Formant frequencies of British English vowels produced by native speakers of Farsi

G. Hunter and H. Kebede

Faculty of Science, Engineering and Computing, Kingston University, Penrhyn Road, KT1 2EE Kingston Upon Thames, UK
g.hunter@kingston.ac.uk

We discuss issues relating to phoneme (in particular, vowel) production in a subject's second language, focusing on the vowel systems of Standard Southern British English (SSBE) and Farsi (Persian). We describe a study wherein first language Farsi speakers who were experienced second language speakers of SSBE were recorded attempting to produce SSBE vowels in words within a standard carrier phrase. The first and second formants of the vowels so produced were measured, and the results compared with measurements of SSBE and Farsi vowels produced by L1 speakers from previous studies.

.

1 Introduction

The problems encountered by non-native speakers pronouncing or perceiving unfamiliar phonemes in a second language (L2) are well-known. Several previous studies have focused on such issues – including the common problem producing or perceiving the /r/-/l/ contrast in English by native speakers of Mandarin Chinese or Japanese [1]. The *perceptual magnet hypothesis* [2] proposes that non-native speakers tend to "attract" phonemes in their L2 to standard exemplar phonemes in their first language (L1). This is believed to be the case both for perception of other people's speech and their own production of sounds in their L2 [3].

The formants of vowel sounds can be particularly important both from the point of view of intelligibility (correct identification of the vowel) and perceived "naturalness" or "native quality" of speech. For the purposes of classification, it is generally believed that the two lowest frequency formants, F1 and F2, are the most important. However, different languages vary considerably in the set of phonetic vowels they use, and in the precise acoustic properties (including formants) of those vowels. For example, Castillian (European) Spanish has just 5 phonetic vowels {/a/, /e/, /i/, /o/, /u/ }, all of which are relatively close together in F1-F2 space (often referred to as the "vowel quadrilateral") [16]. This contrasts with the case of Standard Southern British English, which is normally considered to have 11 phonetic vowels (not including diphthongs), which are widely distributed across F1-F2 space. A good understanding of how the phoneme system of a particular L1 affects a non-native speaker's perception and pronunciation of English can have significant applications to the Teaching of English to Speakers of Other Languages (TESOL), for example, producing teaching aids to help learners of particular L1s speak better English [4], in speech therapy for L2 English language learners with various speech impairments, and in biometric speaker classification or identification – for example, trying to identify a speaker's L1 from the way they attempt to pronounce various English phonemes. Although many previous studies have looked at non-native listeners perception of various sounds in their L2, produced by both native and non-native speakers (e.g. [5] for Norwegian speakers), and some researchers have investigated L2 productions by speakers of some particular L1s - e.g. Latin American Spanish [4], German [4] and Greek [6] - there is still much work which can be done to extend these to a wider range of speakers' native languages. The only previously published work on the pronunciation of English by native speakers of Farsi seems to be that of Hall [7], who investigated the intelligibility to L1 speakers of Australian English of English sentences spoken by L1 Farsi speakers. However, this study did not make any quantitative measurements of quantities such as vowel formants.

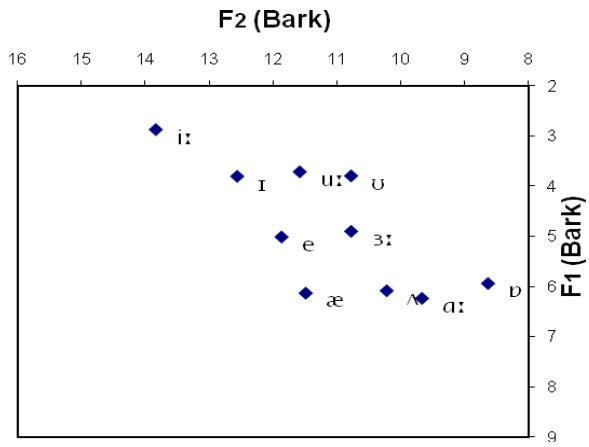
1.1 The Vowel System of SSBE (Standard Southern British English)

It is commonly accepted that Standard Southern British English contains 11 monophthong vowels [9] { /iː/, /ɪ/ , /e/ , /æ/ , /ʌ/ , /ɑː/ , /ɒ/ , /ɔː/ , /ʊ/ , /uː/ , /ɜː/ }. Formant values for each of these have been measured by Wells [8] and more recently by Deterding [9, 10]. Summaries of the latter's findings is given in Table 1 and Figure 1 below.

