MODELLING THE ACOUSTIC DIFFERENCES BETWEEN L1 AND L2 SPEECH: THE SHORT VOWELS OF AFRIKAANS AND SOUTH AFRICAN ENGLISH.

Elizabeth C. Botha^{*} *and Louis C.W. Pols*[•]

*Department of Electrical & Electronic Engineering, University of Pretoria, Pretoria 0002 South Africa.

liesbeth.botha@ee.up.ac.za

• Institute of Phonetic Sciences, University of Amsterdam, Herengracht 338, 1016 CG Amsterdam, The Netherlands. pols@fon.let.uva.nl

Abstract

The acoustic differences between Afrikaans and South African English, spoken as first (L1) and second (L2) language, are measured for nine short vowels. The spoken language data base of 22 male speakers, collected for comparative studies, is described. The features used in an initial comparison of the isolated vowels and vowels in CVC words are the first three formant values and ratios. Significant differences are found between the production of / \mathbf{E} / and /y/ by Afrikaans and English mother-tongue speakers, and to a lesser extent between /i/, / ∂ / and /u/. Several interesting trends that seem to contradict popular beliefs concerning South African accents are observed. Directions for future research and the application of the envisioned L1-L2 model in speech technology are given.

1. INTRODUCTION

In this study[1] we aim to model the acoustic differences between L1 and L2 speech signals. The long-term objectives are to model these acoustic differences and to apply this model to speech recognition systems, so that they can deal with foreign accents in a structured way.

Our short-term goal is to quantify the acoustic differences between two of the many languages spoken in South Africa (SA), namely Afrikaans and English, spoken as L1 and L2. It is generally possible to distinguish between an *Afrikaans* and *English* (mothertongue) accent in SA. It is this distinction that we now want to quantify by studying the production of the nine short vowels /a/, /æ/, / ϵ /, /i/, / θ /, /ce/, /u/, /y/ and / \circ /, when spoken as L1 and as L2. The four conditions are thus Afrikaans spoken as L1 (Afr-L1) and L2 (Afr-L2), and English spoken as L1 (Eng-L1) and L2 (Eng-L2).

The paper is organised as follows: in Section 2 we give some information about the languages used in this study. Sections 3 and 4 contain a description of the database and of the methods used, respectively. In Section 5 we discuss our results and we conclude in Section 6.

2. LANGUAGES USED IN THIS STUDY

White urban South Africans are mostly bilingual in Afrikaans and South African English. This bilingualism is not symmetric, in that Afrikaans mother-tongue speakers are better in their L2 (English) than the SA English mother-tongue speakers are in their L2 (Afrikaans). Both languages are taught for all twelve years of primary and secondary school.

Afrikaans originates from 17th century Dutch. During the past 350 years many other languages such as English, French, German and Malaysian have influenced Afrikaans, which was officially recognised as a separate language (apart from Dutch) in 1925 in SA. Although there are nine other official languages in SA today (from the Nguni, Sotho and Venda language groups), and although these were also spoken in SA throughout the past 350 years, they have had a small influence on Afrikaans.

The literature on acoustic phonetics of both Afrikaans and SA English is sparse, but several introductory texts in Afrikaans phonetics exist, e.g. [2] and [3]. Values for the formants of the Afrikaans vowels were published in [4] and [5]. In [3] the authors remark that at that time (1987) almost no acoustic phonetic research (instrumental, or computer-based) had been performed in SA. This situation was also mentioned in [5], and to the best of our knowledge, still exists today.

3. DATA

A data base of read texts by 11 male Afrikaans and 11 male SA English mother tongue speakers between the ages of 22 and 46 was collected for the comparative study. The texts consist of the "North wind and the sun" passages in both Afrikaans and English, 6 semantically unpredictable Afrikaans sentences, 10 English TIMIT sentences, Afrikaans and English word lists, Afrikaans sound lists with the vowels or consonants in various contexts, an hVt list and a list of isolated vowels. Out of this data base, the isolated vowels of Afrikaans and a list of meaningful words containing the vowels, were used

for the present study. The remainder of the data base will be used in subsequent studies.

