	1624	319.7
$P_{30}\{F_1\}$	635.7	62.99	629.6	60.83	635.2	84.16
$P_{30}\{F_2\}$	1586	128.8	1062	105.1	1608	246.3
$P_{40}\{F_1\}$	619.6	67.09	609.6	62.84	663.0	84.05
$P_{40}\{F_2\}$	1624	151.8	1037	112.8	1612	186.6
$\boldsymbol{P_{50}\{F_1\}}$	**594.2**	**65.58**	**593.4**	**64.19**	**689.8**	**83.48**
$\boldsymbol{P_{50}\{F_2\}}$	**1660**	**150.9**	**1014**	**117.1**	**1618**	**157.0**
$P_{60}\{F_1\}$	566.0	65.77	581.9	69.08	711.3	82.83
$P_{60}\{F_2\}$	1695	162.6	995.9	115.9	1619	129.3
$P_{70}\{F_1\}$	540.4	62.18	573.4	69.62	716.5	85.75
$P_{70}\{F_2\}$	1724	174.8	973.9	108.9	1613	120.3
$P_{80}\{F_1\}$	512.0	62.23	558.4	72.25	722.0	89.23
$P_{80}\{F_2\}$	1758	191.6	961.4	109.5	1619	123.7
$P_{90}\{F_1\}$	488.0	59.90	547.4	72.30	712.0	102.2
$P_{90}\{F_2\}$	1776	217.6	951.8	109.1	1621	124.9
$P_{100}\{F_1\}$	460.1	57.79	536.8	66.19	704.4	115.4
$P_{100}\{F_2\}$	1813	242.2	945.1	109.7	1628	131.8
sec.	0.225	0.086	0.171	0.043	0.280	0.054

STRUT, LOT and FOOT – See § 8.4 on page 331

	STRUT N=204		STRUT$_L$ N=23		LOT N=95		FOOT N=64		FOOT$_L$ N=28	
	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ
$P_0\{F_1\}$	859.3	113.5	704.5	124.0	816.2	122.2	578.6	67.65	549.4	75.68
$P_0\{F_2\}$	1646	152.4	1628	136.2	1319	121.8	1460	198.9	1088	108.6
$P_{10}\{F_1\}$	878.1	108.1	738.7	119.8	829.9	109.6	585.6	69.19	546.5	79.12
$P_{10}\{F_2\}$	1655	143.7	1581	126.2	1317	110.0	1473	190.0	1079	112.4
$P_{20}\{F_1\}$	891.8	108.0	771.5	117.8	839.8	102.5	588.3	68.72	541.2	77.87
$P_{20}\{F_2\}$	1659	139.4	1543	124.1	1319	104.6	1490	179.6	1068	115.9
$P_{30}\{F_1\}$	901.2	109.5	793.0	115.2	845.5	98.86	588.8	70.30	536.3	75.19
$P_{30}\{F_2\}$	1666	138.1	1511	122.2	1327	101.5	1510	174.9	1061	112.7
$P_{40}\{F_1\}$	902.5	111.3	809.3	118.6	848.7	99.17	588.1	71.36	530.0	73.81
$P_{40}\{F_2\}$	1673	136.6	1480	119.5	1337	102.2	1534	171.0	1057	109.5
$\boldsymbol{P_{50}\{F_1\}}$	**899.3**	**118.5**	**821.1**	**125.7**	**851.2**	**103.1**	**585.7**	**71.49**	**522.4**	**70.94**
$\boldsymbol{P_{50}\{F_2\}}$	**1679**	**136.6**	**1446**	**119.8**	**1349**	**105.0**	**1564**	**168.7**	**1047**	**107.1**
$P_{60}\{F_1\}$	892.1	124.2	823.3	121.3	850.5	104.2	582.7	72.13	516.8	68.98
$P_{60}\{F_2\}$	1689	138.9	1421	113.8	1366	108.7	1595	170.9	1038	105.0
$P_{70}\{F_1\}$	882.5	131.0	814.6	113.4	843.3	104.6	579.2	74.12	513.5	68.88
$P_{70}\{F_2\}$	1703	142.2	1396	114.5	1383	117.2	1626	180.0	1030	107.0
$P_{80}\{F_1\}$	870.7	139.8	799.5	104.6	834.5	109.3	572.8	75.68	512.8	69.91
$P_{80}\{F_2\}$	1719	150.5	1358	118.6	1401	133.2	1651	186.3	1021	110.3
$P_{90}\{F_1\}$	849.2	146.2	782.1	89.56	821.0	114.3	565.0	77.52	511.1	67.63
$P_{90}\{F_2\}$	1732	159.2	1333	114.8	1419	154.6	1682	190.6	1015	107.1
$P_{100}\{F_1\}$	826.3	157.3	764.6	88.80	802.1	122.2	554.9	80.57	509.6	67.56
$P_{100}\{F_2\}$	1745	167.3	1314	106.8	1437	171.2	1719	196.0	1004	104.9
sec.	0.153	0.054	0.154	0.045	0.116	0.038	0.104	0.037	0.132	0.036

