Adapting a pronunciation dictionary to Standard South African English for automatic speech recognition

Olga Meruzhanovna Martirosian

Adapting a pronunciation dictionary to Standard South African English for automatic speech recognition

by

O.M. Martirosian

Dissertation submitted in fulfilment of the requirements for the degree

Master of Engineering

at the

Potchefstroom Campus

of the

NORTH-WEST UNIVERSITY

Supervisor: Professor E. Barnard

Co-supervisor: Dr. M.H. Davel

May 2009

ABSTRACT

Die uitspraakwoordeboek is a belangrike inligtingsbron wat benodig word gedurende die ontwikkeling van 'n outomatiese spraakherkenningstelsel (ASR stelsel). In dié tesis pas ons 'n Britse Engels uitspraakwoordeboek aan na Standaard Suid-Afrikaanse Engels (SSAE) as 'n gevallestudie in dialekaanpassing. Ons ondersoek neem ons in drie verskillende rigtings: woordeboekkontrole, foneemoortolligheidsberekening en foneemaanpassing.

'n Uitspraakwoordeboek behoort getoets te word vir juisheid voordat dit in eksperimente of ander toepassings gebruik word. Die proses om 'n mens in diens te neem om 'n volledige uitspraakwoordeboek te kontroleer is toegeeflik en kan nie altyd geakkommodeer word nie. In ons woordeboekkontrole navorsing probeer ons om die mannekrag te verminder wat benodig word om uitspraakwoordeboeke te kontroleer deur outomatiese en halfoutomatiese tegnieke te implementeer wat inskrywings wat moontlik foutief is te vind en isoleer. Ons identifiseer nuwe tegnieke wat foute doeltreffend identifiseer, en wend hulle dan aan op 'n publieke domein Britse Engels uitspraakwoordeboek.

Die ondersoek van foneemoortolligheid verg die oorweging van die moontlikheid dat nie alle foneemonderskydings nodig is in SSAE nie, sowel as die ondersoek van verskillende metodes om die onderskydings te ontleed. Die metodes wat ondersoek word sluit beide data- en kennisgedrewe uitspraak voorstelle in vir 'n uitspraakwoordeboek wat in outomatiese spraakherkenning gebruik word. Hierdie ondersoek gee 'n dieper taalkundige insig in die uitspraak van foneme in SSAE.

Laasteliks kyk ons na foneemaanpassing deur die KIT foneme tussen twee dialekte van Engels aan te pas deur twee stelle aanpassingsreëls te implementeer. Aanpassingsreëls word in die letterkunde bekom maar word ook formuleer na 'n ondersoek van die taalkundige waarnemings in die data. Ons voorspellings is 93% akkuraat, wat aansienlik hoër is as die 71% wat behaalbaar is deur die implementering van reëls wat voorheen geïdentifiseer is. Die aanpassing van 'n Britse uitspraakwoordeboek na SSAE was die finale stap in die ontwikkeling van 'n SSAE uitspraakwoordeboek, wat die doel van die tesis is. 'n Outomatiese spraakherkenningstelsel was ook ontwikkel met die woordeboek en dit het 'n ongedwonge foneemakkuraatheid van 79.7%.

ABSTRACT

The pronunciation dictionary is a key resource required during the development of an automatic speech recognition (ASR) system. In this thesis, we adapt a British English pronunciation dictionary to Standard South African English (SSAE), as a case study in dialect adaptation. Our investigation leads us in three different directions: dictionary verification, phoneme redundancy evaluation and phoneme adaptation.

A pronunciation dictionary should be verified for correctness before its implementation in experiments or applications. However, employing a human to verify a full pronunciation dictionary is an indulgent process which cannot always be accommodated. In our dictionary verification research we attempt to reduce the human effort required in the verification of a pronunciation dictionary by implementing automatic and semi-automatic techniques that find and isolate possible erroneous entries in the dictionary. We identify a number of new techniques that are very efficient in identifying errors, and apply them to a public domain British English pronunciation dictionary.

Investigating phoneme redundancy involves looking into the possibility that not all phoneme distinctions are required in SSAE, and investigating different methods of analysing these distinctions. The methods that are investigated include both data driven and knowledge based pronunciation suggestions for a pronunciation dictionary used in an automatic speech recognition (ASR) system. This investigation facilitates a deeper linguistic insight into the pronunciation of phonemes in SSAE.

Finally, we investigate phoneme adaptation by adapting the KIT phoneme between two dialects of English through the implementation of a set of adaptation rules. Adaptation rules are extracted from literature but also formulated through an investigation of the linguistic phenomena in the data. We achieve a 93% predictive accuracy, which is significantly higher than the 71 % achievable through the implementation of previously identified rules. The adaptation of a British pronunciation dictionary to SSAE represents the final step of developing a SSAE pronunciation dictionary, which is the aim of this thesis. In addition, an ASR system utilising the dictionary is developed, achieving an unconstrained phoneme accuracy of 79.7%.

Keywords: *Pronunciation dictionaries, pronunciation modelling, dictionary verification, KIT vowel, diphthong analysis, South African English, Standard South African English, dialect adaptation, BEEP pronunciation dictionary, CELEX pronunciation dictionary*

TABLE OF CONTENTS

List of Tables

Chapter One

Introduction

The development of an Automatic Speech Recognition (ASR) system for a dialect of a well-known language typically involves the re-use of existing language resources, such as phone sets and dictionaries from different dialects. As the existing dictionaries and phone sets may not quite match the acoustics of a specific dialect, dialect adaptation is required of either the phone set, the pronunciation dictionary or the acoustic data used to construct the system.

Dialect adaptation of an existing pronunciation dictionary requires that the source pronunciation dictionary be as free from errors as possible. The presence of errors in the dictionary can cause the results of experiments to be inaccurate. This is partly due do the scrutiny given to the dictionary during experimentation, errors being identified and the correction of these errors which is then included in positive results, and partly because errors do not behave in a predictable fashion and could thus alter the results of the experiment in an unpredictable manner. Errors reduce the ability of experimenters to analyse their results as their experiments are not implemented as planned.

Once a clean pronunciation dictionary is achieved dialect adaptation can be completed in either a knowledge based or a data-driven manner. Knowledge based methods require expert knowledge to exist on at least the dialect for which one would like to adapt the dictionary, but preferably also the dialect in which the dictionary exists, for the purposes of extracting what transformations are required in order to adapt from one dialect to the other. Data driven methods require data, which can be analysed in order to ascertain the transformations required.

The experiments described in the forthcoming chapters involve the steps taken towards the implementation of a Standard South African English (SSAE) pronunciation dictionary for the purposes of ASR. In this chapter the context of the experiment is described and the research problem that is being addressed is specified. Also, an overview is provided of the remainder of the thesis.

1.1 CONTEXT

Lack of access to information is an important issue in South Africa. There is a lack of access to the Internet and automation is becoming more and more necessary in order to disseminate information by telephone. A recent survey showed that only 7% of the country has access to the Internet at home (Statistics South Africa, 2007).

However, the same survey found that 73% of households have a mobile phone. Thus, if information is to be disseminated to the public, a telephone application would be the most efficient manner in which to do it. Also, many traditional societies have a strong oral culture and are thus more comfortable with voice user interfaces than with graphical or text user interfaces (Sharma *et al.*, 2009). Call centres provide information to those who require it, but high call volumes create necessity for the call centres to become increasingly automated in order to meet demand.

Automation includes the use of speech recognition and speech synthesis to support automatic information dissemination through a telephonic interface. Speech recognition allows the system to understand what the user wants by allowing them to speak to the system. Speech synthesis allows the system to give the user information using speech and thus negates the necessity for the user to be literate. This is important as 18% of South Africa is illiterate (United Nations Educational, Scientific and Cultural Organization, 2001).

Language is also an important part of information dissemination. South Africa has 11 official languages, some are spoken more than others, but each has its own population group in need of information. In order to allow information to reach the highest number of people, it is important to select a language that is most likely to be understood.

Standard South African English is a dialect of English that is spoken widely in South Africa. It is influenced by the 10 other official languages of South Africa. Today only 8.2% of South Africa's population speaks English as a home language (Heugh, 2007). But South African English (SAE) is one of the four languages that are most commonly used, together with isiZulu, isiXhosa and Afrikaans. This suggests that most of the people that are speaking SAE are not first language speakers, and thus exhibit pronunciation differences that are influenced by their home languages. These variants of English all influence Standard South African English, which is the English dialect characteristic of first language South African speakers (Bekker, 2009).

In order to build speech recognition or speech synthesis systems for SSAE, resources such as a pronunciation dictionary and speech data are required. The pronunciations of words for an ASR system are modelled in the pronunciation dictionary. A pronunciation dictionary guides the automatic speech recognition (ASR) system as it analyses speech data and creates acoustic models. Thus, a pronunciation dictionary forms the very basis of pronunciation modelling in an ASR system. However, due to the pervasiveness of British and American English, the pronunciation dictionaries that are available in English tend to model either British or American pronunciations. Thus in order to develop a pronunciation dictionary that is specialised for SSAE, one would need to be adapted for the dialect.

1.2 PROBLEM STATEMENT

The adaptation of a British pronunciation dictionary for SSAE requires the analysis of the original pronunciation dictionary as well as the dialect for which it is being adapted. This thesis investigates three main directions, namely, dictionary verification, acoustic analysis of SSAE diphthongs and the adaptation of the KIT vowel[1] for SSAE. The specific research questions being asked are described in more detail below:

1. **How can lexical analysis techniques be applied during dictionary verification?** The aim here is to remove all errors from a pronunciation dictionary prior to its use for acoustic analysis. Since dictionaries are typically large, automated or semi-automated techniques are of interest.

2. **Are all phonemic distinctions required from an ASR perspective?** Techniques to answer this general question are developed using diphthongs as a case study. The aim here is to analyse the necessity and

[1] The KIT vowel, or the 'short I' is part of Well's lexical sets for describing "The lexical incidence of vowels in all the many accents [of English]" Wells (1982)

acoustic properties of diphthongs in an SSAE ASR system. This process can provide significant linguistic insights into the acoustic properties of an SSAE ASR system, and through that into SSAE itself.

3. **Can the dialect of a pronunciation dictionary be adapted to another dialect through the implementation of a set of rules?** The aim here is to develop techniques that analyse the underlying relationship between British English (BE) and Standard South African English (SSAE) phonemes. As the main source of variation between BE and SSAE is caused by the KIT phoneme, this phoneme is selected as a case study. The main output of this experiment will be a pronunciation dictionary that reflects the SSAE pronunciation of the KIT phoneme.

1.3 OVERVIEW OF THESIS

The thesis is structured as follows:

- Firstly, in Chapter 2 a study of the literature that is relevant to this topic is provided. Decisions that need to be made are discussed and their advantages and disadvantages are pointed out and compared.
- Chapter 3 describes the processes that are followed for dictionary verification, how these processes are implemented and their ability to find and remove errors in a pronunciation dictionary.
- In Chapter 4, the baseline ASR system that is used for experimentation is described in detail.
- Chapter 5 explores the analysis of diphthongs in SSAE. The process of identifying and evaluating diphthong replacements is described. The results of this experiment are provided and discussed. Once the diphthong analysis is complete a data limitation experiment is performed to evaluate the original pronunciation dictionary against one in which there are no diphthongs.
- Chapter 6 describes the process followed for the purpose of adapting a British pronunciation dictionary to SSAE through the adaptation of the KIT phoneme. The linguistic views on the topic are discussed. Adaptation rules are developed, evaluated and finally applied to a British pronunciation dictionary, and the final output dictionary is analysed.
- Finally, Chapter 7 contains a summary and conclusion of the findings in this thesis. This chapter also describes the lessons learnt as well as possible future research directions.

CHAPTER TWO

LITERATURE STUDY

2.1 INTRODUCTION

In this chapter background is provided for the main topics investigated in this thesis through the exploration of related prior work:

- The first section describes **pronunciation**, introduces the concept of **pronunciation variations** that can occur therein, and explored the reasons for these variations. Variations in **non-native speech** are discussed in detail.
- **Standard South African English (SSAE)** is then discussed, the origin of this variety of English is briefly summarised and its pronunciations are explored.
- **Pronunciation modelling in automatic speech recognition (ASR) systems** describes the different representations that are used for pronunciation in ASR systems, in order to help the system analyse and model the speech signal.
- **Pronunciation variance modelling** describes the modelling of pronunciation variance in ASR systems, and what factors influence modelling on each level of the ASR.
- **Information sources** utilised during pronunciation modelling are discussed, including both the extraction and analysis of information about pronunciation, as well as into the verification of this information.
- Finally, **dictionary verification** is explored, describing how errors in pronunciation dictionaries are identified and removed.

2.2 PRONUNCIATION

Pronunciation describes the manner in which sounds or groups of sounds are realised in speech. In ASR systems, pronunciations are modelled using phonemes. Phonemes model semantically distinctive sounds in languages, however, the actual realisations of these sounds differ and are referred to as phones. Each phoneme can thus be realised as one of many phones. The variations that occur in pronunciations can thus be phonetic, which means that the correct phoneme is realised but one or more of the phones are different, meaning that the variation is minimal. The variations can also be phonemic, which implies a higher level of variation and a different phonemic representation.

2.2.1 PRONUNCIATION VARIANCE ORIGINS

The pronunciation of a particular phoneme is influenced by various factors. These include the anatomy of the speaker, whether they have speech impediments or disabilities, how they need to accommodate their listener, their accent, the dialect they are using, their mother tongue, the level of formality of their speech, the amount and importance of the information they are conveying (Jande, 2006), their environment (Lombard effect) and even their emotional state (Strik and Cucchiarini, 1999). Care must also be taken when an automatic method is used in the analysis of speech (as with ASR), as an automatic system does not perceive speech in the same way that a human does, and thus can infer variance through its speech modelling process.

2.2.2 PRONUNCIATION VARIANCE REALISATIONS

Wester *et al.* (1998) define pronunciation variance in speech as having two effects, namely, changes in the number and order of the phonemes in a pronunciation and changes in the pronunciations of those phonemes. However, the simplicity of the effect should not encourage underestimation of its influence on the recogniseability of the resulting pronunciation. The effects described can be devastating for an ASR system which has not been equipped with the tools to reverse the effects or at least reduce their influence.

2.2.3 NON-NATIVE PRONUNCIATION

Non-native speech generally refers to accented speech that does not sound like the native speech used in a specific geographical location. For ASR systems, which are usually designed with a specific nativity of speech in mind, non-native speech refers to any speech that the system was not specifically designed to recognise. The nativity of a person's speech describes the combination of the effects of their mother tongue, the dialect that they are speaking, their accent and their proficiency in the language that they are speaking on their pronunciation.

If an ASR system uses speech and a pronunciation dictionary associated with a certain nativity, non-native speech causes consistently poor system performance (Wang *et al.*, 2003; Oh *et al.*, 2006; Livescu and Glass, 2000) (Lawson *et al.*, 2003). For every different dialect of a language, additional speech recordings are typically required in order to maintain ASR system performance, and pronunciation dictionary adjustments may also be necessary.

The reason non-native pronunciations are so detrimental to an ASR system's performance is because they are variable, depending on the exact variant and dialect of the speaker's native language as well as their proficiency in the language that they are speaking (Benzeghiba *et al.*, 2007). This is partly explained by experiments that have shown that if a person is confronted with a sound that does not exist in their mother tongue, they try to approximate the sound from the sounds they know (Flege, 1987).

If non-native speech can be categorised into dialects or variants, a very simple system can be implemented without adaptation that recognises the dialect or variant being used, and then implements an ASR system that is optimised for that specific task. Beattie *et al.* (1995) implement the same pronunciation dictionary for all the dialects that they test on, but train using this dictionary on each different dialect separately. At run time, the best acoustic models are selected to recognise a specific dialect. A similar study is performed by Lawson *et al.* (2003), also keeping the same pronunciation dictionary and using a single set of accented data, with promising results.

2.3 STANDARD SOUTH AFRICAN ENGLISH

In order to describe the definition that will be used for SSAE in this thesis, South African English (SAE) must first be defined. The definition of SAE used in Bekker (2009) is the dialect used mainly by white speakers in the apartheid past, and which is currently in the process of being acquired by non-white South Africans. This is due mostly to the fact that a large portion of the non-white South African population is only recently learning and making use of English on a first language level.

There are many definitions of the different varieties of SAE as well as varying use of different definitions. The definitions used in this thesis are as follows: We use the definition as proposed by Bekker (2009) to define General SAE, as the pronunciation perceived to be used by the majority of the public (on a first language level). SSAE is defined to be the variant described as the received pronunciation for SAE. A received pronunciation is the 'proper' pronunciation for words, the one that is perceived to be the most correct. It often overlaps with cultivated SAE, which describes the pronunciations taught to children at school. Thus the pronunciation being investigated in this experiment may vary slightly from the pronunciation described in Bekker (2009), but both exist under the SAE umbrella.

SSAE is an English dialect which is influenced by four main SAE variants, namely, White SAE, Black SAE, Indian SAE and Cape Flats English. These names are ethnically motivated, but because each ethnicity is significantly related to a specific variant of SAE, they are seen as accurately descriptive (Kortmann and Schneider, 2004). It should be noted that these variants include extreme, strongly accented English variants that are not included in SSAE, and not addressed in this thesis.

2.3.1 DIPHTHONGS IN SOUTH AFRICAN ENGLISH

A diphthong is a sound that consists of two vowels joined together through a smooth transition. Brink and Botha (2001) perform an analysis of diphthongs in SAE, looking at the formant and pitch tracks to determine the pronunciations of the diphthongs. Although the study does not look at SSAE as a whole, they do specifically look at certain variants (mother-tongue speakers of isiZulu, isiXhosa and Sesotho). The study finds that second language (L2) speakers tend to monophthongise shorter diphthongs and shift the emphasis of long diphthongs so that both elements in the diphthong receive the same emphasis.

Their analysis of diphthongs is continued in Brink and Botha (2002), finding that the /OW/ diphthong is the most strongly affected and that /AW/ and /EY/ are monophthongised. This study concludes that when a person is attempting to pronounce an unfamiliar diphthong, if one of the phonemes that constitutes that diphthong is unfamiliar to them, either that phoneme is replaced with a phoneme from their native language and over-articulated, or it is dropped altogether (resulting in monophthongisation). This is in agreement with the findings of Flege (1987), who say that people try to approximate the sounds in a new language using the sounds they already know from languages they speak (note that these findings relate to non-native speech and not SSAE specifically).

2.3.2 THE KIT VOWEL IN SSAE

The KIT vowel [1] in SSAE exhibits different behaviour under different circumstances. It experiences allophonic variations that overlap with other phonemes, sometimes to the point where it is indistinguishable from them. Much linguistic research has been directed at the analysis of the KIT vowel in SSAE, including research by Lanham and Traill (1962), Lass and Wright (1985) and Webb (1983). Bekker (2009) reviews this literature

[1] The KIT vowel, or the 'short I' is part of Well's lexical sets for describing "The lexical incidence of vowels in all the many accents [of English]"(Wells, 1982). A lexical set is described by a set of words that tend to exhibit similar "within dialect" behaviour, but many differ between dialects.

and analyses acoustic data in order to come up with a set of rules governing the behaviour of the KIT vowel. Bekker (2009) mentions the warning provided by Branford (1994), who noted that, because the history of SAE is so complex, it is unlikely that a single monolithic explanation can be found that will fit the observable facts.

The KIT split- the fact that words such as 'chin' and 'kit' are pronounced with a similar vowel in British English and two different vowels in SSAE- is one of the most distinctive features of SSAE. Additional background with regard to the KIT split is discussed in Chapter 6.

2.4 PRONUNCIATION MODELLING IN ASR SYSTEMS

ASR systems attempt to model the human perception and semantic understanding of speech. In order to do this, the ASR system must model pronunciation in a number of different ways to simulate human processing.

2.4.1 PRONUNCIATION MODELLING LEVELS

Pronunciation modelling in ASR systems takes place on three levels: The pronunciation dictionary, the acoustic models and the language model (Strik and Cucchiarini, 1999).

Pronunciation modelling in the dictionary is represented as mappings between graphemic representations of words to phonemic representations. This is the ASR system's only means of determining the mappings and is thus of importance to the core functionality of the system. The modelling of variant pronunciations for single words needs to be implemented carefully and consistently. The addition of variant pronunciations for words adds to the confusability of pronunciations in the pronunciation dictionary, but, if applied parsimoniously, the benefit to the system can be quite high (Strik and Cucchiarini, 1999). Once the pronunciation dictionary has shown the system the time placement of phonemes, acoustic modelling takes responsibility for modelling the phonemes themselves.

Acoustic modelling describes the behaviour of phonemes as a mathematical model. Neural networks are sometimes used for this purpose (Tebelskis, 1995), due to their flexibility and generalisation potential, however, due to their inability to model temporal characteristics, they are usually used in addition to hidden Markov models (HMMs). HMMs are usually used for acoustic modelling, as they resemble the synchronous and piecewise stationary speech structure quite accurately. They represent speech as a number of sequential states, and transitions between these states. Each state models properties (usually MFCCs, or Mel-frequency cepstrum coefficients) of the piece of speech being modelled at that time as a Gaussian mixture, consisting of one or more Gaussian distributions. As a phoneme changes, it moves through the states of the HMM. The acoustic models learn how to model phonemes better the more they are trained, and can be configured to model the contexts of the phonemes that they represent. Once the acoustic models are fully trained, they require language models to guide recognition.

Language modelling describes statistical relationships between sounds and words. In other words, the language model limits the number of choices that the acoustic model has to decide amongst, thereby reducing the possibility of errors. It extracts a statistical relationships between all the units that occur in the data (which can be a phoneme or a set of phonemes), studying which are likely to occur together and if so in what order. Therefore, at any time it defines what the possible units that should be considered as candidates for recognition, and acoustic models are used to select the optimal one.

2.5 PRONUNCIATION VARIANCE MODELLING

The need for pronunciation variance modelling arises from the shortcomings of the acoustic model. Adda-Decker and Lamel (1999) prove that context dependent acoustic models reduce the requirement for pronunciation variance modelling by performing forced alignment using both context dependent and independent acoustic models and showing that the context independent models align with more pronunciation variants. Also, Holter and Svendsen (1999) find that the need for pronunciation variance modelling decreases with the number of Gaussian mixtures used in the acoustic models. But at the same time, Jurafsky *et al.* (2001) found that acoustic models are quite adapt at modelling phone substitution and vowel reduction. Thus, because the ASR system can model some variance by itself, it is important to train a system to its most optimal functionality before attempting to optimise it further by making use of pronunciation variance modelling. Pronunciation variance modelling background is important to pronunciation adaptation, because both techniques are attempting to model the adaptation of phonemes.

Pronunciation variance can be modelled on any one of the three pronunciation modelling levels. The pronunciation dictionary level, the acoustic modelling level and the language modelling level are described in more detail below.

2.5.1 PRONUNCIATION DICTIONARY

In the pronunciation dictionary, additional pronunciations can simply be added to the existing dictionary. However, before a pronunciation dictionary is expanded a decision must be made with regard to the information representation that will be implemented, as well as whether to model single or multiple pronunciations in a pronunciation dictionary.

2.5.1.1 INFORMATION REPRESENTATION

Pronunciation variance implementation in the pronunciation dictionary can be implemented as one of two methods, known as enumeration and formalisation (Strik and Cucchiarini, 1999).

Enumeration refers to adding specific pronunciation variants to specific words in the dictionary. The variants added to one word are completely independent of variants added to other words in the dictionary. Enumeration allows the pronunciations of all the words in the dictionary to be individual and unconnected to all the other pronunciations. Enumeration is a useful method when one wants to retain pronunciation individuality and is used in Goronzy *et al.* (2004).

Formalisms refer to adding variants in an organised and ordered manner. Usually the formalisms are in the form of rules, which are applied to the whole pronunciation dictionary. Thus the formalisms applied to one word in the pronunciation dictionary will be exactly the same as the formalisms applied to every other word. Because formalisms are so uniform and consistent they are used in many studies including Kessens *et al.* (2003), Tajchman *et al.* (1995), Finke and Waibel (1997) and Wester *et al.* (1998).

Davel and Barnard (2006*a*) offer what can be seen as a compromise between the two methods. They define a concept called *pseudo-phonemes*, which is used along with generation restriction rules to model pronunciation variance in a pronunciation dictionary in a compact and consistent way. This concept makes use of the fact that, when a word does have more than one pronunciation, its pronunciations only tend to differ by one or two phonemes. Pseudo-phonemes incorporate the phonemes that differ in pronunciations, and store them, like variables. Now, multiple pronunciations can be modelled as a single one in the pronunciation dictionary. Also, generation restriction rules monitor that no extra pronunciations are generated when pseudo-phonemes are decoded. Pseudo-phonemes can act as a formalism and always map to a specific set and can then be standardised throughout the dictionary, or they can simply model the pronunciation variation that occurs in individual words. It allows flexibility as well providing a way to monitor the consistency of a pronunciation dictionary.

Using formalisms to implement pronunciation variance in a pronunciation dictionary is the most popular method of information representation. This is so because of the consistency it allows one to have in a pronunciation dictionary, which is important because it allows for the consistent training of acoustic models and thus for a stable system that can be analysed in a structured manner.

Fosler-Lussier (1999) makes use of formalisms to implement dynamic pronunciation dictionaries, which model pronunciations but also the probability of each additional pronunciation, partially merging the pronunciation dictionary and the language model, in order to reduce the effects of the confusability introduced with additional pronunciations.

2.5.1.2 MULTIPLE AND SINGLE PRONUNCIATIONS

Current literature is contradictory with regard to whether a multiple or a single pronunciation dictionary is most effective, and if multiple pronunciations are used, how many pronunciations per word should be allowed. Too many pronunciations per word cause confusion in the ASR system, and too few can limit the pronunciation variance that an ASR system can neutralise. Each time a pronunciation is added to the ASR system, the system becomes more confused between that pronunciation and others which resemble it. Another reason for limiting the amount of pronunciation variance being modelled in a system is because there is not always enough training data to train all the possible options adequately. Adda-Decker and Lamel (1999) find that it is desirable for words of lower frequency of occurrence to have lower variant rates and those of high frequency to have higher ones. This is simply due to the availability of data. If there are only so many examples of a specific word, one should not confuse the system by labelling each example with a different pronunciation.

Hain (2005) investigates taking a pronunciation dictionary that is known to perform well but that has multiple pronunciations per word, and reducing it to a single pronunciation per word pronunciation dictionary that achieves similar or better performance. This investigation is based on the assertion that consistency in phoneme representation may be of higher importance than an improved representation of the training utterances. This assertion is backed up by Saraçlar *et al.* (1999), who say that an acoustic model trained on the most accurate data fails to gain robustness and under-performs as a result.

In contrast to this finding, Wester *et al.* (1998) find that their multiple pronunciation dictionary outperforms their single pronunciation dictionary. In fact Wester *et al.* (1998) conduct experiments with single and multiple pronunciation dictionaries in both the training and test phases of a system, and find that the best performance is achieved when a multiple pronunciation dictionary is used for both the training and the testing phases. Also, Holter and Svendsen (1999) go as far as to experiment with the optimal number of pronunciations per word. They find that the best results are achieved when using between 1.1 and 1.3 pronunciations per word.

Amdal and Fosler-Lussier (2003) contradict both the above findings. They perform an error analysis using single and multiple pronunciation dictionaries. In contrast to the above findings they find the results of the two to be similar, but say that the specific errors made when using the different dictionaries vary.

The studies on the optimal number of pronunciations per word in a pronunciation dictionary vary to such an extent that it seems that the optimal method to follow is to vary the number of pronunciations per word according to one's own data until an optimal balance is achieved.

2.5.2 ACOUSTIC MODELLING

At the acoustic modelling level, varying pronunciations can be implemented by editing the structure of the HMMs, or of the Gaussian mixture distributions inside them. Again, one has to be careful not to cause too much overlap between models, or the benefit of this variance modelling will be reduced.

2.5.3 LANGUAGE MODELLING

Language modelling allows one to compensate somewhat for the confusability introduced into the system through the addition of pronunciation variants. Because it is designed to model how statistically likely the occurrence of units of pronunciation is, it is able to predict the most likely pronunciation variant in different contexts. The prediction of the language model can then be combined with the prediction of the acoustic model, thus reducing the probability of the recognised unit being an error.

2.5.4 COMBINATION MODELLING

The implementation of pronunciation variance is rarely implemented on one level of modelling without implementation or participation on other levels. The pronunciation dictionary is the most fundamental pronunciation modelling layer, influencing the other two layers intrinsically, and can therefore be seen as the primary layer for pronunciation variance implementation.

2.5.5 MODELLING LIMITATIONS

Due to limitations that are imposed on pronunciation variance modelling, when an ASR system is set up with the view of implementing variance modelling, it is usually limited to attempting to model only certain causes of pronunciation variance. Therefore the ASR system usually remains vulnerable to pronunciation variance that has different causes (Benzeghiba *et al.*, 2007).

2.6 INFORMATION SOURCES FOR IDENTIFYING PRONUNCIATION VARIANTS

Different information sources are available for the purposes of identifying varying pronunciations in speech data. Varying pronunciations can be extracted directly from the speech data (data-driven) or one can rely on expert analyses of pronunciation, which are usually based on multiple speech data sets (knowledge-based). The methods are described in more detail below.

2.6.1 DATA-DRIVEN ANALYSIS METHODS

Data driven techniques can be used to identify varying pronunciations. The data-driven approach involves extracting information directly from the speech signal. The acoustic signals are analysed to determine possible pronunciations for a word or phoneme. An ASR system is trained to model acoustic data, therefore, when a linguistic analysis of speech data is required, an ASR system can be used to extract information (Adda-Decker and Lamel, 1999). The output of this approach is a list of pronunciation variants that occur in the data. Specifically, a detailed error analysis can be used to identify possible phonemic variations (Strik and Cucchiarini, 1999). However, because the outputs are so specific to a single data set, the data obtained from it does not generalise well to other data sets.

The are many ways to manipulate acoustic data in such a way as to extract pronunciation information. Outlined below are four techniques, namely, using phone recognisers, using word recognisers but editing word pronunciations, recognising foreign data and analysing phoneme confusability.

2.6.1.1 PHONE RECOGNISERS

Phone recognisers are sometimes used to find variant pronunciations from data. The same set of speech data is transcribed in two different ways. Firstly, the data is transcribed with words. The transcriptions are converted to phone transcriptions using a pronunciation dictionary, which usually contains the canonical pronunciations. Secondly, a phone recogniser is trained and is used to transcribe the speech data. Because the recognition is not limited by the pronunciation dictionary, the phonemes are recognised according to the speech data only. The two different types of transcriptions are compared and variant pronunciations are generated where the phone recogniser differs from the pronunciation dictionary based recogniser. This means that the sounds that are assumed to appear in a word if it were pronounced as predicted by the pronunciation dictionary are compared to the perceived unrestricted pronunciation of the word. Ravishankar and Eskenazi (1997) and Livescu and Glass (2000) make use of a phone recogniser for pronunciation variant suggestion with promising results.

However, phone recognisers are not often very reliable. Their accuracy tends to be low: 50% - 70%, which means that 30% to 50% of their results are unreliable (De Mori, 1998). This means that the variant pronunciations obtained using phone recognisers can not only be slightly incorrect, but because of the low accuracy, the alignment of the two data streams can be affected and variant pronunciations can be generated for non-corresponding words and can themselves be incorrect.

2.6.1.2 WORD RECOGNISERS

Some studies have researched ways to overcome the low accuracy of phone recognisers, usually involving limiting what phonemes can be recognised by using word pronunciations. Kessens *et al.* (2003) use a word recogniser, but edit the word pronunciations, thereby still generating variants but at the same time allowing the variants to be better suited to the data. They investigate phoneme deletions by allowing all phonemes in the canonical pronunciation to be optional. Forced alignment is then used to find actual variant candidates. A similar study is done by Adda-Decker and Lamel (1999) but is slightly more detailed. They investigate phonemic deletions by taking a canonical pronunciation and allowing either all vowels or all consonants to be optional for the variant pronunciations. They also investigate substitutions by defining classes of phonemes (subsets of vowels and consonants that are linguistically similar) and allowing any one phoneme in the class to be substitutable for any other phoneme in that class. Limiting pronunciation recognition does seem to increase the accuracy of found pronunciation variants, and is thus a viable alternative to using phone recognisers.

2.6.1.3 FOREIGN DATA

The recognition of foreign data involves implementing a canonical pronunciation dictionary in an ASR system, training that ASR system on a set of speech data of a certain language/dialect, and then deliberately testing it on selected varying speech data in order to gain insight into the relationship between the data used to train and test the ASR system. Goronzy *et al.* (2004) use just such an approach, training an English phone recogniser and using it to generate English pronunciations for German words. Again this yields two sets of comparable transcriptions from which possible phoneme variations can be derived. The manipulation of data, be it the pronunciation dictionary or the speech data itself, is not the only method of finding varying pronunciations. In fact, pronunciation variants can be found without manipulating the ASR system at all, but through the analysis of the existing system, as discussed in the next section.

2.6.1.4 PHONEME CONFUSABILITY

Phoneme confusability analysis involves analysing an existing system and determining what pronunciation variance modelling is necessary. Although Sloboda and Waibel (1996) only look into adapting a system to recognise spontaneous speech, the variance that exists between non-spontaneous and spontaneous speech can be equated to variants of a language. Their study focuses on the analysis of a phoneme confusion matrix of a full system. Frequent misrecognitions are analysed and tuples are put together for modelling in the dictionary.

When making use of data-driven techniques one must be careful to filter the possible variants in order not to over specialise the system to the test set and the characteristics of the ASR system (Kessens *et al.*, 2003). Usually, as the techniques are not always very accurate, data-driven techniques are complemented by filtering techniques, which are able to ascertain the applicability of the results to a set of data, and filter them accordingly.

2.6.2 DATA-DRIVEN FILTERING TECHNIQUES

Typically data-driven filtering techniques are used to make data-driven analysis techniques more effective, but these techniques can also be used with knowledge-based information sources to find possible variants. The possible variants are filtered according to the acoustic data so that only the most applicable variants are used. The techniques that are most often implemented are frequency counters, acoustic likelihood analysis and classifiers.

2.6.2.1 FREQUENCY COUNTERS

Frequency counters provide a very direct way of evaluating the variant options that are generated from the data. Adda-Decker and Lamel (1999) study the evaluation of acoustic models using forced alignment. They use frequency counters to evaluate the quality of variant suggestions directly. The forced alignment allows them to measure two things, namely, variant practicality and variant requirement. The number of times a variant pronunciation is selected in total measures its practicality. But the number of times it is selected for a specific word, when normalised by that word's frequency, measures the variant's requirement. Forced alignment can be used iteratively, until a dictionary that is accurate enough has been developed, as is done by Wester *et al.* (1998). In fact forced alignment is quite a popular method, it is also implemented by Tajchman *et al.* (1995) and Ravishankar and Eskenazi (1997). However, frequency counters are not only useful for enumeration based variant implementation.

For formalisation based variant implementation techniques the frequency counters can be used in many ways to evaluate the practicality of formalisms. Kessens *et al.* (2003) use plain frequency counters to evaluate their lexical adaptation rules to achieve an improvement. They make use of two counters, the first frequency counter checks how often the conditions for the application of the rule occur, and the second counter checks how often the rule is actually selected for application. They then calculate a ratio between the second and first counters, and together, these three variables (two counters and the ratio between them) measure the requirement for the adaptation rule in the system.

2.6.3 ACOUSTIC LIKELIHOOD ANALYSIS

Acoustic likelihood analysis measures the match between an acoustic model and the data that it is trying to represent. They make use of frequency measures to assert a boundary, which can then be used to measure the quality of a measurement. Badenhorst and Davel (2007) make use of a confidence interval per phoneme that is concerned with the standard deviation of a number of measurements in order to measure the quality of a phonemic model. Williams and Renals (1998) investigate using confidence measures for evaluating the quality of a variant suggestion directly. They make use of acoustic confidence intervals (which measure how well an acoustic model matches acoustic data), to assert a boundary, which can then be used to measure the quality of a measurement. In their experiment, the data used for the construction of acoustic models was altered by editing the baseforms (pronunciation dictionary entries) that the acoustic models make use of for training. Then the conformity between the newly trained acoustic model and the acoustic data was measured to check if it was better than before.

2.6.3.1 CLASSIFIERS

Learning algorithms and classifiers can also be trained to filter generated variants. Goronzy *et al.* (2004) make use of decision trees for this purpose. A number of decision trees are trained with variant pronunciations iteratively, with each tree attempting to fix the mistakes the previous tree made. The trees can then intelligently select variants to implement in a canonical pronunciation dictionary and can thus boost the accuracy of the system much more than when simply applying all the variants.

Fukada and Sagisaka (1997) make use of a neural network for the suggestion of possible pronunciation variances. They implement a phoneme recogniser as described in Section 2.6.1 and use the results to train a neural network, which then predicts alternative pronunciations for a canonical pronunciation dictionary.

Filtering techniques are beneficial for the selection of optimal pronunciation variants for inclusion in a pronunciation dictionary for an ASR system. However, the complexity required for their selection should be measured against the benefit that they are able to provide.

2.6.4 KNOWLEDGE BASED INFORMATION SOURCES

The knowledge-based approach to gaining linguistic information about a language or a dialect of a language involves linguistic experts analysing different sets of data and attempting to generalise possible sources of variance. The formalisms that are uncovered by the experts are usually quite inclusive and can thus be applied to many varying data sets. However, a limited amount of such formalisms exist and the formalisms that do exist are not always directly applicable to a specific data set.

One example of such an implementation is done by Oshika *et al.* (1975), who develop one formalism set for natural continuous American English. The formalisms identified in this study are supported by spectrographic evidence. These formalisms can be used to manipulate a pronunciation dictionary in order to make an ASR system better able to recognise natural continuous American English.

2.7 PRONUNCIATION DICTIONARY VERIFICATION

Strik and Cucchiarini (1999) warn that when constructing an ASR system to be used as a baseline when researching improvement techniques, one must keep in mind that the data used to build the system may contain errors. If these errors are not corrected in the baseline system but are found and corrected in the process of using the system for research, the results from the improvement technique may be overestimated. It is important to validate the baseline system prior to further experimentation, in order to be confident that the method that has been developed for the purpose of improving an ASR system is causing, at the very least, the majority of the improvement observed. This is of specific importance when analysing techniques that are very sensitive to dictionary errors, such as pronunciation variance modelling.

Because pronunciation dictionaries are often compiled from many sources and because automatic means of dictionary extension are sometimes used, the entries in the dictionary can become flawed. In large dictionaries, although a high percentage of the entries are correct, the incorrect entries can detrimentally influence a speech technology system that is developed using the dictionary. If one would like to implement the dictionary to its full potential, the removal of the erroneous entries is required.

Pronunciation dictionary verification can be performed either manually or automatically, depending on the resources available and the outcome required. Manual verification is considered the most accurate method of performing verification on a pronunciation dictionary. However, the effort required is very high and if errors occur, because they are human errors, they are unpredictable. Automatic verification is very efficient and the errors made are more predictable, however, the accuracy tends to be less accurate. Thus a semi automatic approach is often followed to compensate for the downfalls of both methods.

Damper *et al.* (1997) perform dictionary verification but do so by hand and do not follow a specific strategy. Davel and Barnard (2006*b*), however, describe a semi automatic approach to dictionary verification that reduced the human effort required for the process. They make use of the *Default&Refine* algorithm, which is an algorithm that extracts grapheme to phoneme relationships. One of the outputs of this algorithm is that it recognises when a grapheme is mapped to a phoneme that it is not regularly mapped to. This allows it to generate a list of words with exceptional pronunciations. These words are more likely to have erroneous pronunciations than the other words in the dictionary. Thus the human effort required to validate the dictionary is reduced from analysing all words to only the words with exceptional pronunciations.

Once dictionary verification is performed, the dictionary can be used in experiments seeking to improve the performance of an ASR system through manipulation of the dictionary, yielding more reliable results.

2.8 SUMMARY

This chapter provided background on the process of dictionary adaptation, focusing on techniques that are used to identify and analyse pronunciation variants. The main points of the chapter are highlighted below.

SSAE is a recognised dialect of English and exhibits systematic differences when compared to British and American English, the two main dialects of English in which electronic ASR resources exist.

Before attempting to adapt a pronunciation dictionary to a different dialect, that dictionary needs to be verified and all errors found must be fixed. Intuitive techniques and pattern recognition can be implemented to perform rudimentary verification tests on the pronunciation dictionary.

It is important to model pronunciation variation in an ASR system. The more severe the variation is, the more detrimental it is to system performance, and the more important it is to model it explicitly in order to improve the system performance. Non-native speech in particular has a very detrimental effect on a system's performance.

In order to analyse the variance in acoustic data, knowledge-based and data-driven techniques can be implemented. Because the methods are so complementary in implementation (data-driven being so specific to a data set, knowledge-based being so general) the best way to analyse data is to implement them cooperatively. Data-driven analysis methods can be implemented to extract information from the acoustic data. (Word recognition with manipulated pronunciations exploits the higher accuracy achieved with word models but keeps the freedom of phoneme selection). This approach can be combined with knowledge-based information sources to remove pronunciations that are extremely unsuitable. Data-driven filtering techniques complement the above methods and can be used to optimise a resulting set of variants. The variant set that is implemented in the ASR system is be described using a formalism. This way phonemic variations can be analysed instead of individual word pronunciation variations.

Once variants are implemented in an ASR system, the ASR system can be analysed using both its confusion matrix and an error analysis to ascertain the changes made to the system through the implementation of the variants. The number of variants per word can be experimented with at this point, as well as the order of the variants. This optimised system can then be analysed for further improvements.

CHAPTER THREE

DICTIONARY VERIFICATION

3.1 INTRODUCTION

In order to implement optimal and reliable ASR experiments, a reliable pronunciation dictionary is required. Pronunciation dictionaries can contain errors, but verification can be implemented to find these errors, and either eliminate or correct them. However, because pronunciation dictionaries can become quite cumbersome, human verification can become very resource intensive, and thus automatic or semi-automatic verification can be very beneficial.