The data were recorded in the anechoic room of the Language Laboratory at the University of Pretoria, using a Beyer Dynamics M201N(C) microphone. The data were digitised at 11.125 kHz (16 bits). The total database consists of about 350 MByte of data.

The data of each subject was captured in a single, one hour, recording session. The subjects received verbal instructions regarding posture, voice loudness, paper noise, nature of the texts and goals of the research. They were allowed time to scan the texts before recording. A limited amount of stuttering followed by self-correction of the subjects was allowed, otherwise the individual text was rerecorded. In the three experiments conducted and subsequently described, only a small subset of the comprehensive database was used. In the selections described below, data of all 22 speakers were used.

The scope of this study was restricted to the nine short vowels of Afrikaans /a, i, ϑ , \mathfrak{E} , \mathfrak{U} , \mathfrak{U} , \mathfrak{U} , \mathfrak{V} , \mathfrak{V} , \mathfrak{I} , \mathfrak{V} , $\mathfrak{V$

In the first experiment, the vowels spoken short and in isolation were used. Since the speakers were instructed that these are Afrikaans sounds, only the two conditions Afr-L1 and Afr-L2 were tested. The sounds were manually segmented from the background noise on energy changes, by inspection of the waveform/oscillogram only.

In the second experiment, Afrikaans words containing these nine vowels were chosen. These are: *was* [vas], *kiem* [kim], *is* [\exists s], *guns* [x@ns], *roep* [rup], *ses* [s ϵ s], *vel* [fæl], *kos* [k \circ s] and *debuut* [d \exists byt]. Again, since these are Afrikaans words, Afr-L1 and Afr-L2 were investigated. The vowels were again manually segmented by inspecting both the waveform/oscillogram and spectrogram to select the steady-state section of the vowel. In the third experiment, eight English words containing similar vowels (only /y/ does not occur in English) were chosen to test conditions Eng-L1 and Eng-L2, and similarly segmented. These words are: *dart* [dart], *feel* [fi:1], *about* [\exists baut], *pur* [p α r], *fool* [fu:1], *get* [g ϵ t], *cat* [k α t], and *box* [b \exists ks].

4. METHOD AND RESULTS

The objective of the study is to model the acoustic differences between the production of nine short vowels as L1 and L2 by 11 Afrikaans (group 1) and 11 English (group 2) mother-tongue speakers. In order to do this, the first five formant frequencies of each utterance were calculated in consecutive time frames, using the Split-Levinson algorithm [6]. Then the difference between F1 of the group 1 utterances and group 2 utterances was calculated. Subsequently the difference (F2 and F3), as well as for the formant ratios (F1/F2, F2/F3 and F1/F3). A statistical t-test was performed to determine the significance of the differences.

The window and time step sizes of the formant frequency calculations were 25ms and 10ms respectively for the isolated vowels and 12.5ms and 5ms for the vowels in context (within the English and Afrikaans words). The smaller values for the in-context vowels were necessary due to their shorter duration. Pre-emphasis from 50 Hz and a Hamming window were applied.

Since only steady-state vowels were analysed, a true dynamic analysis of the formants was not required. A single value to represent the formant was sought. From duration values and from visual inspection of the formants, three regions were chosen for the isolated vowels (Experiment 1): the first half, second quarter and second third, respectively, of an utterance of average duration. In subsequent experiments, such a detailed analysis was not deemed necessary, and for the Afrikaans word vowels (Experiment 2), formants were calculated in the (1) centre half and (2) first 75% of the vowel. The formant values of the English word vowels (Experiment 3) were only determined in the first 75% of the vowel. In these regions, the average and median of each formant frequency were determined.

The average formant values of the isolated Afrikaans vowels are shown separately for Afrikaans and English mother-tongue speakers on an F1-F2 plane in Figure 1. The values shown are the averages of the median values in the centre 33% of the utterances. The differences between L1 (Afrikaans mother-tongue speakers) and L2 (English mother-tongue speakers) values are clearly visible. The L1 values correspond quite accurately with the formant values published in [5], as well as those in [4].



Figure 1. Average F1 and F2 formant values of the isolated Afrikaans vowels. English mother-tongue speakers indicated with E and Afrikaans mother-tongue speakers indicated with A.