NURSE, NEAR and SQUARE – See § 9.2 on page 363

	NURSE N=44		NURSE$_L$ N=18		NEAR N=60		SQUARE N=61	
	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ
$P_0\{F_1\}$	558.9	75.37	572.7	62.31	482.8	103.7	551.2	93.34
$P_0\{F_2\}$	1876	189.9	1614	136.5	2484	209.5	2355	208.7
$P_{10}\{F_1\}$	570.0	76.01	588.5	65.89	517.4	123.8	566.5	88.52
$P_{10}\{F_2\}$	1891	184.5	1604	128.8	2422	226.2	2347	196.9
$P_{20}\{F_1\}$	572.3	76.53	592.4	70.33	560.6	134.2	581.4	94.43
$P_{20}\{F_2\}$	1912	167.9	1586	126.8	2330	252.6	2363	200.9
$P_{30}\{F_1\}$	574.8	76.92	593.0	69.43	613.1	134.4	595.5	97.04
$P_{30}\{F_2\}$	1917	165.5	1558	129.6	2210	237.8	2356	217.2
$P_{40}\{F_1\}$	574.3	71.20	599.1	66.10	661.8	136.2	599.8	93.96
$P_{40}\{F_2\}$	1927	158.9	1531	131.1	2090	207.2	2342	214.3
$\boldsymbol{P_{50}\{F_1\}}$	**575.7**	**76.37**	**601.4**	**62.35**	**689.3**	**132.5**	**603.9**	**94.29**
$\boldsymbol{P_{50}\{F_2\}}$	**1930**	**152.9**	**1492**	**126.4**	**2002**	**200.5**	**2332**	**212.3**
$P_{60}\{F_1\}$	575.6	81.16	601.1	72.35	709.2	128.2	608.0	94.77
$P_{60}\{F_2\}$	1939	151.8	1444	119.0	1940	181.4	2317	225.3
$P_{70}\{F_1\}$	568.8	73.32	606.8	64.89	719.0	128.8	611.9	91.47
$P_{70}\{F_2\}$	1949	160.0	1402	115.5	1894	166.4	2297	230.3
$P_{80}\{F_1\}$	559.6	67.42	607.0	65.90	719.0	128.2	614.0	94.30
$P_{80}\{F_2\}$	1955	149.7	1363	113.8	1856	157.4	2275	231.0
$P_{90}\{F_1\}$	547.3	66.03	604.1	60.52	711.3	129.2	608.4	88.97
$P_{90}\{F_2\}$	1981	155.5	1316	112.3	1832	162.9	2249	247.4
$P_{100}\{F_1\}$	529.6	75.02	597.8	63.09	689.1	143.2	593.7	85.51
$P_{100}\{F_2\}$	1997	186.5	1273	132.1	1809	161.2	2226	277.0
sec.	0.222	0.056	0.213	0.062	0.300	0.053	0.287	0.053

The Weak Vowels – See § 9.3 on page 394

	lettER N=41		commA N=42		happY N=85		Non-final N=48	
	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ
$P_0\{F_1\}$	740.9	87.39	741.2	126.1	396.4	41.90	558.6	67.07
$P_0\{F_2\}$	1635	148.8	1742	172.2	2755	191.5	1795	147.3
$P_{10}\{F_1\}$	754.5	90.88	751.0	133.8	393.1	41.80	566.0	67.46
$P_{10}\{F_2\}$	1643	144.5	1733	156.2	2796	175.9	1802	146.9
$P_{20}\{F_1\}$	761.6	97.56	757.3	136.5	391.2	45.63	566.4	67.78
$P_{20}\{F_2\}$	1652	134.6	1725	141.3	2808	164.4	1806	140.6
$P_{30}\{F_1\}$	767.7	99.62	760.4	139.8	390.9	46.38	565.3	67.79
$P_{30}\{F_2\}$	1655	127.5	1719	132.5	2824	164.2	1817	136.6
$P_{40}\{F_1\}$	775.0	100.1	761.6	136.9	388.4	48.76	562.7	67.67
$P_{40}\{F_2\}$	1658	122.5	1717	123.5	2835	162.5	1830	141.1
$\boldsymbol{P_{50}\{F_1\}}$	**778.5**	**102.0**	**763.0**	**135.5**	**390.7**	**54.10**	**560.8**	**70.11**
$\boldsymbol{P_{50}\{F_2\}}$	**1666**	**116.2**	**1714**	**116.1**	**2832**	**158.8**	**1837**	**133.7**
$P_{60}\{F_1\}$	775.3	112.0	764.0	143.0	392.2	56.50	557.2	72.99
$P_{60}\{F_2\}$	1667	120.2	1716	124.4	2837	165.6	1848	134.1
$P_{70}\{F_1\}$	767.9	117.8	759.5	148.6	390.3	63.21	552.2	75.09
$P_{70}\{F_2\}$	1671	118.6	1706	119.1	2832	167.5	1862	130.9
$P_{80}\{F_1\}$	766.1	127.5	759.1	153.7	392.2	67.62	549.2	83.98
$P_{80}\{F_2\}$	1673	121.0	1710	119.0	2827	180.0	1873	139.7
$P_{90}\{F_1\}$	755.7	132.0	752.8	150.0	388.4	74.45	540.4	80.97
$P_{90}\{F_2\}$	1679	114.4	1708	105.4	2808	218.7	1874	143.8
$P_{100}\{F_1\}$	747.3	132.8	742.7	153.3	382.1	70.39	525.9	80.77
$P_{100}\{F_2\}$	1680	112.1	1712	113.0	2770	238.9	1879	145.2
sec.	0.159	0.046	0.158	0.044	0.181	0.048	0.129	0.051

Bibliography

Adank, P. (2003), Vowel Normalization: A Perceptual-Acoustic Study of Dutch Vowels, PhD thesis, University of Nijmegen.

Andruski, J. and Nearey, T. (1992), 'On the sufficiency of compound target specification of isolated vowels and vowels in /bVb/ syllables.', *Journal of the Acoustical Society of America* **91**(1), 390–410.

Assmann, P., Nearey, T. and Hogan, J. (1982), 'Vowel identification: Orthographic, perceptual, and acoustic aspects', *Journal of the Acoustical Society of America* **71**(4), 975–989.

Bailey, R. (1984), 'Notes on South African English', *South African Journal of Linguistics* **2/3**, 1–45.

Barnes, J. (2006), *Strength and weakness at the interface: positional neutralization in phonetics and phonology*, Mouton de Gruyter, Berlin and New York.