This experiment focuses on the implementation of mechanisms to filter a pronunciation dictionary that require limited human intervention. Section 3.2 provides background on the topic of dictionary verification, as well as some of the techniques implemented in this chapter. Section 3.3 describes the techniques used in the analysis of the dictionary. Section 3.4 provides a description of the dictionary selected for this study as well as an outline of the process followed in the different parts of the experiment. Section 3.5 describes each of the filtering techniques implemented, how many entries are filtered out using each technique and provides samples of entries that are filtered out using each technique. Section 4.4 describes the ASR system that is used for the purposes of gauging the improvement that the filtering provides. Finally, Section 3.7 summarises the findings and highlights key deductions.

3.2 DICTIONARY VERIFICATION TECHNIQUES BACKGROUND

Our dictionary analysis approach builds on published techniques related to (1) grapheme to phoneme (G2P) alignment, (2) grapheme to phoneme rule extraction and (3) pronunciation variant modelling.

3.2.1 GRAPHEME TO PHONEME ALIGNMENT

Many grapheme to phoneme rule extraction algorithms first require that grapheme to phoneme alignment is performed. Each word in the training dictionary is aligned with its pronunciation on a per-grapheme basis, as illustrated in Table 3.1 where ϕ indicates a null (or empty) grapheme or phoneme. The 44-phoneme BEEP ARPABET set is used (Appendix A). The alignment process involves the insertion of graphemic and phonemic nulls into each of the dictionary's entries of words. A graphemic null is inserted when more than a single phoneme is required to pronounce a single grapheme. A phonemic null is inserted when a single phoneme is realised from more than one grapheme.

Table 3.1: Grapheme to phoneme alignment example

R O S E	$\rightarrow$	/ R OW Z ϕ /
R O W S	$\rightarrow$	/ R OW ϕ Z /
R O O T	$\rightarrow$	/ R UH ϕ T /
M A X ϕ	$\rightarrow$	/ M AE K S /

Viterbi alignment (Viterbi, 1967) is typically used to obtain these mappings, where the alignment algorithm makes use of the probability of each grapheme being mapped to a particular phoneme. The alignment technique described in more detail in Davel and Barnard (2004*b*) is implemented in this experiment:

- Initial probabilities are calculated by selecting the entries in a dictionary that have the same phonemic and orthographic lengths.
- Once these probabilities are calculated, iterative forced Viterbi alignment is performed on the dictionary.
- Graphemic null generator pairs are extracted to be able to insert graphemic nulls while predicting unknown words.
- The probability of any grapheme being aligned to a null phoneme is conditioned on the prior phoneme.
- Phonemic nulls are consistently used to indicate that the prior phoneme is realised by more than one grapheme.

3.2.2 GRAPHEME TO PHONEME RULE EXTRACTION

Various automatic rule extraction techniques exist, including decision trees (Black *et al.*, 1998), pronunciation-by-analogy models (Marchard and Damper, 2000), Dynamically Expanding Context (DEC) (Torkkola, 1993) and IB1-IG, a *k*-nearest neighbour classifier (Daelemans *et al.*, 1999). As these techniques attempt to generalise from learning instances they can be used to identify exceptional instances which may possibly be errors.

In this analysis the *Default&Refine* algorithm is utilised for the extraction of grapheme to phoneme rules (Davel and Barnard, 2004*a*, 2008). This algorithm makes use of two observations: Graphemes are usually realised as one phoneme more often than all the others, and that graphemes have different realisations as phonemes based on their context in a word. The algorithm extracts grapheme-to-phoneme (G2P) rules for each grapheme independently. The following process is applied: all the realisations of a grapheme are considered and the rule that correctly predicts most of the realisations is selected as the default rule. After this, the rule containing the smallest possible context that correctly predicts most of the left over occurrences of a grapheme is selected. This process is applied iteratively until all realisations of a grapheme are correctly predicted. During prediction, a grapheme's context is tested against rules, starting from the rule with the largest context, until a match is found. The final rule, the default one, does not specify a specific context and therefore matches every context in which the grapheme can occur.

3.2.3 VARIANT MODELLING

Most of the G2P rule extraction mechanisms mentioned above can only train on words having single pronunciations (rather than more than one pronunciation for a single word). Pseudo-phonemes and generation restriction rules have been developed as a way to model varying pronunciations of words as a single pronunciation (Davel and Barnard, 2006*a*). Pseudo-phonemes are used to represent two or more phonemes which can appear in a certain place in the pronunciation of a word. When two or more pseudo-phonemes appear in a word, generation restriction rules are applied to limit the combinations of phonemes that can be generated from the set of pseudo-phonemes. This ensures that if the pseudo-phonemes are removed again, nothing will have been added or removed from the original dictionary.

3.3 APPROACH

There are two ways in which a dictionary can be verified: direct observation and indirect analysis. Direct observation of a dictionary is the analysis of a dictionary through direct observation of its context. The techniques include comparing the lengths of the orthographic and phonemic representations, looking at different words that have duplicate pronunciations and the examination of the dictionary for distinguishable errors in both the orthographic and the phonemic transcriptions. Indirect analysis requires the implementation of techniques to transform the dictionary into different formats, each of which allows different errors to become more distinguishable. Indirect analysis techniques include the alignment of the dictionary, extraction of grapheme to phoneme rules and the implementation of pseudo-phonemes along with generation restriction rules.

In this section, a number of novel methods are described and implemented in order to isolate the incorrect entries in a dictionary. Each general method is explained below along with the ways in which it is applied in order to implement verification of the dictionary.

3.3.1 WRITTEN WORD AND PRONUNCIATION LENGTH RELATIONSHIPS

The relationship between a word's orthographic and phonemic representation can be an indicator of whether a word's spelling and pronunciation are consistent. The extraction of words whose orthographic and phonemic transcriptions differ above a certain threshold can allow one to obtain a manageable list of possible erroneous entries from a dictionary.

3.3.2 ALIGNMENT ANALYSIS

The alignment of a word to its pronunciation gives one further insight into the length relationship of a word and its pronunciation, and in addition identifies words which do not match their pronunciation. During alignment, graphemic and phonemic nulls are inserted in order to align every grapheme to a phoneme. Potential errors can be flagged at this stage through the analysis of the placement and the number of nulls inserted into both the orthographic and phonemic representations of a word.

3.3.3 GRAPHEME TO PHONEME RULES

Grapheme to phoneme (G2P) rules are extracted for one grapheme at a time and are sorted such that the number of occurrences that gave rise to any one of the rules is easily obtainable. By inspecting the rules that are generated by the smallest number of occurrences, one can gain insight into potentially incorrect entries.

3.3.4 DUPLICATE PRONUNCIATIONS

Words that have the same pronunciation as other words usually have similar orthographic length. For example, the words CAUSE, CAWS, CORES and CORPS all have the same pronunciation and their spelling consists of four to five letters. One way to isolate problematic entries is to search for words that have the same pronunciation and to compare their orthographic lengths.

3.3.5 VARIANT ANALYSIS

The generation restriction rules that accompany words which contain more than one pseudo-phoneme can allow one to flag possibly incorrect entries in the dictionary. This is done by analysing the generation restriction rules themselves. When pronunciation variants do occur in a dictionary, they usually differ from each other by one or two sounds. If restriction rules are being generated for more than three sounds, it can mean one of three things: (1) The entries are correct and the word truly does allow for vastly different pronunciations, (2) the alignment of the word has not aligned graphemes to the correct phonemes, or (3) that at least some of the variants are incorrect. Once a list of generation restriction rules is obtained, the list of multiple pseudo-phonemes occurring in words is short enough to be evaluated manually.

3.4 EXPERIMENTAL SETUP

3.4.1 DICTIONARY

The BEEP dictionary (Robinson, 1996) is selected for this study. It is a freely available online English pronunciation dictionary that is comparable with other available online dictionaries with regard to its size and content (Damper *et al.*, 1997). It is compiled through the amalgamation of several public domain dictionaries and has not undergone a strict quality control process. However, because it is widely used internationally, especially in pronunciation prediction research (Damper *et al.*, 1997; Davel and Barnard, 2004*a*), the dictionary is expected to be fairly error-free.

3.4.2 PROCESS

A series of steps are followed for the verification of the BEEP dictionary:

1. **Pre-processing** is implemented where correct entries that make further analysis difficult are temporarily removed.
2. **Systematic errors** identified during an initial alignment run and dictionary analysis are removed.
3. The **spelling** of the words in the dictionary is evaluated and a list obtained of allegedly misspelled entries.
4. A list of **pronunciations** that are longer than their orthographic partners is generated. Words which simply need graphemic nulls are systematically removed from the list. The remainder of the list is removed from the dictionary.
5. The list of **graphemic nulls** is examined and graphemic nulls giving rise to errors are identified. The words containing these graphemic nulls are automatically checked and removed from the dictionary if wrong.
6. A list of **spellings** that are longer than their pronunciations is generated. The list is evaluated and manually selected entries are removed from the dictionary.
7. A list of words that have **duplicate pronunciations** but where the orthographic length of the words differs is generated. The list is evaluated manually and entries that are found to be erroneous are removed from the dictionary.
8. The new dictionary is again **aligned**. During alignment, problematic words are found and removed from the dictionary. A list of entries containing a large number of consecutive phonemic nulls is generated, this list is evaluated manually and selected words are removed from the dictionary.

9. **G2P rule extraction** is performed on the dictionary. The rules that predict the phonemic representation of a grapheme from a single word are used to create a list of possible erroneous entries in the dictionary. This list is evaluated manually and entries found to be wrong are removed.

10. **Pseudo-phonemes** are implemented in the dictionary and generation restriction rules are generated. The list of generation restriction rules is searched for strings of pseudo-phonemes occurring in words. These words are evaluated manually for errors and removed from the dictionary if errors are found.

3.5 DICTIONARY ANALYSIS RESULTS

The summary of how many entries are removed by each dictionary verification technique can be found in Table 3.2. The table also indicates whether a verification is automated (requiring no human intervention) or semi-automated (requiring validation of the list of possible errors by a human). Where validation is required, the size of the list requiring validation is also reported.

The separate counts of each of the occurrences removed can be found in Table 3.3. The column 'Verification applied' describes the technique that was applied to verify the pronunciation dictionary. The column 'Verification type' describes the automation of the verification technique applied. The column '# of possible errors' shows the number of errors identified by each technique automatically, and the column '# removed' shows the number of entries that were manually identified as incorrect from the automatically generated list. The column '% possible errors verified' gives the percentage of the automatically generated possible errors that were verified as incorrect. Finally, the column '# remaining' shows the number of entries remaining in the dictionary after the removal of the erroneous entries.

3.5.1 PRE-PROCESSING

For pre-processing, unusual pronunciation patterns are removed. These are removed temporarily, as the entries are not erroneous, but make dictionary analysis difficult. Unusual punctuation removal involved the removal of punctuation which does not occur in general English writing. This process also removed many acronyms from the dictionary. Examples of removed words are: VICU $\sim$ NA, W._R._A._C., and ¡PAUSE¿. A sample of 100 entries from those removed can be found in Table B.1 in Appendix B.

Table 3.2: Results of each step involved in the verification process

Verification applied	Verification type	# possible errors	# removed	% possible errors verified	# remaining
None	N/A	0	0	0%	257 059
Punctuation Removed	Automated	576	576	100%	256 483
Repeated Phonemes	Automated	4 730	4 730	100%	251 753
Lengthened Pronunciations	Semi-automated	1 284	253	19.7%	251 500
Incorrect Graphemic Nulls	Semi-automated	362	189	52.2%	251 311
Lengthened Spelling	Semi-automated	209	69	33%	251 242
Duplicate Pronunciations	Semi-automated	$\simeq$ 305	80	$\simeq$ 26.22%	251 162
Alignment Errors	Semi-automated	9	9	100%	251 153
Consecutive Phonemic Nulls	Semi-automated	42	33	78.57 %	251 120
Singular G2P Rules	Semi-automated	1 450	89	6.14%	251 031
Generation Restriction	Semi-automated	$\simeq$ 90	50	$\simeq$ 55.56%	250 981
Punctuation Replaced	Automated	-576	-576	100%	251 557
Total		9 057	5 502	60.75 %	251 557

3.5.2 REMOVAL OF SYSTEMATIC ERRORS

Through inspecting the result of an initial alignment of the dictionary, a list of systematic errors is compiled, which in this case contains repeated phonemes. It is found that in words where a letter is repeated, the phonemic representation of the letter is usually repeated as well, even where such repetition does not occur. 5 711 instances were originally identified, but minor inspection revealed that some repeated phonemes were legitimate (such as the transcription for ACCOMPANYING being */ AX K AH M P AX N IH IH NG /*), and those entries are left in the dictionary. In total 4 730 entries are removed from the dictionary. A sample of 100 entries from those removed can be found in Table B.2 in Appendix B. Examples of removed entries are:

- ADMITTER, which is transcribed as */ AX D M IH T T ER /*,
- CHIPPIE, which is transcribed as */ CH IH P P AY /*,
- INPUTTED, which is transcribed as */ IH N P UH T T EH D /*,
- SCAMMED, which is transcribed as */ S K AE M M EH D /*, and
- HORSESHOE, which is transcribed as */ HH OH SH SH UH /*.

Table 3.3: Number of entries removed from the dictionary due to repeated phonemes

Repeated Phoneme	Number Removed
AX AX	959
T T	942
N N	586
L L	479
P P	391
D D	275
S S	246
M M	199
K K	182
R R	178
B B	156
G G	56
EY EY	23
F F	14
IY IY	11
SH SH	9
CH CH	8
AA AA	7
OW OW	6
Z Z	3

3.5.3 SPELLING VERIFICATION

In an attempt to verify the spelling used for words in the dictionary, a wordlist is extracted and the spelling checked automatically. However, the list of incorrect spelling contains over 146 000 words, and after a general manual inspection is found to be invalid and discarded. Checking the spelling of the BEEP dictionary may be beneficial, however, the program that would perform the checking would require a more comprehensive coverage of all English words.

3.5.4 LENGTHENED PRONUNCIATIONS

One of the methods that can be used to isolate errors in the dictionary is to verify those entries for which the phonemic representation of the word is longer than the orthographic representation. In order to make this method function correctly, the list output has to be refined by identifying where graphemic nulls should be inserted and taking this into account. This function yielded a list of 1 284 entries. The list is found to contain many proper noun entries, some of whose pronunciations are suspicious but could not be categorised as incorrect. The list is manually filtered down to 253 entries that are removed form the lexicon. A sample of 100 entries from those removed can be found in Table B.3 in Appendix B. Examples of entries removed are:

- The word APRICATION, with the pronunciation */ EY P R IH V AE R IH K EY SH N /*,
- the word EFFECTIVITY, with the pronunciation */ IH F EH K T AX B IH L IH T IY /*,
- the word DECONGESTING, with the pronunciation */ D IH K OH N S IH K R EY T IH NG /*,
- the word OVERCONCERNED, with the pronunciation */ OW V AX K OH N F IH D AX N S ER N D /*,
- the word RECIPROCATORY, with the pronunciation */ R IH S IH P R AX V OH K AX T AX R IH /*,

3.5.5 GRAPHEMIC NULL ANALYSIS

The graphemic nulls identified in Section 3.5.4 are investigated further. The list consists of sequences including the letter 'X' always needing a graphemic null, the letter 'U' needing a graphemic null in certain situations (such as the word ACUTE having the pronunciation */ AX K Y UW T /*) and the letters 'SM' needing a graphemic nulls between them in certain situations (such as the word ADVENTURISM having the pronunciation */ AX D V EH N CH AX R IH Z AX M /*).

However, the graphemic nulls that are in the list are not always applicable, and with manual verification the phonemes */ UW AX /* are found to be invalid in situations where a word contains the letter sequence 'ower'. (An example of valid use of the phonemes being the word BESTOWER with the pronunciation */ B IH S T OW AX R /*.) Once identified, the entries whose pronunciations contains the phonemic sequence are output to a file. This list contains 362 entries, but is manually filtered to 189. A sample of 100 entries from those removed can be found in Table B.4 in Appendix B. Examples of the removed entries are:

- The word DELEGATOR, with the pronunciation */ D EH L IH G AA T OW AX /*, and
- the word VENTOR having the pronunciation */ V EH N T OW AX /*

3.5.6 LENGTHENED SPELLING

Next, the dictionary is analysed to isolate the orthographic representations of words that are more than a selected threshold longer than their phonemic representations. A threshold of four yields a list of 1 366 words, which is judged to contain too many correct entries. A threshold of six yields a list that is too short. Therefore, orthographic representations that are a threshold of at least five characters longer than their pronunciation are flagged as possibly erroneous. A list of 209 entries is extracted, which is analysed manually and filtered down to a list of 69 that is removed from the dictionary. The entries removed can be found in Table B.5 in Appendix B. Examples of the words removed form the dictionary are:

- The word PRESENTIMENTAL, with the pronunciation */ P R IH Z E N T L /*,
- the word SEMITRANSPARENT, with the pronunciation */ S EH M IH T R AX N T /*,

- the word CHANCELLORSHIPS, with the pronunciation */ CH AA N S AX SH IH P S /*,
- the word UNPROSPEROUSLY, with the pronunciation */ AH N P R AX S L IY /*, and
- the word PRIVATIZATIONS, with the pronunciation */ P R AY V EY SH N Z /*.

3.5.7 DUPLICATE PRONUNCIATIONS

This test is implemented using the intuitive principle that words that are pronounced the same should have a similar orthographic length. The dictionary is investigated, specifically looking for words that had the same phonemic representation but orthographic representations varying in length according to a threshold. The algorithm that is implemented calculates the mean length of all the orthographic representations and then works out by how much the length of each of the orthographic representations differs from the mean. The threshold for this value is iteratively tested, and the most applicable value is found to be 1.5. A value of two yields too little output, and a value of one yields too many entries in the output. The list of duplicated pronunciations contains 305 sets of words. A set of words would contain between two and four words with identical pronunciations. This list is manually analysed and a list of 95 erroneous entries is extracted that are removed from the dictionary. The entries removed can be found in Table B.6 in Appendix B. Examples of words that contained pronunciations for other words are:

- The word NONRESPONDENT, with the pronunciation */ N OH N R EH Z IH D AX N T /*,
- the word DISTINGUISED, with the pronunciation */ D IH S G AY Z D /*,
- the word SOCRATISTIC, with the pronunciation */ S AX K R AE T IH K /*,
- the word NARCISSISTIC, with the pronunciation */ N AA S IH S T IH K /*, and
- the word UNSUPPLEMENTED, with the pronunciation */ AH N S AX P L AA N T IH D /*.

3.5.8 ALIGNMENT

Alignment calculates the probabilities of graphemes being realised as certain phonemes, and aligns them accordingly. It can thus be a strong source of information in the search for incorrect entries in the dictionary. For the purpose of flagging incorrect entries, two methods are attempted: listing entries with a high total number of nulls and listing entries with a high number of consecutive nulls.

Listing entries with a high number of total nulls is implemented using different thresholds of how many nulls an entry needs to contain in order to be added to the list. Setting the threshold to four nulls yields a list of over 10 000 entries, a sample of which is verified as mostly correct content. The threshold then is steadily increased to 7 nulls. This yields a list of 82 entries, however, after verification this list is discarded. The list contains incorrect words, however, these words are already listed as incorrect in Section 3.5.6 or listed in the successive nulls list, thus no improvement can be achieved through its implementation.

Listing the entries with a high number of successive nulls provides one insight into where the alignment algorithm had experienced difficulty in aligning a grapheme to the correct phoneme. However, if the number of succesive nulls is larger than the threshold used in Section 3.5.6, the results are redundant as they have been verified. The best threshold for the number of successive nulls is investigated to yield a list of possibly erroneous entries that is short enough for manual verification. Initially set to three nulls, a list of over 3 000 entries is generated. Through verification this list is found to contain too many correct entries and thus discarded. The only other threshold that is lower than the threshold in Section 3.5.6 is four. The threshold is thus set to four successive nulls, and a list of 42 entries is generated, and filtered down to 33 entries through manual verification. This analysis method is thus found to be very efficient. The list of incorrect entries is then removed from the dictionary. The entries removed can be found in Table B.7 in Appendix B. Examples of the removed entries are:

- the word ANTISEPTICISM being aligned to the pronunciation */ AE N T IH S ϕ ϕ ϕ ϕ IH Z AZ M /*,
- the word CAPITIVITY being aligned to the pronunciation */ K AE P IH T ϕ ϕ ϕ ϕ IY /*,
- the word COMPROMITMENT being aligned to the pronunciation */ K OH M ϕ ϕ ϕ ϕ IH T M AX N T /*,
- the word UNALLEVIATED being aligned to the pronunciation */ AH N AX L ϕ ϕ ϕ ϕ OH T IH D /*, and
- the word STOMATOMY being aligned to the pronunciation */ S T ϕ ϕ ϕ ϕ AX M IY /*.

3.5.9 GRAPHEME TO PHONEME RULES

The G2P rules are implemented using the *Default&Refine* algorithm. This algorithm allows one to see the number of instances of a grapheme from which each single rule is extracted. By selecting rules that are extracted from single instances, entries with anomalous pronunciations can be isolated. Rules are extracted from the BEEP dictionary, and rules that are extracted from single instances of a grapheme are extracted. These rules are used to find the instances which gave rise to them. The last 50 rules for each grapheme are analysed, where the set of graphemes included three punctuation marks. This yielded a list of 1450 entries. This list is verified manually and finally, a list of 89 entries is removed from the dictionary. The entries removed can be found in Table B.8 in Appendix B. Examples of entries found are:

- the word HYDROPOLITICS, with the pronunciation */ H AY D R AX P OH N IH K S /*,
- the word UNPRECIPITATED, with the pronunciation */ AH N P R EH S IH D EH N T IH D /*,
- the word DISCREPANCE, with the pronunciation */ D IH S T ER B AX N S /*,
- the word INCONQUERABLE, with the pronunciation */ IH N K AX N S IY L AX B L /*, and
- the word TROUBLESHOOTED, with the pronunciation */ T R AH B L Z HH UW T IH D /*.

3.5.10 PSEUDO-PHONEMES

Pseudo-phonemes are used to implement pronunciation variants in systems that do not allow for variants. The generation restriction rules that are used along with pseudo-phonemes give insight to the relationships between the pronunciation variants of a word. Variants that require more than three pseudo-phonemes are investigated with the expectation that one or more of the variant pronunciations would be incorrect. There are 90 sets of restrictions containing more than three pseudo-phonemes. A list of 50 entries is extracted from these restrictions manually and removed from the dictionary. The list of incorrect entries is then removed from the dictionary. The entries removed can be found in Table B.9 in Appendix B. Examples of the removed entries are:

- The word INAPPRECIABLE, with the pronunciation */ IH N AX P R OW P R IA T /*,
- the word UNATTACHED, with the pronunciation */ AH N AX T EH N D IH D /*,
- the word COMPUNCTION, with the pronunciation */ K OH M P Y UH T EY SH N /*,
- the word LIEUTENANCIES, with the pronunciation */ L EH F T EH N AX N S IH Z /*, and
- the word IRREMOVABLE, with the pronunciation */ IH R EH P AX R AX B L /*.

This method is found to have the most accurate prediction of incorrect entries due to its manual verification percentage being 55.56%.

3.6 EFFECTIVENESS OF ERROR ANALYSIS

Error analysis is performed in order to determine the effectiveness of implementing the above techniques on the BEEP dictionary. 200 entries are randomly selected from the final and initial dictionaries and analysed independently by two researchers. The goal of the exercise is to obtain an estimate of the number of incorrect entries in both, however, it is not always possible to conclusively categorise entries as either correct or not. Some esoteric words are not known to the dictionary verifiers and are not included in any word list consulted. Proper nouns included in the dictionary are exceedingly difficult to evaluate because some seemingly incorrect proper nouns may actually be correct. Thus the category of incorrect is expanded to three categories: Conclusive, Proper noun and Questionable.

For the unfiltered BEEP dictionary, 32 entries are selected as being erroneous. The Conclusive category contains 14 entries, including the word BUMPTY with the pronunciation */ B AH M P IY T IY W AY /*. The Proper noun category contains 10 entries, including the word BUZZY'S, with the pronunciation *B AH Z W AY Z*. The Questionable category contains 7 words, including the word CHADLIN, with the pronunciation */ CH AE D L I N /*. In total 16% of the initial dictionary is found to be incorrect.

For the filtered dictionary, 19 entries are selected as being erroneous. The Conclusive category contains 5 entries, including the word TOURNANT'S having the pronunciation */ T AO N AX M AX N T S /*. The Proper noun category contains 7 entries, including the word MESSA'S, having the pronunciation */ M EH S EY Z /*. The Questionable category contains 6 words, including the word RESCURE, having the pronunciation */ R EH S K Y UA /*. In total, 9.5% of the filtered dictionary is found to be incorrect.

These results imply that approximately 60 % of conclusive errors were removed during a process in which 9 057 words were manually verified (Out of a possible 257 059 words) and 5 502 entries removed. These results indicate a significantly more efficient process than was initially anticipated.

3.7 RESULTS SUMMARY

This chapter focused on identifying algorithms to cater for semi-automated dictionary verification. Several novel methods were implemented and their effectiveness analysed. The techniques that identified the most errors are:

- Searching for repeated phonemes and removing entries whose pronunciations contain incorrect repetitions. 4 730 entries were removed using this method. However, this method is quite dictionary specific and may not generalise well to other dictionaries.
- Pronunciations that were longer than their orthographic representation provided a good source of incorrect entries. 253 entries were identified and removed. This method can be applicable to other dictionaries in English but may be language specific and may not necessarily perform as well with dictionaries in other languages.
- Identifying erroneous graphemic nulls was a method that found many incorrect entries. 189 entries were identified and removed from the dictionary using this technique. The analysis of graphemic nulls may generalise well to other dictionaries, however, the specific nulls that were identified may not.

In addition, the most efficient techniques (identifying the largest number of verified incorrect entries as a percentage of the word list requiring manual verification) are found to be:

- The analysis of words containing multiple consecutive gnulls is found to be the most efficient method, achieving an accuracy of 78.57 %.
- The analysis of the generation restriction rules that accompany the implementation of pseudo-phonemes, concentrating on groups of more than three pseudo-phonemes, was very efficient at identifying truly erroneous entries in the dictionary. This method achieved 55.56% accuracy with its list of potential errors. This method is likely to generalise well to other dictionaries, but is only efficient when one is looking for incorrect pronunciation variants.
- Identifying erroneous graphemic nulls was very efficient as well, achieving a 52.2% accuracy with its predictions.
- Searching for a number of consecutive nulls in pronunciations after alignment is performed on a dictionary was also an efficient technique at finding errors in the dictionary. This technique is 41.18% successful in its prediction of incorrect entries.

In total, 5 502 words are removed from the BEEP dictionary. This is unexpected, as BEEP is a popular dictionary, frequently utilised in a variety of speech technology applications.

3.8 CONCLUSION

Pronunciation dictionaries are fundamental to the functionality of ASR systems and it is prudent to perform verification of them before they are utilised. In this experiment, a number of novel automatic and semi-automatic verification techniques are explored. These are found to be quite efficient, allowing human verifiers to review less words to find the same errors, thus reducing the resources required for dictionary verification.

The number of identified erroneous entries in BEEP, a widely used British English dictionary is found to be surprisingly high, over 5 000 words.

The final results of this experiment is a dictionary that has undergone verification and is thus more reliable for use in ASR experiments.

CHAPTER FOUR

BASELINE ASR SYSTEM DEVELOPMENT

4.1 INTRODUCTION

When experimenting with ASR techniques, it is important to build the most optimal system one can, while attempting to keep the system architecture as close to others studied as possible, to allow for comparability and applicability of results. In view of this similar systems were studied from Hain *et al.* (2001), Hain (2005) and Goronzy *et al.* (2004). Their architecture was studied and selectively implemented in the baseline system that is used throughout this thesis.

This chapter focuses on introducing the baseline system that is used for experiments further on in this thesis. Section 4.2 defines the particulars of the system, describing the speech corpus and pronunciation dictionary in Sections 4.2.2 and 4.2.1, respectively. The technical implementation of the system is described in 4.2.3 and the parameter optimisation undergone for this system is described in Section 4.2.4. The results of the system are shown in Section 4.3, and concluding remarks are given in 4.5.

4.2 ASR SYSTEM PARTICULARS

In this section the baseline ASR system used in our experiments is defined and explored in depth. More specifically, the pronunciation dictionary implemented is described, as well as the speech corpus used to train and test the system. The system architecture and implementation thereof are also defined in full.

4.2.1 PRONUNCIATION DICTIONARY

The pronunciation dictionary consists of a combination of the British English Example Dictionary (BEEP) (Robinson, 1996) and a supplementary pronunciation dictionary that has words contained in the speech corpus but not transcribed in BEEP. (This includes SAE specific words and names of places). The 44-phoneme BEEP ARPABET set is used. The dictionary was put through the verification process described in Chapter 3 but also manually verified to eliminate highly irregular pronunciations. The dictionary has 1 500 entries, 1 319 of which are unique words. The average number of pronunciations per word is 1.14 and the number of words with more than one pronunciation is 181. In further experimentation, this dictionary is referred to as the *default dictionary*.

4.2.2 SPEECH CORPUS

The speech corpus consists of speech recorded using existing interactive voice response systems. The recordings consist of single words and short sentences. There are 19 259 recordings made from 7 329 telephone calls, each of which is expected to contain a different speaker. The sampling rate is 8 kHz and the total length of the calls is 9 hours and 2 minutes. In total, 1 319 words are present in the corpus, but the corpus is rather specialised, with the top 20% of words making up over 90% of the corpus. The relevant phoneme counts are given in Table 4.1. Counts are calculated using forced alignment with the speech corpus and default dictionary.

Table 4.1: Phoneme counts for the speech corpus

Phoneme	Number of Occurrences	Phoneme	Number of Occurrences
AA	2097	**EA**	2566
AW	2037	**DH**	2389
AY	6561	**HH**	2420
CH	1847	**JH**	2385
ZH	75	**D**	5294
NG	2194	**TH**	3929
IY	9634	**B**	4777
AE	5470	**EH**	6158
G	2261	**F**	5502
AH	3883	**K**	6778
M	5261	**L**	5389
AO	3106	**N**	23587
IA	1014	**IH**	9084
S	10701	**R**	9484
EY	4509	**T**	17066
W	4293	**V**	6010
Y	2743	**AX**	14282
Z	5399	**ER**	2499
OH	3232	**UW**	3151
SH	2008	**UH**	1324
OY	39	**OW**	3442
UA	455	**P**	3286

4.2.3 TECHNICAL IMPLEMENTATION

The ASR system implementation is designed to be fairly standard. The system is designed to make use of 39-dimensional Cepstral Mean Normalised vectors (13 Mel Frequency Cepstral Coefficients (MFCCs) ,13 delta coefficients,13 acceleration coefficients). It implements 3 state left-to-right Hidden Markov Models (HMMs) to model context dependent triphones. Each state of the HMM implements a Gaussian Mixture, which is optimised to contain 8 mixtures (see Section 4.2.4.3).

The system makes use of a flat word based language model, which means that the language model attempts to recognise whole words consisting of phonemes instead of individual phonemes, but the sequence of words is assumed to be random (which is a conservative assumption).

The insertion penalty implemented in recognition is experimentally optimised at -60 (See Section 4.2.4.2). Ten fold (90% - 10 %) cross validation is also implemented, in order to verify all results. For cross validation, all the utterances of a single speaker in the speech data are grouped in either the training or the test data, and not allowed to appear in both. The system is implemented using the ASR-Builder software (Zsilavecz, 2008), which is a tool that implements the HMM Toolkit (Young *et al.*, 2000).

4.2.4 OPTIMISING SYSTEM PARAMETERS

4.2.4.1 ACCURACY DEFINITION

Accuracy is calculated as follows (where units can be phonemes or words):

Deletions = Number of times that a unit occurs but is not recognised as itself or any other unit.

Insertions = Number of times that a unit is recognised as occurring, but where neither it nor any other unit occurs.

Substitutions = Number of times that one unit is recognised where another unit occurs.

Total = Total number of occurrences of a specific unit.

$$PercentAccuracy = \frac{Total - Deletions - Substitutions - Insertions}{Total} X100\%$$

4.2.4.2 WORD INSERTION PENALTY TESTING

Penalty testing is performed to ensure that as the ASR system evolves, its results would be directly comparable to other iterations of the ASR system. A grid search is performed to find the optimal penalty. The process is described below and summarised in Table 4.2. Note that word accuracy is used to select the correct penalty for the baseline system. This process is repeated for each individual ASR system trained.

- Firstly, five penalties are tested: -100,-50, 0, 50, 100. These enable instantiating outer limits for penalties, as well as establishing 4 pools into which the penalties can be split.
- The highest accuracy is achieved at the penalty -50. Thus the penalties around it with a radius of 25 need to be tested: -25 and -75.
- The highest accuracy is still achieved at -50. Thus the penalties around around it with a radius of 10 are tested: -40 and -60.
- Now the highest accuracy has moved to -60. Thus the penalties closest to it with a gap of 5 are tested: -65 and -55.
- The highest accuracy is still around -60, thus the radius is decrease to 2.5, and the penalties around it are tested again: -57.5 and -62.5.
- The penalty -60 still achieves the highest accuracy, and is thus determined to be the optimal penalty.

4.2.4.3 GAUSSIAN MIXTURE OPTIMISATION

The number of Gaussian mixtures was optimised experimentally (see Table 4.3), using phoneme accuracy as an optimality measure. The optimal number of mixtures was found to be 8 for the baseline system.

4.3 RESULTS

The system is optimised to achieve a baseline phoneme accuracy of 79.57% and a corresponding word accuracy of 64.50%. As a measure of statistical significance, the standard deviation of the mean is calculated across the 10 cross-validations, resulting in 0.07% and 0.13% for phoneme and word accuracy respectively.

Table 4.2: Results of ASR system using selected penalties

Penalty	Accuracy
-100	62.76 %
-75	64.12 %
-65	64.45 %
-62.5	64.5 %
-60	64.56 %
-57.5	64.5 %
-55	64.49 %
-50	64.41 %
-40	63.9 %
-25	62.39 %
0	54.59 %
50	-144.71 %
100	-530.03 %

Table 4.3: Results of ASR system using selected Gaussian mixture quantities

Number of Mixtures	Phone Accuracy
2	71.29 %
3	72.10 %
4	74.10 %
5	75.70 %
6	76.86 %
7	77.57 %
8	77.81 %
9	77.70 %
10	77.34 %

4.4 COMPARING THE VERIFIED AND UNVERIFIED DICTIONARIES

An ASR system is implemented to test the functionality of the G2P verification process described in Section 3.5.9.

For possible improvement of the ASR system, a dictionary built using BEEP but containing only 1511 entries that appear in the data is verified. The G2P rule extraction technique described in Section 3.5.9 is implemented to isolate entries for manual attention. 498 entries are flagged, of which 33 entries are removed and 3 entries are corrected.

The ASR system is trained with both the unverified dictionary and the verified one. The word accuracy of this system with an unverified dictionary is 64.50%, and the accuracy did not change with verification, even though 2.18% of the dictionary is removed.

4.5 CONCLUSION

The ASR system is implemented to be as optimal as possible, while at the same time implementing standard system architecture. For example, 3 state HMMs are implemented, modelling context-dependent triphones, with 8 Gaussian Mixtures per state. The HMMs and Gaussian mixtures in them are standard, however, the number of Gaussian Mixtures was optimised specifically for this ASR system. The system that is outlined in this chapter is the standard baseline system that will be used as the starting point for experiments for the remainder of this thesis.

CHAPTER FIVE

DIPHTHONG ANALYSIS

5.1 INTRODUCTION

A diphthong is a sound that begins with one vowel and ends with another. Because the transition between the vowels is smooth, it typically is modelled as a single phoneme. However, since it would also have been possible to construct a diphthong using smaller units that are already part of the vowel system, this may be an inefficient representation.

This chapter focuses on the use of diphthongs in SSAE. This is an interesting and challenging starting point to an acoustic analysis of SSAE, but also a good starting point to analysing the requirement for phonemic distinctions. The specific interest in diphthongs is inspired by the scarcity of some of these sounds (such as /OY/ and /UA/, using ARPABET notation (Appendix A)) in spoken English and that large corpora are required to include sufficient samples thereof.

This chapter is aimed at providing a linguistic insight into the role that diphthongs play in the phoneme inventory of SSAE. The need for diphthongs in a lexicon is evaluated by systematically replacing them with selected variants and analysing the system results. Data driven and knowledge based techniques of identifying variants are compared using the ASR system. Finally, once optimal diphthong replacements are found, a data limitation experiment is implemented, to measure the reaction and robustness of the baseline and new system (which does not contain diphthongs) to increasingly limited data sources. These experiments together allow for an understanding of the function of diphthongs in SSAE to be gained.

This chapter describes the steps followed in the analysis of diphthongs in SSAE. Section 5.2 describes how replacement suggestions for diphthongs can be automated and how they can be filtered using an ASR system. Section 5.3 describes how these diphthong replacements are evaluated. Section 5.4 describes the replacement of all diphthongs in the pronunciation dictionary, and this system is used in Section 5.5, to perform data limitation experiments. Concluding remarks can be found in Section 5.6.

5.2 AUTOMATIC SUGGESTION OF VARIANTS

5.2.1 APPROACH

In order to identify possible alternatives (or variants) for a single diphthong, the following process is proposed:

1. An ASR system is trained as described in more detail in Chapter 4. The system is trained using all the data available and the default dictionary containing the original diphthongs.

2. The default dictionary is expanded: variant pronunciations are added to words containing the diphthong in question by replacing the diphthong with all vowels and combinations of two vowels. Two glides (the sounds */W/* and */Y/*) are considered as part of the vowel set for the purpose of this experiment.

3. The original diphthong is removed completely, so that the dictionary only contains possible substitutions. The order of the substitutions is randomised in every word. This ensures that the speech that would represent the diphthong is not consistently labelled as one of the possible substitutions and the training process is therefore not biased in a certain direction.

4. The ASR system is used to force align the data using the options provided by the new dictionary. (Since the diphthong has been removed, the system now has to select the best option from the alternatives that remain.)

5. The forced alignment using the expanded dictionary (alignment B) is compared to the forced alignment using the default dictionary (alignment A):

   - Each time the diphthong in question is found in alignment A, it and its surrounding phonemes are compared to the phonemes recognised at the same time interval in alignment B. The phonemes in alignment B that align with the diphthong in alignment A are noted as possible alternatives to the specific diphthong.
   - The alternatives are counted and sorted by order of frequency.

6. The frequency sorted list is perused and three to five possible replacements for the diphthong are selected by a human verifier from the top 20 candidates. The human verifier is required to assist the system because they are equipped with SSAE and general linguistic knowledge, and are thus able to select replacement candidates that contain vowels or vowel combinations that are more likely to be replacements for the diphthong in question. For example, as a diphthong typically consists of two or more vowels linked together, it is quite likely that the best alternative to a diphthong is a combination of two vowels (diphone). Even though an ASR system may not initially lean towards such a double vowel replacement, including such an alternative may be forced by the human verifier. Also, knowledge-based linguistically motivated choices may be introduced at this stage. These choices are motivated by linguistic definitions of diphthongs as well as SAE variant definitions supplied in Kortmann and Schneider (2004).

Once this process is completed, a list of possible replacements is produced. This list is based on a combination of system suggestions and human selection.

5.2.2 RESULTS

5.2.2.1 DIPHTHONG ANALYSIS: /AY/

The phoneme /AY/ is processed as described in Section 5.2.1. Table 5.1 lists the top 20 candidates and their selection percentage, with the selected replacements marked in bold font. The variants selected for evaluation are all phonetically close or identical to either the standard British English definition of the diphthong /AY/ ((online), 1996; Armstrong, 1996), or the SAE variant definitions of the diphthong (Kortmann and Schneider, 2004). It is interesting to note that both the standard British definition of the diphthong and one of the SAE variant pronunciations both appear in the top 10 replacements. Five sounds are selected for evaluation of their ability to model the phoneme /AY/.

Table 5.1: Results of the automatic variant suggestion experiment for the diphthong /AY/

Replacement	Percentage	Replacement	Percentage
/AE/	0.113	/AA AE/	0.025
/AH/	**0.110**	/UW AH/	0.021
/OH AA/	0.073	**/AE IY/**	**0.018**
/W AA/	0.047	/AH EH/	0.017
/AA/	**0.042**	/W AH/	0.017
/AH AA/	0.041	/OH AH/	0.015
/AH IH/	**0.038**	**/AH IY/**	**0.015**
/OH/	0.037	/ER/	0.014
/AX AA/	0.032	/AE UW/	0.013
/AH AE/	0.028	/AH UW/	0.012

5.2.2.2 DIPHTHONG ANALYSIS: /EY/

The phoneme /EY/ is also processed as described in Section 5.2.1. Table 5.2, similarly to Table 5.1, lists the top 20 candidates and their selection percentage, with the selected replacements marked in bold font. The selected variants for evaluation are again all phonetically close or identical to either the standard British English definition of the diphthong /EY/ ((online), 1996; Armstrong, 1996), or the SAE variant definitions of the diphthong (Kortmann and Schneider, 2004). Although the exact pronunciations of British English and SAE variants are not reflected in the top 20 candidates, sounds that are phonetically very close to them do appear there. Also, five sounds are selected for evaluation of their ability to model the phoneme /EY/. In addition, the linguistically motivated /EH IH/ is added as a sixth sound.