In Table 1, the shading of the rectangles is a subjective judgement, indicating how much the vowels of the L1 and L2 pronunciations differ in each of the experiments (isolated Afrikaans vowels, Afrikaans words and English words). The L1-L2 differences in formant values (F1, F2 and F3) as well as in formant ratios (F1/F2, F2/F3 and F1/F3) are indicated. Since the formant ratios of $\ell \epsilon$ /, /y/ and /i/ seem to differ consistently, at least in the isolated vowels and Afrikaans words, these values are worth closer inspection. The ratios of these vowels (the average median values of the centre 33% of the vowel) are given in Table 2. Also shown for comparison in the table are the formant ratios published by Van der Merwe et al. [5], as well as the values of the neutral vowel / θ / and the vowel / α /.

Formant values											
Vowel	a	æ	3	i	Ð	œ	u	у	С		
Isolated											
Afr word											
Eng word								NA			
Formant ratios											
Vowel	a	æ	3	i	Ð	œ	u	у	С		
Isolated											
Afr word											
Eng word								NA			

Table 1. Degree of difference between L1 and L2 production of vowels in isolation, Afrikaans and English words. Symbolically, the dark rectangle \blacksquare indicates differences in a number of parameters and the clear rectangle \Box indicates a difference in only one or two parameters.

5. DISCUSSION

In terms of modelling the acoustic differences between Afrikaans and English, produced as L1 and L2, the steady-state vowels show few differences. The most prominent differences are in the production of ϵ/ϵ , which were observed in all the tests, namely in isolated vowels, English and Afrikaans words. More differences were

Vowel	Source	Mother	F1/F2	F2/F3	F1/F3
		tongue			
3\	VdM 93	Afr	0.17	0.76	0.13
	/sɛs/	Afr	0.21*	0.75	0.16*
		Eng	0.23*	0.73	0.17*
	average		0.22	0.74	0.16
/i/	VdM 93	Afr	0.11	0.75	0.08
	/kim/	Afr	0.14*	0.75	0.10*
		Eng	0.15*	0.75	0.11*
	average		0.14	0.75	0.11
/y/	VdM 93	Afr	NA	NA	NA
	/dəbyt/	Afr	0.14	0.80*	0.11
		Eng	0.15	0.76*	0.11
	average		0.14	0.78	0.11
/ə/	VdM 93	Afr	0.35	0.62	0.21
	/əs/	Afr	0.29	0.64	0.19
		Eng	0.32	0.61	0.19
	average	-	0.31	0.62	0.19
/æ/	VdM 93	Afr	0.35	0.69	0.23
	/fæl/	Afr	0.35	0.66	0.23
		Eng	0.38	0.63	0.24
	average	-	0.37	0.64	0.23

Table 2. Formant ratio values of the ϵ/ϵ , i/, y/, ∂/a and i/a/v vowels. Significant differences are indicated by the asterisk * and the average of the Afrikaans and English word values are also given.

observed in the production of the isolated vowels (where a number of significant differences between $\frac{\varepsilon}{\sqrt{\theta}}$, $\frac{\theta}{\sqrt{\eta}}$, $\frac{\eta}{\sqrt{\eta}}$ and $\frac{\eta}{\sqrt{\eta}}$ occurred) than in the production of the same vowels in the context of Afrikaans and English words. See Table 1 for a summary of these differences.

The tabulation (Table 2) and graphical representation (not shown) of the formant ratios of the most different

vowels (/ ε /, /i/ and /y/), show some interesting trends. The ratios of these vowels are displayed together with the ratios of the /æ/ and / Θ /. The /æ/ can be viewed as a "stronger" relative of / ε /, because the tongue is lower and tenser and the mouth is open wider. The neutral vowel / Θ / is viewed as a "weaker" relative of all the other vowels. Against this background, the trends for the English mother tongue speakers seem to be that the production of the / ε / is "stronger" than their Afrikaans counterparts, since the values are closer to those of the /æ/. In contrast, the /i/ is pronounced "weaker" by the English speakers, as their /i/-values tend towards the / Θ /-values. Inspecting the values of Table 2, it can be seen that the /y/ is pronounced more like /i/, its unrounded counterpart.