Bauer, L. (2008), 'Lenition revisited', *Journal of Linguistics* **44**, 605–24.

Bauer, L. and Bauer, W. (2002), 'Can we watch regional dialects developing in colonial English?', *English world-wide* **23**(2), 169–193.

Bekker, I. (2007), 'Fronted /s/ in General White South African English', *Language Matters* **38**(1), 46–74.

Bekker, I. and Eley, G. (2007), 'An acoustic analysis of White South African English (WSAfE) monophthongs', *Southern African Linguistics and Applied Language Studies* **25**(1), 107–114.

Blevins, J. (2006), ‘New perspectives on English sound patterns: “natural” and “unnatural” in Evolutionary Phonology’, *Journal of English Linguistics* **34**(1), 6–25.

Boberg, C. (2005), ‘The Canadian shift in Montreal’, *Language variation and change* **17**, 133–154.

Boersma, P. and Weeknink, D. (2007), ‘Praat: Doing phonetics by computer (version 4.6.34)’. Computer Program: Retrieved October 26, 2007, from http://www.praat.org.

Bowerman, S. (2004), White South African English: Phonology, *in* E. Schneider, K. Burridge, B. Kortmann, R. Mesthrie and C. Upton, eds, ‘A Handbook of Varieties of English’, Mouton de Gruyter, Berlin, pp. 931–942.

Branford, W. (1994), English in South Africa, *in* R. Burchfield, ed., ‘The Cambridge History of the English Language, Vol. V: English in Britain and Overseas: Origins and Development’, Cambridge University Press, Cambridge, pp. 430–496.

Breckwoldt, G. (1961), ‘Some aspects of the phonetics of vowels in South African English monophthongs’, *Maître Phonetique* **115**, 5–12.

Britain, D. (2008), ‘When is change not a change? a case study on the dialect origins of New Zealand English’, *Language Variation and Change* **20**(2), 187–223.

Bucholtz, M. and Hall, K. (2008), ‘All of the above: New coalitions in sociocultural linguistics’, *Journal of Sociolinguistics* **12**(4), 401–31.

Burg, J. (1978), A new analysis technique for time series data, *in* D. Childers, ed., ‘Modern Spectrum Analysis’, IEEE Press, New York, pp. 42–8. [Reprint of a paper presented at the NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics, Aug. 12–23, 1968].

Butcher, A. (n.d.), Formant frequencies of /hVd/ vowels in the speech of South Australian females. Paper presented at the Eleventh Australasian International Conference on Speech Science and Technology, University of Auckland, Auckland, New Zealand, December 6 – 8, 2006.

Cameron, D. (2003), ‘Gender issues in language change’, *Annual review of applied linguistics* **23**, 187–201.

Catford, J. (1981), Observations on the recent history of vowel classification, *in* R. Asher and E. Henderson, eds, ‘Towards a History of Phonetics’, Edinburgh University Press, Edinburgh, pp. 19–32.

Chambers, J. (2003), *Sociolinguistic Theory: Linguistic Variation and its Social Significance*, second edn, Blackwell, Oxford.

Cho, T. and McQueen, J. (2008), ‘Not all sounds in assimilation environments are perceived equally: Evidence from Korean’, *Journal of Phonetics* **36**(2), 239–49.

Chomsky, N. and Halle, M. (1968), *The Sound Pattern of English*, Harper & Row, New York.

Cooper, F., Delattre, P., Liberman, A., Borst, J. and Gerstman, L. (1952), ‘Some experiments on the perception of synthetic speech sounds’, *Journal of the Acoustical Society of America* **24**(6), 597–606.

Cox, F. (2006), ‘The acoustic characteristics of /hVd/ vowels in the speech of some Australian teenagers.’, *Australian journal of linguistics* **26**(2), 147–79.

Cruttenden, A. (2001), *Gimson’s Pronunciation of English*, revised sixth edn, Arnold, London.

Da Silva, A. (2007), South African English: A sociolinguistic investigation of an emerging variety, PhD thesis, University of Witwatersrand.

Davenport, T. and Saunders, C. (2000), *South Africa: A Modern History*, 5th edn, Macmillan, Basingstoke, Hampshire.

De Klerk, V. (1996), Introduction, *in* V. De Klerk, ed., ‘Focus on South Africa’, John Benjamins, Amsterdam, pp. 7–17.

De Klerk, V., Adendorff, R., de Vos, M., Hunt, S., Niesler, T., Simango, S. and Todd, L. (2006), ‘Educated mother-tongue South African English: A corpus approach’, *Language Matters* **37**(1), 85–105.

Delattre, P., Liberman, A. and Cooper, F. (1955), ‘Acoustic loci and transitional cues for consonants’, *Journal of The Acoustical Society of America* **27**(4), 769–773.

Deterding, D. (1990), Speaker Normalization for Automatic Speech Recognition, PhD thesis, University of Cambridge.

Deterding, D. (1997), ‘The formants of monophthong vowels in Standard Southern British English pronunciation’, *Journal of the International Phonetic Association* **27**, 47–55.

Deterding, D. (n.d.). Personal Communincation: 23 June 2007.

Deumert, A. (1997), Multivariate clustering techniques in the analysis of linguistic variation, *in* E. Kotzé, ed., 'SA Linguistics 1997: Papers presented at the 33rd Annual Congress of the Linguistic Society of Southern Africa (LSSA) held at the University of Natal (Durban) 2–4 July 1997', LSSA, pp. 159–80.

Di Benedetto, M. (1989), 'Vowel representation: Some observations on temporal and spectral properties of the first formant frequency', *Journal of the Acoustical Society of America* **86**(1), 55–66.