Table 5.2: Results of the automatic variant suggestion experiment for the diphthong /EY/

Replacement	Percentage	Replacement	Percentage
/EH/	**0.088**	/IH/	0.024
/AE IY/	**0.081**	/AE UW/	0.021
/ER/	0.072	/ER UW/	0.016
/AE/	**0.069**	/AH UW/	0.015
/AE IH/	**0.065**	/AE EH/	0.015
/AH IH/	0.045	/IY/	0.015
/EH IY/	**0.032**	/AE UH/	0.015
/ER IY/	0.031	/IH IY/	0.014
/AH EH/	0.029	/Y IY/	0.014
/AH IY/	0.027	**/EH IH/**	0.014

5.2.2.3 DIPHTHONG ANALYSIS: /EA/

The phoneme /EA/ is processed as described in Section 5.2.1. Table 5.3 lists the top 20 candidates and their selection percentage, with the selected replacements marked in bold font. The selected variants for evaluation are all phonetically close or identical to either the standard British English definition of the diphthong /EA/ ((online), 1996; Armstrong, 1996), or the SAE variant definitions of the diphthong (Kortmann and Schneider, 2004). Once again, it is promising to note that some of the sounds appearing in the top 20 list of replacements represent, or are close to representing the definitions of the diphthong /EA/. Only three sounds are selected to evaluation of their ability to model the diphthong /EA/. In addition, the linguistically motivated /EH AX/ is added as a fourth sound.

Table 5.3: Results of the automatic variant suggestion experiment for the diphthong /EA/

Replacement	Percentage	Replacement	Percentage
/EH/	**0.098**	/UW/	0.022
/ER/	0.053	/IH ER/	0.022
/IY/	0.041	/UW Y/	0.020
/EH IY/	0.039	/IY Y/	0.019
/IH IY/	0.037	/Y UH/	0.019
/Y/	0.036	/ER IY/	0.018
/Y EH/	0.033	/Y AX/	0.016
/AE/	**0.027**	/Y IY/	0.015
/IH Y/	0.025	/IY EH/	0.013
/IH EH/	**0.022**	/Y IH/	0.012

5.2.2.4 DIPHTHONG ANALYSIS: /OW/

The phoneme /OW/ is processed as described in Section 5.2.1. Table 5.4 lists the top 20 candidates and their selection percentage, with the selected replacements marked in bold font. The selected variants for evaluation are all phonetically close or identical to either the standard British English definition of the diphthong /OW/ ((online), 1996; Armstrong, 1996), or the SAE variant definitions of the diphthong (Kortmann and Schneider, 2004). It is interesting to note the all of the top four selected sounds are phonetically close to a definition of the diphthong /OW/. All together, five sounds are selected for evaluation of their ability to model the diphthong /OW/.

Table 5.4: Results of the automatic variant suggestion experiment for the diphthong /OW/

Replacement	Percentage	Replacement	Percentage
/ER/	**0.112**	/AO/	0.019
/AE/	**0.061**	/AE UH/	0.017
/ER UW/	**0.050**	/AX/	0.017
/OH/	**0.050**	/UH/	0.016
/AA/	0.033	/AH AA/	0.015
/AX AA/	0.030	/ER UH/	0.015
/AE UW/	**0.025**	/AE ER/	0.015
/Y ER/	0.023	/UW/	0.015
/AH ER/	0.023	/OH AA/	0.015
/OH AO/	0.021	/OH IH/	0.014

5.3 EVALUATING REPLACEMENT OPTIONS

5.3.1 APPROACH

Once a list of three to five possible replacements has been selected for each diphthong, these replacements can be evaluated for their ability to replace the diphthong in question. Per diphthong, the following process is followed:

1. The default dictionary is expanded to include the selected alternatives as variants for the diphthong in question. The pronunciation with the diphthong is removed and the alternative pronunciations are randomised in order not to bias the system towards one pronunciation (as again, the system initially trains on the first occurring pronunciation of every word).

2. Each time the diphthong is replaced by an alternative, a list is kept of all words and pronunciations added.

3. An ASR system is trained on all the data using the expanded dictionary, and the alignments produced during training are analysed.

4. The pronunciations in the forced alignment are compared to each of the lists of added alternatives in turn, calculating the number of times the predicted pronunciation is used in the forced alignment, resulting in an occurrence percentage for each possible replacement.

5. Using these occurrence percentages, the top performing alternatives are selected. The number of selections is not specified, but is rather decided by a human verifier. The human verifier is required to take into account that diphones are likely to replace diphthongs (reported in Section 5.2.1) as well as linguistic knowledge.

6. This process is repeated until only a single alternative remains, or no significant distinction can be made between two alternatives.

7. After each iteration of this process, the ASR phoneme and word accuracies are monitored.

8. In order to test the diphone theory detailed in Section 5.2.1, if a diphone does not constitute one of the final two selected alternatives, the final selected alternative is implemented along with the diphone alternative in order to see whether the ASR system selects the diphone as a better model for a diphthong.

5.3.2 RESULTS

In this section the results of analysing a number of diphthongs individually according to the process described in the previous section (Section 5.2.1) are described.

Since training the full system outlined in Chapter 4 is highly time consuming, a first experiment is performed to determine whether a monophone-based system is sufficient to use during the process to identify and evaluate replacement options. For each diphthong investigated, a dictionary is compiled as described in Section 5.2, a full system is trained using this dictionary, and its forced alignment output when using monophone models with no addition gaussian mixtures is compared with its forced alignment output when using triphone models with 8 mixtures. This comparison always resulted in an equivalence of more than 95%. Therefore, from here onwards, only monophone alignment is used for decision making, while final accuracies, or selection rates, are reported on using the full triphone system.

5.3.2.1 DIPHTHONG ANALYSIS: /AY/

The AY diphthong is first to be analysed. The results of the analysis are summarised in Table 5.5. Each line represents one experiment. For each experiment, the accuracies of each of the included alternatives are noted, as well as the cross validated phoneme and word accuracies of the full ASR system and standard deviation of the mean of their accuracies across the 10 validations (σ_{10}) [1].

The progression of this experiment is outlined below:

- In the first iteration, the alternatives /AH/, /AH IH/ and /AA/ achieve the highest accuracies and are selected for the next round. /AH/ achieves the highest selection rate overall.
- In the second iteration, the alternatives /AH/ and /AA/ achieve the highest accuracies and are selected for the next round. Again, /AH/ has the highest selection rate. All diphones have now been eliminated.

[1] If σ is the standard deviation of n measurements, $\sigma_n = \sigma/\sqrt{n}$ where σ_n measures the standard deviation of the resulting mean calculated over all n experiments

- In the third iteration, /AH/ has the highest selection rate and is therefore selected as the final and best alternative for /AY/.
- In the fourth iteration, /AH/ is tested as a replacement of /AY/. Phoneme accuracy rises to its highest, however, word accuracy suffers. As phoneme accuracy is influenced by the change in number of phonemes (from one experiment to another), word accuracy is the more reliable measure for this experiment.
- The diphone theory, detailed in Section 5.2, suggests that, because diphthongs are made up of two sounds, their replacement must also consist of two sounds in order to have the capacity to model them accurately. In order to test this theory, an iteration is run with /AH/ and /AH IH/ as the alternatives for /AY/. The ASR system still selects the /AH/ alternative over the /AH IH/ alternative. However, the word accuracy increases at this iteration, implying that perhaps having /AH IH/ as an alternative pronunciation for /AY/ fits the acoustic data better than only having /AH/.
- A final iteration is run with the knowledge-based linguistically motivated choice /AH IH/ as the replacement of /AY/. Both phoneme and word accuracy rise to their highest values with this replacement. This shows that the linguistically predicted /AH IH/ is indeed the best replacement for /AY/.

Table 5.5: Results of the variant evaluation experiments for the diphthong /AY/

	/AH/	/AA/	/AH IH/	/AE IY/	/AH IY/	P Acc	P σ_{10}	W Acc	W σ_{10}
1	0.46	0.20	0.18	0.08	0.07	78.51%	0.27	63.88%	0.38
2	0.46	0.36	0.17	N/A	N/A	78.75%	0.23	64.06%	0.40
3	0.56	0.43	N/A	N/A	N/A	79.14%	0.24	64.17%	0.41
4	1	N/A	N/A	N/A	N/A	79.56%	0.23	64.03%	0.40
5	0.62	N/A	0.38	N/A	N/A	79.19%	0.21	64.13%	0.35
6	N/A	N/A	1	N/A	N/A	79.77%	0.23	64.30%	0.38

5.3.2.2 *DIPHTHONG ANALYSIS: /EY/*

The /EY/ diphthong is analysed using the technique outlined in Section 5.3.1. The results are summarised in Table 5.6. In the first iteration, /AE/ and /EH/ are clearly the better candidates, but the diphone (double vowel) scores are lower and very similar. Thus, for the second iteration, all diphones are cut and only /AE/ and /EH/ are tested. But for the third iteration, testing the necessity of including a diphone, two of the diphones are introduced back to be tested again. It should be noted that the highest word accuracy achieved for the suggested variants is achieved in the third iteration, suggesting that diphones are indeed necessary when attempting to replace a diphthong. Again, the highest accuracy achieved overall is for the knowledge-based linguistically suggested alternative /EH IH/.

Table 5.6: Results of the variant evaluation experiments for the diphthong /EY/

	/AE/	/EH/	/AE IY/	/AE IH/	/EH IY/	/EH IH/	P Acc	P σ_{10}	W Acc	W σ_{10}
1	0.24	0.25	0.17	0.17	0.16	N/A	78.97%	0.24	64.27%	0.41
2	0.59	0.41	N/A	N/A	N/A	N/A	79.30%	0.23	64.03%	0.35
3	0.48	N/A	0.26	0.27	N/A	N/A	79.36%	0.25	64.41%	0.42
4	1	N/A	N/A	N/A	N/A	N/A	79.64%	0.26	64.04%	0.35
5	N/A	N/A	N/A	N/A	N/A	1	79.78%	0.24	64.43%	0.45

5.3.2.3 DIPHTHONG ANALYSIS: /EA/

The /EA/ diphthong is now analysed. The results of the experiment are summarised in Table 5.7. These results behave quite differently compared to the other diphthong experiments. The first iteration, where all 3 of the variant options are included, achieves the highest word accuracy, even higher than the iteration which makes use of linguistic knowledge. The phoneme accuracy, however, increases with every iteration, reaching its peak with the use of the linguistic replacement. Again, this may be related to the change in number of phones (in words causing errors) which makes word accuracy a more reliable measure.

Table 5.7: Results of the variant evaluation experiments for the diphthong /EA/

	/EH/	/IH EH/	/AE/	/EH AX/	P Acc	P σ_{10}	W Acc	W σ_{10}
1	0.51	0.34	0.15	N/A	79.22%	0.23	64.49%	0.37
2	0.72	0.28	N/A	N/A	79.51%	0.26	64.43%	0.41
3	1	N/A	N/A	N/A	79.65%	0.24	64.21%	0.45
4	N/A	N/A	N/A	1	79.73%	0.22	64.30%	0.38

5.3.2.4 DIPHTHONG ANALYSIS: /OW/

The experiment is repeated for the diphthong /OW/. The results for the experiment are outlined in Table 5.8. The phoneme accuracy follows a similar pattern to the earlier experiments. The word accuracy is highest at both iteration three, where a diphone is included and iteration five, where the linguistic knowledge-based replacement is implemented. The knowledge-based linguistic replacement once again achieves the highest phoneme and word accuracies.

Table 5.8: Results of the variant evaluation experiments for the diphthong /OW/

	/OH/	/ER/	/ER UW/	/AE/	/AE UW/	/AX UH/	P Acc	P σ_{10}	W Acc	W σ_{10}
1	0.29	0.36	0.14	0.13	0.08	N/A	79.34%	0.25	64.37%	0.38
2	0.52	0.48	N/A	N/A	N/A	N/A	79.57%	0.27	64.41%	0.42
3	0.59	N/A	0.41	N/A	N/A	N/A	79.43%	0.24	64.52%	0.44
4	1	N/A	N/A	N/A	N/A	N/A	79.61%	0.24	64.41%	0.43
5	N/A	N/A	N/A	N/A	N/A	1	79.72%	0.24	64.57%	0.43

5.4 SYSTEMATIC REPLACEMENT OF ALL DIPHTHONGS

5.4.1 ACCURACY RESULTS

Given the results achieved in the earlier experiments, a final experiment is run where all the diphthongs are replaced using a systematic system based on the linguistic definitions of the individual diphthongs.

Two full ASR systems are used, designed as described in Chapter 4. These two systems differ only with regard to their dictionary. One system (system A) uses the baseline dictionary. In the other (system B), the diphthongs in the baseline dictionary are all replaced with their diphone definitions, using British English definitions defined in Table 5.9.

All results are cross-validated and the two systems are compared using their word accuracies. Interestingly word accuracy decreases only very slightly: from 64.53% for system A to 64.35% for system B. The removal of 8 diphthongs is therefore not harmful to the accuracy of the system. This is an interesting result, especially as the detailed analysis is only performed for 4 of the diphthongs and further optimisation may be possible.

Table 5.9: IPA based diphthong replacements

Diphthong	Diphone	Diphthong	Diphone
/AY/	/AH IH/	**/OY/**	/OH IH/
/EY/	/EH IH/	**/AW/**	/AH UH/
/EA/	/EH AX/	**/IA/**	/IH AX/
/OW/	/AX UH/	**/UA/**	/UH AX/

5.4.2 FURTHER ANALYSIS

Additional results can be found in Appendix C, including a confusion matrix for the baseline system (Tables C.1 and C.2) and a confusion matrix for the knowledge-based system (Tables C.3 and C.4). For the analysis of the confusion matrices, the confusion matrix of the knowledge-based ASR system is subtracted from the baseline ASR system's confusion matrix. In order to view only the significant results, all the cells are filtered using a confidence interval equation (equation 5.1). The confidence interval measures the expected deviation of a measurement. This equation is calculated per phoneme, and each change in accuracy in the confusion matrix for that phoneme is tested against the value of the confidence interval and only displayed if it it falls outside this interval.

If n indicates the total number of occurrences of a phoneme, *correct* the number of times a phoneme is correctly recognised and the correctness of a phoneme (p) is calculated as:

$$\text{p} = \frac{correct}{n}$$

Then the confidence interval is calculated as:

$$CI \quad = \quad p \pm \sqrt{\frac{p(1-p)}{n}} \tag{5.1}$$

The results are shown in Table C.5 and Table C.6. The diagonal cells are highlighted in bold font, these are the percentages that each phoneme is recognised correctly. The following points should be noted:

- The accuracy of the phoneme /AH/ decreases considerably, as it becomes more easily confused with other phonemes. This can be explained by the phoneme data becoming more variable due to the diphthongs that now partially contribute to it.
- The phonemes /IH/ and /UH/ become less confusable with other phonemes, probably due to the added data that they are now trained on. However, both still suffer a little due to the extra variance that is introduced into their acoustic models by the new data.
- Although other phoneme accuracies and confusions are affected by the disposal of the diphthongs, the changes are smaller than 5 % of the total occurrences of the phoneme.

5.5 DATA LIMITATION

Two ASR systems have now been developed (system A with diphthongs, and system B without diphthongs) but their difference has only been analysed with regard to their overall accuracy. Data limiting experiments have allowed for insight into the behaviour of different phonemes (Badenhorst and Davel, 2008). The analysis of the two systems with different amounts of data is important because it provides an understanding of the training process of both systems, and their utilisation of the available data.

5.5.1 APPROACH

In this experiment the available training data is artificially limited in order to determine trends during training. Cross validation is not implemented in this experiment due to the standard deviation of the mean in the main results being 0.13, which is low enough to assume a certain standardisation among iterations of the cross validation. In order to ensure comparative results among runs, a larger test set is used than for the cross-validated experiments.

The approach that is followed is described in detail below.

1. The list of speakers that is in the data is randomised.
2. This list is split into two sets, 80% and 20% for training and testing respectively. The specific values for the sets are:
   - Training set: 5 863 calls, 15 566 utterances.
   - Test set: 1 466 calls, 3 693 utterances.
3. Now the training set is split further. Five different training lists are generated, each containing a percentage of the original list: 100%, 75%, 50%, 25% and 10%. The details of the different sets are:
   - 75 % set: 4 398 calls, 11 599 utterances.
   - 50 % set: 2 932 calls, 7 750 utterances.
   - 25 % set: 1 466 calls, 3 892 utterances.
   - 10 % set: 587 calls, 1 587 utterances.
4. Both system A and system B are trained on each training list in turn, and tested on the test set.
5. The results of the two systems are compared.

5.5.2 RESULTS

Table 5.10 details the results achieved in the implementation of data limiting in the ASR systems A (system trained with diphthongs) and B (system trained without diphthongs, using knowledge-based replacements). The table details the phone and word accuracy achieved using each system, as well as the relative difference in error rate between the two systems (system A with diphthongs, and system B without diphthongs). As can be seen from the results, system B maintains a small but constant lead over system A as the amount of data used to train both systems is decreased to 10% of the original training set.

The nature of the diphthong replacement is also investigated through the removal of data. If the replacements of the diphthongs are not correct, and are trained on different amounts of data, the system would be expected to deteriorate with more data, as diphthong data would be added to the replacing phonemes incorrectly, thereby adding noise and instability to those phonemes. However, as more data is added to the system, the accuracy remains stable. This implies that with more of the diphthong data added to the phonemes that have replaced them (which adds some variance to those phonemes) the phoneme models remain stable.

5.6 CONCLUSION

The aim of this chapter is to gain insight into the use of diphthongs in SSAE. A data-driven process is defined through which diphthongs could automatically be replaced with optimal phonemes or phoneme combinations. To complement this process, a knowledge-based experiment is set up using linguistic data for British English.

Table 5.10: Results for data limiting experiment for baseline and knowledge-based ASR systems

Data	Phn Acc Base	Phn Acc Know	Phn Acc Rel Diff	Wrd Acc Base	Wrd Acc Know	Wrd Acc Rel Diff
100	78.42	79.81	6.44 %	61.76	62.32	1.46 %
75	77.66	79.35	7.57 %	60.63	61.56	2.36 %
50	76.80	78.08	5.51 %	59.43	59.47	0.10 %
25	74.03	75.42	5.35 %	55.21	55.61	0.89 %
10	65.48	67.77	6.63 %	45.48	45.76	0.51 %

Although the data-driven method is partially successful in finding the best replacement for diphthongs, the knowledge-based method is superior. Through the implementation of the full knowledge-based system a surprising result is found: diphthongs modelled in the baseline system are found to be superfluous. This result is surprising because diphthongs make up 18 % of the total phoneme count of the baseline system.

The stability of the knowledge-based system is investigated as well, and is found to be sound. However, even though the knowledge-based system is found to model diphthongs better than the data-driven system, the increase in accuracy is small enough that if prior linguistic knowledge is not available, the data-driven technique can be used quite effectively.

It is interesting to consider the South African English variants of diphthongs that are described in (Kortmann and Schneider, 2004). The variants described here or ones close to them always appear on the list of the top candidates of the data-driven selection. This in itself is an interesting observation from a linguistic perspective.

From a technical perspective, the removal of diphthongs simplifies further analysis of SSAE vowels. Our initial investigations were complicated by the confusability between diphthongs and vowel pairs, and this effect can now be circumvented without compromising the precision of the results.

CHAPTER SIX

DIALECT ADAPTATION OF THE KIT VOWEL

6.1 INTRODUCTION

The KIT vowel[1] in British English (BE) exhibits curious behaviour in SSAE, referred to as the 'KIT split'. It experiences allophonic variations that overlap with other phonemes, sometimes to the point where it is indistinguishable from them. Examples of such words include the word ABRUPTNESS, with the British English pronunciation /AX B R AH P T N IH S/, and the SSAE pronunciation of /AX B R AH P T N AX S/. Another example is the word EXACT, with the BE pronunciation /IH G Z AE K T/, and the SSAE pronunciation of /EH G Z AE K T/. However, sometimes the BE pronunciation is the same as the SSAE one, as can be seen with the word IN, with the pronunciation /IH N/ in both BE and SSAE. The analysis of the relationship between SSAE and BE is thus initiated with the analysis of the KIT vowel.

This chapter focuses on extracting a definitive set of rules governing the realisation of the KIT vowel in SSAE by testing known phenomena as well as formulating new rules. The known adaptation environment rules from literature are implemented and analysed initially, following which new rules are formulated to complement the previously implemented ones in order to optimally predict the realisation of the KIT vowel in SSAE.

This chapter describes the experimental procedure followed in adapting a BE pronunciation dictionary to SSAE, through the adaptation of the KIT vowel. Section 6.2 gives some background on the KIT vowel and highlights prior research that has been performed on predicting the behaviour of the KIT vowel. Section 6.3 describes the experimental setup and the approach followed in this experiment. Section 6.4 describes the known and formulated adaptation rules that are implemented on the BE pronunciation dictionary, and the results of implementing these rules. Section 6.5 shows the validation of the results achieved in previous sections using a validation set that is kept separate from the previous experiments. Section 6.6 describes the implementation of the pronunciation dictionaries in an ASR system, to analyse the effect of the adapted differences. Finally, Section 6.7 discusses the results achieved in the experiment.

[1] The KIT vowel, or the 'short I' is part of Well's lexical sets for describing "The lexical incidence of vowels in all the many accents [of English]"(Wells, 1982). A lexical set is described by a set of words that tend to exhibit similar "within dialect " behaviour, but many differ between dialects.

6.2 THE KIT VOWEL IN SSAE

The KIT vowel undergoes a large transformation between British English and SSAE, due to its variable behaviour in SSAE. The source of the variation of the KIT vowel in SSAE, including the time of its introduction to SAE is not definitive. Bekker (2009) finds that there is not a perfect split in the pronunciation of the KIT vowel, but rather a variation known as 'shading'. He does find that there is enough evidence to support a schwa (/AX/) like pronunciation for the KIT vowel under certain circumstances in SSAE. He makes use of the tokens defined by Webb (1983), which are outlined in Table 6.1 with sample words (the ARPABET symbols are not used by Webb, and are added for the sake of clarity. See section 6.4.1 for more detail).

Table 6.1: Known KIT allophones identified by Webb (1983)

Allophone	**Environment(s)**	**Sample Word(s)**	**Closest ARPABET Symbol**
KIT1	disyllabic	**ci**ty,**si**lly	/AX/
KIT2	ʔ_;h_;in context of velars	sing,hit,kit,it	/IH/
KIT3	before palato-alveolars	bitch,dish,fish	/IH/
KIT4	unmarked	chin, sit	/AX/
KIT5	l_; ɹ_;in context of labials	rid,bit,lit,limp,rim,lip	/AX/
KIT6	w_;_ ɫ	till,with,fill,pill	/AX/

6.3 EXPERIMENTAL SETUP

6.3.1 PRONUNCIATION DICTIONARY

The CELEX British English pronunciation dictionary is utilised for this experiment (Baayen *et al.*, 1995). The dictionary is selected due to its accurate syllable boundary and stress markings, linguistic information required during this analysis (BEEP used in earlier chapters does not contain this additional information). The source of the English data in the CELEX dictionary is the Oxford Advanced Learner's Dictionary (1974) and the Longman Dictionary of Contemporary English (1978). The dictionary contains 90 400 unique entries, 49 282 of which contain the /IH/ phoneme. In total there are 68 295 /IH/ phonemes in the whole dictionary.

6.3.2 APPROACH

The approach followed in this experiment is outlined below.

1. **Creating a development dictionary and a validation dictionary**

   - The same steps are followed in putting together the development dictionary and validation dictionary. However, the development set is edited and corrected during the experiment. The validation set is not edited after its initial development.
   - 400 words (200 each for the development and validation sets) that have pronunciations which contain the /IH/ phoneme are randomly selected from the pronunciation dictionary.
   - The /IH/ phonemes in the words are annotated by experienced SSAE speakers (2 Gauteng dwelling L1 English speakers between the ages of 35 and 45) with the perceived SSAE pronunciation. The /IH/ sounds in each word are analysed and either kept, or replaced with either /AX/ or /EH/.
   - This forms a development and validation set, each of 200 words respectively, assumed to be pronounced correctly.

2. **Developing the rule set: Rule implementation**

   - **Known adaptation rules**
     - Known adaptation rules are found in literature and applied to the pronunciation dictionary to form a stable and optimal first layer of adaptation rules on which new experimental rules can be implemented.
     - Each rule is implemented on every /IH/ appearing in the development set.
     - Rule implementation allows for an analysis of the functionality of the rule set. The results contain the initial, predicted and correct pronunciations as well as indications of which rules are applied to that specific /IH/ phoneme and whether the final result is correct or not.
     - All known rules are implemented and verified.
     - This rule set is referred to as the known adaptation rule set.
   - **Selected Adaptation Rules**
     - Each rule is verified and only the top selected rules are implemented. The rule system that consists of these rules is referred to as the selected adaptation rules.
   - **Formulated adaptation rules**
     - Formulated adaptation rules are extracted through a thorough analysis of the data (discussed in more detail in Section 6.4.3).
     - Each rule is implemented on every /IH/ appearing in the development set.
     - Formulated adaptation rules are implemented in addition to the selected adaptation rules. This ensures that the cooperability of the two sets of rules is tested.
     - Rule implementation is again designed to allow for a full analysis of the functionality of the rule set to be performed. The results contain the initial, predicted and correct pronunciations as well as indications of which rules are applied to that specific /IH/ phoneme and whether the final result is correct or not.
     - Each rule is applied individually, and again, only the top performing rules are selected.

3. **Developing the rule set: Rule evaluation**

   - Each rule that requires testing on the set is implemented and analysed with regard to five measures:
     - Number of times applied,
     - Number of entries correctly predicted, and
     - Number of redundant applications. (Entries that are predicted correctly, but are predicted correctly by a different rule as well.)
   - Once the validity of a rule is established on the development set, a manual check is run on random samples from the full pronunciation dictionary to verify that the rule is not over specialised to the development set.

4. **Verifying Results**

   - The final optimal set of rules that is selected is applied to the validation set.
   - The results are analysed in order to measure the utility of all the rules.

6.4 KIT VOWEL ADAPTATION RULES

In order to adapt the KIT vowel to SSAE, an investigation is required to assess the realisation of this vowel in different environments. Two approaches are applied, a knowledge-based approach and a data-driven approach. For the knowledge-based approach, expert suggestions are implemented and analysed. For the data-driven approach, the data is analysed and attempts are made to complement the knowledge-based rules in order to achieve optimal accuracy.

It is assumed for these experiments that the realisation of the KIT vowel will vary between the /IH/, /AX/ and /EH/ phonemes. This assumption is based on an analysis of the input data as well as the acoustic analysis of the KIT vowel performed in Bekker (2009).

The list of 200 entries containing /IH/ from the CELEX dictionary contains 259 /IH/ phonemes. These entries have been filtered manually by splitting all double words in order not to affect the experiment. (Examples of these words include CAPTIVE BALLOON and FAMILY PLANNING.)

6.4.1 KNOWN ADAPTATION RULES

Bekker (2009) performs an acoustic analysis based on the assumption that 6 allophones of the KIT vowel exist. Because allophones fall short of describing a full phoneme, and the pronunciation dictionary must contain full phonemes, the phoneme mapping for each allophone is deduced from the affected words of every environment. Table 6.1 shows the allophones along with sample words, from which the realisation of the KIT vowel in those words can be deduced as follows:

- **KIT1** The sample words given for this environment are ***ci**ty* and ***si**lly*, both of which describe the phoneme /AX/.
- **KIT2** The sample words given for this environment are *sing*, *hit*, *kit* and *it*, all of which contain the phoneme /IH/. The environment describing /IH/ occurring after a glottal stop could not be implemented directly as glottal stops are not marked in any of the pronunciation dictionaries available. As the only glottal stops preceding an /IH/ in both the development and validation sets occurs at the beginning of a word, the rule is implemented indirectly by using the beginning of a word as a marker (to flag a glottal stop). Additional processing would be required to deal with words such as “RE-INSTATE”correctly.
- **KIT3** The sample words given for this environment are *bitch*, *fish* and *dish*, all of which describe the phoneme /IH/.
- **KIT4** The sample words given for this environment are *sit* and *chin*, which describe the phoneme /AX/. However, no environment is given for this realisation thus no specific rules could be implemented. KIT4 acts as a default rule, when no other rules apply.
- **KIT5** The sample words given for this environment are *rid*, *bit*, *lit*, *limp*, *rim* and *lip*, all of which describe the phoneme /AX/.
- **KIT6** The sample words given for this environment are *till*, *with*, *fill* and *pill*, all of which describe the phoneme /AX/.

6.4.1.1 ENVIRONMENT ADAPTATION RULES

The following environments are suggested in Bekker (2009), each controlling a different allophonic realisation of the KIT vowel (only defined environments are listed):

- **KIT1** If an /IH/ occurs in the first syllable of a disyllabic word.
- **KIT2** This rule has three parts:
  - /IH/ occurring after a glottal stop (implemented a /IH/ at the beginning of a word),
  - /IH/ occurring after a voiceless glottal fricative (/H/) , and
  - /IH/ in the context of velars. (A velar is defined in English as one of the following sounds: /K/, /G/, /NG/ and /W/.)
- **KIT3** If an /IH/ occurs before palato-alveolars (a palato-alveolar in English is one of the following sounds: /SH/, /ZH/, /JH/, /CH/), it is governed by this rule.
- **KIT5** This rule has three parts:
  - /IH/ occurring after an alveolar lateral approximant (/L/),
  - /IH/ occurring after an alveolar approximant (/R/), and
  - /IH/ in the context of labials. (A labial is defined in English as one of the following sounds: /M/, /B/, /P/, /V/ and /F/.)
- **KIT6** This rule has two parts:
  - /IH/ occurring before a velarised alveolar lateral approximant (/L/) as a syllable coda, and
  - /IH/ occurring after a voiced labio-velar approximant (/W/).

6.4.1.2 RULE IMPLEMENTATION

These environments are then put to the test on the random development set of 200 words extracted from the pronunciation dictionary. The results are detailed in Table 6.2 and described below. The table depicts the number of times each rules is applied, how many times it correctly predicted the realisation of the /IH/ phoneme, how many times the prediction could have been made by different rules, the percentage of correct prediction and the percentage of correct prediction that could not have been completed by different rules.

Because some of the rules have been adapted for implementation purposes and thus the rule set is not identical to the one in Bekker (2009), the adapted rules have been renamed. It should be noted that the rules are implemented and analysed regardless of possible conflicts. In other words, if two rules do conflict, the effect of each is evaluated separately.

- **KIT1**
  - The environment here dictates that the word in question must have two syllables, and that the KIT vowel must appear in the first syllable. This rule is not modified from the initial definition of KIT1.
  - The implementation of this environment finds 25 words matching the requirements. However, the algorithm only manages a 24 % accuracy.
- **KIT2+**
  - This rule is modelled after the KIT2 environment definitions. However, the glottal stop environment is removed from the environment. Glottal stops cannot be used because they are not labelled in the dictionary. The rest if the rule is the same as the definition of the environment for KIT2. (Note however that all words containing glottal stops prior to the /IH/ are accurately identified by the "Beginning of word " rule.)

Table 6.2: Results of /IH/ adaptation for knowledge-based rules

Rule Implemented	# entries	# correct	# redundant	% correct	% correct with redundancy
KIT1:					
Disyllabic	25	6	18	24 %	N/A
KIT2+:					
Beginning of word	24	23	6	100 %	71 %
Following a /H/	5	5	2	100 %	60 %
In the context of /K/	15	13	7	87 %	40 %
In the context of /NG/	41	41	15	100 %	63 %
In the context of /G/	7	6	4	86 %	29 %
In the context of /W/	6	0	2	0 %	0 %
KIT3+:					
Preceding a /SH/	3	3	1	100 %	67 %
Preceding a /N SH/	0	0	0	N/A	N/A
Preceding a /ZH/	0	0	0	N/A	N/A
Preceding a /N ZH/	0	0	0	N/A	N/A
Preceding a /CH/	0	0	0	N/A	N/A
Preceding a /N CH/	0	0	0	N/A	N/A
Preceding a /JH/	1	1	1	100 %	0 %
Preceding a /N JH/	2	2	1	100 %	50 %
KIT5:					
Following a /L/	26	7	1	27 %	23 %
Following a /R/	28	5	4	18 %	4 %
In the context of /M/	11	7	0	64 %	64 %
In the context of /B/	12	5	3	42 %	17 %
In the context of /P/	16	11	0	69 %	69 %
In the context of /F/	10	7	4	70 %	30 %
In the context of /V/	11	7	1	64 %	55 %
KIT6:					
Before tautosyllabic /L/	4	2	0	50 %	50 %
Following a /W/	6	6	0	100 %	100 %

  - A preceding /H/ and beginning of word rules are found to be very good predictors, achieving a 100% accuracy.
  - Contextual velars are mostly very effective, all achieving accuracies higher than 85 %, except for the /W/ phoneme, which is fully inaccurate.

- **KIT3+**
  - This rule is modelled after the KIT3 environment definitions.
  - However, the environments that are described in the literature do not have many occurrences in the data. In fact, only two palato-alveolars featured: /SH/, which had three occurrences, all correctly predicted (only one of which is predicted redundantly), and /JH/, which had only one occurrence, which is also correctly predicted.
  - Lanham and Traill (1962) mention that there may be an intervening /N/ before a palato-alveolar, and that in that case the palato-alveolar would still have the same effect. Thus, one adjustment is made to the KIT3 definition: The addition of a possible intervening /N/ phoneme (described below), which required the the full set (the /IH/ phoneme, and intervening /N/ and the palato-alveolar) to occur within one syllable. This change causes the number of occurrence of /JH/ to increase to three,

which are still fully correct.

- **KIT5**
  - This rule is modelled after the KIT5 environment definitions, no adjustments are made to the known definition.
  - A following /L/ and /R/ do not perform well, neither managing an accuracy over 25 %.
  - Contexts of labials do not perform well either, all achieving between 30 % and 70 % accuracy.
- **KIT6**
  - This rule is modelled after the KIT6 environment definitions, no adjustments are made to the known definition.
  - A following tautosyllabic /L/ achieved quite a low accuracy of 50 %.
  - A preceding /W/ achieved a 100 % accuracy.

6.4.2 SELECTED ADAPTATION RULES

Once the known rules have been analysed, the most accurate and applicable environments need to be selected and adapted to form the known adaptation rule set. This system is required to be as accurate as possible because it would then be utilised in order to find new rules that complement the known ones. If too many inaccurate rules are implemented, the resulting entries that the new rule system must manipulate would be more difficult to analyse, because the adaptations that the /IH/ phoneme instances require will not be natural. The accuracy is not the only selection factor. However, if the accuracy is high the applicability of a rule is also manually tested on the bigger set to ascertain whether the smaller set results may be deceiving.

The rules that are selected as the strongest candidates are listed below. The accuracy and applicability of the other rules are not high enough to warrant implementation.

- **KIT2+** This rule achieves very high accuracies, with the exception of the velar /W/, which is removed from the rule definition. Although the redundance free accuracies are not always that high, this simply means that with other rules eliminated, **KIT2+** will be able to make up for their absence.
- **KIT3+** This rule, although having few occurrences, is quite accurate. Also, the application of this rule on the full pronunciation dictionary is verified manually and found to be applicable.

The final knowledge-based system is implemented using the rules defined above. When no rules are applicable, the KIT vowel is more likely to be realised as an /AX/ than as an /IH/ or /EH/. Therefore instances that do not match any of the rules are changed to /AX/, analogous to the original KIT4 rule. The final known adaptation rule system correctly predicts the realisation of 190 KIT vowels, which gives the system an accuracy of 71 %. A table presenting the detailed results that show the application of every rule to every word can be found in Tables D.1 to D.4 in Appendix D.

6.4.3 FORMULATED ADAPTATION RULES

The formulated rules are designed to complement the selected adaptation rules in building a system to adapt the British KIT vowel to the SSAE KIT vowel. The formulated rules that are selected from an investigation of the analysis set of entries are listed below.

- **Graphemic Structure Analysis.** In order to predict the pronunciations of some KIT vowels correctly, it seems that a morphological analysis is required. The rules implemented here attempt to approximate a morphological analysis. The graphemic structures and the phonemes they map to are analysed in the data set and a set of adaptation environments are created.
  - **Word starting with EB, EM, EN or EX.** In words beginning with the graphemes EB, EM, EN and EX, the KIT vowel appears at the beginning of the word. If the the syllable containing this combination (the first syllable) is closed, the KIT vowel may be pronounced as an /EH/ phoneme.
  - **Word ending with IES, IED and EYS.** Word ending is also a problematic area if morphological analysis is not applied. However, words ending in the grapheme combinations IES, IED, EYS contain the KIT vowel as their second last phoneme, and the realisation of the KIT vowel in these situations is always the /IH/ phoneme.
  - **Ending of word.** When a word ends with the /IH/ phoneme, the adaptation of the phoneme tends to keep it as an /IH/ phoneme.
  - **/IH/ phoneme source.** When the KIT2+ rule is applied, the /IH/ phoneme can be brought about by more than one grapheme (usually I, Y or E), and this grapheme seems to have an effect on the realisation of the /IH/ phoneme in SSAE. When the graphemes I or Y bring about the /IH/ phoneme, the realisation is usually /IH/. When a different grapheme brings about the /IH/ phoneme, the realisation tends to lean towards /AX/. This rule is only implemented as a refinement to **KIT2+**, in order to make **KIT2+** more accurate.

- **Vowel Harmony.** KIT vowels at the end of syllables (open syllables) tend to lean towards a realisation of /IH/. However, simply being at the end of an open syllable does not seem to guarantee this realisation. Vowel harmony occurs when a vowel is assimilated with the vowel following it, even though there may be intervening consonants. English is not a language that is known to contain vowel harmony. However, an analysis of the development set suggests that vowel harmony is at least partially present in SSAE. Entries that are already adapted are not included in this analysis. (This means that only those KIT vowels that have not yet triggered any rules are considered for further adaptation.)
  - **Vowel Harmony /AX/ due to /AX/.** Words containing an /AX/ phoneme as the closest following vowel to the KIT vowel, tends to make the KIT vowel become an /AX/ phoneme as well.
  - **Vowel Harmony /IH/ due to /EY/ adjacent.** Words containing a /EY/ phoneme adjacent to the KIT vowel, tend to make the KIT vowel realise as /IH/.
  - **Vowel Harmony /AX/ due to /EY/ not adjacent.** If an /EY/ phoneme is not adjacent, it makes the KIT vowel tend to be realised as an /AX/ phoneme.
  - **Non Vowel Harmony /AX/.** Words that are not edited by the other vowel harmony rules are mapped to /IH/. This is because /IH/ phonemes at the end of open syllables lean towards a realisation of /IH/, and most of the vowel harmony rules are aimed towards a realisation of /AX/.

6.4.4 FINAL ADAPTATION RULES

The formulated rules are implemented together with the selected adaptation rules to form the final rule system. The rules are implemented in the following order for each instance of an /IH/ phoneme (the order establishes interactions):

1. Firstly, the **Word starting with with EB,EM,EN or EX** rule is applied. If this rules is applied to a word, no other rule is applied.

2. A group of rules is now applied, and their output predictions are compared. If any single rule predicts a /IH/ phoneme realisation of /IH/, the phoneme is kept as /IH/. If any of these rules are applied, no further rule is evaluated. These rules are:
   - **KIT2+** and within it **/IH/ phoneme source**,
   - **KIT3+**, and
   - **First in word**.
   - **Word ending with IES, IED or EYS** is then applied.
   - **Ending of a word** is applied.
3. The vowel harmony rules are applied to open syllables containing the KIT vowel. **Vowel Harmony /AX/ due to /AX/**, **Vowel Harmony /IH/ due to /EY/ Adjacent** and **Vowel Harmony /AX/ due to /EY/ not adjacent** are applied together.
4. Finally, any remaining KIT vowels are realised as /AX/.

6.4.4.1 RULE SET ANALYSIS

The formulated rules were implemented together with the selected adaptation rules to form the final adaptation rule system. The final adaptation system implements both selected adaptation rules and formulated adaptation rules. The system correctly adapts 243 pronunciations of the KIT vowel, making it 95 % accurate. The detailed results for every entry in the analysis list as well as every adaptation can be found in Tables D.5 to D.8 in Appendix D. A full table showing all the results for the different adaptation rule systems is shown in Table 6.3. The table lists the number of entries that the rules are applied to (the /IH/ phonemes not edited by rules are adapted to the /AX/ phoneme), the number of correct entries, and percentage of correct entries. (This percentage is calculated using the total number of /IH/ phonemes, which is 259.)

- The accuracy of the original known rules is slightly lower than the accuracy of the adapted known rules. This is due to the fact that only the rules with the highest accuracies and applicability rates are selected for implementation in the adapted known rule set.
- The formulated rule system and the adapted rule system obtain similar accuracies.
- However, together the formulated rule system and the adapted rule system complement each other and cooperate to adapt the /IH/ phoneme with a 95 % accuracy.

Table 6.3: Results comparison of /IH/ adaptation for known, selected and final adaptation rule systems

Rule system Adaptation	**# entries**	**# correct**	**% correct**
Known Rules Original	259	184	71 %
Selected Adaptation Rules	259	190	73 %
Final Adaptation rules	259	245	95 %

6.4.4.2 ANALYSIS OF ERRORS

Sixteen errors are made by the final adaptation rule system. All the erroneous words and their correct pronunciations are listed below, and if the cause of the error is known it is discussed as well.