The trend concerning /i/ and $/\partial/$ is contrary to the popular notion that Afrikaans speakers tend to neutralise the /i/, relative to their English neighbours. This seems to be worth an in depth separate study.

The former trend concerning the $/\mathscr{R}/$ and $/\varepsilon/$ is interesting in relation to a general belief about Afrikaans speakers in the northern part of SA, formerly *Transvaal*. It is typical of Transvaal speakers to pronounce the word *ek* as $[\mathscr{R}k]$, while Southern Cape speakers would pronounce it as $[\varepsilon k]$. The interesting part is that the Transvaal English speakers of this study produce an even stronger $/\mathscr{R}/$ than the Afrikaans speakers in the same region.

The third interesting observation relates to the theory of *equivalence classification* [7], which states that L2 speakers learn to produce unknown sounds similar to the L1 speakers of the same sounds, because these sounds form a new category for L2-speakers. The results of this study contradict this theory, as the /y/ which is unknown in English, shows a number of significant differences in both the isolated vowels as well as the Afrikaans word [dəbyt]. (See Table 1.)

6. CONCLUSIONS

It is generally believed that an Afrikaans or English accent (for the language spoken as L1 or L2) can be easily detected in the bilingual society of SA. The results of this study show that the steady-state vowel is not the only place to look if this accent is to be quantified. We did learn, however, that $\langle \epsilon \rangle$ and $\langle y \rangle$ are consistently pronounced differently, and to a lesser extent also $\langle i/, \langle \Theta \rangle$ and $\langle u/.$

Since it is common knowledge that prosody contributes largely to foreign accentedness, an obvious continuation of this study would be to investigate dynamic differences, such as found in the diphthongs and controversial/ambiguous sounds /e/, /Ø/ and /o/ in Afrikaans. If the acoustic differences of Afrikaans and English are to be modelled comprehensively, a statistical

approach is necessary. This will have to include all the phones, and in many contexts, such as found for American or British English in the large corpora used in ASR research.

Finally it remains to be seen whether and how this model can be applied to speech technology systems. Two avenues for investigation are currently envisioned:

1. If a rule-based speech synthesis system for English is available, the model is applied to change the rules so that when the sound production rules are used for Afrikaans words (the new language), these are pronounced without an English accent.

2. In ASR systems that use explicit phonetic features (not the frame-based acoustic features of the state-of-the-art systems), the recognition parameters are to be adjusted to fit the new language.

ACKNOWLEDGEMENTS

The first author executed most of the measurements while on sabbatical at the Institute of Phonetic Sciences at the University of Amsterdam. The support for this period by the Foundation for Research Development in SA and the University of Pretoria is gratefully acknowledged.

REFERENCES

[1] Botha, Liesbeth (1996): "Towards modelling acoustic differences between L1 and L2 speech: the short vowels of Afrikaans and South-African English", *Proceedings of the Institute of Phonetic Sciences* 20, University of Amsterdam: 65-80.

[2] Coetzee, A.E. (1981): *Fonetiek vir eerstejaars*, Academica, Pretoria.

[3] De Villiers, M. and Ponelis, F.A. (1987): *Afrikaanse klankleer*, Tafelberg, Kaapstad.

[4] Taylor, J.R. and Uys, J.Z. (1988): "Notes on the Afrikaans vowel system", *Leuvense Bijdragen* **77**(2): 129-149.

[5] Van der Merwe, A., Groenewald, E., van Aardt, D., Tesner, H.E.C. and Grimbeeck, R.J. (1993): "Die formantpatrone van Afrikaanse vokale soos geproduseer deur manlike sprekers", *Suid-Afrikaanse Tydskrif vir Taalkunde* **11**(2): 71-79.

[6] Willems, L.F. (1986): "Robust formant analysis", *Technical Report no. 529*, Institute for Perception Research, Eindhoven, The Netherlands.

[7] Flege, J.E. (1987): "The production of new and similar phones in a foreign language: evidence for the effect of equivalence classification", *Journal of Phonetics* **15**: 47-65.