Docherty, G. and Foulkes, P. (1999), Derby and Newcastle: Instrumental phonetics and variationist studies, *in* P. Foulkes and G. Docherty, eds, 'Urban Voices: Accent Studies in the British Isles', Arnold, London, pp. 47–71.

Dodsworth, R. (2008), 'Sociological consciousness as a component of linguistic variation', *Journal of Sociolinguistics* **12**(1), 5–33.

Dubois, S. and Horvath, B. (2000), 'When the music changes, you change too: Gender and language change in Cajun English.', *Language variation and change* **11**, 287–313.

Easton, A. and Bauer, L. (2000), 'An acoustic study of the vowels of New Zealand English', *Australian journal of linguistics* **20**(2), 93–117.

Eckert, P. (1989), 'The whole woman: Sex and gender differences in variation', *Language variation and change* **1**, 245–267.

Escudero, P. and Bion, R. (2007), Modeling vowel normalization and sound perception as sequential processes, *in* 'Proceedings of the 16th International Congress of Phonetic Sciences', Saarbrücken, pp. 1413–1416.

Evans, Z. and Watson, C. (2004), An acoustic comparison of Australian and New Zealand English vowel change, *in* 'Proceedings of the 10th Australian International Conference on Speech Science and Technology, Macquarie University, Sydney, December 8-10, 2004', pp. 195–200.

Fabricius, A. (2002), 'Weak vowels in modern RP: An acoustic study of happy-tensing and KIT/schwa shift.', *Language variation and change* **14**, 211–237.

Fabricius, A. (2007*a*), 'Variation and change in the TRAP and STRUT vowels of RP: A real time comparison of five acoustic data sets', *Journal of the International Phonetic Association* **37**(3), 293–320.

Fabricius, A. (2007*b*), Vowel formants and angle measurements in diachronic sociophonetic studies: *FOOT*-fronting in RP. Proceedings of ICPhS XVI, Saarbrüken, 6–10 August, pgs 1477–1480.

Fant, G. (1960), *Acoustic Theory of Speech Production*, Mouton, The Hague.

Fant, G. (1970), Analysis and synthesis of speech processes, *in* B. Malmberg, ed., 'Manual of Phonetics', North–Holland, Amsterdam, pp. 173–277.

Flemming, E. and Johnson, S. (2007), '*Rosa's roses*: reduced vowels in American English', *Journal of the International Phonetic Association* **37**(1), 83–96.

Foulkes, P. (2005), Sociophonetics, *in* K. Brown, ed., 'Encyclopedia of Language and Linguistics', 2nd edn, Elsevier Press, Amsterdam, pp. 495–500.

Foulkes, P. and Docherty, G. (1999), *Urban Voices* – overview, *in* P. Foulkes and G. Docherty, eds, 'Urban Voices: Accent Studies in the British Isles', Arnold, London, pp. 1–24.

Foulkes, P. and Docherty, G. (2006), 'The social life of phonetics and phonology', *Journal of Phonetics* **34**, 409–38.

Fridland, V. (2008), 'Regional differences in perceiving vowel tokens on Southerness, education, and pleasantness ratings', *Language Variation and Change* pp. 67–83.

Fujimura, O. and Erickson, D. (1997), Acoustic phonetics, *in* W. Hardcastle and J. Laver, eds, 'The Handbook of Phonetic Sciences', Blackwell, Oxford, pp. 65–115.

Gerstman, L. (1968), 'Classification of self-normalized vowels', *IEEE Transcations on Audio and Electroacoustics* **AU–16**, 78–80.

Giudici, R. (n.d.), An acoustic analysis of /l/ in general south african english. Honours Project in the English Language and Linguistics Department at Rhodes University.

Gordon, E., Campbell, L., Hay, J., Maclagan, M., Sudbury, A. and Trudgill, P. (2004), *New Zealand English: Its Origins and Evolution*, Cambridge University Press, Cambridge.

Gordon, M. (2006), 'Interview with william labov', *Journal of English linguistics* **34**(4), 332–351.

Gussenhoven, C. (2004), *The Phonology of Tone and Intonation*, Cambridge University Press, Cambridge.

Halle, M. and Stevens, K. (1964), Speech recognition: A model and a program for research, *in* J. Fodor and J. Katz, eds, 'The Structure of Language: Readings in the Philosophy of Language', Prentice-Hall, Englewood Cliffs, New Jersey, pp. 604–12.

Harrington, J. and Cassidy, S. (1994), 'Dynamic and target theories of vowel classification: Evidence from monophthongs and diphthongs in Australian English', *Language and Speech* **37**(4), 357–73.

Harrington, J. and Cassidy, S. (1999), *Techniques in Speech Acoustics*, Kluwer Academic, Dordrecht.

Harrington, J., Cox, F. and Evans, Z. (1997), 'An acoustic phonetic study of Broad, General, and Cultivated Australian English vowels', *Australian Journal of Linguistics* **17**, 155–84.

Harrington, J., Kleber, F. and Reubold, U. (2008), 'Compensation for coarticulation: /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study', *Journal of the Acoustical Society of America* **123**(5), 2825–35.

Harris, Z. (1964), Distributional structure, *in* J. Fodor and J. Katz, eds, 'The Structure of Language: Readings in the Philosophy of Language', Prentice-Hall, Englewood Cliffs, New Jersey, pp. 33–49. [First published as Z. Harris, "Distributional Structure," *Word*, 10, No. 2–3 (August-December 1954), 146–62.].

Hartmann, D. and Zerbian, S. (n.d.), Rhoticity in Black South African English. To be published in *Southern African Linguistics and Applied Language Studies*.

Hattersley, A. (1950), *The British Settlement of Natal : A Study in Imperial Migration*, Cambridge University Press, Cambridge.

Hawkins, S. and Midgley, J. (2005), 'Formant frequencies of RP monophthongs in four age groups of speakers', *Journal of the International Phonetic Association* **35**(2), 183–199.