Table 1: Mean Formant frequencies of the eleven phonetic monophthong vowels of SSBE, as produced by three L1 speakers [10].

	Formants	
SSBE Vowel	F1 / Hz	F2 / Hz
iː	296	2241
ɪ	396	1839
e	532	1656
æ	667	1565
ʌ	661	1296
ɑː	680	1193
ɒ	643	1019
ɔː	480	857
ʊ	395	1408
uː	386	1587
ɜː	519	1408

Figure 1 : Data on SSBE vowels in Table 1 (from [10]) plotted as a "vowel quadrilateral", with formant frequencies converted to Bark scale [17, 18]



1.2 Farsi and its Vowel System

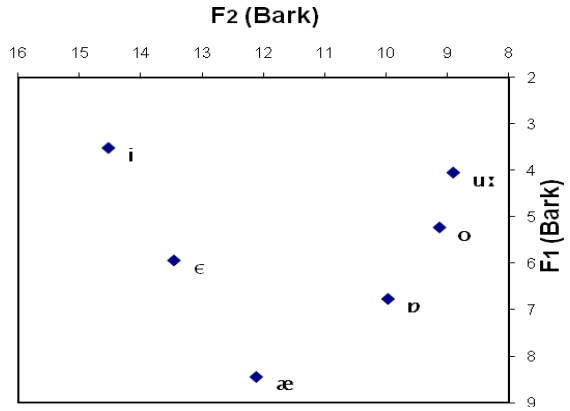
Farsi (Persian) is, if various dialects such as Dari and Tajiki are included, L1 for over 30 million people [11] (some estimates are considerably higher), placing it amongst the 40 most widely spoken of the World's languages. It is the official and principal language of Iran, suggesting that it is spoken as either an L1 or L2 by over 60 million people. Farsi is one of the Indo-Iranian subgroup of the Indo-European family of languages, and is closely related to Kurdish, Pashtoo, Hindi and Urdu. It is generally accepted that it contains 6 phonetic vowels { /i/, /e/, /æ/, /ɒ/, /o/, /u/} [12]. The first and second formants of productions of these Farsi vowels by young adult female L1 Farsi speakers were measured by Ansarin [13], and his results are summarized in Table 2 and Figure 2 below.

There are rather more vowels in SSBE than in Farsi, and it was expected that Farsi speakers may fail to pronounce some of the SSBE vowels correctly [7]. Anecdotally, it is said that some L1 Farsi speakers tend to pronounce "ship" as "sheep". In this paper, we describe a study where recordings were made of L1 Farsi speakers attempting to produce British English vowels, within a controlled context. These recordings were analysed to measure the formants, in order to investigate the hypothesis that the speakers' productions of the English vowels would be influenced by the properties of the nearest equivalent Farsi vowel (when there was one such), and to study how well the speakers could produce vowel sounds which were not present in their L1.

Table 2. Mean Formant frequencies of the six phonetic vowels of Farsi, as produced by 12 young female L1 Farsi speakers [13].

	Formants	
Farsi Vowel	F1 / Hz	F2 / Hz
i	365	2508
ɛ	644	2115
æ	990	1722
ɒ	750	1251
o	558	1102
u	423	1065

Figure 2: Data on Farsi vowels in Table 2 (from [13]) plotted as a "vowel quadrilateral", with formant frequencies converted to Bark scale [17, 18]



2 Experimental Procedure

The study was performed by recording a number of adult L1 Farsi speakers reading a number of English sentences aloud. Each sentence included a /bVd/ word, for example, "bad", where V is a British English phonetic vowel, within a standard carrier phrase. The carrier phrase was used to control the phonetic environment for each vowel. The recordings were analysed to allow the measurement of the formants of each vowel. The project was carried out within the guidelines of our institutions ethical policies for research involving human subjects.