Table 6.4: Incorrectly predicted words in final adaptation rule system

Number	Predicted Pronunciation	Correct Pronunciation	Comments
1	B IY AE T IH F AY	B IY AE T AX F AY	The pronunciation of this word is ambiguous. (It is not used in everyday language and is thus difficult to annotate.)
3	K L AH M Z AX N AX S	K L AH M Z IH N AX S	Morphological snalysis is required to predict this pronunciation correctly. (Baseword: Clumsy)
4	D IY M AE G N IH T AY Z D	D IY M AE G N AX T AY Z D	Morphological analysis is required to predict this pronunciation correctly. (Baseword: Magnet)
5	EH L IH F AE N T AY N	EH L AX F AE N T AY N	-
6	JH IH N EH R IH K AX L IH	JH AX N EH R IH K AX L IH	-
7	G IH L IH M OH T S	G IH L AX M OH T S	The pronunciation of this word is ambiguous.
8	M IH T AE L IH K	M AX T AE L IH K	Morphological analysis is required to predict this pronunciation correctly. (Baseword: Metal)
9	M AH Z IH L IH	M AH Z AX L IH	Morphological analysis is required to predict this pronunciation correctly. (Baseword: Muzzle)
10	N AA S T IH L IH	N AA S T AX L IH	-
11	P ER S IH V IA D	P ER S AX V IA D	-
12	P IH T IH NG	P AX T IH NG	Morphological analysis is required to predict this pronunciation correctly. (Baseword: Pit)
13	R IH S T OH R AX T AX V Z	R AX S T OH R AX T AX V Z	The pronunciation of this word is ambiguous.
14	S IH D Y UH S IH NG	S AX D Y UH S IH NG	-
15	S IH L EH K T	S AX L EH K T	The pronunciation of this word is ambiguous.
16	W IH P IA R	W AX P IA R	Morphological analysis is required to predict this pronunciation correctly. (Baseword: Whip)

6.5 VERIFYING RESULTS USING THE VALIDATION SET

Once the final set of rules is developed, the full adaptation rule set is implemented on the validation set of the CELEX pronunciation dictionary. Once the selection is complete, double and hyphenated words are split. The list contains 259 instances of the /IH/ phoneme. After inspection, one of the two sets is selected as the validation set.

6.5.1 INTER-SPEAKER AGREEMENT

This list is annotated by two informants, both of whom are first language English speakers with linguistic backgrounds between the ages of 35 and 45, who reside in Gauteng. Out of a total of 198 words, the speakers disagreed on 5 pronunciations. The speakers thus agreed on 97.5 % of the pronunciations that they labelled.(Labelling was performed independently.)

6.5.2 RESULTS

Table 6.5 depicts the results of the full adaptation rule set when applied to the validation set of the pronunciation dictionary. The detailed table of results can be found in Tables D.9 to D.12 in Appendix D. The final number of correct entries is 241, which equals an accuracy of 93 %.

- The adaptation rules perform similarly to their performance on the development set.
- Many rules achieve 100 % correct rate (excluding redundancies), including **KIT3+**, **Word starting with EB, EM, EN or EX**, **Word ending with IES, IED or EYS** and **Ending of a word**.
- Some rules do still require (further) refinement. These include the **Beginning of word** and **/IH/ phoneme source**.
- One rule is not tested on the validation set, **Vowel Harmony /IH/ due to /EY/ Adjacent**. This rule therefore remains unverified. However, it achieved perfect performance on the development set, and can thus be assumed to be trustworthy.

Table 6.5: Results of /IH/ adaptation using the final adaptation rule set

Rule Implemented	**# entries**	**# corr**	**# redun**	**% corr**	**% corr - redun**
KIT2+	86	82	7	95 %	87 %
KIT3+	7	7	1	100 %	86 %
Word starting with with EB,EM,EN or EX	5	5	0	100 %	100 %
Beginning of word	25	19	4	76 %	60 %
Word ending with IES, IED or EYS	6	6	1	100 %	83 %
Ending of a word	29	29	1	100 %	97 %
/IH/ phoneme source	8	5	0	63 %	63 %
Vowel Harmony /AX/ due to /AX/	18	16	0	89 %	89 %
Vowel Harmony /IH/ due to /EY/ Adjacent	0	0	0	N/A	N/A
Vowel Harmony /AX/ due to /EY/ not adjacent	6	5	0	83 %	83 %
Non Vowel Harmony /IH/	24	20	0	83 %	83 %

All the erroneous words and their correct pronunciations are listed below, and if the error cause is found it is included as well.

Table 6.6: Incorrectly predicted words in validation set

Num	Predicted Pronunciation	Correct Pronunciation	Comments
1	H AH S K IH AX S T	H AH S K IH EH S T	Possible reference set error.
2	H AA B AX N JH AX R	H AA B IH N JH AX R	The pronunciation of this word is ambiguous. (It is not used in everyday language and is thus difficult to annotate.)
3	AE P IH T AY Z IH NG	AE P AX T AY Z IH NG	-
4	IH Z R EY L AX Z	IH Z R EY L IH Z	Morphological analysis is required to predict this pronunciation correctly. (Baseword: Israel)
5	D AX S AX N T ER M AX N T	D AX S IH N T ER M AX N T	Morphological analysis is required to predict this pronunciation correctly. (Baseword: In)
6	IH P IH T AX M IH	EH P AX T AX M IH	The first KIT vowel pronunciation is ambiguous. The second KIT vowel has been made an /AX/ by the proceeding /AX/ incorrectly.
7	D IH S T IH NG K SH AX N	D AX S T IH NG K SH AX N	-
8	R AX M EY N IH NG	R IH M EY N IH NG	The pronunciation of this word is ambiguous.
9	R EH P T IH L IA N Z	R EH P T AX L IA N Z	-
10	P AX N CH IH NG	P IH N CH IH NG	-
11	W AY F L AX AX S T	W AY F L IH EH S T	Morphological analysis is required to predict this pronunciation correctly. (Baseword: Wife)
12	W AX N D Z	W IH N D Z	-
13	B IH CH AX N AX S	B IH CH IH N AX S	Morphological analysis is required to predict this pronunciation correctly. (Baseword: Bitchy)
14	S P R AY T L AX N AX S	S P R AY T L IH N AX S	Morphological analysis is required to predict this pronunciation correctly. (Baseword: Sprightly)
15	M AE S IH D OW N Y AX N Z	M AE S AX D OW N Y AX N Z	-

6.6 ASR SYSTEM RESULTS

An ASR system, as described in Chapter 4, is trained and tested using the adapted dictionary. The phone accuracy of the system is 79.7 % and the word accuracy is 64.5 %. These accuracies are comparable to that achieved by the baseline system. (The phoneme accuracy is slightly but significantly higher than that achieved with the baseline system.)

6.7 CONCLUSION

In this chapter, a set of adaptation rules is developed and implemented on a British English pronunciation dictionary in order to adapt the KIT vowel to Standard South African English. The adaptations developed are 93 % effective in predicting the SSAE pronunciations. The simple environmental rules that are implemented are not always sufficient in modelling the phenomena in the data, and it is expected that more complex techniques such as morphological analysis could result in adaptation that is close to perfect.

One issue encountered with SSAE is that with so many allophones of the KIT vowel existing, the pronunciations of certain words can become ambiguous. Examples of such words include DECEIT and REMAINING that may have two valid pronunciations, each. SSAE is constantly mutating, affected by the varieties of SAE. It is thus not always possible for an annotator to label the pronunciations of words unambiguously. However, the adaptations that are implemented successfully adapted most pronunciations to what is judged to be the closest phoneme to the pronounced allophone.

The final output of the process of adaptation is a pronunciation dictionary with accurate and consistent SSAE pronunciations for the KIT vowel. This dictionary is implemented in an ASR system, and results in a slight improvement of the phoneme accuracy of the system.

CHAPTER SEVEN

CONCLUSION

7.1 INTRODUCTION

The aim of the experiments undertaken in this thesis was the development of a SSAE pronunciation dictionary using a British pronunciation dictionary as a source. The steps undertaken in order to reach this aim were the verification of a British pronunciation dictionary, the analysis of the phoneme distinctions in the pronunciation dictionary, and finally the adaptation of the KIT vowel from British English to SSAE. In this chapter each of these steps is discussed as well as their ability to reach the goals for which they were designed.

7.2 SUMMARY OF CONTRIBUTION

This thesis was able to develop a SSAE pronunciation dictionary for ASR system implementation. Three separate sets of techniques were developed for the verification, validation and adaptation of a pronunciation dictionary. These techniques are described below.

7.2.1 DICTIONARY VERIFICATION

Automatic and semi-automatic pronunciation dictionary verification techniques were investigated and a number of novel techniques were developed and implemented (Martirosian and Davel, 2007). Methods were analysed with regard to their effectiveness in verifying the BEEP dictionary as well as their predicted ability to generalise to other dictionaries. This experiment used both simple and complex analysis techniques: Simple techniques such as comparing the length of words and their pronunciations and complex techniques including the alignment of words and pronunciations and the analysis of their phonemic and graphemic nulls. Although the analysis of generation restriction rules was very efficient, it is only applicable when looking for erroneous pronunciation variants. The analysis of graphemic nulls was the best method investigated, as it is able to identify the most errors, perform efficiently and is likely to generalise well to other pronunciation dictionaries.

The number of errors identified (5 553) was surprisingly large for a fairly widely used pronunciation dictionary. Manual analysis of a random selection from the dictionary indicated that approximately 60% of all conclusive errors were removed.

7.2.2 DIPHTHONG ANALYSIS

Phoneme distinctions were analysed from an ASR point of view through the analysis of diphthongs in SSAE (Martirosian and Davel, 2008). This study analysed knowledge-based and data-driven methods of splitting diphthongs into the phonemes that they are constructed from. The study found that splitting the diphthongs into their defined British English pronunciation phonemes did not deteriorate system performance. Thus the phoneme distinction requiring diphthongs to exist in SSAE was found to be unnecessary from an ASR perspective. This was a surprising result because diphthongs make up a substantial part of the phoneme system.

Our initial investigations were complicated by the confusability between diphthongs and vowel pairs, and this effect can now be circumvented without compromising the precision of the results.

7.2.3 KIT VOWEL ADAPTATION

Finally, a set of rules was developed in order to adapt a British English (BE) pronunciation dictionary to SSAE. The experiment was performed through the implementation of the SSAE 'KIT split' on the British English pronunciation dictionary. The KIT vowel was specifically selected as the 'KIT split' is one of the most significant differences between BE and SSAE. The study found that a set of rules can be implemented to accurately adapt the KIT vowel BE to SSAE. Also, the experiment produced new ideas with regard to how the KIT vowel is affected by the phonemic structure of a word due to vowel harmony, and the graphemic structure thereof due to structural combinations. The set of adaptation rules that was developed, consisted of linguistically motivated rules (these only achieved a 71 % accuracy in development) which were complimented by a new formulated rule set, that brings the total accuracy up to over 90 % on both the development set and an independent validation set.

The final product of this experiment is a BE dictionary with all KIT vowels adapted to SSAE. While initially developed from an ASR perspective, it is expected that this pronunciation dictionary will be particularly useful in the text-to-speech domain, as well.

7.3 FUTURE WORK

This thesis has only begun the exploration into the topics of dictionary verification, phoneme distinction analysis and pronunciation dictionary adaptation. Further directions in these fields of research are discussed below.

- Dictionary verification techniques were analysed in this thesis. However, with additional data the techniques that were developed could undergo additional analysis, especially with regard to their adaptiveness to new environments. Through the implementation of the verification techniques on different pronunciation dictionaries, possibly in different languages, the typical mistakes made by pronunciation dictionary developers can be found and prevented, thus allowing the process of the development of pronunciation dictionaries to be more efficient with little extra effort.

- Phoneme distinction analysis was also implemented in this thesis, and this method can also be utilised well in other languages. For example, the affricates that exist in the Bantu languages can definitely be analysed in this manner, and if the method is successful in splitting the affricates the development of comparable ASR systems for these languages will require less data to achieve the same accuracy, which is important for resource scarce languages.

- The adaptation of the KIT vowel for SSAE was implemented in this thesis. Although the KIT vowel experiences the most variation and is characteristic of SSAE, other phonemes can be analysed to attempt to improve the accuracy of the SSAE pronunciation dictionary further. Specific examples include the merging of the LOT and CLOTH vowels, as well as the splitting of the GOOSE vowel. In addition, morphological analysis may be useful in refining the adaptation rule produced for the KIT vowel. Furthermore, once a dictionary is successfully adapted from British English to SSAE, methods of semi automatic adaptation can be implemented, thus allowing dialects with small usage groups to develop ASR systems for their own native dialect.

7.4 CONCLUSION

This thesis has implemented a number of techniques required for the successful adaptation of a British English pronunciation dictionary to SSAE. Specifically, new results were obtained related to dictionary verification, phoneme redundancy analysis and phoneme adaptation. Final results include both a number of novel techniques and a practically usable SSAE pronunciation dictionary.

ASR technologies are very effective in disseminating information and providing services to people that may not have access to alternative infrastructure. It is thus important to develop techniques that support researchers in rapid dialect adaptation, in order to provide information services to a larger user group in an affordable way. This thesis provides insight into the process of dialect adaptation of a pronunciation dictionary and discusses some additional avenues for future exploration.

APPENDIX A

THE ARPABET PHONE SET

The ARPAbet phone set was developed as part of the ARPA Speech Understanding project (1971-1976), and is included with the TIMIT speech corpus (Garofolo *et al.*, 1993). A more limited set using only 44 of the ARPABET phonemes was used for the BEEP pronunciation dictionary.

Table A.1: The BEEP ARPAbet phone set

phoneme	**example**	**phoneme**	**example**
IY	beat	L	led
IH	bit	R	red
EH	bet	Y	yet
AE	bat	W	wet
AX	the	M	mom
AH	butt	N	non
UW	boot	NG	sing
UH	book	CH	church
AO	horse	JH	judge
OH	cot	ZH	measure
AA	bath	SH	shoe
ER	bird	B	bob
EY	bait	P	pop
AY	bite	D	dad
OY	boy	T	tot
AW	bout	G	gag
OW	goat	K	kick
EA	air	Z	zoo
UA	cure	S	sis
IA	near	V	very
		F	fief
		DH	they
		TH	thief
		HH	hay

APPENDIX B

DICTIONARY VERIFICATION RESULTS

Table B.1: Sample of removed entries in pre-processing step

Word	Pronunciation
G._B.	JH IY B IY
DÉTENTE	D EY T AA N T
DÉMARCHES	D EY M AA SH IH Z
A._R.	EY AX
L._E._A.	EH L IY EY
N._S._P._C._C.	EH N EH S P IY S IY S IY
B._A.	B IY EY
DR.	D OH K T AX R
J._P.	JH EY P IY
G._C._H._Q.	JH IY S IY Y EY CH K Y UW
G._M._B.	JH IY Y EH M B IY
M._P._G.	EH M P IY JH IY
C._I._D.	S IY AY D IY
SÉANCES	S EY AA N S IH Z
BÊTES_NOIRES	B EH T N W AA
PROX.	P R OH K S
D._N._A.	D IY EH N EY
CRÈME_DE_MENTHE	K R EH M D AX M OH N TH
BLASÉ	B L AA Z EY
INGÉNUES	AE N ZH EY N Y UW Z
Y._M._C._A.	W AY EH M S IY EY
P._L._O._D.	P IY Y EH L OW D IY
P._O._W.S	P IY OW D AH B L Y UW Z
N._T.	EH N T IY
I._T._V.	AY T IY V IY
I._R._A.	AY AA R EY
M._A._I._R.	EH M EY Y AY AX
RÉGIMES	R EY ZH IY M Z
G._K._N.	JH IY K EY Y EH N
&AMPERSAND	AE M P ER S AE N D
L._P.	EH L P IY
U._V._F.	Y UW V IY EH F
CRÊPE	K R EY P
C._O._D.	S IY OW D IY
DO.	D IH T OW
F._E._R.	EH F AX
SOUP¡CONS	S UW P S OH N Z
MÉSALLIANCE	M EY Z AE L IH AA N S
G._N._P.	JH IH EH N P IY
S.	EH S
PURÉES	P Y UA R EY Z
J.	JH EY
P._G._A.	P IY JH IY EY
CHARGÉ	SH AA ZH EY
U._N.'S	Y UW EH N Z
LAMÉ	L AA M EY
VIS-À-VIS	V IY Z AA V IY
CURA¡COA	K Y UA R AX S OW
DÉBÂCLES	D EY B AA K L Z
APPLIQUÉ	AE P L IY K EY
V._H._F.	V IY EY CH EH F
H._V._O.	EY CH V IY Y OW
COMPÈRES	K OH M P EA Z
LYCÉES	L IY S EY Z
SAUTÉING	S OW T EY IH NG
B._S._E.	B IY Y EH S IY
MRS.	M IH S IH Z
D'ÊTRE	D EH T R AX
V._P.	V IY P IY
Z.'S	Z EH D Z
MÉLANGES	M EY L AA N ZH IH Z
N._C._O.	EH N S IH OW
S._S.	EH S EH S
CORTÈGE	K AO T EY ZH
L._L._B.	EH L EH L B IY
VICU NAS	V IH K Y UW N AX Z
TÊTE-À-TÊTE	T EY T AA T EY T
SE NOR	S EH N Y AO R
M._B.	EH M B IY
R._S._I.	AA R EH S AY
FA¡CADE	F AX S AA D
RECHERCHÉ	R AX SH EA SH EY
PROTÉGÉES	P R OH T IH ZH EY Z
PREMIÈRE	P R EH M IH EA R
G._P.'S	JH IY P IY Z
P._A.	P IY EY
R._A.S	AA R EY Z
A._N._C.	EY Y EH N S IY
E._R._M.'S	IY Y AA R EH M Z
V._I._P.S	V IY AY P IY Z
FIANCÉES	F IH OH N S EY Z
P._O._W.	P IY OW D AH B L Y UW
SCÈNE	S EY N
P._L._O.	P IY Y EH L OW
SAUTÉS	S OW T EY Z
L._E._A._R.	EH L IY EY AX R
AUTOS-DA-FÉ	AO T OW Z D AA F EY
DÉBUTS	D EY B Y UW Z
B._S.	B IY EH S
C._J._D.	S IY JH EY D IY
O._T.	OW T IY
ABBÉ	AE B EY
ARRIÈRE_PENSÉES	AE R IH EA P OH N S EY Z
B._B._C.'S	B IY B IY S IY Z
PÂTÉS	P AE T EY Z
P._T._A.	P IY T IY EY
F._A._O.	EH F EY OW
PORTIÈRES	P AO T IH EA Z
CLICHÉ	K L IY SH EY
S._N._P.	EH S EH N P IY

Table B.2: Sample of removed entries during removal of repeated phonemes

Word	Pronunciation
MAMMALIANS	M AX M M AA L IA N Z
MAXSON'S	M AE K S S AH N Z
INVESTTECH	IH N V EH S T T EH K
TIPPER'S	T IH P P ER Z
COAT-TAILS	K OW T T EY L S
UNKNOWN	AH N N OW N
SHIRER	SH AY AX AX R
BELLINO	B EH L L AY N OW
CHITTEL	CH IH T T EH L
MADERER	M EY D AX AX
LATHERER	L AA DH AX AX
LEPERER	L EH P AX AX
SPATULAR	S P AE T Y UH L AX AX
SOULLESSLY	S OW L L AX S L IY
PINNELLS	P IH N N EH L Z
WITTENHAM'S	W IH T T EH N HH AE M Z
POPPEA'S	P OH P P IY Z
HUBBART	HH AH B B AA T
FETTERERS	F EH T AX AX Z
HIBBARD	HH IH B B AA D
BOULDERERS	B OW L D AX AX R Z
HAMMACHER'S	HH AE M M AE K HH ER R Z
CHAFFERER	CH AA F AX AX
MACKEY	M AE K K IY
TESTAR	T EH S T AX AX
HUCKSTERER	HH AH K S T AX AX
PAXSON'S	P AE K S S AH N Z
VULGARER	V AH L G AX AX
HALLIER'S	HH AE L L AY AA Z
RUTTER'S	R AH T T ER Z
POT-TRAINED	P OH T T R EY N D
TORRINGFORD	T AO R R IH NG F AO D
TAIL-LIGHT	T EY L L AY T
BRANNON'S	B R AE N N OH N Z
HIBBARDS	HH IH B B AA D Z
TAMPERER	T AE M P AX R AX AX
WHITTINGALE	W IH T T IH NG G EY L
DIMWITTEDNESS	D IH M W IH T T EH D N EH S
TROWELLED	T R AW AX L L EH D
VALLEE'S	V AE L IY IY Z
HODDING	HH OH D D IH NG
PUTTERER	P AH T AX AX R
SHOTTING	SH OH T T IH NG
ADMITTER	AX D M IH T T ER
FOPPIANO	F OH P P IH AE N OW
LIPPERTS	L IH P P ER T S
UNPLAINNESS	AH N P L EY N N AX S
NILLED	N IH L L EH D
PENGASSAN	P EH N G AE S S AE N
GADDIE	G AE D D AY IY
WARRANTECH'S	W AO R R AE N T EH K S
WAITT'S	W EY T T IY Z
GUNNELLS	G AH N N EH L Z
LOCKETT'S	L OH K IH T T IY Z
LIPPARD'S	L IH P P AA D Z
REVICTUALLING	R IY V IH T L L IH NG
LETTIER	L EH T T IA
POSTURER	P OH S CH AX AX R
WHITTON'S	W IH T T AH N Z
HATTING	HH AE T T IH NG
DADDIE	D AE D D AY IY
KANSAS_CITY	K AE N Z AX S S IH T IY
SENNETT	S EH N N EH T
KILOMETRER	K IH L AX M IY T AX AX R
WOOLLIES	W UH L L AY Z
GAULLISTS'	G AO L L IH S T S
TAPPET	T AE P P EH T
SUPPORTOR	S AH P P AO T OW AX R
SOFAER	S OW F AX AX
UNNOTIFIED	AH N N OW T IH F AY D
SITTIN	S IH T T IH N
PANTERER	P AE N T AX AX
MODERNNESS	M OH D N N EH S
CADDELL	K AE D D EH L
PARCELLATED	P AA S L L EY T IH D
CURETTING	K Y UA R EH T T IH NG
REAPPORTIONMENT'S	R IY P P AO SH AX N M AX N T S
HEMINGER	M M IH NG AX
SLAYABLE	S L EY EY B L
NETTIE	N EH T T AY
HALLENSTEIN	HH AE L L EH N S T AY N
LOBBAN	L OH B B AE N
PICTURER	P IH K CH AX AX
HAMMAD'S	HH AE M M AE D Z
STRUTTER'S	S T R AH T T ER Z
BITTEL	B IH T T EH L
SHAMMAR	SH AE M M AA CH
RICKETT'S	R IH K IH T T IY Z
JUSTICAR	JH AH S T IH K AX AX
COPPLED	K OH P P L EH D
FITTON	F IH T T AH N
CLUBBIER	K L AH B B IA
VOLLACK	V OH L L AE K
MILLERANDAGE	M IH L AX AX N EY JH
SCATTED	S K AE T T EH D
FIGURERS	F IH G AX AX Z
SCAMPERER	S K AE M P AX AX
CHITTENDEN	CH IH T T EH N D EH N
RENTAR	R EH N T AX AX R
PARKARD	P AA K AX AX D

Table B.3: Sample of removed entries during the analysis of lengthened pronunciation

Word	Pronuciation
INSTRUCTD	IH N S T R AH K T IH D
NAZIS	N AA T S IH Z
METZ’S	M EH T S IH Z
GRANN’S	G R AE N EH N Z
BUMPTY	B AH M P IY T IY W AY
WARK’S	W AO R K EY Z
KOBLENZ	K AX B L EH N T S
SECULARIZES	S EH K Y AX L AX R AY Z IH Z
SENORITAS	S EH N Y AO R IY T AX Z
SKUNKY’S	S K AH NG K W AY Z
PUMPS’S	P AH M P S IH Z
MONOD’S	M OH N OW D IY Z
AHAB’S	AA HH AA B IY Z
LEMONS’S	L EH M AX N Z IH Z
STEMM’S	S T EH M EH M Z
SONY’S	S AH N W AY Z
SUNY’S	S AH N W AY Z
RUGG’S	R AH G JH IY Z
SUPREMACISTS’S	S UH P R EH M AX S IH S T S IH Z
PUBLICIS’S	P AH B L IH S IH S IH Z
UNIVER	Y UW N IH V ER S IH T IY AX
CRUX’S	K R AH K S IH Z
TUNG’S	T AH N JH IY Z
LOCUS’S	L OW K AX S IH Z
ANTES’S	AE N T IH Z IH Z
SPUDS’S	S P AH D Z IH Z
CARICOM’S	K AE R IH K OW EH M Z
STANDS’	S T AE N D AX D Z
BOLING’S	B AX L IH N JH IY Z
NIMBUS’S	N IH M B AX S IH Z
AKINS’S	AX K IH N Z IH Z
SIND’S	S IH N D IY Z
BOLDT’S	B OW L D T IY Z
PROST’S	P R OW Z T IY Z
CANTV’S	K AE N T V IY Z
MARV’S	M AA CH V IY Z
DEPT’S	D EH P T IY Z
FLAMM’S	F L AE M EH M Z
ABRAHAMS’S	EY B R AX HH AE M Z IH Z
MAXX’S	M AE K S EH K S IH Z
TALBOTS’S	T OH L B AX T S IH Z
TRICON’S	T R IH K OW EH N Z
ERIS’S	EH R IH S IH Z
UNCLASSABLY	AH N K L AE S IH F AY AX B L IY
MARONEY’S	M AA CH W AH N W AY Z
MINSKY’S	M IH N S K W AY Z
WOLFERT’S	W UH L F AX R T IY Z
GRISTY’S	G R IH S T W AY Z
SABA’S	S AE B AX B AX Z
ECOSTAR’S	IY K OW S IH S T AX AA Z
ALSOP’S	AO L S OW P IY Z
DATAPRODUCTS’S	D EY T AX P R OH D AH K T S IH Z
LOFTUS’S	L OH F T AX S IH Z
ANTRIX’S	AE N T R IH K S IH Z
LANDIS’S	L AE N D IH S IH Z
GOLDY’S	G OW L D W AY Z
UNIVER	Y UW N IH V ER S IH T IY AX R
RADOVAN’S	R AE D OW V AX EH N Z
SENORA	S EH N Y AO R AX
SOLOW’S	S OW L OW D AH B L Y UW Z
SILICONIX’S	S IH L IH K OW N IH K S IH Z
BURL’S	B ER R EH L Z
DAGON’S	D EY G OW EH N Z
COTY’S	K OH T W AY Z
COLOMBUS’S	K AX L OH M B AX S IH Z
HASKINS’S	HH AE Z K IH N Z IH Z
DONEGALL’S	D OH N IH G AO L EH L Z
Y’S	W AY Z
POPS’S	P OH P S IH Z
GOLDMANN’S	G OW L D M AE N EH N Z
SPECTOGRAM	S P EH K T R OW G R AE M
REGULATORY’S	R EH G Y UH L EY T AX R W AY Z
PELVIS’S	P EH L V IH S IH Z
IT’S’S	IH T S IH Z
LUPUS’S	L UW P AX S IH Z
FRIZ’S	F R IH Z IH Z
ERADES’S	IA R AX D EH Z IH Z
ABACUS’S	AE B AX K AX S IH Z
MINISTR	M IH N IH S T AX R
DELEON’S	D IY L IY OW EH N Z
PROPAGANDISE	P R OH P AX G AE N D AX D AY Z
ALFREDOR	AE L F R IY D UW AX R
LINSKY’S	L IH N S K W AY Z
UNGERER’S	AH N JH EH N AX R AX Z
CARY’S	K AA R W AY Z
PENN’S	P EH N EH N Z
DESANTIS’S	D EH Z AE N T IH S IH Z
GENDIS’S	JH EH N D IH Z IH Z
JOSEF’S	HH AX Z EY EH F S
ELECTD	IH L EH K T IH D
MITRER	M AY T AX R AX R
ROSING’S	R OH Z IH N JH IY Z
APRICATION	EY P R IH V AE R IH K EY SH N
IMPRECATIVES	IH M P R IH D IH K AX T IH V Z
BURB’S	B ER R B IY Z
SALZ’S	S OH L T S IH Z
WALRUS’S	W AO L R AX S IH Z
LANDS’S	L AE N D Z IH Z
STANDRIN	S T AE N D AX R IH N
BIOMETR	B AY OH M IH T AX R

Table B.4: Sample of removed entries during graphemic null analysis

Word	Pronunciation
VERSOR	V ER S OW AX
SO-AND-SOS	S OW AX N S OW Z
ZOOPHYTES	Z OW AX F AY T S
PANDOR	P AE N D OW AX
VIGOR	V IY G OW AX
ECHOER	EH K OW AX
ROWAND	R OW AX N D
SALVORS	S AE L V OW AX R Z
COALITION'S	K OW AX L IH SH N Z
PRATOR	P R AA T OW AX
STUCCOERS	S T AH K OW AX Z
LEGATOR	L IH G AA T OW AX
INCENTOR	IH NG K EH N T OW AX
VERSOR	V ER S OW AX R
FEATHER-BOA	F EH DH AX B OW AX
MELLOR	M EH L OW AX R
CARLINOR	K AA R L AY N OW AX R
TRANTOR	T R AE N T OW AX R
MELLOR	M EH L OW AX
HEYBOER	HH EY B OW AX R
OBLIGATORS	OH B L IH G AA T OW AX Z
BOAK'S	B OW AX K EY Z
VETOER	V IY T OW AX R
THEATRE-GOERS	TH IA T AX G OW AX Z
CLAMBOR	K L AE M B OW AX R
BOERINGER	B OW AX IH NG AX
PSYCHOANALYSISTER	S AY K OW AX N AE L AX S IH S T ER
SANDOR	S AE N D OW AX R
JEROBOAMS	JH EH R AX B OW AX M Z
POTATOR	P AX T EY T OW AX
BOA-CONSTRICTOR	B OW AX K AX N S T R IH K T AX
BOALES	B OW AX L EH Z
SAMOA'S	S AX M OW AX Z
ALLOWABLE	AX L OW AX B L
AFRO-AMERICANS	AE F R OW AX M EH R IH K AX N Z
ZOOLOGIC	Z OW AX L OH JH IH K
LENTOR	L EH N T OW AX
TAMBOR	T AE M B OW AX
BOER	B OW AX R
TRUSTOR	T R AH S T OW AX
GOERGENS	G OW AX JH EH N Z
SALVORS	S AE L V OW AX Z
FORGOER	F AO G OW AX R
LEGATORS	L IH G AA T OW AX R Z
VETOER	V IY T OW AX
GOER'S	G OW AX Z
BOA-CONSTRICTOR	B OW AX K AX N S T R IH K T AX R
MOABITES	M OW AX B AY T S
PROTOZOA	P R OW T AX Z OW AX
PATIOR	P AE T IY OW AX
THEATRE-GOER	TH IA T AX G OW AX
CHURCHGOERS	CH ER CH G OW AX Z
BOA	B OW AX
HEMATOZOA	HH EH M AX T AX Z OW AX
HEMATOZOON	HH EH M AX T AX Z OW AX N
BANCOR	B AE NG K OW AX R
PSYCHOANALYSIS	S AY K OW AX N AE L AX S IH S
SAMOAN	S AX M OW AX N
COMBUSTORS	K AX M B AH S T OW AX Z
SANTOR	S AE N T OW AX
BLONDOR	B L OH N D OW AX R
TEMPOR	T EH M P OW AX R
ZOOPHYTE	Z OW AX F AY T
GO-AS-YOU-PLEASE	G OW AX Z Y UW P L IY Z
ADMINISTRATIORS	AX D M IH N IH S T R EY SH IY OW AX Z
PRATOR	P R AA T OW AX R
DOUGHER	D OW AX
ZOOLITE	Z OW AX L AY T
SAMBORSKI	S AE M B OW AX R S K IY
ROWAN'S	R OW AX N Z
OPERATIORS	OH P AX R EY SH IY OW AX Z
AVOWAL	AX V OW AX L
ADJUSTOR	AX JH AH S T OW AX
ECHOERS	EH K OW AX R Z
PROABORTION	P R OW AX B AO SH N
COADJACENTLY	K OW AX JH EY S AX N T L IY
TANGOR	T AE NG G OW AX R
LEXICOR	L EH K S IH K OW AX
VIGORS	V IY G OW AX R Z
CANTICOR	K AE N T IH K OW AX
COALESCENT	K OW AX L EH S N T
O'ER	OW AX
COALESCE'S	K OW AX L EH S IH Z
SANDORD	S AE N D OW AX R D
BOATABLE	B OW AX T EY B L
ALLOA	AE L OW AX
VIGOR	V IY G OW AX R
COADJACENT	K OW AX JH EY S N T
AVOWALS	AX V OW AX L Z
SPERMATOZOONS	S P ER M AX T AX Z OW AX N Z
BOER	B OW AX
SANDOR	S AE N D OW AX
COALESCENCES	K OW AX L EH S AX N S IH Z
FALDOR	F AE L D OW AX
TROPER	T R OW AX R
FOREGOER	F AO G OW AX
FRESCOER	F R EH S K OW AX R
LEGATORS	L IH G AA T OW AX Z
CHURCHGOER	CH ER CH G OW AX
ROWAN-TREES	R OW AX N T R IY Z

Table B.5: All entries removed during the analysis of lengthened spelling

Word	**Pronunciation**
COUCHETTES	K UW SH EH T S
PHENOMENOLOGIES	F IH N OH L AX JH IH Z
ELECTIONEERED	EH L EH K SH NG IA D
TELECOMMUNICATINS	T EH L IH K AX T IH N Z
CONTRIBUTORILY	K AX N T R IH L IY
SUPERCONSEQUENCY	S UW P AX R K AX N S IY
BUCKINGHAMSHIRE'S	B AH K IH NG AX M SH AX Z
ESCHSCHOLTZIA	IH S K OH L SH AX
PRIVATIZATIONS	P R AY V EY SH N Z
LARROQUETTE'S	L AA R OW K EH T S
INCONSISTENCES	IH N K AX N S IH Z
CHARGES_D'AFFAIRES	SH AA ZH EY D AE F EA
TWOPENNY-HALFPENNY	T AH P N IH HH EY P N IY
UNPROSPEROUSNESS	AH N P R AX S N AX S
INTRAVENTRICULAR	IH N T R IH K Y UH L AX
DERIVATIZATION	D EH R IH V EY SH N
SEMITRANSPARENCY	S EH M IH T R AX N S IY
UNCOUNTERACTED	AH N K AW N T IH D
PROPRIVATIZATION	P R OW P R IH V EY SH N
POLYSYNTHETICISM	P OH L IH S IH Z AX M
CHARLOTTESVILLE	SH AA L AX T S V IH L
SUBMERSIBILITY	S AX B IH L IH T IY
UNCONSCIENTIOUSNESS	AH N K OH N SH AX S N AX S
PRESENTIMENTAL	P R IH Z EH N T L
SEMITRANSPARENT	S EH M IH T R AX N T
NOUVEAUX_RICHES	N UW V OW R IY SH
HUMANITARIANISES	HH Y UW M AE N IH S IH Z
BUCKINGHAMSHIRE	B AH K IH NG AX M SH AX
PSEUDOSCIENTIFICALLY	S Y UW D OW S AY AX N T IH F IH K L IY
CHAMOIS-LEATHER	SH AE M IH L EH DH AX
CHAISE_LONGUE	SH EY Z L OH NG
INSTITUTIONALISATION	IH N S T IH T Y UW SH AX N
MAISONNETTES	M EY Z AX N EH T S
AUCTIONEERING	AO K SH AX N IA IH NG
PREPRIVATIZATION	P R EH P R IH V EY SH N
UNCONSCIENTIOUSLY	AH N K OH N SH AX S L IY
MUNICIPALISES	M Y UW N IH S IH Z
UNPROSPEROUSLY	AH N P R AX S L IY
CHAISE_LONGUES	SH EY Z L OH NG Z
IRREPRESENTABLENESS	IH R IH P R EH S AX B L N IH S
CHAMOIS-LEATHERS	SH AE M IH L EH DH AX Z
DISPROPORTIONATION	D IH S P R AX P AO SH AX N
COWPERTHWAITES	K AW P ER TH W EY T S
CHANCELLORSHIPS	CH AA N S AX SH IH P S
MACROMANAGEMENT	M AE K R OW M AX N T
COURGETTES	K UA ZH EH T S
CORPS_DE_BALLETS	K AO D AX B AE L EY
CAOUTCHOUCS	K AW CH UH K S
ANTISEPTICIZING	AE N T IH S AY Z IH NG
SCHOTTISCHE	SH OH T IY SH
INTERNATIONATIONAL	IH N T AX N AE SH AX N AX L
COUNTERCOMMERCIAL	K AW N T AX K AX M ER SH L
NARCISSISTICALLY	N AA S IH S T IH K L IY
FORECASTLE	F OW K S L
SILICATIZATION	S IH L IH K EY SH N
LAISSEZ-FAIRE	L EY S EY F EA
HYPERSENSITIVITIES	HH AY P ER S EH N S IH T IH Z
NIGHTSHIRTS	N AY CH ER T S
INTRANSIGENCIES	IH N T R AE N S IH Z
ANTISEPTICIZED	AE N T IH S AY Z D
SONSONATE'S	S AH N EH T S
IRRECONCILIABILITY	IH R EH K AX N S IH L IH T IY
UNAPPREHENSIBLENESS	AH N AX P R IY SH AX B L N AX S
PSYCHIATRISTRIC	S IH K AY AX T R IH K
PHENOMENOLOGISTS	F IH N OH L AX JH IH S T S
PROTESTANTISES	P R OH T IH S IH Z
UNAPPREHENDABLENESS	AH N AX P R IY SH AX B L N AX S
DEPRIVATIZATION	D EH P R IH V EY SH N
NIGHTSHIRT'S	N AY CH ER T S

Table B.6: Sample of removed entries during analysis of duplicate pronunciations

Word	Pronunciation
UNCHALLENGEABLY	AH N CH EY N JH AX B L IY
SUPERINDENTENT	S UW P AX R IH N D EH N T
NARCISSIST	N AA S IH S T
ANTARCHISTICAL	AE N T AA K T IH K AX L
SPECTR'S	S P EH K T AX Z
UNACCELERATED	AH N AE K S EH N T IH D
TORTORELLO	T AO R EH L OW
THURS	TH ER Z D EY
BOATSWAIN	B OW S N
DISTINGUISED	D IH S G AY Z D
UNSUPPURATED	AH N S AX P L AA N T IH D
ENERGETISTIC	EH N AX JH EH T IH K
RAISEED	R EY Z D
REPRESENTENTED	R EH P R IH Z EH N T IH D
UNINFILTRATED	AH N IH N F L EY T IH D
PATHETICLY	P AX TH EH T IH K L IY
BOATSWAIN'S	B OW S N Z
STICH'S	S T IH K S
LATERILISATION	L AE T AX R AY Z EY SH AX N
UNCLASSABLE	AH N K L AE S IH F AY AX B L
DECOLONIALISATION	D IY K OH L AX N AY Z EY SH AX N
NEGOTIATIATING	N IH G OW SH IH EY T IH NG
NONRESPONDENTS	N OH N R EH Z IH D AX N T S
UNSUPPLEMENTED	AH N S AX P L AA N T IH D
NONRESISTANT	N OH N R EH Z IH D AX N T
UNFORESEEABLENESS	AH N F AO S IY AX B L N AX S
UNASCENDABLE	AH N AE S AX T EY N AX B L
THRUWAY	TH R UW W EY
UNFORDABLE	AH N F AO S IY AX B L
UNDILAPIDATED	AH N D AY L Y UW T IH D
THRUPUT	TH R UW P UH T
REPREPARATION	R EH P AX R EY SH N
CONTENTESTED	K AX N T EH S T IH D
IMPERISHABLENESS	IH M P R AE K T IH K AX B L N AX S
UNREPENTENT	AH N R IH P EH N T
UNASCENDABLENESS	AH N AE S AX T EY N AX B L N AX S
INDEPENDENDENCE	IH N D IH P EH N D AX N S
UNFORESEEABLE	AH N F AO S IY AX B L
WRITEES	R AY T S
NON-RESPONDENTS	N OH N R EH Z IH D AX N T S
SOVRANTY	S OH V R AX N T IY
NEGOTIATIATED	N IH G OW SH IH EY T IH D
SPECTACULUM	S P EH K Y UH L AX M
UNFORFEITABLE	AH N F AO S IY AX B L
HIREES	HH AY AX Z
UNMEMORIALIZED	AH N M EH M AX R AY Z D
ANTARCHISTIC	AE N T AA K T IH K
TORTORICE	T AO R AY S
REPRESENTENTIVE	R EH P R IH Z EH N T IH V
IMPLAUSIBLENESS	IH M P R AE K T IH K AX B L N AX S
NONRESPONDENT	N OH N R EH Z IH D AX N T
SPECTR	S P EH K T AX
SPECTACTOR	S P EH K T AX
OPPOINTMENTS	AX P OW N AX N T S
UNRECANTED	AH N R IH S IY T IH D
MISSIPPI	M IH S IH S IH P IY
UNFORCIBLE	AH N F AO S IY AX B L
NONRESISTANTS	N OH N R EH Z IH D AX N T S
NARCISSISTIC	N AA S IH S T IH K
UNACCREDITATED	AH N AE K S EH N T IH D
UNRECREATED	AH N R IH S IY T IH D
NEGOTIATIATORS	N IH G OW SH IH EY T AX Z
BREAKTHRU	B R EY K TH R UW
SANTANTA	S AE N T AX
UNACTUATED	AH N AE K T IH D
UNDIVERSIFIABLE	AH N D IH V AY D AX B L
UNRECUPERATED	AH N R IH S IY T IH D
HIREES	HH AY AX R Z
ROBOTISTIC	R OW B OH T IH K
NON-RESPONDENT	N OH N R EH Z IH D AX N T
SUPERVISORS'	S UW P IA R IA Z
CONSISTENTEDLY	K AX N T EH N T IH D L IY
SOCRATISTIC	S AX K R AE T IH K
IMPRESSIBLENESS	IH M P R AE K T IH K AX B L N AX S
UNMEMORIALISED	AH N M EH M AX R AY Z D
PATHEMATICALLY	P AX TH EH T IH K L IY
SANTER	S AE N T AX
UNACCLIMATED	AH N AE K S EH N T IH D
UNDIVERTIBLE	AH N D IH V AY D AX B L
PHALANSTERIANISM	F AE L AX N S T AX R IH Z AX M
TORTORELLI	T AO R EH L AY
STICHES	S T IH K S
UNSITUATED	AH N S IH T IH D
DETENTENTION	D IH T EH N SH N
SIMPLETONIANISM	S IH M P L T AX N IH Z AX M
GRAMMATISTICAL	G R AX M AE T IH K L
NEGATIATING	N IH G EY T IH NG
ORIENTALIS	AO R IH EY L IH S
HYDRODROME	HH AY D R OW M
KANSANS	K AE N Z
UNFORDABLENESS	AH N F AO S IY AX B L N AX S
ANTISEPTICISES	AE N T IH S AY S IH Z
LE	L IY
UNFORESEEABLY	AH N F AO S IY AX B L IY
REPENTENT	R IH P EH N T