Hindle, D. (1978), Approches to vowel normalization in the study of natural speech, *in* D. Sankoff, ed., 'Linguistic Variation: Models and Methods', Academic Press, New York, pp. 161–171.

Hooper, A. (1945), 'A preliminary report of an investigation of spoken English in South Africa', *South African Journal of Science* **41**, 476–84.

Hooper, A. (1951), English in South Africa, *in* E. Partridge and J. Clark, eds, 'British and American English Since 1900', Andrew Dakers, London, pp. 80–4.

Hooper, A. (1952), ‘The deterioration of English’, *Trek* **16**(4), 6–7. One of several contributors to the piece.

Hopwood, D. (1928), *South African English Pronunciation*, Juta, Cape Town.

Hopwood, D. (1952), ‘The deterioration of English’, *Trek* **16**(4), 7–8. One of several contributers to the piece.

Hopwood, D. (1961), ‘South African English pronunciation’, *English Studies in Africa* **4**(1), 68–74.

Horvath, B. (1985), *Variation in Australian English: The Sociolects of Sydney*, Cambridge University Press, Cambridge.

International Phonetic Association (1999), *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*, Cambridge University Press, Cambridge.

Jeffery, C. (1982), ‘Review of *The Standard in South African English and its social history (Varieties of English around the world: General series volume 1)*. By L.W. Lanham and C.A. Macdonald. Heidelberg: Julius Groot Verlag, 1979’, *Folia linguistica historica* **3**(2), 251–63.

Johnson, K. (1997), *Acoustic and Auditory Phonetics*, first edn, Blackwell, Cambridge, Mass.

Johnstone, B. and Kiesling, S. (2008), ‘Indexicality and experience: Exploring the meanings of /aw/-monophthongization in Pittsburgh’, *Journal of Sociolinguistics* **12**(1), 5–33.

Jones, C. (2006), *English Pronunciation in the Eighteenth and Nineteenth Centuries*, Palgrave Macmillan, Basingstoke, Hampshire.

Jones, D. (1924), *An English Pronouncing Dictionary*, J.M. Dent, New York. 1928 reprinting with revised introduction.

Jones, D. (1934), *An Outline of English Phonetics*, fourth edn, W. Heffer and Sons., Cambridge.

Jongman, A., Wayland, R. and Wong, S. (2000), ‘Acoustic characteristics of English fricatives’, *Journal of the Acoustical Society of America* **108**(3), 1252–1263.

Joos, M. (1948), ‘Acoustic phonetics’, *Language* **24**(2), 5–136. Language Monograph no. 23.

Kaplan, M. and Robertson, M. (1991), Introduction, *in* M. Kaplan and M. Robertson, eds, ‘Founders and Followers : Johannesburg Jewry, 1887-1915’, Vlaeberg, Cape Town, pp. 10–16.

Kerswill, P. (2006), Socio-economic class, *in* C. Llamas and P. Stockwell, eds, ‘The Routledge Companion to Sociolinguistics’, Routledge, London, pp. 51–61.

Kewley-Port, D. (1983), ‘Time-varying features as correlates of place of articulation in stop consonants’, *Journal of the Acoustical Society of America* **73**(1), 322–335.

Kierkegaard, S. (1959), *Either/Or*, Vol. 1, Anchor Books, Garden City, New York. Translated into English by David F. Swenson, originally published in Danish in 1843.

Koenig, W. (1949), ‘A new frequency scale for acoustic measurements’, *Bell Laboratories Record* **27**, 299–301.

Labov, W. (1963), ‘The social motivation of a sound change’, *Word* **19**, 273–309.

Labov, W. (1966), *The Social Stratification of English in New York City*, 1st edn, Center for Applied Linguistics, Washington D.C.

Labov, W. (1972), *Sociolinguistic Patterns*, University of Pennsylvania Press, Philadelphia.

Labov, W. (1980), The social origins of sound change, *in* W. Labov, ed., ‘Locating Language in Time and Space’, Academic Press, New York, pp. 251–266.

Labov, W. (1990), ‘The intersection of sex and social class in the course of linguistic change’, *Language variation and change* **2**(2), 205–254.

Labov, W. (1991), The three dialects of English, *in* P. Eckert, ed., ‘New Ways of Analyzing Sound Change’, Academic Press, San Diego, pp. 1–44.

Labov, W. (1994), *Principles of Linguistic Change*, Vol. 1: internal factors, Blackwell, Oxford.

Labov, W. (2003), Some sociolinguistic principles, *in* C. Paulston and G. Tucker, eds, ‘Sociolinguistics: The Essential Readings’, Blackwell, Malden, MA., pp. 234–50.

Labov, W. (2007), ‘Transmission and diffusion’, *Language* **83**(2), 344–387.

Labov, W., Ash, S. and Boberg, C. (2006), *The Atlas of North American English: Phonetics, Phonology, and Sound Change: A Multimedia Reference Tool*, Mouton de Gruyter, Berlin and New York.

Labov, W. and Baranowski, M. (2006), '50 msec', *Language variation and change* **18**(3), 223–40.

Ladd, D. (1980), *The Structure of Intonational Meaning: Evidence from English*, Indiana University Press, Bloomington and London.

Ladefoged, P. (1967), *Three Areas of Experimental Phonetics*, Oxford University Press, London.

Ladefoged, P. (1996), *Elements of Acoustic Phonetics*, second edn, University of Chicago Press, Chicago.

Ladefoged, P. (2001), *Vowels and Consonants: An Introduction to the Sounds of Languages*, Blackwell, Malden, Mass.

Ladefoged, P. (2006), *A Course in Phonetics*, fifth edn, Thomson Wadsworth, Boston, MA.