2.1 Subjects

All subjects were L1 Farsi speaking Iranians aged between 20 and 35 who had lived and studied in the U.K. for between 2 and 6 years, and were still so doing at the time of taking part in the study. Of these, 4 were female and 2 male. All had spent their formative years in Iran, but all were competent speakers of conversational English and were used to reading English text on a daily (or near daily) basis. All subjects were volunteers who had been informed of what the experiment involved, told that the recordings of their voices and their identities would not made public and that they had the right to withdraw from the study at any point.

2.2 Materials

Each subject was asked to read aloud a sequences of printed English sentences, each one containing a /bVd/ word within a standard carrier phrase, "Could you say the word /bVd/ please ?", pausing briefly between successive sentences. Prior to the subject reading the sentences, each had been able to read the sentences silently whilst listening to a native SSBE speaker reading all the sentences aloud, in order to ensure the set did not contain any words with which the subject was completely unfamiliar. Examples of /bVd/ words where the vowel was an SSBE diphthong were also recorded, but the results are not presented here.

2.3 Recording Procedure

Audio recordings of the subjects' read utterances were made, and a electroglottograph (EGG) signal, to monitor the vibration of the subject's vocal folds and facilitate pitch (fundamental frequency) calculation for further analysis, recorded simultaneously using a *Laryngograph* processor [14].

2.4 Analysis

The *Speech Filing System* (SFS) software suite [15] was used to analyse the recordings, including measuring the formants of each vowel. This system allows, amongst many other features, the display of both waveforms and spectrograms and the playback of sections of the recording to enable the easy identification of individual words and phonemes within an utterance.

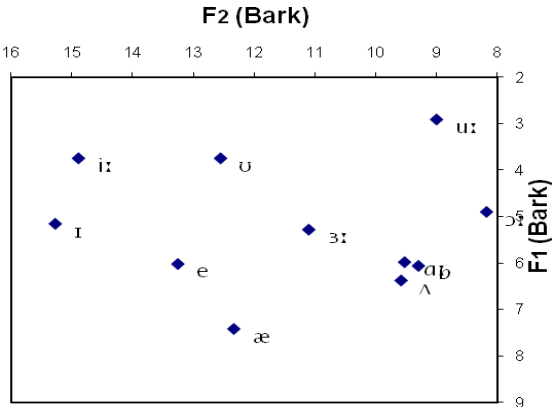
3 Results and Discussion

The formants of the L1 Farsi speakers' attempts at producing SSBE monophthong vowels are give in tables 3 (for female subjects) and 4 (for male subjects) below. The values marked with * for /ʊ/ were obtained from pronunciations of the word "good", since there is no valid word /bʊd/ in SSBE.

Table 3. Mean Formant frequencies of attempts at producing the eleven phonetic monophthong vowels of SSBE, by typical female L1 speaker of Farsi.

	Formants	
SSBE Vowel	F1 / Hz	F2 / Hz
iː	390	2665
ɪ	550	2840
e	655	2050
æ	840	1780
ʌ	700	1180
ɑː	650	1170
ɒ	660	1130
ɔː	520	950
ʊ	390*	1840*
uː	300	1080
ɜː	565	1480

Figure 3: Data on attempts by Iranian female speaker to produce SSBE vowels (from Table 3) plotted as a "vowel quadrilateral", with formant frequencies converted to Bark scale [17, 18].

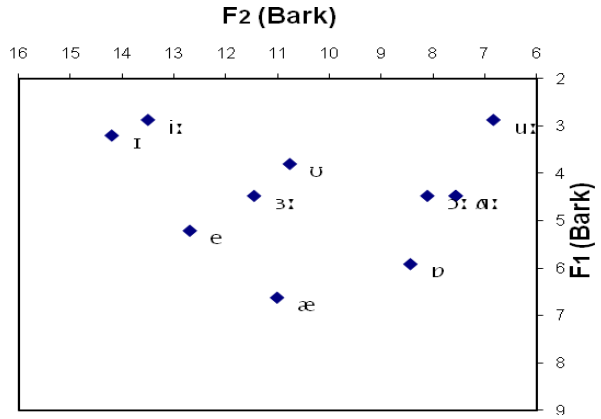


As can be seen from the tables and figures, the L1 Farsi speakers tend to produce good approximations (in terms of F1 and F2 values) to SSBE vowel when there is a closely equivalent vowel in Farsi. Their attempts at pronunciation of /i/ and /ɪ / tended to get confused, as did /ʌ/ and /ɑː/, and some participants also confused the latter with /ɒ/.