Table B.7: All removed entries during alignment

BRUNSON’S	B R AH N Z
BUREAUCRATISES	B Y UA R AX T IH S IH Z
DWINDLIND	D W IH N D
IMPUNCTUALITY	IH M P Y UW N IH T IY
BUREAUCRATIST	B Y UA R AX T IH S T
COMPROMITMENT	K OH M IH T M AX N T
HORSER	AO AX
PREMOVEMENT	P R IY M AX N T
OBANDONED	OW B AX N D
TRAGICOMICALITY	T R AE JH IH K AE L IH T IY
HORSED	AO D
UNALLEVIATEDLY	AH N AX L OH T IH D L IY
UNALLEVIATED	AH N AX L OH T IH D
MASTERSTON	M AA S T AX N
PREDISMISSORY	P R IY D IH S AX R IY
NEOPLASTICISM	N IY OW P L AE Z AX M
CAPITIVITY	K AE P IH T IY
BERIBERIC	B EH R IH K
SYMBIOGENESIS	S IH M B IH OW S IH S
VATICANICAL	V AE T IH K AX L
BENTSENTO	B EH N T OW
FACIATION	F EY SH AX N
ABSENTMENT	AE B S AX N T
PLATINATED	P L EY T IH D
UNIVERSITARIO	Y UW N IH T EA R IY OW
SEMICOMIC	S EH M IH K
CUTICULARIZE	K Y UW L AX R AY Z
QUINQUIVALENT	K W IH V AX L AX N T
INSTANTANEITY	IH N S T AX N IA T IY
STOMATOMY	S T AX M IY
ANTISEPTICISM	AE N T IH S IH Z AX M
POLYSYNTHESISM	P OH L IH P S IH Z AX M
CONSTANST	K OH N S T

Table B.8: All entries removed due to grapheme to phoneme rule analysis

Word	Pronunciation
TOWELL'S	T AW AX L EH L Z
HONEYMOONSHINE	HH AH N IY M UW N S T OW N
TROUBLESHOOTED	T R AH B L Z HH UW T IH D
MOONSHINER'S	M UW N S T OW N AA Z
UNINTERJECTED	AH N IH N T R AX S T IH D
DENATURIZATION	D IY N AE CH R AX L AY Z EY SH N
IRREDUCIBLES	IH R IH F Y UW T AX B L Z
SELF-REFERENCE	S EH L F R IH L AY AX N S
DENATURISATION	D IY N AE CH R AX L AY Z EY SH AX N
COLD-BLOODED	K OW L D HH AA T IH D
PROFESSIONISES	P R AX S EH SH AX N IH S IH Z
UNAPPLICABLE	AH N AX P L AY AX B L
MOONSHINE'S	M UW N S T OW N Z
MOONSHINERS	M UW N S T OW N AX R Z
HORSING	AO IH NG
STATISTICIZE	S T AE T IH S AY Z
DENUMERATION	D IY N Y UW D EY SH N
STATISTICISM	S T AE T IH S IH Z AX M
NICKELL'S	N IH K L EH L Z
MOONSHINER	M UW N S T OW N AX R
HORSMANS	AO M AE N Z
STETHOSCOPIES	S T EH R IA S K OH P IH Z
BLACKINGS	B L AE K L EH G Z
HORSFALL'S	AO F AO L Z
PRESTATION	P R EH Z T AE T AY AX N
STATISTICISE	S T AE T IH S AY S
HORSFALL	AO F AO L
UNPATROLLABLE	AH N P AE T R AX N IH Z AX B L
BLACKING	B L AE K L EH G
TWO-EDGED	T UW F EY S T
STATISTICIZES	S T AE T IH S AY Z IH Z
UNSTAMPEDED	AH N S T AE M P T IH D
INERASABLENESS	IH N IH R AE D IH K AX B L N IH S
PROFESSIONISE	P R AX S EH SH AX N IH S
SEMITICISM	S IH M IH T IH K IH Z AX M
UNCOMBUSTIBLE	AH N K AH M F T AX B L
UNSYMPTOMATIC	AH N S IH M P AX TH EH T IH K
HEATHERLANDS'	HH EH DH AX R L AE N D AX D Z
VORTICES	V OW T AX R IH Z
SANDFORDS'	S AE N D AX D Z
CHECKERIST	CH EH K AX R IA R IH S T
BAREFOOTED	B EA L EH G IH D
MOONSHINED	M UW N S T OW N D
ELLERY'S	EH L AX R W AY Z
UNPUTREFIABLE	AH N P UH T D AW N AX B L
HYPERSPECULATIVE	HH AY P AX S EH N S IH T IH V
PURPOSIVENESS	P ER P AX S W EY S IH V N AX S
EXPERIENCEING	IH K S P IA R IA N S IH NG
STANFORDS'	S T AE N D AX D Z
UNCOMBINABLE	AH N K AH M F T AX B L
UNPALPABLE	AH N P AE L AX T AX B L
UNATTACHED	AH N AX T EH N D IH D
HORSER	AO AX R
UNINTERDICTED	AH N IH N T R AX S T IH D
HYDROPOLITICS	HH AY D R AX P OH N IH K S
UNTREADABLE	AH N T R IY Z AX N AX B L
PROFESSIONIST	P R AX S EH SH AX N IH S T
UNRECOVERABLENESS	AH N R EH K AX G N AY Z AX B L N IH S
UNCHARTED	AH N CH EH K T
PERMISSIVITY	P AX M IH S AX B IH L IH T IY
UNACCREDITED	AH N AE K S EH N T IH D
NONCONFORMITANT	N OH N K AX N K OH M IH T AX N T
CYBERNETICIAN	S AY B AX N EH T IH K IY AX N
MOONSHINER	M UW N S T OW N AX
INTERPAYMENT	IH N T ER P EH L AX N T
DISCREPANCE	D IH S T ER B AX N S
POETICISMS	P OW EH T IH K IH Z AX M Z
INCONQUERABLE	IH N K AX N S IY L AX B L
UNPRECIPITATED	AH N P R EH S IH D EH N T IH D
DELIVERERS'	D IH L IH JH AX Z
MOONSHINERS	M UW N S T OW N AX Z
UNRECOVERABLE	AH N R EH K AX G N AY Z AX B L
CONTRAPUNTALLY	K OH N S AX N AX N T AE L AY
CHARGES	SH AA ZH EY
STEPHANOVA	S T EH P HH AE N OW V AX
STEPHAS	S T EH P HH AE Z
STATISTICISES	S T AE T IH S AY S IH Z
IRREMOVABLY	IH R EH P AX R AX B L IY
INDEMNIZATION	IH N D EH M N IH F IH K EY SH N
IRREDUCIBLE	IH R IH F Y UW T AX B L
UNRECRUITED	AH N R IH S IY T IH D
MOONSHINE	M UW N S T OW N
UNAFFILIATED	AH N AX F EH K T IH D
UNINTERCEPTED	AH N IH N T R AX S T IH D
HORSELL	AO EH L
HORSMAN	AO M AE N
RELICTION	R EH L IH K T AY AX N
UNCOMBINABLENESS	AH N K AH M F T AX B L N EH S
INTERCATION	IH N T ER R K AE T AY AX N

Table B.9: Sample of removed entries during the analysis of pseudo-phonemes and generation restriction rules

Word	**Pronunciation**
LIEUTENANCY	L EH F T EH N AX N S IY
VIRILITY	V AY AX R OH L AX JH IY
COUPER	K UW AX
OBSOLETED	OH B S T AX K L D
UNPROVIDED	AH N P R AX V OW K T
AGENTS'S	AE ZH OH N Z
INCIVILITY	IH N K L EH M AX N S IY
UNREASONING	AH N R IH L EH N T IH NG
IRRESPONSIBLE	IH R IH T R IY V AX B L
LIEUTENANT'S	L EH F T EH N AX N T S
OBSOLETES	OH B S T AX K L Z
INESCAPABLE	IH N EH S T IH M AX B L
AGENTES	AE ZH OH N Z
AGENTS	AE ZH OH N
AGENTS_PROVOCATEURS	AE ZH OH N P R AX V OH K AX T ER
OBSOLETENESS	OH B S T AX K L N EH S
RETRIBUTIVE	R IH T R IY V AX B L
AGENT_PROVOCATEUR	AE ZH OH N P R AX V OH K AX T ER R
ENTOURAGES	OH N T R AE K T S
AGENTED	AE ZH OH N D
INDISPUTABLE	IH N D IH S OH L Y UH B L
AGENT	AE ZH OH N
LANGSYNE	L AE NG G W IH JH
UNRESTRAINED	AH N R IH S T R IH K T IH D
COMPUNCTION	K OH M P Y UH T EY SH N
COMPUNCTIONS	K OH M P Y UH T EY SH N Z
INCOMPATIBLE	IH N K OH M P IH T AX N S
LIEUTENANCIES	L EH F T EH N AX N S IH Z
SENEGALESE	S IH N EH S N S
AGENT_PROVOCATEUR	AE ZH OH N P R AX V OH K AX T ER
HOMOEOPATHY'S	HH OW M AX JH IH N IA T IY Z
AGENTS_PROVOCATEURS	AE ZH OH N P R AX V OH K AX T ER R
SINER	S IH N EY AX
AGENTING	AE ZH OH N IH NG
OBSOLETE	OH B S T AX K L
INCOMPATIBLES	IH N K OH M P IH T AX N S IH Z
INDESTRUCTIBLE	IH N D IH T ER M IH N AX B L
HOMOEOPATHY	HH OW M AX JH IH N IA T IY
LIEUTENANT	L EH F T EH N AX N T
INOPERATIVE	IH N OH P AX T Y UW N
AGENT'S	AE ZH OH N Z
IRREMOVABLE	IH R EH P AX R AX B L
LIEUTENANTS	L EH F T EH N AX N T S
LIEUTENANTS'	L EH F T EH N AX N T S
AGENTSHIP	AE ZH OH N SH IH P
CRETONNES	K R IH V AE S IH Z
LANGSYNES	L AE NG G W IH JH IH Z
INDESTRUCTIBLES	IH N D IH T ER M IH N AX B L Z
CRETONNE	K R IH V AE S

APPENDIX C

DIPHTHONG ANALYSIS RESULTS

Table C.1: Baseline System Confusion Matrix Part 1

	AA	AE	AH	AO	AW	AX	AY	B	CH	D	DH	EA	EH	ER	EY	F	G	HH	IA	IH	IY	JH	K
AA	**0.80**	0.00	0.02	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AE	0.00	**0.85**	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
AH	0.01	0.01	**0.77**	0.00	0.01	0.01	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
AO	0.01	0.00	0.00	**0.90**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AW	0.00	0.00	0.01	0.00	**0.87**	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AX	0.00	0.01	0.00	0.00	0.00	**0.77**	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
AY	0.01	0.01	0.01	0.00	0.00	0.01	**0.89**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
B	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.88**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.87**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00
D	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	**0.79**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
DH	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	**0.46**	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
EA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	**0.90**	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00
EH	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	**0.86**	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
ER	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.01	**0.85**	0.01	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00
EY	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	**0.83**	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	**0.89**	0.00	0.00	0.00	0.00	0.00	0.00	0.00
G	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.82**	0.00	0.00	0.00	0.00	0.00	0.02
HH	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	**0.82**	0.00	0.00	0.00	0.00	0.01
IA	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	**0.82**	0.00	0.01	0.00	0.00
IH	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.00	**0.84**	0.02	0.00	0.00
IY	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	**0.90**	0.00	0.00
JH	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	**0.89**	0.00
K	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	**0.85**
L	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
M	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
N	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NG	0.00	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
OH	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01
OW	0.00	0.00	0.01	0.00	0.00	0.02	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OY	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.17	0.00	0.33
P	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
T	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
TH	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
UA	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
UH	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
UW	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.05	0.00	0.00
V	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Y	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
Z	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
ZH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table C.2: Baseline System Confusion Matrix Part 2

	L	M	N	NG	OH	OW	P	R	S	SH	T	TH	UA	UH	UW	V	W	Y	Z	ZH
AA	0.01	0.00	0.01	0.00	0.01	0.01	0.01	0.00	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AH	0.00	0.00	0.01	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00
AO	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AW	0.01	0.00	0.00	0.00	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
AX	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AY	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
B	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CH	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DH	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.02	0.01	0.00	0.03	0.03	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00
EA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
EH	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ER	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
EY	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
G	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
HH	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
IA	0.00	0.00	0.03	0.02	0.00	0.02	0.01	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.01	0.00
IH	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
IY	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
JH	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
L	**0.84**	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
M	0.01	**0.83**	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
N	0.00	0.00	**0.92**	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NG	0.01	0.00	0.04	**0.82**	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00
OH	0.01	0.00	0.00	0.00	**0.78**	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OW	0.02	0.00	0.03	0.00	0.00	**0.73**	0.01	0.01	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00
OY	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P	0.00	0.00	0.01	0.00	0.00	0.00	**0.86**	0.00	0.01	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R	0.01	0.00	0.01	0.00	0.00	0.00	0.00	**0.87**	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	**0.87**	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
SH	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	**0.93**	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
T	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.00	**0.89**	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TH	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.02	**0.80**	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00
UA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.92**	0.00	0.00	0.00	0.00	0.00	0.00	0.00
UH	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.01	0.01	0.01	**0.84**	0.00	0.00	0.00	0.00	0.00	0.00
UW	0.01	0.00	0.02	0.00	0.00	0.01	0.01	0.01	0.01	0.00	0.02	0.01	0.00	0.00	**0.71**	0.00	0.00	0.00	0.02	0.00
V	0.01	0.01	0.02	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.00	**0.78**	0.00	0.00	0.00	0.00
W	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.89**	0.00	0.00	0.00
Y	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	**0.80**	0.00	0.00
Z	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	**0.83**	0.00
ZH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.93**

Table C.3: Knowledge-based system confusion matrix (part 1)

	AA	**AE**	**AH**	**AO**	**AX**	**B**	**CH**	**D**	**DH**	**EH**	**ER**	**F**	**G**	**HH**	**IH**	**IY**	**JH**	**K**
AA	**0.78**	0.00	0.03	0.01	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00
AE	0.00	**0.84**	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.02	0.00	0.00	0.00
AH	0.00	0.00	**0.89**	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AO	0.00	0.00	0.01	**0.90**	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
AX	0.00	0.00	0.01	0.00	**0.82**	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
B	0.00	0.00	0.01	0.00	0.00	**0.86**	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
CH	0.00	0.00	0.00	0.00	0.01	0.00	**0.86**	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00
D	0.00	0.00	0.00	0.00	0.01	0.01	0.00	**0.80**	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
DH	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.01	**0.42**	0.02	0.02	0.01	0.00	0.00	0.03	0.00	0.00	0.01
EH	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00	**0.89**	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
ER	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.02	**0.85**	0.01	0.00	0.00	0.03	0.00	0.00	0.00
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	**0.88**	0.00	0.00	0.00	0.00	0.00	0.00
G	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.82**	0.00	0.00	0.00	0.01	0.01
HH	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.01	**0.82**	0.00	0.00	0.00	0.00
IH	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	**0.90**	0.01	0.00	0.00
IY	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	**0.89**	0.00	0.00
JH	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	**0.89**	0.00
K	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.00	**0.85**
L	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
M	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
NG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.01	0.01	0.00	0.00
OH	0.01	0.00	0.01	0.01	0.02	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00
P	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00
R	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
S	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00
SH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
T	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
TH	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
UH	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00
UW	0.00	0.00	0.00	0.01	0.04	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.04	0.01	0.00
V	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00
W	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Y	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.02	0.01	0.00	0.00
Z	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
ZH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table C.4: Knowledge-based system confusion matrix (part 2)

	L	M	N	NG	OH	P	R	S	SH	T	TH	UH	UW	V	W	Y	Z	ZH
AA	0.01	0.00	0.01	0.00	0.02	0.01	0.00	0.00	0.00	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00
AE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AH	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
AO	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
AX	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
B	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00
CH	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.02	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
D	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DH	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.04	0.03	0.01	0.00	0.01	0.00	0.01	0.00	0.00
EH	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ER	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
G	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HH	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.02	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00
IH	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
IY	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
JH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
K	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
L	**0.82**	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00
M	0.00	**0.83**	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
N	0.00	0.00	**0.91**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NG	0.00	0.00	0.05	**0.81**	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.01	0.00	0.01	0.01	0.00	0.00	0.00
OH	0.01	0.00	0.00	0.00	**0.77**	0.00	0.00	0.01	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00
P	0.00	0.00	0.01	0.00	0.00	**0.86**	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R	0.01	0.00	0.01	0.00	0.00	0.00	**0.86**	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S	0.00	0.00	0.01	0.00	0.00	0.00	0.01	**0.87**	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00
SH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	**0.93**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
T	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	**0.89**	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TH	0.00	0.00	0.01	0.00	0.00	0.01	0.01	0.01	0.00	0.02	**0.80**	0.00	0.00	0.00	0.00	0.00	0.01	0.00
UH	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	**0.86**	0.00	0.00	0.00	0.00	0.00	0.00
UW	0.01	0.00	0.02	0.00	0.00	0.01	0.01	0.00	0.00	0.01	0.01	0.01	**0.70**	0.00	0.00	0.00	0.02	0.00
V	0.01	0.01	0.02	0.00	0.00	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.00	**0.78**	0.00	0.00	0.00	0.00
W	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.89**	0.00	0.00	0.00
Y	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	**0.79**	0.00	0.00
Z	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	**0.83**	0.00
ZH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	**0.87**

Table C.5: Knowledge-based system confusion matrix subtracted from baseline confusion matrix (part 1)

	/AA/	/AE/	/AH/	/AO/	/AW/	/AX/	/AY/	/B/	/CH/	/D/	/DH/	/EA/	/EH/	/ER/	/EY/	/F/	/G/	/HH/	/IA/	/IH/	/IY/	/JH/	/K/
/AA/	**0.014**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.024	0.000	0.000	0.000
/AE/	0.000	**0.015**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.014	0.000	0.000	0.000
/AH/	0.004	0.005	**-0.108**	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	-0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/AO/	0.000	0.000	0.000	**0.007**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.007	0.000	0.000	0.000
/AW/	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/AX/	0.000	0.003	-0.005	0.000	0.000	**-0.037**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.005	0.000	0.000	0.000
/AY/	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/B/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.007**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/CH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.012**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/D/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/DH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.011**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/EA/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/EH/	0.000	0.000	-0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**-0.032**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.005	0.000	0.000
/ER/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.016**	0.000	0.000	0.000	0.000	0.000	-0.017	0.000	0.000	0.000
/EY/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/F/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/G/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000	0.000
/HH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.017**	0.000	0.000	0.000	0.000	0.000
/IA/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000
/IH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.003	0.000	0.000	0.000	0.000	0.000	**-0.040**	0.005	0.000	0.000
/IY/	0.000	0.000	0.000	0.000	0.000	-0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.007	**0.004**	0.000	0.000
/JH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000
/K/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.007	0.000	0.000	**0.000**
/L/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/M/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/N/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.005	0.000	0.000	0.000
/NG/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/OH/	0.000	0.000	-0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.008	0.000	0.000	0.000
/OW/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/OY/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/P/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/R/	0.000	0.000	-0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.008	0.000	0.000	0.000	0.000	0.000	0.000	-0.005	0.000	0.000	0.000
/S/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/SH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/T/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.003	0.000	0.000	0.000
/TH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/UA/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/UH/	0.000	0.000	-0.006	0.005	0.000	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.010	0.006	0.000	0.000
/UW/	0.000	0.000	0.000	0.000	0.000	-0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/V/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/W/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/Y/	0.000	0.000	0.000	0.000	0.000	-0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.014	0.000	0.000	0.000
/Z/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/ZH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table C.6: Knowledge-based system confusion matrix subtracted from baseline confusion matrix (part 2)

	/L/	/M/	/N/	/NG/	/OH/	/OW/	/OY/	/P/	/R/	/S/	/SH/	/T/	/TH/	/UA/	/UH/	/UW/	/V/	/W/	/Y/	/Z/	/ZH/
/AA/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.022	0.000	0.000	0.000	0.000	0.000	0.000
/AE/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.006	0.000	0.000	0.000	0.000	0.000	0.000
/AH/	0.000	0.000	0.006	0.000	0.012	0.000	0.000	0.000	0.000	0.000	0.000	0.005	0.004	0.000	-0.006	0.000	0.003	0.000	0.000	0.000	0.000
/AO/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/AW/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/AX/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/AY/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/B/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/CH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/D/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/DH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/EA/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/EH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/ER/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/EY/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/F/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/G/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/HH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/IA/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/IH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.004	0.000	0.000	0.000	0.000	0.000	0.000
/IY/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/JH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/K/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/L/	**0.007**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.011	0.000	0.000	0.000	0.000	0.000	0.000
/M/	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.006	0.000	0.000	0.000	0.000	0.000	0.000
/N/	0.000	-0.002	**0.006**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.002	0.000	0.000	0.000	0.000	0.000	0.000
/NG/	0.000	0.000	0.000	**0.018**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/OH/	0.000	0.000	0.000	0.000	**0.029**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.016	0.000	0.000	0.000	0.000	0.000	0.000
/OW/	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/OY/	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/P/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/R/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.013**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/S/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.009**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/SH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/T/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.008**	0.000	0.000	-0.003	0.000	0.000	0.000	0.000	0.000	0.000
/TH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.010**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/UA/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000
/UH/	-0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**-0.007**	0.000	0.000	0.000	0.000	0.000	0.000
/UW/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.010	**0.000**	0.000	0.000	0.000	0.000	0.000
/V/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.006	0.000	**0.009**	0.000	0.000	0.000	0.000
/W/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000
/Y/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.012**	0.000	0.000
/Z/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000
/ZH/	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**

Appendix D

KIT vowel adaptation results

The entries in the tables included in this appendix are sorted first according to their accuracy (incorrect entries are first on the list) and subsequently according to their spelling. The number that appears next to the graphemic representation of every word indicates the location of the /IH/ phoneme in the word that was considered.

Table D.1: Results when applying selected adaptation rules (part 1)

word	pronun_base	pronun_predict	pronun_corr	Rule 2	Rule 3	corr
abbreviations 5	AX -"B R IY -V IH -'EY -SH AX N Z	AX -"B R IY -V AX -'EY -SH AX N Z	AX -"B R IY -V IH -'EY -SH AX N Z	0	0	0
babies 3	"B EY -B IH Z	"B EY -B AX Z	"B EY -B IH Z	0	0	0
bedevilled 1	B IH -'D EH -V L D	B AX -'D EH -V L D	B IH -'D EH -V L D	0	0	0
belies 1	B IH -'L AY Z	B AX -'L AY Z	B IH -'L AY Z	0	0	0
beloved 1	B IH -'L AH V D	B AX -'L AH V D	B IH -'L AH V D	0	0	0
blanket 5	B L AE NG -K IH T	B L AE NG -K IH T	B L AE NG -K AX T	K	0	0
carrying 3	"K AE -R IH -IH NG	"K AE -R AX -IH NG	"K AE -R IH -IH NG	0	0	0
clumsiness 5	K L AH M -Z IH -N IH S	K L AH M -Z AX -N IH S	K L AH M -Z IH -N AX S	0	0	0
cockney 4	K OH K -N IH	K OH K -N AX	K OH K -N IH	0	0	0
continue 4	K AX N -'T IH -N Y UH	K AX N -'T AX -N Y UH	K AX N -'T IH -N Y UH	0	0	0
cordiality 3	"K AO -D IH -'AE -L AX -T IH	"K AO -D AX -'AE -L AX -T IH	"K AO -D IH -'AE -L AX -T IH	0	0	0
cordiality 8	"K AO -D IH -'AE -L AX -T IH	"K AO -D AX -'AE -L AX -T AX	"K AO -D IH -'AE -L AX -T IH	0	0	0
coterminously 10	"K OW -'T ER -M IH -N AX S -L IH	"K OW -'T ER -M AX -N AX S -L AX	"K OW -'T ER -M AX -N AX S -L IH	0	0	0
descended 1	D IH -'S EH N -D IH D	D AX -'S EH N -D IH D	D IH -'S EH N -D AX D	0	0	0
dignitary 8	D IH G -N IH -T AX -R IH	D IH G -N AX -T AX -R AX	D IH G -N AX -T AX -R IH	0	0	0
dilapidated 1	D IH -'L AE -P IH -D EY -T IH D	D AX -'L AE -P IH -D EY -T IH D	D IH -'L AE -P AX -D EY -T AX D	0	0	0
dowdy 3	D AW -D IH	D AW -D AX	D AW -D IH	0	0	0
emergencies 7	IH -'M ER -JH AX N -S IH Z	AX -'M ER -JH AX N -S AX Z	IH -'M ER -JH AX N -S IH Z	0	0	0
everything 3	EH -V R IH -TH IH NG	EH -V R AX -TH IH NG	EH -V R IH -TH IH NG	0	0	0
examinations 0	IH G -"Z AE -M IH -'N EY -SH AX N Z	IH G -"Z AE -M IH -'N EY -SH AX N Z	EH G -"Z AE -M AX -'N EY -SH AX N Z	G	0	0
extremely 0	IH K -'S T R IY M -L IH	IH K -'S T R IY M -L IH	EH K -'S T R IY M -L IH	K	0	0
extremely 8	IH K -'S T R IY M -L IH	IH K -'S T R IY M -L AX	EH K -'S T R IY M -L IH	0	0	0
factory 6	"F AE K -T AX -R IH	"F AE K -T AX -R AX	"F AE K -T AX -R IH	0	0	0
family 5	"F AE -M AX -L IH	"F AE -M AX -L AX	"F AE -M AX -L IH	0	0	0
felonies 5	F EH -L AX -N IH Z	F EH -L AX -N AX Z	F EH -L AX -N IH Z	0	0	0
fishy 3	F IH -SH IH	F IH -SH AX	F IH -SH IH	0	0	0
formalities 7	F AO -'M AE -L AX -T IH Z	F AO -'M AE -L AX -T AX Z	F AO -'M AE -L AX -T IH Z	0	0	0
generically 9	JH IH -'N EH -R IH -K AX -L IH	JH AX -'N EH -R IH -K AX -L AX	JH AX -'N EH -R IH -K AX -L IH	0	0	0
gillies 3	G IH -L IH Z	G IH -L AX Z	G IH -L IH Z	0	0	0
gustatory 8	G AH -S T AX -T AX -R IH	G AH -S T AX -T AX -R AX	G AH -S T AX -T AX -R IH	0	0	0
honeymooners 3	H AH -N IH -"M UH -N AX Z	H AH -N AX -"M UH -N AX Z	H AH -N IH -"M UH -N AX Z	0	0	0
inanities 6	IH N -'AE -N AX -T IH Z	AX N -'AE -N AX -T AX Z	IH N -'AE -N AX -T IH Z	0	0	0
insipidity 9	"IH N -S IH -'P IH -D AX -T IH	"AX N -S AX -'P AX -D AX -T AX	"IH N -S AX -'P AX -D AX -T IH	0	0	0
jiminy 5	JH IH -M AX -N IH	JH AX -M AX -N AX	JH AX -M AX -N IH	0	0	0
meteor 3	M IY -T IH -AO R	M IY -T AX -AO R	M IY -T IH -AO R	0	0	0
minority 7	M AY -"N OH -R AX -T IH	M AY -"N OH -R AX -T AX	M AY -"N OH -R AX -T IH	0	0	0
muzzily 5	M AH -Z IH -L IH	M AH -Z AX -L AX	M AH -Z AX -L IH	0	0	0
nastily 6	N AA -S T IH -L IH	N AA -S T AX -L AX	N AA -S T AX -L IH	0	0	0
nonconformity 11	"N OH N -K AX N -'F AO -M AX -T IH	"N OH N -K AX N -'F AO -M AX -T AX	"N OH N -K AX N -'F AO -M AX -T IH	0	0	0
Parisians 3	P AX -'R IH -Z Y AX N Z	P AX -'R AX -Z Y AX N Z	P AX -'R IH -Z Y AX N Z	0	0	0
pathologically 11	"P AE -TH AX -'L OH -JH IH -K AX -L IH	"P AE -TH AX -'L OH -JH IH -K AX -L AX	"P AE -TH AX -'L OH -JH IH -K AX -L IH	0	0	0
patriarchal 4	"P EY -T R IH -'AA -K L	"P EY -T R AX -'AA -K L	"P EY -T R IH -'AA -K L	0	0	0
permanency 8	P ER -M AX -N AX N -S IH	P ER -M AX -N AX N -S AX	P ER -M AX -N AX N -S IH	0	0	0
pirouetting 1	"P IH -R UH -'EH -T IH NG	"P AX -R UH -'EH -T IH NG	"P IH -R UH -'EH -T IH NG	0	0	0
prefectural 2	P R IH -'F EH K -CH AX -R AX L	P R AX -'F EH K -CH AX -R AX L	P R IH -'F EH K -CH AX -R AX L	0	0	0
primacy 6	P R AY -M AX -S IH	P R AY -M AX -S AX	P R AY -M AX -S IH	0	0	0
primary 6	"P R AY -M AX -R IH	"P R AY -M AX -R AX	"P R AY -M AX -R IH	0	0	0
recited 1	R IH -'S AY -T IH D	R AX -'S AY -T IH D	R IH -'S AY -T AX D	0	0	0
redundancies 1	R IH -'D AH N -D AX N -S IH Z	R AX -'D AH N -D AX N -S IH Z	R IH -'D AH N -D AX N -S IH Z	0	0	0
redundancies 9	R IH -'D AH N -D AX N -S IH Z	R AX -'D AH N -D AX N -S AX Z	R IH -'D AH N -D AX N -S IH Z	0	0	0
remarries 5	"R IY -'M AE -R IH Z	"R IY -'M AE -R AX Z	"R IY -'M AE -R IH Z	0	0	0
repatriations 6	"R IY -P AE -T R IH -'EY -SH AX N Z	"R IY -P AE -T R AX -'EY -SH AX N Z	"R IY -P AE -T R IH -'EY -SH AX N Z	0	0	0
reproached 1	R IH -'P R OW CH T	R AX -'P R OW CH T	R IH -'P R OW CH T	0	0	0
repudiation 1	R IH -"P Y UH -D IH -'EY -SH AX N	R AX -"P Y UH -D IH -'EY -SH AX N	R IH -"P Y UH -D IH -'EY -SH AX N	0	0	0
repudiation 6	R IH -"P Y UH -D IH -'EY -SH AX N	R AX -"P Y UH -D AX -'EY -SH AX N	R IH -"P Y UH -D IH -'EY -SH AX N	0	0	0
resolved 1	R IH -'Z OH L V D	R AX -'Z OH L V D	R IH -'Z OH L V D	0	0	0
saleslady 7	S EY L Z -"L EY -D IH	S EY L Z -"L EY -D AX	S EY L Z -"L EY -D IH	0	0	0
scariest 4	S K EA -R IH -IH S T	S K EA -R AX -IH S T	S K EA -R IH -AX S T	0	0	0
sedulously 9	S EH -D Y UH -L AX S -L IH	S EH -D Y UH -L AX S -L AX	S EH -D Y UH -L AX S -L IH	0	0	0
sensuously 8	S EH N -S Y UA S -L IH	S EH N -S Y UA S -L AX	S EH N -S Y UA S -L IH	0	0	0
servility 7	S ER -'V IH -L AX -T IH	S ER -'V AX -L AX -T AX	S ER -'V AX -L AX -T IH	0	0	0
sporty 4	S P AO -T IH	S P AO -T AX	S P AO -T IH	0	0	0
stereo 4	S T EH -R IH -OW	S T EH -R AX -OW	S T EH -R IH -OW	0	0	0
superfluities 9	"S UH -P AX -'F L UH -AX -T IH Z	"S UH -P AX -'F L UH -AX -T AX Z	"S UH -P AX -'F L UH -AX -T IH Z	0	0	0
surefootedly 8	"SH AO -'F UH -T IH D -L IH	"SH AO -'F UH -T AX D -L AX	"SH AO -'F UH -T AX D -L IH	0	0	0
theory 3	"TH IA -R IH	"TH IA -R AX	"TH IA -R IH	0	0	0
trolleybus 4	T R OH -L IH -B AH S	T R OH -L AX -B AH S	T R OH -L IH -B AH S	0	0	0
weeny 3	W IY -N IH	W IY -N AX	W IY -N IH	0	0	0

Table D.2: Results when applying selected adaptation rules (part 2)

word	pronun_base	pronun_predict	pronun_corr	Rule 2	Rule 3	corr
wordy 3	W ER -D IH	W ER -D AX	W ER -D IH	0	0	0
abstinent 4	AE B -S T IH -N AX N T	AE B -S T AX -N AX N T	AE B -S T AX -N AX N T	0	0	1
addresses 5	AX -’D R EH -S IH Z	AX -’D R EH -S AX Z	AX -’D R EH -S AX Z	0	0	1
affidavits 2	”AE -F IH -’D EY -V IH T S	”AE -F AX -’D EY -V IH T S	”AE -F AX -’D EY -V AX T S	0	0	1
affidavits 6	”AE -F IH -’D EY -V IH T S	”AE -F AX -’D EY -V AX T S	”AE -F AX -’D EY -V AX T S	0	0	1
aggrieving 5	AX -’G R IY -V IH NG	AX -’G R IY -V IH NG	AX -’G R IY -V IH NG	NG	0	1
amulet 5	AE -M Y UH -L IH T	AE -M Y UH -L AX T	AE -M Y UH -L AX T	0	0	1
apposing 4	AE -’P OW -Z IH NG	AE -’P OW -Z IH NG	AE -’P OW -Z IH NG	NG	0	1
assuming 5	AX -’S Y UH -M IH NG	AX -’S Y UH -M IH NG	AX -’S Y UH -M IH NG	NG	0	1
auditor 2	AO -D IH -T AX R	AO -D AX -T AX R	AO -D AX -T AX R	0	0	1
badness 4	B AE D -N IH S	B AE D -N AX S	B AE D -N AX S	0	0	1
baffling 4	B AE -F L IH NG	B AE -F L IH NG	B AE -F L IH NG	NG	0	1
beatify 4	B IY -’AE -T IH -F AY	B IY -’AE -T AX -F AY	B IY -’AE -T AX -F AY	0	0	1
big 1	”B IH G	”B IH G	”B IH G	G	0	1
billets 1	B IH -L IH T S	B AX -L IH T S	B AX -L AX T S	0	0	1
billets 3	B IH -L IH T S	B AX -L AX T S	B AX -L AX T S	0	0	1
blandishes 5	B L AE N -D IH -SH IH Z	B L AE N -D IH -SH IH Z	B L AE N -D IH -SH AX Z	0	SH	1
blandishes 7	B L AE N -D IH -SH IH Z	B L AE N -D IH -SH AX Z	B L AE N -D IH -SH AX Z	0	0	1
blooding 4	B L AH -D IH NG	B L AH -D IH NG	B L AH -D IH NG	NG	0	1
bonnets 3	B OH -N IH T S	B OH -N AX T S	B OH -N AX T S	0	0	1
brings 2	”B R IH NG Z	”B R IH NG Z	”B R IH NG Z	NG	0	1
bristols 2	B R IH -S T L Z	B R AX -S T L Z	B R AX -S T L Z	0	0	1
captive 4	”K AE P -T IH V	”K AE P -T AX V	”K AE P -T AX V	0	0	1
carbonate 5	K AA -B AX -N IH T	K AA -B AX -N AX T	K AA -B AX -N AX T	0	0	1
carrying 4	”K AE -R IH -IH NG	”K AE -R AX -IH NG	”K AE -R IH -IH NG	NG	0	1
carving 3	K AA -V IH NG	K AA -V IH NG	K AA -V IH NG	NG	0	1
circumscription 8	”S ER -K AX M -’S K R IH P -SH AX N	”S ER -K AX M -’S K R AX P -SH AX N	”S ER -K AX M -’S K R AX P -SH AX N	0	0	1
clumsiness 7	K L AH M -Z IH -N IH S	K L AH M -Z AX -N AX S	K L AH M -Z IH -N AX S	0	0	1
contagiousness 9	K AX N -’T EY -JH AX S -N IH S	K AX N -’T EY -JH AX S -N AX S	K AX N -’T EY -JH AX S -N AX S	0	0	1
contaminates 6	K AX N -’T AE -M IH -N EY T S	K AX N -’T AE -M AX -N EY T S	K AX N -’T AE -M AX -N EY T S	0	0	1
coterminously 5	”K OW -’T ER -M IH -N AX S -L IH	”K OW -’T ER -M AX -N AX S -L IH	”K OW -’T ER -M AX -N AX S -L IH	0	0	1
counterpoising 8	K AW N -T AX -P OY -Z IH NG	K AW N -T AX -P OY -Z IH NG	K AW N -T AX -P OY -Z IH NG	NG	0	1
curetted 6	K Y UA -’R EH -T IH D	K Y UA -’R EH -T AX D	K Y UA -’R EH -T AX D	0	0	1
demagnetized 6	”D IY -’M AE G -N IH -T AY Z D	”D IY -’M AE G -N AX -T AY Z D	”D IY -’M AE G -N AX -T AY Z D	0	0	1
descended 6	D IH -’S EH N -D IH D	D AX -’S EH N -D AX D	D IH -’S EH N -D AX D	0	0	1
dignitary 1	D IH G -N IH -T AX -R IH	D IH G -N IH -T AX -R IH	D IH G -N AX -T AX -R IH	G	0	1
dignitary 4	D IH G -N IH -T AX -R IH	D IH G -N AX -T AX -R IH	D IH G -N AX -T AX -R IH	0	0	1
dilapidated 5	D IH -’L AE -P IH -D EY -T IH D	D AX -’L AE -P AX -D EY -T IH D	D IH -’L AE -P AX -D EY -T AX D	0	0	1
dilapidated 9	D IH -’L AE -P IH -D EY -T IH D	D AX -’L AE -P AX -D EY -T AX D	D IH -’L AE -P AX -D EY -T AX D	0	0	1
din 1	D IH N	D AX N	D AX N	0	0	1
dint 1	D IH N T	D AX N T	D AX N T	0	0	1
dwellings 4	”D W EH -L IH NG Z	”D W EH -L IH NG Z	”D W EH -L IH NG Z	NG	0	1
eaglet 3	IY G -L IH T	IY G -L AX T	IY G -L AX T	0	0	1
ebullience 0	IH -’B AH L -Y AX N S	IH -’B AH L -Y AX N S	IH -’B AH L -Y AX N S		0	1
edit 2	”EH -D IH T	”EH -D AX T	”EH -D AX T	0	0	1
elaborateness 0	IH -’L AE -B AX -R AX T -N IH S	IH -’L AE -B AX -R AX T -N IH S	IH -’L AE -B AX -R AX T -N AX S		0	1
elaborateness 9	IH -’L AE -B AX -R AX T -N IH S	AX -’L AE -B AX -R AX T -N AX S	IH -’L AE -B AX -R AX T -N AX S	0	0	1
elaborating 0	IH -’L AE -B AX -R EY -T IH NG	IH -’L AE -B AX -R EY -T IH NG	IH -’L AE -B AX -R EY -T IH NG	NG	0	1
elaborating 8	IH -’L AE -B AX -R EY -T IH NG	IH -’L AE -B AX -R EY -T IH NG	IH -’L AE -B AX -R EY -T IH NG	NG	0	1
elephantine 2	”EH -L IH -’F AE N -T AY N	”EH -L AX -’F AE N -T AY N	”EH -L AX -’F AE N -T AY N	0	0	1
emergencies 0	IH -’M ER -JH AX N -S IH Z	IH -’M ER -JH AX N -S IH Z	IH -’M ER -JH AX N -S IH Z		0	1
enable 0	IH -’N EY -B L	IH -’N EY -B L	IH -’N EY -B L		0	1
entrap 0	IH N -’T R AE P	IH N -’T R AE P	IH N -’T R AE P		0	1
evacuate 0	IH -’V AE -K Y UH -EY T	IH -’V AE -K Y UH -EY T	IH -’V AE -K Y UH -EY T		0	1
everything 5	EH -V R IH -TH IH NG	EH -V R AX -TH IH NG	EH -V R IH -TH IH NG	NG	0	1
examinations 5	IH G -”Z AE -M IH -’N EY -SH AX N Z	IH G -”Z AE -M AX -’N EY -SH AX N Z	EH G -”Z AE -M AX -’N EY -SH AX N Z	0	0	1
farming 3	F AA -M IH NG	F AA -M IH NG	F AA -M IH NG	NG	0	1
filtertipped 1	”F IH L -T AX -’T IH P T	”F AX L -T AX -’T IH P T	”F AX L -T AX -’T AX P T	0	0	1
filtertipped 6	”F IH L -T AX -’T IH P T	”F AX L -T AX -’T AX P T	”F AX L -T AX -’T AX P T	0	0	1
fishy 1	F IH -SH IH	F IH -SH IH	F IH -SH IH	0	SH	1
forbidden 3	F AX -”B IH -D AX N	F AX -”B AX -D AX N	F AX -”B AX -D AX N	0	0	1
friendship 6	F R EH N D -SH IH P	F R EH N D -SH AX P	F R EH N D -SH AX P	0	0	1
gaining 3	G EY -N IH NG -AX	G EY -N IH NG -AX	G EY -N IH NG -AX	NG	0	1
generically 1	JH IH -’N EH -R IH -K AX -L IH	JH AX -’N EH -R IH -K AX -L IH	JH AX -’N EH -R IH -K AX -L IH	0	0	1
generically 5	JH IH -’N EH -R IH -K AX -L IH	JH AX -’N EH -R IH -K AX -L IH	JH AX -’N EH -R IH -K AX -L IH	K	0	1
gillies 1	G IH -L IH Z	G IH -L IH Z	G IH -L IH Z	G	0	1
gist 1	JH IH S T	JH AX S T	JH AX S T	0	0	1
given 1	”G IH -V AX N	”G IH -V AX N	”G IH -V AX N	G	0	1
guillemots 1	G IH -L IH -M OH T S	G IH -L IH -M OH T S	G IH -L AX -M OH T S	G	0	1
guillemots 3	G IH -L IH -M OH T S	G IH -L AX -M OH T S	G IH -L AX -M OH T S	0	0	1