Ladefoged, P. (2007), 'Articulatory features for describing lexical distinctions', *Language* **83**(1), 161–80.

Lahiri, A., Gewirth, L. and Blumstein, S. (1984), 'A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study', *Journal of the Acoustical Society of America* pp. 391–404.

Langstrof, C. (2006), 'Acoustic evidence for a push-chain shift in the Intermediate Period of New Zealand English', *Language variation and change* **18**, 141–164.

Lanham, L. (1964?), *English in South Africa: Its History, Nature and Social Role*, The Institute for the Study of Man in Africa, Johannesburg.

Lanham, L. (1965), 'South African English Pronunciation 2065 A.D.', *Taalfasette* **1**, 89–95.

Lanham, L. (1967), *The Pronunciation of South African English: A Phonetic-Phonemic Introduction*, A.A. Balkema, Cape Town.

Lanham, L. (1978), South African English, *in* L. Lanham and K. Prinsloo, eds, 'Language and Communication Studies in South Africa: Current Issues and Directions in Research', Oxford University Press, Cape Town, pp. 138–165.

Lanham, L. (1982), English in South Africa, *in* R. Bailey and M. Görlach, eds, 'English as a world language', Cambridge University Press, Cambridge, pp. 324–52.

Lanham, L. and Macdonald, C. (1979), *The Standard in South African English and its Social History*, Julius Groos Verlag, Heidelberg.

Lanham, L. and Traill, A. (1962), *South African English Pronunciation*, Witwatersrand University Press, Johannesburg. Reprinted from *English Studies in Africa* 5(2): 171–206, 1962.

Lass, R. (1980), *On Explaining Language Change*, Cambridge University Press, Cambridge.

Lass, R. (1981), Undigested history and synchronic 'structure', *in* D. Goyvaerts, ed., 'Phonology in the 1980's', E.Story–Scientia, Ghent, Belgium, pp. 525–44.

Lass, R. (1984), *Phonology: An Introduction to Basic Concepts*, Cambridge University Press, Cambridge.

Lass, R. (1986), 'On schwa', *Stellenbosch papers in Linguistics* **15**, 1–30.

Lass, R. (1990), A 'standard' South African vowel system, *in* S. Ramsaran, ed., 'Studies in the pronunciation of English: a commerative volume in honour of A.C. Gimson', Routledge, London and New York, pp. 272–285.

Lass, R. (1995), South African English, *in* R. Mesthrie, ed., 'Language and Social History: Studies in South African Sociolinguistics', David Philip, Cape Town, pp. 89–106.

Lass, R. (2002), South African English, *in* R. Mesthrie, ed., 'Language in South Africa', Cambridge University Press, Cambridge, pp. 104–126.

Lass, R. (2004), South African English, *in* R. Hickey, ed., 'Legacies of Colonial English: Studies in Transported Dialects', Cambridge University Press, Cambridge, pp. 363–386.

Lass, R. and Wright, S. (1985), The South African chain shift: Order out of chaos?, *in* R. Eaton, O. Fischer, W. Koopman and F. Van der Leek, eds, 'Papers from the 4th International Conference on English Historical Linguistics', John Benjamins, Amsterdam, pp. 137–161.

Lass, R. and Wright, S. (1986), 'Endogeny vs. contact: 'Afrikaans influence' on South African English', *English world-wide* **7**(2), 201–223.

Laver, J. (1994), *Principles of Phonetics*, Cambridge University Press, Cambridge.

Le Roux, T. and De Villiers Pienaar, P. (1927), *Afrikaanse Fonetiek*, Juta, Cape Town and Johannesburg.

Lehiste, I. and Peterson, G. (1961), 'Transitions, glides and diphthongs', *Journal of the Acoustical Society of America* **33**(3), 268–77.

Lennes, M. (2002), 'label_from_text_file.praat (latest version: 25.01.2002)'. Computer programe: Retrieved 26 October 2007, from http://www.helsinki.fi/ lennes/praat-scripts/.

Lennes, M. (2006*a*), 'mark_pauses.praat (latest version: 23.01.2006)'. Computer program: Retrieved 26 October, 2007 from http://www.helsinki.fi/ lennes/praat-scripts/.

Lennes, M. (2006*b*), 'Mietta's Praat scripts (lastest update: 11.10.2006)'. Website: Accessed on 26 October 2007 at http://www.helsinki.fi/ lennes/praat-scripts/.

Leyds, G. (1964), *A History of Johannesburg : The Early Years*, Nasionale Boekhandel, Cape Town.

Lieberman, P. and Blumstein, S. (1988), *Speech Physiology, Speech Perception, and Acoustic Phonetics*, Cambridge University Press, Cambridge.

Limbrick, V. (1989), 'Palatalization and affrication in South African English – a sample study', *English Usage in South Africa* **20**(2), 20–38.

Lobanov, B. (1971), 'Classification of Russian vowels spoken by different speakers', *Journal of the Acoustical Society of America* **49**, 606–8.

Mair, C. (2006), *Twentieth-Century English: History, Variation and Standardization*, Cambridge University Press, Cambridge.

Marshall, C. (n.d.), An acoustic phonetic analysis of White Zimbabwean English (WZimE). English Language and Linguistics Department, Rhodes University, Grahamstown, South Africa.

Mesthrie, R. (1993), 'English in South Africa', *English Today* **9**(1), 27–33.

Mesthrie, R. (2008), "death of the mother tongue' – is English a glottophagic language in South Africa?', *English Today* **24**(2), 13–9.

Mesthrie, R., Swann, J., Deumert, A. and Leap, W. (2000), *Introducing Sociolinguistics*, Edinburgh University Press, Edinburgh.

Miller, J. (1989), 'Auditory-perceptual interpretation of the vowel', *Journal of the Acoustical Society of America* **85**(5), 2114–2134.