Table 4. Mean Formant frequencies of attempts at producing the eleven phonetic monophthong vowels of SSBE, by typical male L1 speaker of Farsi.

	Formants	
SSBE Vowel	F1 / Hz	F2 / Hz
iː	295	2130
ɪ	330	2380
e	555	1880
æ	730	1460
ʌ	470	860
ɑː	470	860
ɒ	640	990
ɔː	470	940
ʊ	395*	1408*
uː	295	760
ɜː	470	1560

Figure 4: Data on attempts by Iranian male speaker to produce SSBE vowels (from Table 4) plotted as a "vowel quadrilateral", with formant frequencies converted to Bark scale [17, 18].



The Iranian female speaker (Figure 3) has an English vowel space which is less tightly-clustered than that for SSBE speakers (Figure 1), and some of her productions of English vowels do seem to be very close to the nearest Farsi equivalent (Figure 2). The male Iranian shows a broadly similar pattern, with his English vowel space appearing to be "bound" more by the "triangular" boundary of Farsi vowel space than the approximate "trapezium" for SSBE. Both produced / uː/ more as a close back vowel (as it is in Farsi [13]) rather than a close-mid vowel (as it is in SSBE [10]).

Example sets of SFS outputs, showing speech signal waveform, wideband spectrogram and laryngograph signal are shown in Figure 5 (for a male Farsi speaker) and Figure 6 (for a female Farsi speaker) pronouncing some /bVd/ English words. In Figure 6, it can be noted that the speaker pronounced the word pairs "beard" and "bird", and "bod" and "bud" in a very similar manner.

Figure 5. Example SFS output, showing speech waveform, wideband spectrogram, laryngograph signal and formant track for productions of /bVd/ SSBE words by a male L1 Farsi speaker.

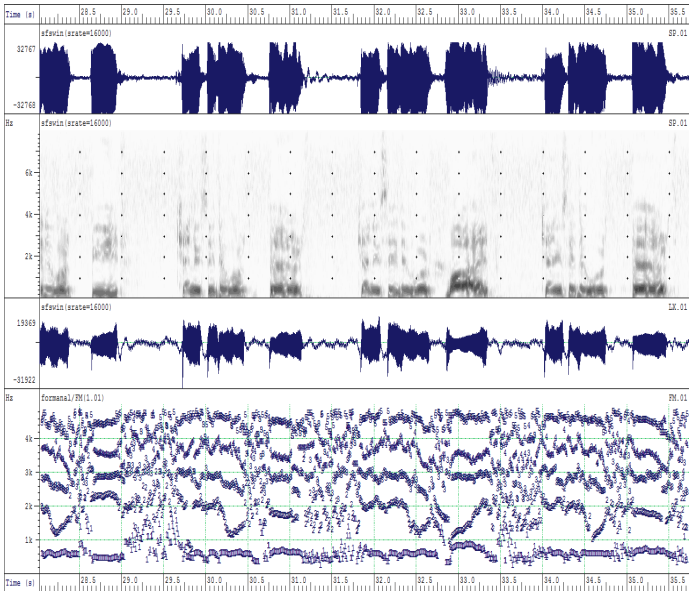
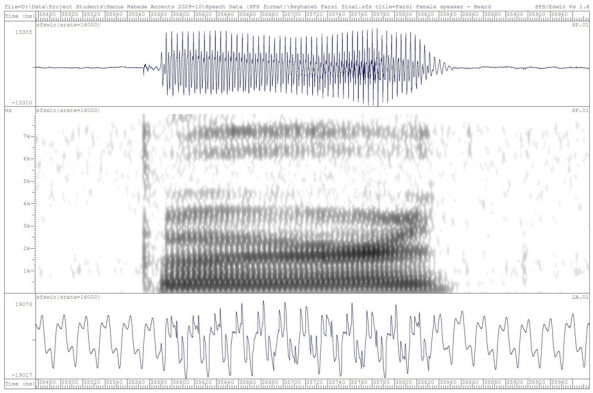
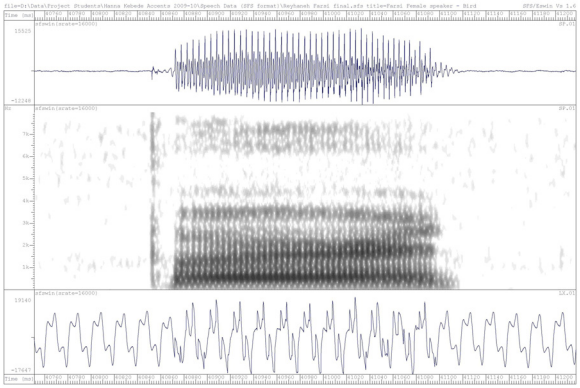