Table D.3: Results when applying selected adaptation rules (part 3)

word	pronun_base	pronun_predict	pronun_corr	Rule 2	Rule 3	corr
guiltier 1	G IH L -T IA R	G IH L -T IA R	G IH L -T IA R	G	0	1
gyrating 6	"JH AY -AX -'R EY -T IH NG	"JH AY -AX -'R EY -T IH NG	"JH AY -AX -'R EY -T IH NG	NG	0	1
hilt 1	H IH L T	H IH L T	H IH L T	H	0	1
historical 1	H IH -"S T OH -R IH -K L	H IH -"S T OH -R IH -K L	H IH -"S T OH -R IH -K L	H	0	1
historical 6	H IH -"S T OH -R IH -K L	H IH -"S T OH -R IH -K L	H IH -"S T OH -R IH -K L	K	0	1
hypnotize 1	H IH P -N AX -T AY Z	H IH P -N AX -T AY Z	H IH P -N AX -T AY Z	H	0	1
illusionists 0	IH -'L UH -ZH AX -N IH S T S	IH -'L UH -ZH AX -N IH S T S	IH -'L UH -ZH AX -N AX S T S		0	1
illusionists 6	IH -'L UH -ZH AX -N IH S T S	AX -'L UH -ZH AX -N AX S T S	IH -'L UH -ZH AX -N AX S T S	0	0	1
in 0	IH N	IH N	IH N		0	1
inanities 0	IH N -'AE -N AX -T IH Z	IH N -'AE -N AX -T IH Z	IH N -'AE -N AX -T IH Z		0	1
infectiousness 0	IH N -'F EH K -SH AX S -N IH S	IH N -'F EH K -SH AX S -N IH S	IH N -'F EH K -SH AX S -N AX S		0	1
infectiousness 9	IH N -'F EH K -SH AX S -N IH S	AX N -'F EH K -SH AX S -N AX S	IH N -'F EH K -SH AX S -N AX S	0	0	1
infusions 0	IH N -'F Y UH -ZH AX N Z	IH N -'F Y UH -ZH AX N Z	IH N -'F Y UH -ZH AX N Z		0	1
inner 0	IH -N AX R	IH -N AX R	IH -N AX R		0	1
insipidity 0	"IH N -S IH -'P IH -D AX -T IH	"IH N -S IH -'P IH -D AX -T IH	"IH N -S AX -'P AX -D AX -T IH		0	1
insipidity 3	"IH N -S IH -'P IH -D AX -T IH	"AX N -S AX -'P IH -D AX -T IH	"IH N -S AX -'P AX -D AX -T IH	0	0	1
insipidity 5	"IH N -S IH -'P IH -D AX -T IH	"AX N -S AX -'P AX -D AX -T IH	"IH N -S AX -'P AX -D AX -T IH	0	0	1
instrument 0	IH N -S T R UH -M AX N T	IH N -S T R UH -M AX N T	IH N -S T R UH -M AX N T		0	1
interceptions 0	"IH N -T AX -'S EH P -SH AX N Z	"IH N -T AX -'S EH P -SH AX N Z	"IH N -T AX -'S EH P -SH AX N Z		0	1
intermediate 0	"IH N -T AX -'M IY -D Y AX T	"IH N -T AX -'M IY -D Y AX T	"IH N -T AX -'M IY -D Y AX T		0	1
intern 0	IH N -T ER N	IH N -T ER N	IH N -T ER N		0	1
into 0	"IH N -T UH	"IH N -T UH	"IH N -T UH		0	1
introspects 0	"IH N -T R OW -'S P EH K T S	"IH N -T R OW -'S P EH K T S	"IH N -T R OW -'S P EH K T S		0	1
inundated 0	IH -N AH N -D EY -T IH D	IH -N AH N -D EY -T IH D	IH -N AH N -D EY -T AX D		0	1
invites 0	IH N -'V AY T S	IH N -'V AY T S	IH N -'V AY T S		0	1
inundated 7	IH -N AH N -D EY -T IH D	AX -N AH N -D EY -T AX D	IH -N AH N -D EY -T AX D	0	0	1
jiminy 1	JH IH -M AX -N IH	JH AX -M AX -N IH	JH AX -M AX -N IH	0	0	1
junkies 4	JH AH NG -K IH Z	JH AH NG -K IH Z	JH AH NG -K IH Z	K	0	1
kicking 1	K IH -K IH NG	K IH -K IH NG	K IH -K IH NG	K	0	1
kicking 3	K IH -K IH NG	K IH -K IH NG	K IH -K IH NG	NG	0	1
laminate 3	L AE -M IH -N EY T	L AE -M AX -N EY T	L AE -M AX -N EY T	0	0	1
larynxes 3	L AE -R IH NG K -S IH Z	L AE -R IH NG K -S IH Z	L AE -R IH NG K -S AX Z	NG	0	1
larynxes 7	L AE -R IH NG K -S IH Z	L AE -R IH NG K -S AX Z	L AE -R IH NG K -S AX Z	0	0	1
licensing 6	L AY -S AX N -S IH NG	L AY -S AX N -S IH NG	L AY -S AX N -S IH NG	NG	0	1
linguists 1	L IH NG -G W IH S T S	L IH NG -G W IH S T S	L IH NG -G W AX S T S	NG	0	1
linguists 5	L IH NG -G W IH S T S	L IH NG -G W AX S T S	L IH NG -G W AX S T S	0	0	1
liquidates 1	L IH -K W IH -D EY T S	L IH -K W IH -D EY T S	L IH -K W AX -D EY T S	K	0	1
liquidates 4	L IH -K W IH -D EY T S	L IH -K W AX -D EY T S	L IH -K W AX -D EY T S	0	0	1
littlest 1	L IH -T L -IH S T	L AX -T L -IH S T	L AX -T L -AX S T	0	0	1
littlest 4	L IH -T L -IH S T	L AX -T L -AX S T	L AX -T L -AX S T	0	0	1
looking 3	"L UH -K IH NG	"L UH -K IH NG	"L UH -K IH NG	NG	0	1
Lucifer 3	L UH -S IH -F AX R	L UH -S AX -F AX R	L UH -S AX -F AX R	0	0	1
mainsprings 6	M EY N -S P R IH NG Z	M EY N -S P R IH NG Z	M EY N -S P R IH NG Z	NG	0	1
metallic 1	M IH -'T AE -L IH K	M AX -'T AE -L IH K	M AX -'T AE -L IH K	0	0	1
metallic 5	M IH -'T AE -L IH K	M AX -'T AE -L IH K	M AX -'T AE -L IH K	K	0	1
muzzily 3	M AH -Z IH -L IH	M AH -Z AX -L IH	M AH -Z AX -L IH	0	0	1
nastily 4	N AA -S T IH -L IH	N AA -S T AX -L IH	N AA -S T AX -L IH	0	0	1
neckerchief 5	N EH -K AX -CH IH F	N EH -K AX -CH AX F	N EH -K AX -CH AX F	0	0	1
Orangemen 2	OH -R IH N JH -M AX N	OH -R IH N JH -M AX N	OH -R IH N JH -M AX N	0	N JH	1
overdresses 7	"OW -V AX -'D R EH -S IH Z	"OW -V AX -'D R EH -S AX Z	"OW -V AX -'D R EH -S AX Z	0	0	1
particularize 3	P AX -'T IH -K Y UH -L AX -R AY Z	P AX -'T IH -K Y UH -L AX -R AY Z	P AX -'T IH -K Y UH -L AX -R AY Z	K	0	1
pathologically 7	"P AE -TH AX -'L OH -JH IH -K AX -L IH	"P AE -TH AX -'L OH -JH IH -K AX -L IH	"P AE -TH AX -'L OH -JH IH -K AX -L IH	K	0	1
pausing 3	P AO -Z IH NG	P AO -Z IH NG	P AO -Z IH NG	NG	0	1
perpetrate 3	P ER -P IH -T R EY T	P ER -P AX -T R EY T	P ER -P AX -T R EY T	0	0	1
persevered 3	"P ER -S IH -'V IA D	"P ER -S AX -'V IA D	"P ER -S AX -'V IA D	0	0	1
piling 3	"P AY -L IH NG	"P AY -L IH NG	"P AY -L IH NG	NG	0	1
pill 1	P IH L	P AX L	P AX L	0	0	1
pingpong 1	P IH NG -P OH NG	P IH NG -P OH NG	P IH NG -P OH NG	NG	0	1
pirouetting 6	"P IH -R UH -'EH -T IH NG	"P AX -R UH -'EH -T IH NG	"P IH -R UH -'EH -T IH NG	NG	0	1
pitting 1	P IH -T IH NG	P AX -T IH NG	P AX -T IH NG	0	0	1
pitting 3	P IH -T IH NG	P AX -T IH NG	P AX -T IH NG	NG	0	1
planning 4	P L AE -N IH NG	P L AE -N IH NG	P L AE -N IH NG	NG	0	1
plaything 4	P L EY -TH IH NG	P L EY -TH IH NG	P L EY -TH IH NG	NG	0	1
pleasantness 8	P L EH -Z AX N T -N IH S	P L EH -Z AX N T -N AX S	P L EH -Z AX N T -N AX S	0	0	1
plucky 4	P L AH -K IH	P L AH -K IH	P L AH -K IH	K	0	1
politeness 6	P AX -'L AY T -N IH S	P AX -'L AY T -N AX S	P AX -'L AY T -N AX S	0	0	1
polyps 3	P OH -L IH P S	P OH -L AX P S	P OH -L AX P S	0	0	1
postulated 9	P OH S -T Y UH -L EY -T IH D	P OH S -T Y UH -L EY -T AX D	P OH S -T Y UH -L EY -T AX D	0	0	1
potholing 6	P OH T -"H OW -L IH NG	P OH T -"H OW -L IH NG	P OH T -"H OW -L IH NG	NG	0	1

Table D.4: Results when applying selected adaptation rules (part 4)

word	pronun_base	pronun_predict	pronun_corr	Rule 2	Rule 3	corr
preconcerted 9	”P R IY -K AX N -’S ER -T IH D	”P R IY -K AX N -’S ER -T AX D	”P R IY -K AX N -’S ER -T AX D	0	0	1
presuppositions 8	”P R IY -S AH -P AX -’Z IH -SH AX N Z	”P R IY -S AH -P AX -’Z IH -SH AX N Z	”P R IY -S AH -P AX -’Z IH -SH AX N Z	0	SH	1
prognosticating 12	P R OH G -’N OH -S T IH -K EY -T IH NG	P R OH G -’N OH -S T IH -K EY -T IH NG	P R OH G -’N OH -S T IH -K EY -T IH NG	NG	0	1
prognosticating 8	P R OH G -’N OH -S T IH -K EY -T IH NG	P R OH G -’N OH -S T IH -K EY -T IH NG	P R OH G -’N OH -S T IH -K EY -T IH NG	K	0	1
prohibited 4	P R AX -’H IH -B IH -T IH D	P R AX -’H IH -B IH -T IH D	P R AX -’H IH -B AX -T AX D	H	0	1
prohibited 6	P R AX -’H IH -B IH -T IH D	P R AX -’H IH -B AX -T IH D	P R AX -’H IH -B AX -T AX D	0	0	1
prohibited 8	P R AX -’H IH -B IH -T IH D	P R AX -’H IH -B AX -T AX D	P R AX -’H IH -B AX -T AX D	0	0	1
protruded 7	P R AX -’T R UH -D IH D	P R AX -’T R UH -D AX D	P R AX -’T R UH -D AX D	0	0	1
public 4	”P AH -B L IH K	”P AH -B L IH K	”P AH -B L IH K	K	0	1
purism 4	P Y UA -R IH -Z AX M	P Y UA -R AX -Z AX M	P Y UA -R AX -Z AX M	0	0	1
races 3	R EY -S IH Z	R EY -S AX Z	R EY -S AX Z	0	0	1
recited 5	R IH -’S AY -T IH D	R AX -’S AY -T AX D	R IH -’S AY -T AX D	0	0	1
recommended 8	”R EH -K AX -’M EH N -D IH D	”R EH -K AX -’M EH N -D AX D	”R EH -K AX -’M EH N -D AX D	0	0	1
rejoices 1	R IH -’JH OY -S IH Z	R IH -’JH OY -S IH Z	R IH -’JH OY -S AX Z	0	JH	1
rejoices 5	R IH -’JH OY -S IH Z	R IH -’JH OY -S AX Z	R IH -’JH OY -S AX Z	0	0	1
replaying 5	”R IY -’P L EY -IH NG	”R IY -’P L EY -IH NG	”R IY -’P L EY -IH NG	NG	0	1
restoratives 1	R IH -’S T OH -R AX -T IH V Z	R AX -’S T OH -R AX -T IH V Z	R AX -’S T OH -R AX -T AX V Z	0	0	1
restoratives 8	R IH -’S T OH -R AX -T IH V Z	R AX -’S T OH -R AX -T AX V Z	R AX -’S T OH -R AX -T AX V Z	0	0	1
rib 1	R IH B	R AX B	R AX B	0	0	1
right-minded 7	”R AY T -’M AY N -D IH D	”R AY T -’M AY N -D AX D	”R AY T -’M AY N -D AX D	0	0	1
rotted 3	”R OH -T IH D	”R OH -T AX D	”R OH -T AX D	0	0	1
sacristan 4	S AE -K R IH -S T AX N	S AE -K R AX -S T AX N	S AE -K R AX -S T AX N	0	0	1
scariest 5	S K EA -R IH -IH S T	S K EA -R AX -AX S T	S K EA -R IH -AX S T	0	0	1
seducing 1	S IH -’D Y UH -S IH NG	S AX -’D Y UH -S IH NG	S AX -’D Y UH -S IH NG	0	0	1
seducing 6	S IH -’D Y UH -S IH NG	S AX -’D Y UH -S IH NG	S AX -’D Y UH -S IH NG	NG	0	1
select 1	S IH -’L EH K T	S AX -’L EH K T	S AX -’L EH K T	0	0	1
servility 3	S ER -’V IH -L AX -T IH	S ER -’V AX -L AX -T IH	S ER -’V AX -L AX -T IH	0	0	1
shingled 1	SH IH NG -G L D	SH IH NG -G L D	SH IH NG -G L D	NG	0	1
sits 1	”S IH T S	”S AX T S	”S AX T S	0	0	1
sphincter 2	S F IH NG K -T AX R	S F IH NG K -T AX R	S F IH NG K -T AX R	NG	0	1
squawking 5	S K W AO -K IH NG	S K W AO -K IH NG	S K W AO -K IH NG	NG	0	1
squeegeeing 6	”S K W IY -’JH IY -IH NG	”S K W IY -’JH IY -IH NG	”S K W IY -’JH IY -IH NG	NG	0	1
squirrels 3	S K W IH -R AX L Z	S K W AX -R AX L Z	S K W AX -R AX L Z	0	0	1
surefootedly 5	”SH AO -’F UH -T IH D -L IH	”SH AO -’F UH -T AX D -L IH	”SH AO -’F UH -T AX D -L IH	0	0	1
swiftest 2	S W IH F -T IH S T	S W AX F -T IH S T	S W AX F -T AX S T	0	0	1
swiftest 5	S W IH F -T IH S T	S W AX F -T AX S T	S W AX F -T AX S T	0	0	1
tailless 4	T EY L -L IH S	T EY L -L AX S	T EY L -L AX S	0	0	1
taking 3	T EY -K IH NG	T EY -K IH NG	T EY -K IH NG	NG	0	1
tarbooshes 5	T AA -’B UH -SH IH Z	T AA -’B UH -SH AX Z	T AA -’B UH -SH AX Z	0	0	1
tennis 3	T EH -N IH S	T EH -N AX S	T EH -N AX S	0	0	1
thalidomide 3	TH AX -’L IH -D AX -M AY D	TH AX -’L AX -D AX -M AY D	TH AX -’L AX -D AX -M AY D	0	0	1
tormenting 6	T AO -’M EH N -T IH NG	T AO -’M EH N -T IH NG	T AO -’M EH N -T IH NG	NG	0	1
trickle 2	T R IH -K L	T R IH -K L	T R IH -K L	K	0	1
trochaics 5	T R OW -’K EY -IH K S	T R OW -’K EY -IH K S	T R OW -’K EY -IH K S	K	0	1
twit 2	T W IH T	T W AX T	T W AX T	0	0	1
unhinging 3	”AH N -’H IH N -JH IH NG	”AH N -’H IH N -JH IH NG	”AH N -’H IH N -JH IH NG	H	0	1
unhinging 6	”AH N -’H IH N -JH IH NG	”AH N -’H IH N -JH IH NG	”AH N -’H IH N -JH IH NG	NG	0	1
unseated 5	”AH N -’S IY -T IH D	”AH N -’S IY -T AX D	”AH N -’S IY -T AX D	0	0	1
untinged 3	”AH N -’T IH N JH D	”AH N -’T IH N JH D	”AH N -’T IH N JH D	0	N JH	1
visible 1	V IH -Z AX -B L	V AX -Z AX -B L	V AX -Z AX -B L	0	0	1
warning 3	”W AO -N IH NG	”W AO -N IH NG	”W AO -N IH NG	NG	0	1
whippier 1	W IH -P IA R	W AX -P IA R	W AX -P IA R	0	0	1

Table D.5: Full adaptation rule system results (part 1)

word	pronun_base	pronun_predict	pronun_corr	strt eb,m,n,x	rule 2	rule 3	frst in wrd	end eys,ies,ied,ey	vwl hrmny ax	vwl hrmny ih	adj vwl hrmny ih	fnl in wrd	phnm src	corr
beatify 4	B IY -'AE -T IH -F AY	B IY -'AE -T IH -F AY	B IY -'AE -T AX -F AY	0	0	0	0	0	0	1	0	0	0	0
clumsiness 5	K L AH M -Z IH -N IH S	K L AH M -Z AX -N AX S	K L AH M -Z IH -N AX S	0	0	0	0	0	AX	0	0	0	0	0
demagnetized 6	"D IY -'M AE G -N IH -T AY Z D	"D IY -'M AE G -N IH -T AY Z D	"D IY -'M AE G -N AX -T AY Z D	0	0	0	0	0	0	1	0	0	0	0
elephantine 2	"EH -L IH -'F AE N -T AY N	"EH -L IH -'F AE N -T AY N	"EH -L AX -'F AE N -T AY N	0	0	0	0	0	0	1	0	0	0	0
entrap 0	IH N -'T R AE P	EH N -'T R AE P	IH N -'T R AE P	1	0	0	1	0	0	0	0	0	0	0
generically 1	JH IH -'N EH -R IH -K AX -L IH	JH IH -'N EH -R IH -K AX -L IH	JH AX -'N EH -R IH -K AX -L IH	0	0	0	0	0	0	1	0	0	0	0
guillemots 3	G IH -L IH -M OH T S	G IH -L IH -M OH T S	G IH -L AX -M OH T S	0	0	0	0	0	0	1	0	0	0	0
metallic 1	M IH -'T AE -L IH K	M IH -'T AE -L IH K	M AX -'T AE -L IH K	0	0	0	0	0	0	1	0	0	0	0
muzzily 3	M AH -Z IH -L IH	M AH -Z IH -L IH	M AH -Z AX -L IH	0	0	0	0	0	0	1	0	0	0	0
nastily 4	N AA -S T IH -L IH	N AA -S T IH -L IH	N AA -S T AX -L IH	0	0	0	0	0	0	1	0	0	0	0
persevered 3	"P ER -S IH -'V IA D	"P ER -S IH -'V IA D	"P ER -S AX -'V IA D	0	0	0	0	0	0	1	0	0	0	0
pitting 1	P IH -T IH NG	P IH -T IH NG	P AX -T IH NG	0	0	0	0	0	0	1	0	0	0	0
restoratives 1	R IH -'S T OH -R AX -T IH V Z	R IH -'S T OH -R AX -T AX V Z	R AX -'S T OH -R AX -T AX V Z	0	0	0	0	0	0	1	0	0	0	0
seducing 1	S IH -'D Y UH -S IH NG	S IH -'D Y UH -S IH NG	S AX -'D Y UH -S IH NG	0	0	0	0	0	0	1	0	0	0	0
select 1	S IH -'L EH K T	S IH -'L EH K T	S AX -'L EH K T	0	0	0	0	0	0	1	0	0	0	0
whippier 1	W IH -P IA R	W IH -P IA R	W AX -P IA R	0	0	0	0	0	0	1	0	0	0	0
abbreviations 5	AX -"B R IY -V IH -'EY -SH AX N Z	AX -"B R IY -V IH -'EY -SH AX N Z	AX -"B R IY -V IH -'EY -SH AX N Z	0	0	0	0	0	0	0	EY	0	0	1
abstinent 4	AE B -S T IH -N AX N T	AE B -S T AX -N AX N T	AE B -S T AX -N AX N T	0	0	0	0	0	AX	0	0	0	0	1
addresses 5	AX -'D R EH -S IH Z	AX -'D R EH -S AX Z	AX -'D R EH -S AX Z	0	0	0	0	0	0	0	0	0	0	1
affidavits 2	"AE -F IH -'D EY -V IH T S	"AE -F AX -'D EY -V AX T S	"AE -F AX -'D EY -V AX T S	0	0	0	0	0	EY	0	0	0	0	1
affidavits 6	"AE -F IH -'D EY -V IH T S	"AE -F AX -'D EY -V AX T S	"AE -F AX -'D EY -V AX T S	0	0	0	0	0	0	0	0	0	0	1
aggrieving 5	AX -'G R IY -V IH NG	AX -'G R IY -V IH NG	AX -'G R IY -V IH NG	0	NG	0	0	0	0	0	0	0	0	1
amulet 5	AE -M Y UH -L IH T	AE -M Y UH -L AX T	AE -M Y UH -L AX T	0	0	0	0	0	0	0	0	0	0	1
apposing 4	AE -'P OW -Z IH NG	AE -'P OW -Z IH NG	AE -'P OW -Z IH NG	0	NG	0	0	0	0	0	0	0	0	1
assuming 5	AX -'S Y UH -M IH NG	AX -'S Y UH -M IH NG	AX -'S Y UH -M IH NG	0	NG	0	0	0	0	0	0	0	0	1
auditor 2	AO -D IH -T AX R	AO -D AX -T AX R	AO -D AX -T AX R	0	0	0	0	0	AX	0	0	0	0	1
babies 3	"B EY -B IH Z	"B EY -B IH Z	"B EY -B IH Z	0	0	0	0	1	0	0	0	0	0	1
badness 4	B AE D -N IH S	B AE D -N AX S	B AE D -N AX S	0	0	0	0	0	0	0	0	0	0	1
baffling 4	B AE -F L IH NG	B AE -F L IH NG	B AE -F L IH NG	0	NG	0	0	0	0	0	0	0	0	1
bedevilled 1	B IH -'D EH -V L D	B IH -'D EH -V L D	B IH -'D EH -V L D	0	0	0	0	0	0	1	0	0	0	1
belies 1	B IH -'L AY Z	B IH -'L AY Z	B IH -'L AY Z	0	0	0	0	0	0	1	0	0	0	1
beloved 1	B IH -'L AH V D	B IH -'L AH V D	B IH -'L AH V D	0	0	0	0	0	0	1	0	0	0	1
big 1	"B IH G	"B IH G	"B IH G	0	G	0	0	0	0	0	0	0	0	1
billets 1	B IH -L IH T S	B AX -L AX T S	B AX -L AX T S	0	0	0	0	0	AX	0	0	0	0	1
billets 3	B IH -L IH T S	B AX -L AX T S	B AX -L AX T S	0	0	0	0	0	0	0	0	0	0	1
blandishes 5	B L AE N -D IH -SH IH Z	B L AE N -D IH -SH IH Z	B L AE N -D IH -SH AX Z	0	0	SH	0	0	0	0	0	0	0	1
blandishes 7	B L AE N -D IH -SH IH Z	B L AE N -D IH -SH AX Z	B L AE N -D IH -SH AX Z	0	0	0	0	0	0	0	0	0	0	1
blanket 5	B L AE NG -K IH T	B L AE NG -K AX T	B L AE NG -K AX T	0	K	0	0	0	0	0	0	0	1	1
blooding 4	B L AH -D IH NG	B L AH -D IH NG	B L AH -D IH NG	0	NG	0	0	0	0	0	0	0	0	1
bonnets 3	B OH -N IH T S	B OH -N AX T S	B OH -N AX T S	0	0	0	0	0	0	0	0	0	0	1
brings 2	"B R IH NG Z	"B R IH NG Z	"B R IH NG Z	0	NG	0	0	0	0	0	0	0	0	1
bristols 2	B R IH -S T L Z	B R AX -S T L Z	B R AX -S T L Z	0	0	0	0	0	AX	0	0	0	0	1
captive 4	"K AE P -T IH V	"K AE P -T AX V	"K AE P -T AX V	0	0	0	0	0	0	0	0	0	0	1
carbonate 5	K AA -B AX -N IH T	K AA -B AX -N AX T	K AA -B AX -N AX T	0	0	0	0	0	0	0	0	0	0	1
carrying 3	"K AE -R IH -IH NG	"K AE -R IH -IH NG	"K AE -R IH -IH NG	0	0	0	0	0	0	1	0	0	0	1
carrying 4	"K AE -R IH -IH NG	"K AE -R AX -IH NG	"K AE -R IH -IH NG	0	NG	0	0	0	0	0	0	0	0	1
carving 3	K AA -V IH NG	K AA -V IH NG	K AA -V IH NG	0	NG	0	0	0	0	0	0	0	0	1
circumscription 8	"S ER -K AX M -'S K R IH P -SH AX N	"S ER -K AX M -'S K R AX P -SH AX N	"S ER -K AX M -'S K R AX P -SH AX N	0	0	0	0	0	0	0	0	0	0	1
clumsiness 7	K L AH M -Z IH -N IH S	K L AH M -Z AX -N AX S	K L AH M -Z IH -N AX S	0	0	0	0	0	0	0	0	0	0	1
cockney 4	K OH K -N IH	K OH K -N IH	K OH K -N IH	0	0	0	0	0	0	0	0	1	0	1
contagiousness 9	K AX N -'T EY -JH AX S -N IH S	K AX N -'T EY -JH AX S -N AX S	K AX N -'T EY -JH AX S -N AX S	0	0	0	0	0	0	0	0	0	0	1
contaminates 6	K AX N -'T AE -M IH -N EY T S	K AX N -'T AE -M AX -N EY T S	K AX N -'T AE -M AX -N EY T S	0	0	0	0	0	EY	0	0	0	0	1
continue 4	K AX N -'T IH -N Y UH	K AX N -'T IH -N Y UH	K AX N -'T IH -N Y UH	0	0	0	0	0	0	1	0	0	0	1
cordiality 3	"K AO -D IH -'AE -L AX -T IH	"K AO -D IH -'AE -L AX -T IH	"K AO -D IH -'AE -L AX -T IH	0	0	0	0	0	0	1	0	0	0	1
cordiality 8	"K AO -D IH -'AE -L AX -T IH	"K AO -D AX -'AE -L AX -T IH	"K AO -D IH -'AE -L AX -T IH	0	0	0	0	0	0	0	0	1	0	1
coterminously 10	"K OW -'T ER -M IH -N AX S -L IH	"K OW -'T ER -M AX -N AX S -L IH	"K OW -'T ER -M AX -N AX S -L IH	0	0	0	0	0	0	0	0	1	0	1
coterminously 5	"K OW -'T ER -M IH -N AX S -L IH	"K OW -'T ER -M AX -N AX S -L IH	"K OW -'T ER -M AX -N AX S -L IH	0	0	0	0	0	AX	0	0	0	0	1
counterpoising 8	K AW N -T AX -P OY -Z IH NG	K AW N -T AX -P OY -Z IH NG	K AW N -T AX -P OY -Z IH NG	0	NG	0	0	0	0	0	0	0	0	1
curetted 6	K Y UA -'R EH -T IH D	K Y UA -'R EH -T AX D	K Y UA -'R EH -T AX D	0	0	0	0	0	0	0	0	0	0	1
descended 1	D IH -'S EH N -D IH D	D IH -'S EH N -D AX D	D IH -'S EH N -D AX D	0	0	0	0	0	0	1	0	0	0	1
descended 6	D IH -'S EH N -D IH D	D AX -'S EH N -D AX D	D IH -'S EH N -D AX D	0	0	0	0	0	0	0	0	0	0	1
dignitary 1	D IH G -N IH -T AX -R IH	D IH G -N IH -T AX -R IH	D IH G -N AX -T AX -R IH	0	G	0	0	0	0	0	0	0	0	1
dignitary 4	D IH G -N IH -T AX -R IH	D IH G -N AX -T AX -R IH	D IH G -N AX -T AX -R IH	0	0	0	0	0	AX	0	0	0	0	1
dignitary 8	D IH G -N IH -T AX -R IH	D IH G -N AX -T AX -R IH	D IH G -N AX -T AX -R IH	0	0	0	0	0	0	0	0	1	0	1

Table D.6: Full adaptation rule system results (part 2)

word	pronun_base	pronun_predict	pronun_corr	strt eb,m,n,x	rule 2	rule 3	frst in wrd	end eys,ies,ied,ey	vwl hrmny ax	vwl hrmny ih	adj vwl hrmny ih	fnl in wrd	phnm src	corr
dilapidated 1	D IH -'L AE -P IH -D EY -T IH D	D IH -'L AE -P AX -D EY -T AX D	D IH -'L AE -P AX -D EY -T AX D	0	0	0	0	0	0	1	0	0	0	1
dilapidated 5	D IH -'L AE -P IH -D EY -T IH D	D AX -'L AE -P AX -D EY -T AX D	D IH -'L AE -P AX -D EY -T AX D	0	0	0	0	0	EY	0	0	0	0	1
dilapidated 9	D IH -'L AE -P IH -D EY -T IH D	D AX -'L AE -P AX -D EY -T AX D	D IH -'L AE -P AX -D EY -T AX D	0	0	0	0	0	0	0	0	0	0	1
din 1	D IH N	D AX N	D AX N	0	0	0	0	0	0	0	0	0	0	1
dint 1	D IH N T	D AX N T	D AX N T	0	0	0	0	0	0	0	0	0	0	1
dowdy 3	D AW -D IH	D AW -D IH	D AW -D IH	0	0	0	0	0	0	0	0	1	0	1
dwellings 4	”D W EH -L IH NG Z	”D W EH -L IH NG Z	”D W EH -L IH NG Z	0	NG	0	0	0	0	0	0	0	0	1
eaglet 3	IY G -L IH T	IY G -L AX T	IY G -L AX T	0	0	0	0	0	0	0	0	0	0	1
ebullience 0	IH -'B AH L -Y AX N S	IH -'B AH L -Y AX N S	IH -'B AH L -Y AX N S	0	0	0	1	0	0	0	0	0	0	1
edit 2	”EH -D IH T	”EH -D AX T	”EH -D AX T	0	0	0	0	0	0	0	0	0	0	1
elaborateness 0	IH -'L AE -B AX -R AX T -N IH S	IH -'L AE -B AX -R AX T -N IH S	IH -'L AE -B AX -R AX T -N AX S	0	0	0	1	0	0	0	0	0	0	1
elaborateness 9	IH -'L AE -B AX -R AX T -N IH S	IH -'L AE -B AX -R AX T -N AX S	IH -'L AE -B AX -R AX T -N AX S	0	0	0	0	0	0	0	0	0	0	1
elaborating 0	IH -'L AE -B AX -R EY -T IH NG	IH -'L AE -B AX -R EY -T IH NG	IH -'L AE -B AX -R EY -T IH NG	0	NG	0	1	0	0	0	0	0	0	1
elaborating 8	IH -'L AE -B AX -R EY -T IH NG	IH -'L AE -B AX -R EY -T IH NG	IH -'L AE -B AX -R EY -T IH NG	0	NG	0	0	0	0	0	0	0	0	1
emergencies 0	IH -'M ER -JH AX N -S IH Z	IH -'M ER -JH AX N -S IH Z	IH -'M ER -JH AX N -S IH Z	0	0	0	1	0	0	0	0	0	0	1
emergencies 7	IH -'M ER -JH AX N -S IH Z	IH -'M ER -JH AX N -S IH Z	IH -'M ER -JH AX N -S IH Z	0	0	0	0	1	0	0	0	0	0	1
enable 0	IH -'N EY -B L	IH -'N EY -B L	IH -'N EY -B L	0	0	0	1	0	0	0	0	0	0	1
evacuate 0	IH -'V AE -K Y UH -EY T	IH -'V AE -K Y UH -EY T	IH -'V AE -K Y UH -EY T	0	0	0	1	0	0	0	0	0	0	1
everything 3	EH -V R IH -TH IH NG	EH -V R IH -TH IH NG	EH -V R IH -TH IH NG	0	0	0	0	0	0	1	0	0	0	1
everything 5	EH -V R IH -TH IH NG	EH -V R AX -TH IH NG	EH -V R IH -TH IH NG	0	NG	0	0	0	0	0	0	0	0	1
examinations 0	IH G -”Z AE -M IH -'N EY -SH AX N Z	EH G -”Z AE -M IH -'N EY -SH AX N Z	EH G -”Z AE -M AX -'N EY -SH AX N Z	1	G	0	1	0	0	0	0	0	0	1
examinations 5	IH G -”Z AE -M IH -'N EY -SH AX N Z	EH G -”Z AE -M AX -'N EY -SH AX N Z	EH G -”Z AE -M AX -'N EY -SH AX N Z	0	0	0	0	0	EY	0	0	0	0	1
extremely 0	IH K -'S T R IY M -L IH	EH K -'S T R IY M -L IH	EH K -'S T R IY M -L IH	1	K	0	1	0	0	0	0	0	0	1
extremely 8	IH K -'S T R IY M -L IH	EH K -'S T R IY M -L IH	EH K -'S T R IY M -L IH	0	0	0	0	0	0	0	0	1	0	1
factory 6	”F AE K -T AX -R IH	”F AE K -T AX -R IH	”F AE K -T AX -R IH	0	0	0	0	0	0	0	0	1	0	1
family 5	”F AE -M AX -L IH	”F AE -M AX -L IH	”F AE -M AX -L IH	0	0	0	0	0	0	0	0	1	0	1
farming 3	F AA -M IH NG	F AA -M IH NG	F AA -M IH NG	0	NG	0	0	0	0	0	0	0	0	1
felonies 5	F EH -L AX -N IH Z	F EH -L AX -N IH Z	F EH -L AX -N IH Z	0	0	0	0	1	0	0	0	0	0	1
filtertipped 1	”F IH L -T AX -'T IH P T	”F AX L -T AX -'T IH P T	”F AX L -T AX -'T AX P T	0	0	0	0	0	0	0	0	0	0	1
filtertipped 6	”F IH L -T AX -'T IH P T	”F AX L -T AX -'T AX P T	”F AX L -T AX -'T AX P T	0	0	0	0	0	0	0	0	0	0	1
fishy 1	F IH -SH IH	F IH -SH IH	F IH -SH IH	0	0	SH	0	0	0	0	0	0	0	1
fishy 3	F IH -SH IH	F IH -SH IH	F IH -SH IH	0	0	0	0	0	0	0	0	1	0	1
forbidden 3	F AX -”B IH -D AX N	F AX -”B AX -D AX N	F AX -”B AX -D AX N	0	0	0	0	0	AX	0	0	0	0	1
formalities 7	F AO -'M AE -L AX -T IH Z	F AO -'M AE -L AX -T IH Z	F AO -'M AE -L AX -T IH Z	0	0	0	0	1	0	0	0	0	0	1
friendship 6	F R EH N D -SH IH P	F R EH N D -SH AX P	F R EH N D -SH AX P	0	0	0	0	0	0	0	0	0	0	1
gaining 3	G EY -N IH NG -AX	G EY -N IH NG -AX	G EY -N IH NG -AX	0	NG	0	0	0	0	0	0	0	0	1
generically 5	JH IH -'N EH -R IH -K AX -L IH	JH AX -'N EH -R IH -K AX -L IH	JH AX -'N EH -R IH -K AX -L IH	0	K	0	0	0	0	0	0	0	0	1
generically 9	JH IH -'N EH -R IH -K AX -L IH	JH AX -'N EH -R IH -K AX -L IH	JH AX -'N EH -R IH -K AX -L IH	0	0	0	0	0	0	0	0	1	0	1
gillies 1	G IH -L IH Z	G IH -L IH Z	G IH -L IH Z	0	G	0	0	0	0	0	0	0	0	1
gillies 3	G IH -L IH Z	G IH -L IH Z	G IH -L IH Z	0	0	0	0	1	0	0	0	0	0	1
gist 1	JH IH S T	JH AX S T	JH AX S T	0	0	0	0	0	0	0	0	0	0	1
given 1	”G IH -V AX N	”G IH -V AX N	”G IH -V AX N	0	G	0	0	0	0	0	0	0	0	1
guillemots 1	G IH -L IH -M OH T S	G IH -L IH -M OH T S	G IH -L AX -M OH T S	0	G	0	0	0	0	0	0	0	0	1
guiltier 1	G IH L -T IA R	G IH L -T IA R	G IH L -T IA R	0	G	0	0	0	0	0	0	0	0	1
gustatory 8	G AH -S T AX -T AX -R IH	G AH -S T AX -T AX -R IH	G AH -S T AX -T AX -R IH	0	0	0	0	0	0	0	0	1	0	1
gyrating 6	”JH AY -AX -'R EY -T IH NG	”JH AY -AX -'R EY -T IH NG	”JH AY -AX -'R EY -T IH NG	0	NG	0	0	0	0	0	0	0	0	1
hilt 1	H IH L T	H IH L T	H IH L T	0	H	0	0	0	0	0	0	0	0	1
historical 1	H IH -”S T OH -R IH -K L	H IH -”S T OH -R IH -K L	H IH -”S T OH -R IH -K L	0	H	0	0	0	0	0	0	0	0	1
historical 6	H IH -”S T OH -R IH -K L	H IH -”S T OH -R IH -K L	H IH -”S T OH -R IH -K L	0	K	0	0	0	0	0	0	0	0	1
honeymooners 3	H AH -N IH -”M UH -N AX Z	H AH -N IH -”M UH -N AX Z	H AH -N IH -”M UH -N AX Z	0	0	0	0	0	0	1	0	0	0	1
hypnotize 1	H IH P -N AX -T AY Z	H IH P -N AX -T AY Z	H IH P -N AX -T AY Z	0	H	0	0	0	0	0	0	0	0	1
illusionists 0	IH -'L UH -ZH AX -N IH S T S	IH -'L UH -ZH AX -N IH S T S	IH -'L UH -ZH AX -N AX S T S	0	0	0	1	0	0	0	0	0	0	1
illusionists 6	IH -'L UH -ZH AX -N IH S T S	IH -'L UH -ZH AX -N AX S T S	IH -'L UH -ZH AX -N AX S T S	0	0	0	0	0	0	0	0	0	0	1
in 0	IH N	IH N	IH N	0	0	0	1	0	0	0	0	0	0	1
inanities 0	IH N -'AE -N AX -T IH Z	IH N -'AE -N AX -T IH Z	IH N -'AE -N AX -T IH Z	0	0	0	1	0	0	0	0	0	0	1
inanities 6	IH N -'AE -N AX -T IH Z	IH N -'AE -N AX -T IH Z	IH N -'AE -N AX -T IH Z	0	0	0	0	1	0	0	0	0	0	1
infectiousness 0	IH N -'F EH K -SH AX S -N IH S	IH N -'F EH K -SH AX S -N IH S	IH N -'F EH K -SH AX S -N AX S	0	0	0	1	0	0	0	0	0	0	1
infectiousness 9	IH N -'F EH K -SH AX S -N IH S	IH N -'F EH K -SH AX S -N AX S	IH N -'F EH K -SH AX S -N AX S	0	0	0	0	0	0	0	0	0	0	1
infusions 0	IH N -'F Y UH -ZH AX N Z	IH N -'F Y UH -ZH AX N Z	IH N -'F Y UH -ZH AX N Z	0	0	0	1	0	0	0	0	0	0	1
inner 0	IH -N AX R	IH -N AX R	IH -N AX R	0	0	0	1	0	0	0	0	0	0	1
insipidity 0	”IH N -S IH -'P IH -D AX -T IH	”IH N -S IH -'P IH -D AX -T IH	”IH N -S AX -'P AX -D AX -T IH	0	0	0	1	0	0	0	0	0	0	1
insipidity 3	”IH N -S IH -'P IH -D AX -T IH	”IH N -S AX -'P AX -D AX -T IH	”IH N -S AX -'P AX -D AX -T IH	0	0	0	0	0	AX	0	0	0	0	1
insipidity 5	”IH N -S IH -'P IH -D AX -T IH	”IH N -S AX -'P AX -D AX -T IH	”IH N -S AX -'P AX -D AX -T IH	0	0	0	0	0	AX	0	0	0	0	1
insipidity 9	”IH N -S IH -'P IH -D AX -T IH	”IH N -S AX -'P AX -D AX -T IH	”IH N -S AX -'P AX -D AX -T IH	0	0	0	0	0	0	0	0	1	0	1
instrument 0	IH N -S T R UH -M AX N T	IH N -S T R UH -M AX N T	IH N -S T R UH -M AX N T	0	0	0	1	0	0	0	0	0	0	1