Mitchell, A. and Delbridge, A. (1965), *The Pronunciation of English in Australia*, Angus and Robertson, Sydney.

Moore, B. and Glasberg, B. (1983), 'Suggested formulae for calculating auditory-filter bandwidths and excitation patterns', *Journal of the Acoustical Society of America* pp. 750–3.

Mufwene, S. (1991), Pidgins, creoles, typology, and markedness, *in* F. Byrne and T. Huebner, eds, 'Development and Structures of Creole Languages: essays in honour of Derek Bickerton', Benjamin, Amsterdam, pp. 123–43.

Mufwene, S. (2008), 'Colonization, population contacts, and the emergence of new language varieties: A response to Peter Trudgill', *Language in Society* **37**, 254–8.

Nearey, T. (1977), Phonetic Feature Systems for Vowels, PhD thesis, University of Connecticut, Storrs, CT.

Nearey, T. (1989), 'Static, dynamic, and relational properties in vowel perception', *Journal of the Acoustical Society of America* **85**(5), 2088–2113.

Nearey, T. (1997), 'Speech perception as pattern recognition', *Journal of the Acoustical Society of America* **101**(6), 3241–54.

Nearey, T. and Assmann, P. (1986), 'Modeling the role of inherent spectral change in vowel identification', *Journal of the Acoustical Society of America* **80**(5), 1297–308.

Nguyen, N. and Fagyal, Z. (2008), 'Acoustic aspects of vowel harmony in French', *Journal of Phonetics* **36**(1), 1–27.

Norström, P. and Lindblom, B. (1975), A normalization procedure for vowel formant data, *in* 'Proceedings of the 8th International Congress of Phonetic Sciences', Leeds. Paper 212.

Odden, D. (2005), *Introducing Phonology*, Cambridge University Press, Cambridge.

Oh, E. (2008), 'Coarticulation in non-native speakers of English and French: An acoustic study', *Journal of Phonetics* **36**(2), 361–84.

Orpen, J. (n.d.), The second FORCE merger in South African and Zimbabwean English. English Language and Linguistics Department, Rhodes University, South Africa.

Palethorpe, S. and Cox, F. (2003), Vowel modification in pre-lateral environments. Poster presented at the International Seminar on Speech Production, December, Sydney.

Peterson, G. and Barney, H. (1952), ‘Control methods used in the study of the vowels’, *Journal of the Acoustical Society of America* pp. 175–184.

Peterson, G. and Lehiste, I. (1960), ‘Duration of syllable nuclei in English’, *Journal of the Acoustical Society of America* **32**(6), 693–703.

Pisoni, D. and Luce, P. (1986), Speech perception: Research, theory, and the principal issues, *in* E. Schwab and H. Nusbaum, eds, ‘Pattern Recognition by Humans and Machines’, Academic Press, New York, pp. 1–50.

Potter, R., Kopp, G. and Kopp, H. (1966), *Visible Speech*, second edn, Dover, New York. First published 1947.

Pullum, G. and Ladusaw, W. (1996), *Phonetic Symbol Guide*, The University of Chicago Press, Chicago and London.

Rosewarne, D. (1984), ‘Estuary English’, *Times educational supplement* p. 29. 19 October 1984.

Rosewarne, D. (1994), ‘Estuary English: Tommorrow’s RP?’, *English today* **10**(1), 3–8.

Rosner, B. and Pickering, J. (1994), *Vowel Perception and Production*, Oxford University Press, Oxford.

Russell, B. (1928), *Sceptical Essays*, George Allen & Unwin, London.

Schmitt, H. (2007), ‘The case for the epsilon symbol (ε) in RP DRESS’, *Journal of the International Phonetic Association* **37**(3), 321–7.

Schneider, E. (2003), ‘The dynamics of New Englishes: From identity construction to dialect birth.’, *Language* **79**(2), 233–281.

Schneider, E. (2007), *Postcolonial Englishes: Varieties Around the World*, Cambridge University Press, Cambridge.

Schneider, E. (2008), ‘Accommodation versus identity? a response to trudgill’, *Language in Society* **37**(2), 262–67.

Statistics South Africa (2006), ‘South African Statistics, 2006’, Pretoria, Statistics South Africa. Available at http://www.statssa.gov.za.

Stevens, K. (1997), Articulatory–acoustic–auditory relationships, *in* W. Hardcastle and J. Laver, eds, ‘The Handbook of Phonetic Sciences’, Blackwell, Oxford, pp. 462–506.

Stevens, K. (1998), *Acoustic Phonetics*, MIT Press, Cambridge, Massachusetts.

Stevens, K. and Blumstein, S. (1981), The search for invariant acoustic correlates of phonetic features, *in* P. Eimas and J. Miller, eds, 'Perspectives on the Study of Speech', Erlbaum, Hillsdale, NJ., pp. 1–38.

Stevens, S. and Volkmann, J. (1940), 'The relation of pitch to frequency: A revised scale', *American journal of psychology* **53**(3), 329–53.

Story, B. (2008), 'Comparison of magnetic resonance imaging-based vocal tract area functions obtained from the same speaker in 1994 and 2002', *Journal of the Acoustical Society of America* **123**(1), 327–335.

Strange, W., Jenkins, J. and Johnson, T. (1983), 'Dynamic specification of coarticulated vowels', *Journal of the Acoustical Society of America* **74**(3), 695–705.

Sussman, H., Fruchter, D. and Cable, A. (1995), 'Locus equations derived from compensatory articulation', *Journal of the Acoustical Society of America* **97**(5), 3112–24.

Sussman, H., McCaffrey, H. and Matthews, S. (1991), 'An investigation of locus equations as a source of relational invariance for stop place categorization', *Journal of the Acoustical Society of America* **90**(3), 1309–1325.

Syrdal, A. and Gopal, H. (1986), 'A perceptual model of vowel recognition based on the auditory representation of american english vowels', *Journal of the Acoustical Society of America* **79**(4), 1086–100.

Taylor, J. (1991), 'Remarks on the KIN-PIN vowels in South African English', *English world-wide* **12**(1), 75–85.

Thomas, E. and Kendall, T. (2007), 'NORM: The vowel normalization and plotting suite'. Online resource at http://ncslaap.lib.ncsu.edu/tools/norm/. Accessed on 11 June 2008.

Todd, L. (n.d.), Personal communication. 16 October 2007.

Torgersen, E. and Kerswill, P. (2004), 'Internal and external motivations in phonetic change: Dialect leveling outcomes for an English vowel shift.', *Journal of sociolinguistics* **8**(1), 23–53.

Trudgill, P. (1974), *The Social Differentiation of English in Norwich*, Cambridge University Press, Cambridge.

Trudgill, P. (1986), *Dialects in Contact*, Blackwell, Oxford.

Trudgill, P. (1999*a*), Norwich: Endogenous and exogenous linguistic change, *in* P. Foulkes and G. Docherty, eds, 'Urban Voices: Accent Studies in the British Isles', Arnold, London, pp. 124–40.

Trudgill, P. (1999*b*), 'A window on the past: "colonial lag" and New Zealand evidence for the phonology of nineteenth-century English', *American Speech* **74**(3), 227–239.

Trudgill, P. (2004), *New-Dialect Formation: The Inevitability of Colonial Englishes*, Edinburgh University Press, Edinburgh.

Trudgill, P. (2008), 'Colonial dialect contact in the history of European languages: On the irrelevance of identity to new-dialect formation', *Language in Society* **37**, 241–54.

Van Onselen, C. (1982), *Studies in the Social and Economic History of the Witwatersrand 1886–1914. 1: New Babylon.*, Raven Press, Johannesburg.

van Rooy, A. (1997), The generative paradigm in phonology: is it worth the while?, *in* E. Kotzé, ed., 'SA Linguistics 1997: Papers presented at the 33rd Annual Congress of the Linguistic Society of Southern Africa (LSSA) held at the University of Natal (Durban) 2–4 July 1997', LSSA, pp. 376–89.

Van Rooy, B. and Pienaar, M. (2006), 'Trends in recent South African linguistic research', *South African Linguistics and Applied Language Studies* **24**(2), 191–216.

Verbrugge, R. and Rakerd, B. (1986), 'Evidence of talker-independent information for vowels', *Language and Speech* **29**(1), 39–57.

Verbrugge, R., Strange, W., Shankweiler, D. and Edman, T. (1976), 'What information enables a listener to map a talker's vowel space?', *Journal of the Acoustical Society of America* **60**(1), 198–212.

Watson, C., Harrington, J. and Evans, Z. (1998), 'An acoustic comparison between New Zealand and Australian English vowels', *Australian journal of linguistics* **18**(2), 185–207.

Watson, C., Harrington, J. and Palethorpe, S. (1998), A kinematic analysis of New Zealand and Australian English vowel spaces, *in* 'Proceedings of the Fifth International Conference on Spoken Language Processing', pp. 2363–2366.

Watson, C., Maclagan, M. and Harrington, J. (2000), ‘Acoustic evidence for vowel change in New Zealand English’, *Language variation and change* **12**, 51–68.

Watt, D. and Fabricus, A. (2003), ‘Evaluation of a technique for improving the mapping of multiple speakers’ vowel spaces in the F1–F2 plane.’, *Leeds Working Papers in Linguistics and Phonetics* **9**, 159–173.

Watt, D. and Milroy, L. (1999), Patterns of variation and change in three Newcastle vowels: Is this dialect levelling?, *in* P. Foulkes and G. Docherty, eds, ‘Urban Voices: Accent Studies in the British Isles’, Arnold, London, pp. 25–46.

Watt, D. and Tillotson, J. (2001), ‘A spectrographic analysis of vowel fronting in Bradford English’, *English World-Wide* **22**(2), 269–302.

Watts, H. (1976), A social and demographic portrait of English-speaking White South Africans, *in* A. De Villiers, ed., ‘English-Speaking South African Today: Proceedings of the National Conference, July 1974’, Oxford University Press, Cape Town, pp. 41–89.

Webb, V. (1983), ‘On a South African English vowel system’, *Stellenbosch Papers in Linguistics* **10**, 134–164.

Webb, V. (2008). Personal communication.

Wells, J. (1982), *Accents of English*, Vol. 1–3, Cambridge University Press, Cambridge.

Welsh, F. (1998), *A History of South Africa*, HarperCollins, London.

Wissing, D. (2006), ‘Het jou mô en jou pô ’n strôndhuis by Hôrtenbos? Feit of fiksie?’, *Southern African linguistics and applied language studies* **24**(1), 87–100.

Wolfram, W., Carter, P. and Moriello, B. (2004), ‘Emerging Hispanic English: new dialect formation in the American South’, *Journal of Sociolinguistics* **8**(3), 339–58.

Yang, C. (2008), ‘The great number crunch: review of Rens Bod, Jennifer Hay & Stefanie Jannedy (eds.), *Probablistic linguistics*. Cambridge, MA.: MIT Press, 2003’, *Journal of Linguistics* **44**, 205–28.

Zhang, Q. (2008), ‘Rhotacization and the ‘Beijing Smooth Operator’: the social meaning of a linguistic variable’, *Journal of Sociolinguistics* **12**(2), 201–222.

Zwicker, E. (1961), ‘Subdivision of the audible frequency range into critical bands (frequenzgruppen)’, *Journal of the Acoustical Society of America* **33**, 248.