Table D.7: Full adaptation rule system results (part 3)

word	pronun_base	pronun_predict	pronun_corr	strt eb,m,n,x	rule 2	rule 3	frst in wrd	end eys,ies,ied,ey	vwl hrmny ax	vwl hrmny ih	adj vwl hrmny ih	fnl in wrd	phnm src	corr
interceptions 0	"IH N -T AX -'S EH P -SH AX N Z	"IH N -T AX -'S EH P -SH AX N Z	"IH N -T AX -'S EH P -SH AX N Z	0	0	0	1	0	0	0	0	0	0	1
intermediate 0	"IH N -T AX -'M IY -D Y AX T	"IH N -T AX -'M IY -D Y AX T	"IH N -T AX -'M IY -D Y AX T	0	0	0	1	0	0	0	0	0	0	1
intern 0	IH N -T ER N	IH N -T ER N	IH N -T ER N	0	0	0	1	0	0	0	0	0	0	1
into 0	"IH N -T UH	"IH N -T UH	"IH N -T UH	0	0	0	1	0	0	0	0	0	0	1
introspects 0	"IH N -T R OW -'S P EH K T S	"IH N -T R OW -'S P EH K T S	"IH N -T R OW -'S P EH K T S	0	0	0	1	0	0	0	0	0	0	1
inundated 0	IH -N AH N -D EY -T IH D	IH -N AH N -D EY -T IH D	IH -N AH N -D EY -T AX D	0	0	0	1	0	0	0	0	0	0	1
inundated 7	IH -N AH N -D EY -T IH D	IH -N AH N -D EY -T AX D	IH -N AH N -D EY -T AX D	0	0	0	0	0	0	0	0	0	0	1
invites 0	IH N -'V AY T S	IH N -'V AY T S	IH N -'V AY T S	0	0	0	1	0	0	0	0	0	0	1
jiminy 1	JH IH -M AX -N IH	JH AX -M AX -N IH	JH AX -M AX -N IH	0	0	0	0	0	AX	0	0	0	0	1
jiminy 5	JH IH -M AX -N IH	JH AX -M AX -N IH	JH AX -M AX -N IH	0	0	0	0	0	0	0	0	1	0	1
junkies 4	JH AH NG -K IH Z	JH AH NG -K IH Z	JH AH NG -K IH Z	0	K	0	0	1	0	0	0	0	0	1
kicking 1	K IH -K IH NG	K IH -K IH NG	K IH -K IH NG	0	K	0	0	0	0	0	0	0	0	1
kicking 3	K IH -K IH NG	K IH -K IH NG	K IH -K IH NG	0	NG	0	0	0	0	0	0	0	0	1
laminate 3	L AE -M IH -N EY T	L AE -M AX -N EY T	L AE -M AX -N EY T	0	0	0	0	0	EY	0	0	0	0	1
larynxes 3	L AE -R IH NG K -S IH Z	L AE -R IH NG K -S IH Z	L AE -R IH NG K -S AX Z	0	NG	0	0	0	0	0	0	0	0	1
larynxes 7	L AE -R IH NG K -S IH Z	L AE -R IH NG K -S AX Z	L AE -R IH NG K -S AX Z	0	0	0	0	0	0	0	0	0	0	1
licensing 6	L AY -S AX N -S IH NG	L AY -S AX N -S IH NG	L AY -S AX N -S IH NG	0	NG	0	0	0	0	0	0	0	0	1
linguists 1	L IH NG -G W IH S T S	L IH NG -G W IH S T S	L IH NG -G W AX S T S	0	NG	0	0	0	0	0	0	0	0	1
linguists 5	L IH NG -G W IH S T S	L IH NG -G W AX S T S	L IH NG -G W AX S T S	0	0	0	0	0	0	0	0	0	0	1
liquidates 1	L IH -K W IH -D EY T S	L IH -K W IH -D EY T S	L IH -K W AX -D EY T S	0	K	0	0	0	0	0	0	0	0	1
liquidates 4	L IH -K W IH -D EY T S	L IH -K W AX -D EY T S	L IH -K W AX -D EY T S	0	0	0	0	0	EY	0	0	0	0	1
littlest 1	L IH -T L -IH S T	L AX -T L -AX S T	L AX -T L -AX S T	0	0	0	0	0	AX	0	0	0	0	1
littlest 4	L IH -T L -IH S T	L AX -T L -AX S T	L AX -T L -AX S T	0	0	0	0	0	0	0	0	0	0	1
looking 3	"L UH -K IH NG	"L UH -K IH NG	"L UH -K IH NG	0	NG	0	0	0	0	0	0	0	0	1
Lucifer 3	L UH -S IH -F AX R	L UH -S AX -F AX R	L UH -S AX -F AX R	0	0	0	0	0	AX	0	0	0	0	1
mainsprings 6	M EY N -S P R IH NG Z	M EY N -S P R IH NG Z	M EY N -S P R IH NG Z	0	NG	0	0	0	0	0	0	0	0	1
metallic 5	M IH -'T AE -L IH K	M AX -'T AE -L IH K	M AX -'T AE -L IH K	0	K	0	0	0	0	0	0	0	0	1
meteor 3	M IY -T IH -AO R	M IY -T IH -AO R	M IY -T IH -AO R	0	0	0	0	0	0	1	0	0	0	1
minded 4	M AY N -D IH D	M AY N -D AX D	M AY N -D AX D	0	0	0	0	0	0	0	0	0	0	1
minority 7	M AY -"N OH -R AX -T IH	M AY -"N OH -R AX -T IH	M AY -"N OH -R AX -T IH	0	0	0	0	0	0	0	0	1	0	1
muzzily 5	M AH -Z IH -L IH	M AH -Z AX -L IH	M AH -Z AX -L IH	0	0	0	0	0	0	0	0	1	0	1
nastily 6	N AA -S T IH -L IH	N AA -S T AX -L IH	N AA -S T AX -L IH	0	0	0	0	0	0	0	0	1	0	1
neckerchief 5	N EH -K AX -CH IH F	N EH -K AX -CH AX F	N EH -K AX -CH AX F	0	0	0	0	0	0	0	0	0	0	1
nonconformity 11	"N OH N -K AX N -'F AO -M AX -T IH	"N OH N -K AX N -'F AO -M AX -T IH	"N OH N -K AX N -'F AO -M AX -T IH	0	0	0	0	0	0	0	0	1	0	1
Orangemen 2	OH -R IH N JH -M AX N	OH -R IH N JH -M AX N	OH -R IH N JH -M AX N	0	0	JH	0	0	0	0	0	0	0	1
overdresses 7	"OW -V AX -'D R EH -S IH Z	"OW -V AX -'D R EH -S AX Z	"OW -V AX -'D R EH -S AX Z	0	0	0	0	0	0	0	0	0	0	1
Parisians 3	P AX -'R IH -Z Y AX N Z	P AX -'R IH -Z Y AX N Z	P AX -'R IH -Z Y AX N Z	0	0	0	0	0	0	1	0	0	0	1
particularize 3	P AX -'T IH -K Y UH -L AX -R AY Z	P AX -'T IH -K Y UH -L AX -R AY Z	P AX -'T IH -K Y UH -L AX -R AY Z	0	K	0	0	0	0	0	0	0	0	1
pathologically 11	"P AE -TH AX -'L OH -JH IH -K AX -L IH	"P AE -TH AX -'L OH -JH IH -K AX -L IH	"P AE -TH AX -'L OH -JH IH -K AX -L IH	0	0	0	0	0	0	0	0	1	0	1
pathologically 7	"P AE -TH AX -'L OH -JH IH -K AX -L IH	"P AE -TH AX -'L OH -JH IH -K AX -L IH	"P AE -TH AX -'L OH -JH IH -K AX -L IH	0	K	0	0	0	0	0	0	0	0	1
patriarchal 4	"P EY -T R IH -'AA -K L	"P EY -T R IH -'AA -K L	"P EY -T R IH -'AA -K L	0	0	0	0	0	0	1	0	0	0	1
pausing 3	P AO -Z IH NG	P AO -Z IH NG	P AO -Z IH NG	0	NG	0	0	0	0	0	0	0	0	1
permanency 8	P ER -M AX -N AX N -S IH	P ER -M AX -N AX N -S IH	P ER -M AX -N AX N -S IH	0	0	0	0	0	0	0	0	1	0	1
perpetrate 3	P ER -P IH -T R EY T	P ER -P AX -T R EY T	P ER -P AX -T R EY T	0	0	0	0	0	EY	0	0	0	0	1
piling 3	"P AY -L IH NG	"P AY -L IH NG	"P AY -L IH NG	0	NG	0	0	0	0	0	0	0	0	1
pill 1	P IH L	P AX L	P AX L	0	0	0	0	0	0	0	0	0	0	1
pingpong 1	P IH NG -P OH NG	P IH NG -P OH NG	P IH NG -P OH NG	0	NG	0	0	0	0	0	0	0	0	1
pirouetting 1	"P IH -R UH -'EH -T IH NG	"P IH -R UH -'EH -T IH NG	"P IH -R UH -'EH -T IH NG	0	0	0	0	0	0	1	0	0	0	1
pirouetting 6	"P IH -R UH -'EH -T IH NG	"P AX -R UH -'EH -T IH NG	"P IH -R UH -'EH -T IH NG	0	NG	0	0	0	0	0	0	0	0	1
pitting 3	P IH -T IH NG	P AX -T IH NG	P AX -T IH NG	0	NG	0	0	0	0	0	0	0	0	1
planning 4	P L AE -N IH NG	P L AE -N IH NG	P L AE -N IH NG	0	NG	0	0	0	0	0	0	0	0	1
plaything 4	P L EY -TH IH NG	P L EY -TH IH NG	P L EY -TH IH NG	0	NG	0	0	0	0	0	0	0	0	1
pleasantness 8	P L EH -Z AX N T -N IH S	P L EH -Z AX N T -N AX S	P L EH -Z AX N T -N AX S	0	0	0	0	0	0	0	0	0	0	1
plucky 4	P L AH -K IH	P L AH -K IH	P L AH -K IH	0	K	0	0	0	0	0	0	1	0	1
politeness 6	P AX -'L AY T -N IH S	P AX -'L AY T -N AX S	P AX -'L AY T -N AX S	0	0	0	0	0	0	0	0	0	0	1
polyps 3	P OH -L IH P S	P OH -L AX P S	P OH -L AX P S	0	0	0	0	0	0	0	0	0	0	1
postulated 9	P OH S -T Y UH -L EY -T IH D	P OH S -T Y UH -L EY -T AX D	P OH S -T Y UH -L EY -T AX D	0	0	0	0	0	0	0	0	0	0	1
potholing 6	P OH T -"H OW -L IH NG	P OH T -"H OW -L IH NG	P OH T -"H OW -L IH NG	0	NG	0	0	0	0	0	0	0	0	1
preconcerted 9	"P R IY -K AX N -'S ER -T IH D	"P R IY -K AX N -'S ER -T AX D	"P R IY -K AX N -'S ER -T AX D	0	0	0	0	0	0	0	0	0	0	1
prefectural 2	P R IH -'F EH K -CH AX -R AX L	P R IH -'F EH K -CH AX -R AX L	P R IH -'F EH K -CH AX -R AX L	0	0	0	0	0	0	1	0	0	0	1
presuppositions 8	"P R IY -S AH -P AX -'Z IH -SH AX N Z	"P R IY -S AH -P AX -'Z IH -SH AX N Z	"P R IY -S AH -P AX -'Z IH -SH AX N Z	0	0	SH	0	0	0	0	0	0	0	1
primacy 6	P R AY -M AX -S IH	P R AY -M AX -S IH	P R AY -M AX -S IH	0	0	0	0	0	0	0	0	1	0	1
primary 6	"P R AY -M AX -R IH	"P R AY -M AX -R IH	"P R AY -M AX -R IH	0	0	0	0	0	0	0	0	1	0	1
prognosticating 12	P R OH G -'N OH -S T IH -K EY -T IH NG	P R OH G -'N OH -S T IH -K EY -T IH NG	P R OH G -'N OH -S T IH -K EY -T IH NG	0	NG	0	0	0	0	0	0	0	0	1
prognosticating 8	P R OH G -'N OH -S T IH -K EY -T IH NG	P R OH G -'N OH -S T IH -K EY -T IH NG	P R OH G -'N OH -S T IH -K EY -T IH NG	0	K	0	0	0	0	0	0	0	0	1

Table D.8: Full adaptation rule system results (part 4)

word	pronun_base	pronun_predict	pronun_corr	strt eb,m,n,x	rule 2	rule 3	frst in wrd	end eys,ies,ied,ey	vwl hrmny ax	vwl hrmny ih	adj vwl hrmny ih	fnl in wrd	phnm src	corr
prohibited 4	P R AX -'H IH -B IH -T IH D	P R AX -'H IH -B IH -T IH D	P R AX -'H IH -B AX -T AX D	0	H	0	0	0	0	0	0	0	0	1
prohibited 6	P R AX -'H IH -B IH -T IH D	P R AX -'H IH -B AX -T AX D	P R AX -'H IH -B AX -T AX D	0	0	0	0	0	AX	0	0	0	0	1
prohibited 8	P R AX -'H IH -B IH -T IH D	P R AX -'H IH -B AX -T AX D	P R AX -'H IH -B AX -T AX D	0	0	0	0	0	0	0	0	0	0	1
protruded 7	P R AX -'T R UH -D IH D	P R AX -'T R UH -D AX D	P R AX -'T R UH -D AX D	0	0	0	0	0	0	0	0	0	0	1
public 4	"P AH -B L IH K	"P AH -B L IH K	"P AH -B L IH K	0	K	0	0	0	0	0	0	0	0	1
purism 4	P Y UA -R IH -Z AX M	P Y UA -R AX -Z AX M	P Y UA -R AX -Z AX M	0	0	0	0	0	AX	0	0	0	0	1
races 3	R EY -S IH Z	R EY -S AX Z	R EY -S AX Z	0	0	0	0	0	0	0	0	0	0	1
recited 1	R IH -'S AY -T IH D	R IH -'S AY -T AX D	R IH -'S AY -T AX D	0	0	0	0	0	0	1	0	0	0	1
recited 5	R IH -'S AY -T IH D	R AX -'S AY -T AX D	R IH -'S AY -T AX D	0	0	0	0	0	0	0	0	0	0	1
recommended 8	"R EH -K AX -'M EH N -D IH D	"R EH -K AX -'M EH N -D AX D	"R EH -K AX -'M EH N -D AX D	0	0	0	0	0	0	0	0	0	0	1
redundancies 1	R IH -'D AH N -D AX N -S IH Z	R IH -'D AH N -D AX N -S IH Z	R IH -'D AH N -D AX N -S IH Z	0	0	0	0	0	0	1	0	0	0	1
redundancies 9	R IH -'D AH N -D AX N -S IH Z	R AX -'D AH N -D AX N -S IH Z	R IH -'D AH N -D AX N -S IH Z	0	0	0	0	1	0	0	0	0	0	1
rejoices 1	R IH -'JH OY -S IH Z	R IH -'JH OY -S IH Z	R IH -'JH OY -S AX Z	0	0	JH	0	0	0	0	0	0	0	1
rejoices 5	R IH -'JH OY -S IH Z	R IH -'JH OY -S AX Z	R IH -'JH OY -S AX Z	0	0	0	0	0	0	0	0	0	0	1
remarries 5	"R IY -'M AE -R IH Z	"R IY -'M AE -R IH Z	"R IY -'M AE -R IH Z	0	0	0	0	1	0	0	0	0	0	1
repatriations 6	"R IY -P AE -T R IH -'EY -SH AX N Z	"R IY -P AE -T R IH -'EY -SH AX N Z	"R IY -P AE -T R IH -'EY -SH AX N Z	0	0	0	0	0	0	0	EY	0	0	1
replaying 5	"R IY -'P L EY -IH NG	"R IY -'P L EY -IH NG	"R IY -'P L EY -IH NG	0	NG	0	0	0	0	0	0	0	0	1
reproached 1	R IH -'P R OW CH T	R IH -'P R OW CH T	R IH -'P R OW CH T	0	0	0	0	0	0	1	0	0	0	1
repudiation 1	R IH -"P Y UH -D IH -'EY -SH AX N	R IH -"P Y UH -D AX -'EY -SH AX N	R IH -"P Y UH -D IH -'EY -SH AX N	0	0	0	0	0	0	1	0	0	0	1
repudiation 6	R IH -"P Y UH -D IH -'EY -SH AX N	R AX -"P Y UH -D IH -'EY -SH AX N	R IH -"P Y UH -D IH -'EY -SH AX N	0	0	0	0	0	0	0	EY	0	0	1
resolved 1	R IH -'Z OH L V D	R IH -'Z OH L V D	R IH -'Z OH L V D	0	0	0	0	0	0	1	0	0	0	1
restoratives 8	R IH -'S T OH -R AX -T IH V Z	R AX -'S T OH -R AX -T AX V Z	R AX -'S T OH -R AX -T AX V Z	0	0	0	0	0	0	0	0	0	0	1
rib 1	R IH B	R AX B	R AX B	0	0	0	0	0	0	0	0	0	0	1
rotted 3	"R OH -T IH D	"R OH -T AX D	"R OH -T AX D	0	0	0	0	0	0	0	0	0	0	1
sacristan 4	S AE -K R IH -S T AX N	S AE -K R AX -S T AX N	S AE -K R AX -S T AX N	0	0	0	0	0	AX	0	0	0	0	1
saleslady 7	S EY L Z -"L EY -D IH	S EY L Z -"L EY -D IH	S EY L Z -"L EY -D IH	0	0	0	0	0	0	0	0	1	0	1
scariest 4	S K EA -R IH -IH S T	S K EA -R IH -AX S T	S K EA -R IH -AX S T	0	0	0	0	0	0	1	0	0	0	1
scariest 5	S K EA -R IH -IH S T	S K EA -R AX -AX S T	S K EA -R IH -AX S T	0	0	0	0	0	0	0	0	0	0	1
seducing 6	S IH -'D Y UH -S IH NG	S AX -'D Y UH -S IH NG	S AX -'D Y UH -S IH NG	0	NG	0	0	0	0	0	0	0	0	1
sedulously 9	S EH -D Y UH -L AX S -L IH	S EH -D Y UH -L AX S -L IH	S EH -D Y UH -L AX S -L IH	0	0	0	0	0	0	0	0	1	0	1
sensuously 8	S EH N -S Y UA S -L IH	S EH N -S Y UA S -L IH	S EH N -S Y UA S -L IH	0	0	0	0	0	0	0	0	1	0	1
servility 3	S ER -'V IH -L AX -T IH	S ER -'V AX -L AX -T IH	S ER -'V AX -L AX -T IH	0	0	0	0	0	AX	0	0	0	0	1
servility 7	S ER -'V IH -L AX -T IH	S ER -'V AX -L AX -T IH	S ER -'V AX -L AX -T IH	0	0	0	0	0	0	0	0	1	0	1
shingled 1	SH IH NG -G L D	SH IH NG -G L D	SH IH NG -G L D	0	NG	0	0	0	0	0	0	0	0	1
sits 1	"S IH T S	"S AX T S	"S AX T S	0	0	0	0	0	0	0	0	0	0	1
sphincter 2	S F IH NG K -T AX R	S F IH NG K -T AX R	S F IH NG K -T AX R	0	NG	0	0	0	0	0	0	0	0	1
sporty 4	S P AO -T IH	S P AO -T IH	S P AO -T IH	0	0	0	0	0	0	0	0	1	0	1
squawking 5	S K W AO -K IH NG	S K W AO -K IH NG	S K W AO -K IH NG	0	NG	0	0	0	0	0	0	0	0	1
squeegeeing 6	"S K W IY -'JH IY -IH NG	"S K W IY -'JH IY -IH NG	"S K W IY -'JH IY -IH NG	0	NG	0	0	0	0	0	0	0	0	1
squirrels 3	S K W IH -R AX L Z	S K W AX -R AX L Z	S K W AX -R AX L Z	0	0	0	0	0	AX	0	0	0	0	1
stereo 4	S T EH -R IH -OW	S T EH -R IH -OW	S T EH -R IH -OW	0	0	0	0	0	0	1	0	0	0	1
superfluities 9	"S UH -P AX -'F L UH -AX -T IH Z	"S UH -P AX -'F L UH -AX -T IH Z	"S UH -P AX -'F L UH -AX -T IH Z	0	0	0	0	1	0	0	0	0	0	1
surefootedly 5	"SH AO -'F UH -T IH D -L IH	"SH AO -'F UH -T AX D -L IH	"SH AO -'F UH -T AX D -L IH	0	0	0	0	0	0	0	0	0	0	1
surefootedly 8	"SH AO -'F UH -T IH D -L IH	"SH AO -'F UH -T AX D -L IH	"SH AO -'F UH -T AX D -L IH	0	0	0	0	0	0	0	0	1	0	1
swiftest 2	S W IH F -T IH S T	S W AX F -T IH S T	S W AX F -T AX S T	0	0	0	0	0	0	0	0	0	0	1
swiftest 5	S W IH F -T IH S T	S W AX F -T AX S T	S W AX F -T AX S T	0	0	0	0	0	0	0	0	0	0	1
tailless 4	T EY L -L IH S	T EY L -L AX S	T EY L -L AX S	0	0	0	0	0	0	0	0	0	0	1
taking 3	T EY -K IH NG	T EY -K IH NG	T EY -K IH NG	0	NG	0	0	0	0	0	0	0	0	1
tarbooshes 5	T AA -'B UH -SH IH Z	T AA -'B UH -SH AX Z	T AA -'B UH -SH AX Z	0	0	0	0	0	0	0	0	0	0	1
tennis 3	T EH -N IH S	T EH -N AX S	T EH -N AX S	0	0	0	0	0	0	0	0	0	0	1
thalidomide 3	TH AX -'L IH -D AX -M AY D	TH AX -'L AX -D AX -M AY D	TH AX -'L AX -D AX -M AY D	0	0	0	0	0	AX	0	0	0	0	1
theory 3	"TH IA -R IH	"TH IA -R IH	"TH IA -R IH	0	0	0	0	0	0	0	0	1	0	1
tormenting 6	T AO -'M EH N -T IH NG	T AO -'M EH N -T IH NG	T AO -'M EH N -T IH NG	0	NG	0	0	0	0	0	0	0	0	1
trickle 2	T R IH -K L	T R IH -K L	T R IH -K L	0	K	0	0	0	0	0	0	0	0	1
trochaics 5	T R OW -'K EY -IH K S	T R OW -'K EY -IH K S	T R OW -'K EY -IH K S	0	K	0	0	0	0	0	0	0	0	1
trolleybus 4	T R OH -L IH -B AH S	T R OH -L IH -B AH S	T R OH -L IH -B AH S	0	0	0	0	0	0	1	0	0	0	1
twit 2	T W IH T	T W AX T	T W AX T	0	0	0	0	0	0	0	0	0	0	1
unhinging 3	"AH N -'H IH N -JH IH NG	"AH N -'H IH N -JH IH NG	"AH N -'H IH N -JH IH NG	0	H	0	0	0	0	0	0	0	0	1
unhinging 6	"AH N -'H IH N -JH IH NG	"AH N -'H IH N -JH IH NG	"AH N -'H IH N -JH IH NG	0	NG	0	0	0	0	0	0	0	0	1
unseated 5	"AH N -'S IY -T IH D	"AH N -'S IY -T AX D	"AH N -'S IY -T AX D	0	0	0	0	0	0	0	0	0	0	1
untinged 3	"AH N -'T IH N JH D	"AH N -'T IH N JH D	"AH N -'T IH N JH D	0	0	JH	0	0	0	0	0	0	0	1
visible 1	V IH -Z AX -B L	V AX -Z AX -B L	V AX -Z AX -B L	0	0	0	0	0	AX	0	0	0	0	1
warning 3	"W AO -N IH NG	"W AO -N IH NG	"W AO -N IH NG	0	NG	0	0	0	0	0	0	0	0	1
weeny 3	W IY -N IH	W IY -N IH	W IY -N IH	0	0	0	0	0	0	0	0	1	0	1
wordy 3	W ER -D IH	W ER -D IH	W ER -D IH	0	0	0	0	0	0	0	0	1	0	1

Table D.9: Results of full rule set applied to validation set (part 1)

word	pronun_base	pronun_predict	pronun_corr	strt eb,m,n,x	rule 2	rule 3	first	end eys,ies,ied	vwl hrmny ax	vwl hrmny ih	adj vwl hrmny ih	last	phnm src	corr
aggregation 3	"AE -G R IH -'G EY -SH AX N	"AE -G R AX -'G EY -SH AX N	"AE -G R IH -'G EY -SH AX N	0	G	0	0	0	0	0	0	0	1	0
appetizing 2	AE -P IH -T AY -Z IH NG	AE -P IH -T AY -Z IH NG	AE -P AX -T AY -Z IH NG	0	0	0	0	0	0	1	0	0	0	0
behead 1	B IH -'H EH D	B AX -'H EH D	B IH -'H EH D	0	H	0	0	0	0	0	0	0	1	0
bitchiness 3	B IH -CH IH -N IH S	B IH -CH AX -N AX S	B IH -CH IH -N AX S	0	0	0	0	0	AX	0	0	0	0	0
disinterment 3	"D IH S -IH N -'T ER -M AX N T	"D AX S -AX N -'T ER -M AX N T	"D AX S -IH N -'T ER -M AX N T	0	0	0	0	0	0	0	0	0	0	0
distinction 1	D IH -'S T IH NG K -SH AX N	D IH -'S T IH NG K -SH AX N	D AX -'S T IH NG K -SH AX N	0	0	0	0	0	0	1	0	0	0	0
enamouring 0	IH -'N AE -M AX -R IH NG	IH -'N AE -M AX -R IH NG	EH -'N AE -M AX -R IH NG	0	0	0	1	0	0	0	0	0	0	0
epitome 0	IH -'P IH -T AX -M IH	IH -'P IH -T AX -M IH	EH -'P AX -T AX -M IH	0	0	0	1	0	0	0	0	0	0	0
harbinger 3	H AA -B IH N -JH AX R	H AA -B AX N -JH AX R	H AA -B IH N -JH AX R	0	0	0	0	0	0	0	0	0	0	0
huskiest 5	H AH S -K IH -IH S T	H AH S -K IH -AX S T	H AH S -K IH -EH S T	0	0	0	0	0	0	0	0	0	0	0
irrecoverable 2	"IH -R IH -'K AH -V AX -R AX -B L	"IH -R AX -'K AH -V AX -R AX -B L	"IH -R IH -'K AH -V AX -R AX -B L	0	K	0	0	0	0	0	0	0	1	0
Israelis 5	IH Z -'R EY -L IH Z	IH Z -'R EY -L AX Z	IH Z -'R EY -L IH Z	0	0	0	0	0	0	0	0	0	0	0
Macedonians 3	"M AE -S IH -'D OW -N Y AX N Z	"M AE -S IH -'D OW -N Y AX N Z	"M AE -S AX -'D OW -N Y AX N Z	0	0	0	0	0	0	1	0	0	0	0
pinching 1	P IH N -CH IH NG	P AX N -CH IH NG	P IH N -CH IH NG	0	0	0	0	0	0	0	0	0	0	0
remaining 1	R IH -'M EY -N IH NG	R AX -'M EY -N IH NG	R IH -'M EY -N IH NG	0	0	0	0	0	EY	0	0	0	0	0
reptilians 4	R EH P -'T IH -L IA N Z	R EH P -'T IH -L IA N Z	R EH P -'T AX -L IA N Z	0	0	0	0	0	0	1	0	0	0	0
sprightliness 6	S P R AY T -L IH -N IH S	S P R AY T -L AX -N AX S	S P R AY T -L IH -N AX S	0	0	0	0	0	AX	0	0	0	0	0
wifeliest 5	W AY F -L IH -IH S T	W AY F -L AX -AX S T	W AY F -L IH -EH S T	0	0	0	0	0	0	0	0	0	0	0
winds 1	W IH N D Z	W AX N D Z	W IH N D Z	0	0	0	0	0	0	0	0	0	0	0
acknowledge 5	AX K -'N OH -L IH JH	AX K -'N OH -L IH JH	AX K -'N OH -L IH JH	0	0	JH	0	0	0	0	0	0	0	1
agnostic 6	AE G -'N OH -S T IH K	AE G -'N OH -S T IH K	AE G -'N OH -S T IH K	0	K	0	0	0	0	0	0	0	0	1
allowances 6	AX -'L AW -AX N -S IH Z	AX -'L AW -AX N -S AX Z	AX -'L AW -AX N -S AX Z	0	0	0	0	0	0	0	0	0	0	1
antiquities 3	AE N -'T IH -K W AX -T IH Z	AE N -'T IH -K W AX -T IH Z	AE N -'T IH -K W AX -T IH Z	0	K	0	0	0	0	0	0	0	0	1
antiquities 8	AE N -'T IH -K W AX -T IH Z	AE N -'T IH -K W AX -T IH Z	AE N -'T IH -K W AX -T IH Z	0	0	0	0	1	0	0	0	0	0	1
anybody 2	EH -N IH -"B OH -D IH	EH -N IH -"B OH -D IH	EH -N IH -"B OH -D IH	0	0	0	0	0	0	1	0	0	0	1
anybody 6	EH -N IH -"B OH -D IH	EH -N AX -"B OH -D IH	EH -N IH -"B OH -D IH	0	0	0	0	0	0	0	0	1	0	1
aphoristic 4	"AE -F AX -'R IH S -T IH K	"AE -F AX -'R AX S -T IH K	"AE -F AX -'R AX S -T IH K	0	0	0	0	0	0	0	0	0	0	1
aphoristic 7	"AE -F AX -'R IH S -T IH K	"AE -F AX -'R AX S -T IH K	"AE -F AX -'R AX S -T IH K	0	K	0	0	0	0	0	0	0	0	1
appetizing 6	AE -P IH -T AY -Z IH NG	AE -P AX -T AY -Z IH NG	AE -P AX -T AY -Z IH NG	0	NG	0	0	0	0	0	0	0	0	1
Arabist 4	AE -R AX -B IH S T	AE -R AX -B AX S T	AE -R AX -B AX S T	0	0	0	0	0	0	0	0	0	0	1
automatisms 6	AO -'T OH -M AX -T IH -Z AX M Z	AO -'T OH -M AX -T AX -Z AX M Z	AO -'T OH -M AX -T AX -Z AX M Z	0	0	0	0	0	AX	0	0	0	0	1
beefing 3	"B IY -F IH NG	"B IY -F IH NG	"B IY -F IH NG	0	NG	0	0	0	0	0	0	0	0	1
belongings 1	B IH -'L OH -NG IH NG Z	B IH -'L OH -NG IH NG Z	B IH -'L OH -NG IH NG Z	0	0	0	0	0	0	1	0	0	0	1
belongings 5	B IH -'L OH -NG IH NG Z	B AX -'L OH -NG IH NG Z	B IH -'L OH -NG IH NG Z	0	NG	0	0	0	0	0	0	0	0	1
bevy 3	B EH -V IH	B EH -V IH	B EH -V IH	0	0	0	0	0	0	0	0	1	0	1
bickers 1	B IH -K AX Z	B IH -K AX Z	B IH -K AX Z	0	K	0	0	0	0	0	0	0	0	1
billfolds 1	B IH L -F OW L D Z	B AX L -F OW L D Z	B AX L -F OW L D Z	0	0	0	0	0	0	0	0	0	0	1
bitchiness 1	B IH -CH IH -N IH S	B IH -CH IH -N IH S	B IH -CH IH -N AX S	0	0	CH	0	0	0	0	0	0	0	1
bitchiness 5	B IH -CH IH -N IH S	B IH -CH AX -N AX S	B IH -CH IH -N AX S	0	0	0	0	0	0	0	0	0	0	1
bleeding 4	"B L IY -D IH NG	"B L IY -D IH NG	"B L IY -D IH NG	0	NG	0	0	0	0	0	0	0	0	1
blinking 2	"B L IH NG -K IH NG	"B L IH NG -K IH NG	"B L IH NG -K IH NG	0	NG	0	0	0	0	0	0	0	0	1
blinking 5	"B L IH NG -K IH NG	"B L IH NG -K IH NG	"B L IH NG -K IH NG	0	NG	0	0	0	0	0	0	0	0	1
blowing 3	"B L OW -IH NG	"B L OW -IH NG	"B L OW -IH NG	0	NG	0	0	0	0	0	0	0	0	1
bogey 3	B OW -G IH	B OW -G IH	B OW -G IH	0	G	0	0	0	0	0	0	1	0	1
boundless 5	B AW N D -L IH S	B AW N D -L AX S	B AW N D -L AX S	0	0	0	0	0	0	0	0	0	0	1
brazening 6	"B R EY -Z AX N -IH NG	"B R EY -Z AX N -IH NG	"B R EY -Z AX N -IH NG	0	NG	0	0	0	0	0	0	0	0	1
bruises 4	B R UH -Z IH Z	B R UH -Z AX Z	B R UH -Z AX Z	0	0	0	0	0	0	0	0	0	0	1
brushing 4	"B R AH -SH IH NG	"B R AH -SH IH NG	"B R AH -SH IH NG	0	NG	0	0	0	0	0	0	0	0	1
bucketing 3	B AH -K IH -T IH NG	B AH -K AX -T IH NG	B AH -K AX -T IH NG	0	K	0	0	0	0	0	0	0	1	1
bucketing 5	B AH -K IH -T IH NG	B AH -K IH -T IH NG	B AH -K AX -T IH NG	0	NG	0	0	0	0	0	0	0	0	1
buoyantly 6	B OY -AX N T -L IH	B OY -AX N T -L IH	B OY -AX N T -L IH	0	0	0	0	0	0	0	0	1	0	1
calvaries 6	K AE L -V AX -R IH Z	K AE L -V AX -R IH Z	K AE L -V AX -R IH Z	0	0	0	0	1	0	0	0	0	0	1
caricaturist 10	K AE -R IH -K AX -"T Y UA -R IH S T	K AE -R IH -K AX -"T Y UA -R AX S T	K AE -R IH -K AX -"T Y UA -R AX S T	0	0	0	0	0	0	0	0	0	0	1
caricaturist 3	K AE -R IH -K AX -"T Y UA -R IH S T	K AE -R IH -K AX -"T Y UA -R IH S T	K AE -R IH -K AX -"T Y UA -R AX S T	0	K	0	0	0	0	0	0	0	0	1
carrying 3	K AE -R IH -IH NG	K AE -R IH -IH NG	K AE -R IH -IH NG	0	0	0	0	0	0	1	0	0	0	1
carrying 4	K AE -R IH -IH NG	K AE -R AX -IH NG	K AE -R IH -IH NG	0	NG	0	0	0	0	0	0	0	0	1
cashes 3	K AE -SH IH Z	K AE -SH AX Z	K AE -SH AX Z	0	0	0	0	0	0	0	0	0	0	1
cicala 1	S IH -'K AA -L AX	S IH -'K AA -L AX	S IH -'K AA -L AX	0	K	0	0	0	0	0	0	0	0	1
comforting 6	K AH M -F AX -T IH NG	K AH M -F AX -T IH NG	K AH M -F AX -T IH NG	0	NG	0	0	0	0	0	0	0	0	1
commensurably 11	K AX -'M EH N -SH AX -R AX -B L IH	K AX -'M EH N -SH AX -R AX -B L IH	K AX -'M EH N -SH AX -R AX -B L IH	0	0	0	0	0	0	0	0	1	0	1
commodiously 9	K AX -'M OW -D Y AX S -L IH	K AX -'M OW -D Y AX S -L IH	K AX -'M OW -D Y AX S -L IH	0	0	0	0	0	0	0	0	1	0	1
compressing 7	K AX M -'P R EH -S IH NG	K AX M -'P R EH -S IH NG	K AX M -'P R EH -S IH NG	0	NG	0	0	0	0	0	0	0	0	1
containerizing 10	K AX N -'T EY -N AX -R AY -Z IH NG	K AX N -'T EY -N AX -R AY -Z IH NG	K AX N -'T EY -N AX -R AY -Z IH NG	0	NG	0	0	0	0	0	0	0	0	1
corpses 4	K AO P -S IH Z	K AO P -S AX Z	K AO P -S AX Z	0	0	0	0	0	0	0	0	0	0	1

Table D.10: Results table of full rule set applied to validation set (part 2)

word	pronun_base	pronun_predict	pronun_corr	strt eb,m,n,x	rule 2	rule 3	first	end eys,ies,ied	vwl hrmny ax	vwl hrmny ih	adj vwl hrmny ih	last	phnm src	corr
corroded 5	K AX -'R OW -D IH D	K AX -'R OW -D AX D	K AX -'R OW -D AX D	0	0	0	0	0	0	0	0	0	0	1
crossing 4	K R OH -S IH NG	K R OH -S IH NG	K R OH -S IH NG	0	NG	0	0	0	0	0	0	0	0	1
cutlasses 6	K AH T -L AX -S IH Z	K AH T -L AX -S AX Z	K AH T -L AX -S AX Z	0	0	0	0	0	0	0	0	0	0	1
deceit 1	D IH -'S IY T	D IH -'S IY T	D IH -'S IY T	0	0	0	0	0	0	1	0	0	0	1
dedicate 3	D EH -D IH -K EY T	D EH -D IH -K EY T	D EH -D IH -K EY T	0	K	0	0	0	0	0	0	0	0	1
defalcated 8	D IY -F AE L -K EY -T IH D	D IY -F AE L -K EY -T AX D	D IY -F AE L -K EY -T AX D	0	0	0	0	0	0	0	0	0	0	1
deforested 5	"D IY -'F OH -R IH -S T IH D	"D IY -'F OH -R AX -S T AX D	"D IY -'F OH -R AX -S T AX D	0	0	0	0	0	AX	0	0	0	0	1
deforested 8	"D IY -'F OH -R IH -S T IH D	"D IY -'F OH -R AX -S T AX D	"D IY -'F OH -R AX -S T AX D	0	0	0	0	0	0	0	0	0	0	1
demerits 5	D IY -'M EH -R IH T S	D IY -'M EH -R AX T S	D IY -'M EH -R AX T S	0	0	0	0	0	0	0	0	0	0	1
demimondaine 3	"D EH -M IH -M OH N -'D EY N	"D EH -M IH -M OH N -'D EY N	"D EH -M IH -M OH N -'D EY N	0	0	0	0	0	0	1	0	0	0	1
deodars 1	D IH -OW -D AA Z	D IH -OW -D AA Z	D IH -OW -D AA Z	0	0	0	0	0	0	1	0	0	0	1
detestable 1	D IH -'T EH -S T AX -B L	D IH -'T EH -S T AX -B L	D IH -'T EH -S T AX -B L	0	0	0	0	0	0	1	0	0	0	1
detrimentally 11	"D EH -T R IH -'M EH N -T AX -L IH	"D EH -T R AX -'M EH N -T AX -L IH	"D EH -T R IH -'M EH N -T AX -L IH	0	0	0	0	0	0	0	0	1	0	1
detrimentally 4	"D EH -T R IH -'M EH N -T AX -L IH	"D EH -T R IH -'M EH N -T AX -L IH	"D EH -T R IH -'M EH N -T AX -L IH	0	0	0	0	0	0	1	0	0	0	1
did 1	"D IH D	"D AX D	"D AX D	0	0	0	0	0	0	0	0	0	0	1
dining 3	D AY -N IH NG	D AY -N IH NG	D AY -N IH NG	0	NG	0	0	0	0	0	0	0	0	1
disappearance 1	"D IH S -AX -'P IA -R AX N S	"D AX S -AX -'P IA -R AX N S	"D AX S -AX -'P IA -R AX N S	0	0	0	0	0	0	0	0	0	0	1
disinterment 1	"D IH S -IH N -'T ER -M AX N T	"D AX S -IH N -'T ER -M AX N T	"D AX S -IH N -'T ER -M AX N T	0	0	0	0	0	0	0	0	0	0	1
dismisses 1	D IH S -'M IH -S IH Z	D AX S -'M IH -S IH Z	D AX S -'M AX -S AX Z	0	0	0	0	0	0	0	0	0	0	1
dismisses 4	D IH S -'M IH -S IH Z	D AX S -'M AX -S AX Z	D AX S -'M AX -S AX Z	0	0	0	0	0	AX	0	0	0	0	1
dismisses 6	D IH S -'M IH -S IH Z	D AX S -'M AX -S AX Z	D AX S -'M AX -S AX Z	0	0	0	0	0	0	0	0	0	0	1
distinction 4	D IH -'S T IH NG K -SH AX N	D AX -'S T IH NG K -SH AX N	D AX -'S T IH NG K -SH AX N	0	NG	0	0	0	0	0	0	0	0	1
dogmatics 6	D OH G -'M AE -T IH K S	D OH G -'M AE -T IH K S	D OH G -'M AE -T IH K S	0	K	0	0	0	0	0	0	0	0	1
dowering 4	D AW -AX -R IH NG	D AW -AX -R IH NG	D AW -AX -R IH NG	0	NG	0	0	0	0	0	0	0	0	1
economical 6	"IY -K AX -'N OH -M IH -K L	"IY -K AX -'N OH -M IH -K L	"IY -K AX -'N OH -M IH -K L	0	K	0	0	0	0	0	0	0	0	1
electronics 0	"IH -L EH K -'T R OH -N IH K S	"IH -L EH K -'T R OH -N IH K S	"IH -L EH K -'T R OH -N IH K S	0	0	0	1	0	0	0	0	0	0	1
electronics 8	"IH -L EH K -'T R OH -N IH K S	"IH -L EH K -'T R OH -N IH K S	"IH -L EH K -'T R OH -N IH K S	0	K	0	0	0	0	0	0	0	0	1
elongating 7	IY -L OH NG -G EY -T IH NG	IY -L OH NG -G EY -T IH NG	IY -L OH NG -G EY -T IH NG	0	NG	0	0	0	0	0	0	0	0	1
emotive 0	IH -'M OW -T IH V	IH -'M OW -T IH V	IH -'M OW -T AX V	0	0	0	1	0	0	0	0	0	0	1
emotive 4	IH -'M OW -T IH V	IH -'M OW -T AX V	IH -'M OW -T AX V	0	0	0	0	0	0	0	0	0	0	1
enamouring 6	IH -'N AE -M AX -R IH NG	IH -'N AE -M AX -R IH NG	EH -'N AE -M AX -R IH NG	0	NG	0	0	0	0	0	0	0	0	1
enlarged 0	IH N -'L AA JH D	EH N -'L AA JH D	EH N -'L AA JH D	1	0	0	1	0	0	0	0	0	0	1
entailed 0	IH N -'T EY L D	EH N -'T EY L D	EH N -'T EY L D	1	0	0	1	0	0	0	0	0	0	1
entomological 9	"EH N -T AX -M AX -'L OH -JH IH -K L	"EH N -T AX -M AX -'L OH -JH IH -K L	"EH N -T AX -M AX -'L OH -JH IH -K L	0	K	0	0	0	0	0	0	0	0	1
epitome 2	IH -'P IH -T AX -M IH	IH -'P AX -T AX -M IH	EH -'P AX -T AX -M IH	0	0	0	0	0	AX	0	0	0	0	1
epitome 6	IH -'P IH -T AX -M IH	IH -'P AX -T AX -M IH	EH -'P AX -T AX -M IH	0	0	0	0	0	0	0	0	1	0	1
equivalence 0	IH -'K W IH -V AX -L AX N S	IH -'K W IH -V AX -L AX N S	IH -'K W AX -V AX -L AX N S	0	K	0	1	0	0	0	0	0	0	1
equivalence 3	IH -'K W IH -V AX -L AX N S	IH -'K W AX -V AX -L AX N S	IH -'K W AX -V AX -L AX N S	0	0	0	0	0	AX	0	0	0	0	1
eruditely 7	EH -R UH -D AY T -L IH	EH -R UH -D AY T -L IH	EH -R UH -D AY T -L IH	0	0	0	0	0	0	0	0	1	0	1
essential 0	IH -'S EH N -SH L	IH -'S EH N -SH L	IH -'S EH N -SH L	0	0	0	1	0	0	0	0	0	0	1
estimation 3	"EH -S T IH -'M EY -SH AX N	"EH -S T AX -'M EY -SH AX N	"EH -S T AX -'M EY -SH AX N	0	0	0	0	0	EY	0	0	0	0	1
excusable 0	IH K -'S K Y UH -Z AX -B L	EH K -'S K Y UH -Z AX -B L	EH K -'S K Y UH -Z AX -B L	1	K	0	1	0	0	0	0	0	0	1
explorative 9	EH K S -'P L OH -R AX -T IH V	EH K S -'P L OH -R AX -T AX V	EH K S -'P L OH -R AX -T AX V	0	0	0	0	0	0	0	0	0	0	1
extinguished 0	IH K -'S T IH NG -G W IH SH T	EH K -'S T IH NG -G W IH SH T	EH K -'S T IH NG -G W IH SH T	1	K	0	1	0	0	0	0	0	0	1
extinguished 4	IH K -'S T IH NG -G W IH SH T	EH K -'S T IH NG -G W IH SH T	EH K -'S T IH NG -G W IH SH T	0	NG	0	0	0	0	0	0	0	0	1
extinguished 8	IH K -'S T IH NG -G W IH SH T	EH K -'S T IH NG -G W IH SH T	EH K -'S T IH NG -G W IH SH T	0	0	SH	0	0	0	0	0	0	0	1
family 5	"F AE -M AX -L IH	"F AE -M AX -L IH	"F AE -M AX -L IH	0	0	0	0	0	0	0	0	1	0	1
fiddler 1	F IH -D L -AX R	F AX -D L -AX R	F AX -D L -AX R	0	0	0	0	0	AX	0	0	0	0	1
figures 1	"F IH -G AX Z	"F IH -G AX Z	"F IH -G AX Z	0	G	0	0	0	0	0	0	0	0	1
filching 1	F IH L -CH IH NG	F AX L -CH IH NG	F AX L -CH IH NG	0	0	0	0	0	0	0	0	0	0	1
filching 4	F IH L -CH IH NG	F AX L -CH IH NG	F AX L -CH IH NG	0	NG	0	0	0	0	0	0	0	0	1
finishing 1	F IH -N IH -SH IH NG	F IH -N IH -SH IH NG	F IH -N IH -SH IH NG	0	0	0	0	0	0	1	0	0	0	1
finishing 3	F IH -N IH -SH IH NG	F AX -N IH -SH IH NG	F IH -N IH -SH IH NG	0	0	SH	0	0	0	0	0	0	0	1
finishing 5	F IH -N IH -SH IH NG	F AX -N IH -SH IH NG	F IH -N IH -SH IH NG	0	NG	0	0	0	0	0	0	0	0	1
fizzles 1	F IH -Z L Z	F AX -Z L Z	F AX -Z L Z	0	0	0	0	0	AX	0	0	0	0	1
flatfishes 5	F L AE T -F IH -SH IH Z	F L AE T -F IH -SH IH Z	F L AE T -F IH -SH AX Z	0	0	SH	0	0	0	0	0	0	0	1
flatfishes 7	F L AE T -F IH -SH IH Z	F L AE T -F IH -SH AX Z	F L AE T -F IH -SH AX Z	0	0	0	0	0	0	0	0	0	0	1
fosses 3	F OH -S IH Z	F OH -S AX Z	F OH -S AX Z	0	0	0	0	0	0	0	0	0	0	1
freckling 5	F R EH -K L IH NG	F R EH -K L IH NG	F R EH -K L IH NG	0	NG	0	0	0	0	0	0	0	0	1
frivolousness 2	F R IH -V AX -L AX S -N IH S	F R AX -V AX -L AX S -N AX S	F R AX -V AX -L AX S -N AX S	0	0	0	0	0	AX	0	0	0	0	1
frivolousness 9	F R IH -V AX -L AX S -N IH S	F R AX -V AX -L AX S -N AX S	F R AX -V AX -L AX S -N AX S	0	0	0	0	0	0	0	0	0	0	1
glimmered 2	G L IH -M AX D	G L AX -M AX D	G L AX -M AX D	0	0	0	0	0	AX	0	0	0	0	1
gouges 3	G AW -JH IH Z	G AW -JH AX Z	G AW -JH AX Z	0	0	0	0	0	0	0	0	0	0	1
heavies 3	H EH -V IH Z	H EH -V IH Z	H EH -V IH Z	0	0	0	0	1	0	0	0	0	0	1
Hippocratic 1	"H IH -P OW -K R AE -T IH K	"H IH -P OW -K R AE -T IH K	"H IH -P OW -K R AE -T IH K	0	0	0	0	0	0	1	0	0	0	1

Table D.11: Results table of full rule set applied to validation set (part 3)

word	pronun_base	pronun_predict	pronun_corr	strt eb,m,n,x	rule 2	rule 3	first	end eys,ies,ied	vwl hrmny ax	vwl hrmny ih	adj vwl hrmny ih	last	phnm src	corr
Hippocratic 8	”H IH -P OW -K R AE -T IH K	”H AX -P OW -K R AE -T IH K	”H IH -P OW -K R AE -T IH K	0	K	0	0	0	0	0	0	0	0	1
huskiest 4	H AH S -K IH -IH S T	H AH S -K IH -AX S T	H AH S -K IH -EH S T	0	K	0	0	0	0	0	0	0	1	1
idiomatically 0	”IH -D IA -’M AE -T IH -K AX -L IH	”IH -D IA -’M AE -T IH -K AX -L IH	”IH -D IA -’M AE -T IH -K AX -L IH	0	0	0	1	0	0	0	0	0	0	1
idiomatically 10	”IH -D IA -’M AE -T IH -K AX -L IH	”IH -D IA -’M AE -T IH -K AX -L IH	”IH -D IA -’M AE -T IH -K AX -L IH	0	0	0	0	0	0	0	0	1	0	1
idiomatically 6	”IH -D IA -’M AE -T IH -K AX -L IH	”IH -D IA -’M AE -T IH -K AX -L IH	”IH -D IA -’M AE -T IH -K AX -L IH	0	K	0	0	0	0	0	0	0	0	1
igloos 0	IH -G L UH Z	IH -G L UH Z	IH -G L UH Z	0	G	0	1	0	0	0	0	0	0	1
imitations 0	”IH -M IH -’T EY -SH AX N Z	”IH -M IH -’T EY -SH AX N Z	”IH -M AX -’T EY -SH AX N Z	0	0	0	1	0	0	0	0	0	0	1
imitations 2	”IH -M IH -’T EY -SH AX N Z	”IH -M AX -’T EY -SH AX N Z	”IH -M AX -’T EY -SH AX N Z	0	0	0	0	0	EY	0	0	0	0	1
impinge 0	IH M -’P IH N JH	IH M -’P IH N JH	IH M -’P IH N JH	0	0	0	1	0	0	0	0	0	0	1
impinge 3	IH M -’P IH N JH	IH M -’P IH N JH	IH M -’P IH N JH	0	0	JH	0	0	0	0	0	0	0	1
implicitly 0	IH M -’P L IH -S IH T -L IH	IH M -’P L IH -S IH T -L IH	IH M -’P L AX -S AX T -L IH	0	0	0	1	0	0	0	0	0	0	1
implicitly 4	IH M -’P L IH -S IH T -L IH	IH M -’P L AX -S AX T -L IH	IH M -’P L AX -S AX T -L IH	0	0	0	0	0	AX	0	0	0	0	1
implicitly 6	IH M -’P L IH -S IH T -L IH	IH M -’P L AX -S AX T -L IH	IH M -’P L AX -S AX T -L IH	0	0	0	0	0	0	0	0	0	0	1
implicitly 9	IH M -’P L IH -S IH T -L IH	IH M -’P L AX -S AX T -L IH	IH M -’P L AX -S AX T -L IH	0	0	0	0	0	0	0	0	1	0	1
in 0	IH N	IH N	IH N	0	0	0	1	0	0	0	0	0	0	1
Indian 0	IH N -D Y AX N	IH N -D Y AX N	IH N -D Y AX N	0	0	0	1	0	0	0	0	0	0	1
ingeniously 0	IH N -’JH IY -N Y AX S -L IH	IH N -’JH IY -N Y AX S -L IH	IH N -’JH IY -N Y AX S -L IH	0	0	0	1	0	0	0	0	0	0	1
ingeniously 9	IH N -’JH IY -N Y AX S -L IH	IH N -’JH IY -N Y AX S -L IH	IH N -’JH IY -N Y AX S -L IH	0	0	0	0	0	0	0	0	1	0	1
interfered 0	”IH N -T AX -’F IA D	”IH N -T AX -’F IA D	”IH N -T AX -’F IA D	0	0	0	1	0	0	0	0	0	0	1
into 0	”IH N -T UH	”IH N -T UH	”IH N -T UH	0	0	0	1	0	0	0	0	0	0	1
intruders 0	IH N -’T R UH -D AX Z	IH N -’T R UH -D AX Z	IH N -’T R UH -D AX Z	0	0	0	1	0	0	0	0	0	0	1
irrecoverable 0	”IH -R IH -’K AH -V AX -R AX -B L	”IH -R IH -’K AH -V AX -R AX -B L	”IH -R IH -’K AH -V AX -R AX -B L	0	0	0	1	0	0	0	0	0	0	1
irregular 0	IH -’R EH -G Y UH -L AX R	IH -’R EH -G Y UH -L AX R	IH -’R EH -G Y UH -L AX R	0	0	0	1	0	0	0	0	0	0	1
Israelis 0	IH Z -’R EY -L IH Z	IH Z -’R EY -L IH Z	IH Z -’R EY -L IH Z	0	0	0	1	0	0	0	0	0	0	1
it 0	IH T	IH T	IH T	0	0	0	1	0	0	0	0	0	0	1
kibbutznik 1	K IH -’B UH T S -N IH K	K IH -’B UH T S -N IH K	K IH -’B UH T S -N IH K	0	0	0	0	0	0	1	0	0	0	1
kibbutznik 7	K IH -’B UH T S -N IH K	K AX -’B UH T S -N IH K	K IH -’B UH T S -N IH K	0	K	0	0	0	0	0	0	0	0	1
kyrie 1	”K IH -R IH	”K IH -R IH	”K IH -R IH	0	0	0	0	0	0	1	0	0	0	1
kyrie 3	”K IH -R IH	”K AX -R IH	”K IH -R IH	0	0	0	0	0	0	0	0	1	0	1
lapis 3	”L AE -P IH S	”L AE -P AX S	”L AE -P AX S	0	0	0	0	0	0	0	0	0	0	1
lingua 1	”L IH NG -G W AX	”L IH NG -G W AX	”L IH NG -G W AX	0	NG	0	0	0	0	0	0	0	0	1
lisped 1	L IH S P T	L AX S P T	L AX S P T	0	0	0	0	0	0	0	0	0	0	1
lives 1	”L IH V Z	”L AX V Z	”L AX V Z	0	0	0	0	0	0	0	0	0	0	1
lovey 3	L AH -V IH	L AH -V IH	L AH -V IH	0	0	0	0	0	0	0	0	1	0	1
matinees 3	M AE -T IH -N EY Z	M AE -T AX -N EY Z	M AE -T AX -N EY Z	0	0	0	0	0	EY	0	0	0	0	1
meetinghouse 3	M IY -T IH NG -H AW S	M IY -T IH NG -H AW S	M IY -T IH NG -H AW S	0	NG	0	0	0	0	0	0	0	0	1
microscopically 13	”M AY -K R AX -’S K OH -P IH -K AX -L IH	”M AY -K R AX -’S K OH -P IH -K AX -L IH	”M AY -K R AX -’S K OH -P IH -K AX -L IH	0	0	0	0	0	0	0	0	1	0	1
microscopically 9	”M AY -K R AX -’S K OH -P IH -K AX -L IH	”M AY -K R AX -’S K OH -P IH -K AX -L IH	”M AY -K R AX -’S K OH -P IH -K AX -L IH	0	K	0	0	0	0	0	0	0	0	1
misfiring 1	”M IH S -’F AY -AX -R IH NG	”M AX S -’F AY -AX -R IH NG	”M AX S -’F AY -AX -R IH NG	0	0	0	0	0	0	0	0	0	0	1
misfiring 7	”M IH S -’F AY -AX -R IH NG	”M AX S -’F AY -AX -R IH NG	”M AX S -’F AY -AX -R IH NG	0	NG	0	0	0	0	0	0	0	0	1
misnomer 1	”M IH S -’N OW -M AX R	”M AX S -’N OW -M AX R	”M AX S -’N OW -M AX R	0	0	0	0	0	0	0	0	0	0	1
moggies 3	M OH -G IH Z	M OH -G IH Z	M OH -G IH Z	0	G	0	0	1	0	0	0	0	0	1
monstrosity 10	M OH N S -’T R OH -S AX -T IH	M OH N S -’T R OH -S AX -T IH	M OH N S -’T R OH -S AX -T IH	0	0	0	0	0	0	0	0	1	0	1
motivating 3	M OW -T IH -V EY -T IH NG	M OW -T AX -V EY -T IH NG	M OW -T AX -V EY -T IH NG	0	0	0	0	0	EY	0	0	0	0	1
motivating 7	M OW -T IH -V EY -T IH NG	M OW -T AX -V EY -T IH NG	M OW -T AX -V EY -T IH NG	0	NG	0	0	0	0	0	0	0	0	1
mountain 4	M AW N -T IH N	M AW N -T AX N	M AW N -T AX N	0	0	0	0	0	0	0	0	0	0	1
naivety 5	N AY -’IY V -T IH	N AY -’IY V -T IH	N AY -’IY V -T IH	0	0	0	0	0	0	0	0	1	0	1
newfangledness 10	N Y UH -”F AE NG -G L D -N IH S	N Y UH -”F AE NG -G L D -N AX S	N Y UH -”F AE NG -G L D -N AX S	0	0	0	0	0	0	0	0	0	0	1
ninepenny 6	N AY N -”P EH -N IH	N AY N -”P EH -N IH	N AY N -”P EH -N IH	0	0	0	0	0	0	0	0	1	0	1
nosing 3	N OW -Z IH NG	N OW -Z IH NG	N OW -Z IH NG	0	NG	0	0	0	0	0	0	0	0	1
ordinal 2	”AO -D IH -N L	”AO -D AX -N L	”AO -D AX -N L	0	0	0	0	0	AX	0	0	0	0	1
overrated 6	”OW -V AX -’R EY -T IH D	”OW -V AX -’R EY -T AX D	”OW -V AX -’R EY -T AX D	0	0	0	0	0	0	0	0	0	0	1
package 3	P AE -K IH JH	P AE -K AX JH	P AE -K IH JH	0	K	JH	0	0	0	0	0	0	1	1
passing 3	P AA -S IH NG	P AA -S IH NG	P AA -S IH NG	0	NG	0	0	0	0	0	0	0	0	1
peaches 3	P IY -CH IH Z	P IY -CH AX Z	P IY -CH AX Z	0	0	0	0	0	0	0	0	0	0	1
peninsula 3	P AX -’N IH N -S Y UH -L AX	P AX -’N AX N -S Y UH -L AX	P AX -’N AX N -S Y UH -L AX	0	0	0	0	0	0	0	0	0	0	1
perfunctorily 11	P AX -’F AH NG K -T AX -R AX -L IH	P AX -’F AH NG K -T AX -R AX -L IH	P AX -’F AH NG K -T AX -R AX -L IH	0	0	0	0	0	0	0	0	1	0	1
pinching 4	P IH N -CH IH NG	P AX N -CH IH NG	P IH N -CH IH NG	0	NG	0	0	0	0	0	0	0	0	1
pinpoints 1	P IH N -P OY N T S	P AX N -P OY N T S	P AX N -P OY N T S	0	0	0	0	0	0	0	0	0	0	1
prearranging 8	”P R IY -AX -’R EY N -JH IH NG	”P R IY -AX -’R EY N -JH IH NG	”P R IY -AX -’R EY N -JH IH NG	0	NG	0	0	0	0	0	0	0	0	1
presides 2	P R IH -’Z AY D Z	P R IH -’Z AY D Z	P R IH -’Z AY D Z	0	0	0	0	0	0	1	0	0	0	1
public 4	”P AH -B L IH K	”P AH -B L IH K	”P AH -B L IH K	0	K	0	0	0	0	0	0	0	0	1
puritans 4	P Y UA -R IH -T AX N Z	P Y UA -R AX -T AX N Z	P Y UA -R AX -T AX N Z	0	0	0	0	0	AX	0	0	0	0	1
queening 4	K W IY -N IH NG	K W IY -N IH NG	K W IY -N IH NG	0	NG	0	0	0	0	0	0	0	0	1
racing 3	R EY -S IH NG	R EY -S IH NG	R EY -S IH NG	0	NG	0	0	0	0	0	0	0	0	1

Table D.12: Results table of full rule set applied to validation set (part 4)

word	pronun_base	pronun_predict	pronun_corr	strt eb,m,n,x	rule 2	rule 3	first	end eys,ies,ied	vwl hrmny ax	vwl hrmny ih	adj vwl hrmny ih	last	phnm src	corr
radicalizing 3	R AE -D IH -K AX -L AY -Z IH NG	R AE -D IH -K AX -L AY -Z IH NG	R AE -D IH -K AX -L AY -Z IH NG	0	K	0	0	0	0	0	0	0	0	1
radicalizing 9	R AE -D IH -K AX -L AY -Z IH NG	R AE -D IH -K AX -L AY -Z IH NG	R AE -D IH -K AX -L AY -Z IH NG	0	NG	0	0	0	0	0	0	0	0	1
radio 3	R EY -D IH -OW	R EY -D IH -OW	R EY -D IH -OW	0	0	0	0	0	0	1	0	0	0	1
raggedness 3	R AE -G IH D -N IH S	R AE -G AX D -N AX S	R AE -G AX D -N AX S	0	G	0	0	0	0	0	0	0	1	1
raggedness 6	R AE -G IH D -N IH S	R AE -G IH D -N AX S	R AE -G AX D -N AX S	0	0	0	0	0	0	0	0	0	0	1
ready 3	”R EH -D IH	”R EH -D IH	”R EH -D IH	0	0	0	0	0	0	0	0	1	0	1
receiving 1	R IH -’S IY -V IH NG	R IH -’S IY -V IH NG	R IH -’S IY -V IH NG	0	0	0	0	0	0	1	0	0	0	1
receiving 5	R IH -’S IY -V IH NG	R AX -’S IY -V IH NG	R IH -’S IY -V IH NG	0	NG	0	0	0	0	0	0	0	0	1
refashioning 7	”R IY -’F AE -SH AX N -IH NG	”R IY -’F AE -SH AX N -IH NG	”R IY -’F AE -SH AX N -IH NG	0	NG	0	0	0	0	0	0	0	0	1
reflected 1	R IH -’F L EH K -T IH D	R IH -’F L EH K -T AX D	R IH -’F L EH K -T AX D	0	0	0	0	0	0	1	0	0	0	1
reflected 7	R IH -’F L EH K -T IH D	R AX -’F L EH K -T AX D	R IH -’F L EH K -T AX D	0	0	0	0	0	0	0	0	0	0	1
remaining 5	R IH -’M EY -N IH NG	R AX -’M EY -N IH NG	R IH -’M EY -N IH NG	0	NG	0	0	0	0	0	0	0	0	1
retches 3	R EH -CH IH Z	R EH -CH AX Z	R EH -CH AX Z	0	0	0	0	0	0	0	0	0	0	1
reticule 3	R EH -T IH -K Y UH L	R EH -T IH -K Y UH L	R EH -T IH -K Y UH L	0	K	0	0	0	0	0	0	0	0	1
revise 1	R IH -’V AY Z	R IH -’V AY Z	R IH -’V AY Z	0	0	0	0	0	0	1	0	0	0	1
rocket 3	”R OH -K IH T	”R OH -K AX T	”R OH -K AX T	0	K	0	0	0	0	0	0	0	1	1
rooted 3	R UH -T IH D	R UH -T AX D	R UH -T AX D	0	0	0	0	0	0	0	0	0	0	1
running 3	”R AH -N IH NG	”R AH -N IH NG	”R AH -N IH NG	0	NG	0	0	0	0	0	0	0	0	1
safety 4	S EY F -T IH	S EY F -T IH	S EY F -T IH	0	0	0	0	0	0	0	0	1	0	1
safety 4	S EY F -T IH	S EY F -T IH	S EY F -T IH	0	0	0	0	0	0	0	0	1	0	1
saucy 3	S AO -S IH	S AO -S IH	S AO -S IH	0	0	0	0	0	0	0	0	1	0	1
screeches 5	S K R IY -CH IH Z	S K R IY -CH AX Z	S K R IY -CH AX Z	0	0	0	0	0	0	0	0	0	0	1
scripts 3	S K R IH P T S	S K R AX P T S	S K R AX P T S	0	0	0	0	0	0	0	0	0	0	1
settling 4	”S EH -T L -IH NG	”S EH -T L -IH NG	”S EH -T L -IH NG	0	NG	0	0	0	0	0	0	0	0	1
songstresses 6	S OH NG -S T R IH -S IH Z	S OH NG -S T R AX -S AX Z	S OH NG -S T R AX -S AX Z	0	0	0	0	0	AX	0	0	0	0	1
songstresses 8	S OH NG -S T R IH -S IH Z	S OH NG -S T R AX -S AX Z	S OH NG -S T R AX -S AX Z	0	0	0	0	0	0	0	0	0	0	1
soviet 3	S OW -V IH -EH T	S OW -V IH -EH T	S OW -V IH -EH T	0	0	0	0	0	0	1	0	0	0	1
sprightliness 8	S P R AY T -L IH -N IH S	S P R AY T -L AX -N AX S	S P R AY T -L IH -N AX S	0	0	0	0	0	0	0	0	0	0	1
squaring 5	”S K W EA -R IH NG	”S K W EA -R IH NG	”S K W EA -R IH NG	0	NG	0	0	0	0	0	0	0	0	1
stenography 9	S T AX -’N OH -G R AX -F IH	S T AX -’N OH -G R AX -F IH	S T AX -’N OH -G R AX -F IH	0	0	0	0	0	0	0	0	1	0	1
Styx 2	S T IH K S	S T IH K S	S T IH K S	0	K	0	0	0	0	0	0	0	0	1
superscription 7	”S UH -P AX -’S K R IH P -SH AX N	”S UH -P AX -’S K R AX P -SH AX N	”S UH -P AX -’S K R AX P -SH AX N	0	0	0	0	0	0	0	0	0	0	1
swindled 2	S W IH N -D L D	S W AX N -D L D	S W AX N -D L D	0	0	0	0	0	0	0	0	0	0	1
syncopating 1	S IH NG -K AX -P EY -T IH NG	S IH NG -K AX -P EY -T IH NG	S IH NG -K AX -P EY -T IH NG	0	NG	0	0	0	0	0	0	0	0	1
syncopating 8	S IH NG -K AX -P EY -T IH NG	S IH NG -K AX -P EY -T IH NG	S IH NG -K AX -P EY -T IH NG	0	NG	0	0	0	0	0	0	0	0	1
tacking 3	T AE -K IH NG	T AE -K IH NG	T AE -K IH NG	0	NG	0	0	0	0	0	0	0	0	1
tactlessly 5	T AE K T -L IH S -L IH	T AE K T -L AX S -L IH	T AE K T -L AX S -L IH	0	0	0	0	0	0	0	0	0	0	1
tactlessly 8	T AE K T -L IH S -L IH	T AE K T -L AX S -L IH	T AE K T -L AX S -L IH	0	0	0	0	0	0	0	0	1	0	1
taxi 4	T AE K -S IH	T AE K -S IH	T AE K -S IH	0	0	0	0	0	0	0	0	1	0	1
telling 3	T EH -L IH NG	T EH -L IH NG	T EH -L IH NG	0	NG	0	0	0	0	0	0	0	0	1
tepid 3	T EH -P IH D	T EH -P AX D	T EH -P AX D	0	0	0	0	0	0	0	0	0	0	1
tittering 1	T IH -T AX -R IH NG	T AX -T AX -R IH NG	T AX -T AX -R IH NG	0	0	0	0	0	AX	0	0	0	0	1
tittering 5	T IH -T AX -R IH NG	T AX -T AX -R IH NG	T AX -T AX -R IH NG	0	NG	0	0	0	0	0	0	0	0	1
tragedies 6	T R AE -JH AX -D IH Z	T R AE -JH AX -D IH Z	T R AE -JH AX -D IH Z	0	0	0	0	1	0	0	0	0	0	1
tragically 4	T R AE -JH IH -K AX -L IH	T R AE -JH IH -K AX -L IH	T R AE -JH IH -K AX -L IH	0	K	0	0	0	0	0	0	0	0	1
tragically 8	T R AE -JH IH -K AX -L IH	T R AE -JH IH -K AX -L IH	T R AE -JH IH -K AX -L IH	0	0	0	0	0	0	0	0	1	0	1
tutting 3	T AH -T IH NG	T AH -T IH NG	T AH -T IH NG	0	NG	0	0	0	0	0	0	0	0	1
twitterers 2	T W IH -T AX -R AX Z	T W AX -T AX -R AX Z	T W AX -T AX -R AX Z	0	0	0	0	0	AX	0	0	0	0	1
unnoticed 5	”AH N -’N OW -T IH S T	”AH N -’N OW -T AX S T	”AH N -’N OW -T AX S T	0	0	0	0	0	0	0	0	0	0	1
unpretentious 4	”AH N -P R IH -’T EH N -SH AX S	”AH N -P R IH -’T EH N -SH AX S	”AH N -P R IH -’T EH N -SH AX S	0	0	0	0	0	0	1	0	0	0	1
vindications 1	”V IH N -D IH -’K EY -SH AX N Z	”V AX N -D IH -’K EY -SH AX N Z	”V AX N -D IH -’K EY -SH AX N Z	0	0	0	0	0	0	0	0	0	0	1
vindications 4	”V IH N -D IH -’K EY -SH AX N Z	”V AX N -D IH -’K EY -SH AX N Z	”V AX N -D IH -’K EY -SH AX N Z	0	K	0	0	0	0	0	0	0	0	1
vivaciously 1	V IH -’V EY -SH AX S -L IH	V AX -’V EY -SH AX S -L IH	V AX -’V EY -SH AX S -L IH	0	0	0	0	0	EY	0	0	0	0	1
vivaciously 8	V IH -’V EY -SH AX S -L IH	V AX -’V EY -SH AX S -L IH	V AX -’V EY -SH AX S -L IH	0	0	0	0	0	0	0	0	1	0	1
vivid 1	V IH -V IH D	V AX -V AX D	V AX -V AX D	0	0	0	0	0	AX	0	0	0	0	1
vivid 3	V IH -V IH D	V AX -V AX D	V AX -V AX D	0	0	0	0	0	0	0	0	0	0	1
watching 3	W OH -CH IH NG	W OH -CH IH NG	W OH -CH IH NG	0	NG	0	0	0	0	0	0	0	0	1
wifeliest 4	W AY F -L IH -IH S T	W AY F -L IH -AX S T	W AY F -L IH -EH S T	0	0	0	0	0	0	1	0	0	0	1
wiggling 1	W IH -G L IH NG	W IH -G L IH NG	W IH -G L IH NG	0	G	0	0	0	0	0	0	0	0	1
wiggling 4	W IH -G L IH NG	W IH -G L IH NG	W IH -G L IH NG	0	NG	0	0	0	0	0	0	0	0	1
Wimpies 1	W IH M -P IH Z	W AX M -P IH Z	W AX M -P IH Z	0	0	0	0	0	0	0	0	0	0	1
Wimpies 4	W IH M -P IH Z	W AX M -P IH Z	W AX M -P IH Z	0	0	0	0	1	0	0	0	0	0	1
win 1	W IH N	W AX N	W AX N	0	0	0	0	0	0	0	0	0	0	1
wisest 3	W AY -Z IH S T	W AY -Z AX S T	W AY -Z AX S T	0	0	0	0	0	0	0	0	0	0	1
with 1	W IH DH	W AX DH	W AX DH	0	0	0	0	0	0	0	0	0	0	1

References

Adda-Decker, M. and Lamel, L. (1999). Pronunciation variants across system configuration, language and speaking style. *Speech Communication*, vol. 29, pp. 83–98.

Amdal, I. and Fosler-Lussier, E. (2003). Pronunciation variation modeling in sutomatic speech recognition. *Telektronikk*, pp. 70–82.

Armstrong, E. (1996). International phonetic alphabet (online).
Available at: `http://www.yorku.ca/earmstro/ipa/diphthongs.html`

Baayen, R., Piepenbrock, R. and Gulikers, L. (1995). The CELEX lexical database. Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Badenhorst, J. and Davel, M. (2007 November). The effect of speech rate variation on acoustic phone stablility in Afrikaans speech recognition. In: *Proceedings of the eighteenth international symposium of the Pattern Recognition Association of South Africa*, pp. 6–9.

Badenhorst, J. and Davel, M. (2008 November). Data requirements for speaker independent acoustic models. In: *Nineteenth international symposium of the Pattern Recognition Association of South Africa*, pp. 147–152.

Beattie, V., Edmondson, S., Miller, D., Patel, Y. and Talvola, G. (1995 September). An integrated multi-dialect speech recognition system with optional speaker adaptation. In: *Fourth European Conference on Speech Communication and Technology*, pp. 1123–1126. ISCA, Madrid, Spain.

Bekker, I. (2009). *The Vowels of South African English (online)*. Ph.D. thesis, North-West University.
Available at: `http://oldwww.ru.ac.za/academic/departments/linguistics/staff/ian/phdmain.pdf`

Benzeghiba, M., De Mori, R., Deroo, O., Dupont, S., Erbes, T. and Jouvet, D.e.a. (2007). Automatic speech recognition and speech variability: A review. *Speech communication*, vol. 49, no. 10-11, pp. 763–786.

Black, A., Lenzo, K. and Pagel, V. (1998 November). Issues in building general letter to sound rules. In: *3rd ESCA Workshop on Speech Synthesis*, pp. 77–80. Jenolan Caves, Australia.

Branford, W. (1994). *The Cambridge History of the English Language, Vol. V: English in Britain and Overseas: Origins and Development*. Cambridge University Press, Cambridge.

Brink, J. and Botha, E. (2001). An acoustic comparison of the vowels and diphthongs of first and second language South African English. In: *Proceedings of the Pattern Recognition Association of South Africa*, pp. 103–108.

Brink, J. and Botha, E. (2002 September). A comparison of L1 and African-mother-tongue acoustic models for South African English speech recognition. In: *Seventh International Conference on Spoken Language Processing*, pp. 1241–1244. ISCA, Colorado, USA.

Daelemans, W., van den Bosch, A. and Zavrel, J. (1999). Forgetting exceptions is harmful in language learning. *Machine Learning*, vol. 34, no. 1-3, pp. 11–41.

Damper, R., Marchand, Y., Marsters, J.-D. and Bazin, A. (1997 June). Aligning text and phonemes for speech technology applications using an EM-like algorithms. *Journal of Sol-Gel Science and Technology*, vol. 8, no. 2, pp. 147–160.

Davel, M. and Barnard, E. (2004*a*). A default-and-refinement approach to pronunciation prediction. In: *Proceedings of the Pattern Recognition Association of South Africa*, pp. 119–123.

Davel, M. and Barnard, E. (2004*b*). The efficient creation of pronunciation dictionaries: machine learning factors in bootstrapping. In: *Proceedings of Interspeech*, pp. 2781–2784. Jeju, Korea.

Davel, M. and Barnard, E. (2006 September*a*). Developing consistent pronunciation models for phonemic variants. In: *Ninth International Conference on Spoken Language Processing*. Pittsburgh, PA, USA.

Davel, M. and Barnard, E. (2006*b*). Effort and accuracy during language resource generation: a pronunciation prediction case study. In: *Seventeenth Annual Symposium of the Pattern Recognition Association of South Africa*, vol. 29. Parys, South Africa.

Davel, M. and Barnard, E. (2008 October). Pronunciation predication with Default&Refine. *Computer Speech and Language*, vol. 22, pp. 374–393.

De Mori, R. (1998). *Spoken Dialogue With Computers*. Academic Press.

Finke, M. and Waibel, A. (1997 September). Speaking mode dependent pronunciation modeling in large vocabulary conversational speech recognition. In: *Fifth European Conference on Speech Communication and Technology*, pp. 2379–2382. ISCA, Rhodes, Greece.

Flege, J. (1987). The production of 'new' and 'similar' phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*, vol. 15, pp. 47–65.

Fosler-Lussier, J. (1999). *Dynamic pronunciation models for automatic speech recognition*. Ph.D. thesis, University of California, Berkeley Fall 1999.

Fukada, T. and Sagisaka, Y. (1997 September). Automatic generation of a pronuciation dictionary based on a pronunciation network. In: *Fifth European Conference on Speech Communication and Technology*, pp. 2471–2474. ISCA, Rhodes, Greece.

Garofolo, J., Lamel, L., Fisher, W., Fiscus, J., Pallett, D. and Dahlgren, N. (1993 February). The DARPA TIMIT acoustic-phonetic continuous speech corpus, NIST order number PB91-100354.

Goronzy, S., Rapp, S. and Kompe, R. (2004). Generating non-native pronunciation variants for lexicon adaptation. *Speech Communication*, vol. 42, pp. 109–123.

Hain, T. (2005). Implicit modelling of pronunciation variation in automatic speech recognition. *Speech communication*, vol. 46, no. 2, pp. 171–188.

Hain, T., Woodland, P., Evermann, G. and Povey, D. (2001). New features in the CU-HTK system for transcription of conversational telephone speech. In: *ICASSP*, pp. 57–60. Salt Lake City, Utah.

Heugh, K. (2007). Language and literacy issues in South Africa. In: *Global Issues in Language, Education and Development. Perspectives from Postcolonial Countries.*, pp. 187–217.

Holter, T. and Svendsen, T. (1999 November). Maximum likelihood modelling of pronunciation variation. *Speech Communication*, vol. 29, no. 2-4, pp. 177–191.

Jande, P. (2006). *Modelling phone level pronunciation in discourse context*. Ph.D. thesis, KTH Computer Science and Communication, Stockholm, Sweden.

Jurafsky, D., Ward, W., Jiaping, Z., Herold, K., Xiuyang, Y. and Sen, Z. (2001). What kind of pronunciation variations is hard for triphones to model? In: *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 1, pp. 577–580. Salt Lake City, UT, USA.

Kessens, J., Cucchiarini, C. and Strik, H. (2003). A data-driven method for modelling pronunciation variation. *Speech Communication*, vol. 40, pp. 517–534.

Kortmann, B. and Schneider, E. (2004). *A Handbook of Varieties of English*, vol. 1. Mouton de Gruyter New York.

Lanham, L. and Traill, A. (1962). South African English pronunciation. *Witwatersrand University Press, Johannesburg*. Reprinted from English Studies in Africa 5(2): 171 206.

Lass, R. and Wright, S. (1985). *The South African chain shift: Order out of chaos*. John Benjamins, Amsterdam.

Lawson, A., Harris, D. and Grieco, J. (2003). Effect of foreign accent on speech recognition in the NATO N-4 corpus. In: *Eighth European Conference on Speech Communication and Technology*, pp. 1505–1508. ISCA, Geneva, Switzerland.

Livescu, K. and Glass, J. (2000). Lexical modeling of non-native speech for automatic speech recognition. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 3, pp. 1683–1686. Istanbul, Turkey.

Marchard, Y. and Damper, R. (2000). A multistrategy approach to improving pronunciation by analogy. *Computational Linguistics*, vol. 26, pp. 195–219.

Martirosian, O. and Davel, M. (2007 November). Error analysis of a public domain pronunciation dictionary. In: *PRASA 2007: Eighteenth Annual Symposium of the Pattern Recognition Association of South Africa*, pp. 13–18. Pietermaritzburg, South Africa.

Martirosian, O. and Davel, M. (2008 November). Acoustic analysis of diphthongs in Standard South African English. In: *PRASA 2008: Nineteenth Annual Symposium of the Pattern Recognition Association of South Africa*, pp. 153–158. Cape Town, South Africa.

Oh, Y., Yoon, J. and Kim, H. (2006). Acoustic models adaptation based on pronunciation variability analysis for non-native speech recognition. *Speech Communication*, vol. 49, pp. 59–70.

(online), I.P.A. (1996). The international phonetic association.
Available at: `http://www.arts.gla.ac.uk/IPA/fullchart.html`

Oshika, B., Zue, V., Weeks, R., Neu, H. and Aurbach, J. (1975 February). The role of phonological rules in speech understanding research. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 23, pp. 104–112.

Ravishankar, M. and Eskenazi, M. (1997 September). Automatic generation of context-dependent pronunciations. In: *5th European Conference on Speech Communication and Technology*, pp. 2467–2470. ISCA, Rhodes, Greece.

Robinson, T. (1996). The British English Example Pronunciation (BEEP) dictionary. (online) ftp://svr-ftp.eng.cam.ac.uk/pub/comp.speech/dictionaries.

Saraçlar, M., Nock, H. and Khudanpur, S. (1999 September). Pronunciation modeling by sharing Gaussian densities across phonetic models. In: *Sixth European Conference on Speech Communication and Technology*. ISCA, Budapest, Hungary.

Sharma, A., M., P., Barnard, E. and Kuun, C. (2009 April). HIV health information access using spoken dialogue systems: Touchtone vs. speech. In: *International Conference on Information and Communication Technologies and Development (to be published)*. Doha, Qatar.

Sloboda, T. and Waibel, A. (1996 October). Dictionary learning for spontaneous speech recognition. In: *Proceedings of the International Conference on Spoken Language Processing*, vol. 4, pp. 2328–2331. Philadelphia, PA, USA.

Statistics South Africa (2007). Community survey 2007 (online).
Available at: http://www.statssa.gov.za/publications/P0301/P0301.pdf

Strik, H. and Cucchiarini, C. (1999). Modeling pronunciation variation in ASR: A survey of the literature. *Speech Communication*, vol. 29, pp. 225–246.

Tajchman, G., Jurafsky, D. and Fosler, E. (1995). Learning phonological rule probabilities from speech corpora with exploratory computational phonology. In: *Proceedings of the 33rd annual meeting on Association for Computational Linguistics*. Cambridge, Massachusetts.

Tebelskis, J. (1995 May). *Speech recognition using neural networks*. Ph.D. thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania.

Torkkola, K. (1993 April). An efficient way to learn English grapheme-to-phoneme rules automatically. In: *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, vol. 2, pp. 199–202. Minneapolis, USA.

United Nations Educational, Scientific and Cultural Organization (2001). Education for all 2000 assessment. (online) http://www.unicef.org/infobycountry/southafrica_statistics.html.

Viterbi, A. (1967). Error bounds for convolutional codes and a asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, vol. 13, pp. 260–269.

Wang, Z., Schultz, T. and Waibel, A. (2003 April). Comparison of acoustic model adaptation techniques of non-native speech. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 1, pp. 540 – 543. Hong Kong.

Webb, V. (1983). On a South African English vowel system. *Stellenbosch Papers in Linguistics*, vol. 10, pp. 134–164.

Wells, F. (1982). *Accents of English*, vol. 1-3. Cambridge University Press, Cambridge.

Wester, M., Kessens, J. and Strik, H. (1998 May). Improving the performance of a Dutch CSR by modelling pronunciation variation. In: *Proceedings of the Workshop Modeling Pronunciation Variation for Automatic Speech Recognition*, pp. 145–150. Rolduc, The Netherlands.

Williams, G. and Renals, S. (1998 May). Confidence measures for evaluating pronunciation models. In: *Proceedings of the Workshop Modeling Pronunciation Variation for Automatic Speech Recognition*, pp. 151–156. Rolduc, The Netherlands.

Young, S., Kershaw, D., Odell, J., Ollason, D., Valtchev, V. and Woodland, P. (2000 July). The HTK book. revised for HTK version 3.0. (online) `http://htk.eng.cam.ac.uk/`.

Zsilavecz, M. (2008 January). ASR-Builder. (online) `http://sourceforge.net/projects/asr-builder`.