Figure 6. SFS output, showing speech waveform, wideband spectrogram and laryngograph signal for productions of /bVd/ SSBE words by an L1 female Farsi speaker.

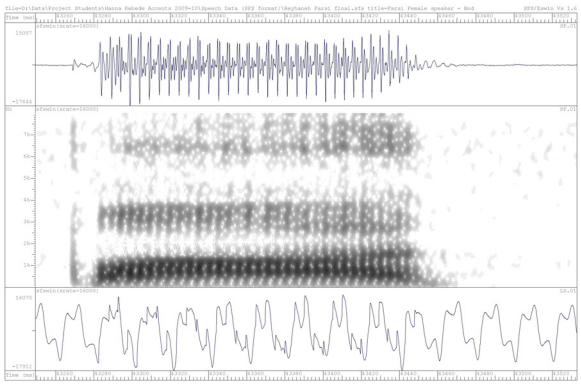
(a) “beard”



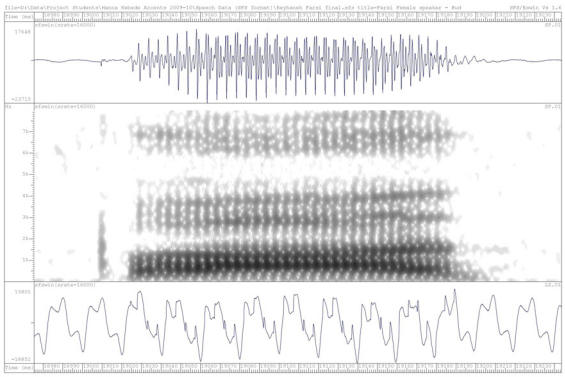
(b) “bird”



(c) “bod”



(d) “bud”



4 Conclusions and Future Work

We have recorded attempts at producing SSBE vowel sounds by experienced L2 speakers of English whose L1 is Farsi, and measured the first two formants of the resulting phonemes. As expected, the participants generally produced “good” examples of SSBE vowels when there was a Farsi vowel which was close in F1-F2 space to the required SSBE one. However, their productions of SSBE vowels for which there was no close equivalent in Farsi were considerably poorer, but the results had formants which were often rather close to those for some other Farsi vowel. As a consequence, the subjects frequently confused certain SSBE vowels. These findings are consistent with the perceptual magnet hypothesis [2].

It is planned to extend this study to investigate productions of a wider range of SSBE phonemes, and possibly perception of such phonemes, by L1 Farsi speakers, and also to diversify this study to other L2 speakers of English, such as Punjabi, whose L1 has received relatively little attention in this context.

Acknowledgments

We would like to thank all the subjects for their participation in the study, and Dr. Abbas Heydari of the University of Durham and Dr. Nasrollah Saebi of Kingston University for some advice about Farsi pronunciation.

References

[1] K. Hattori, K. & P. Iverson (2009) "English /r/-/l/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy". *Journal of the Acoustical Society of America* 125(1), 469-479

[2] P. Iverson & P. Kuhl (1995) "Mapping the Perceptual Magnet Effect for Speech Using Signal Detection Theory and Multidimensional Scaling" *Journal of the Acoustical Society of America*, Vol. 97 (1), pp 553-562

[3] P. Iverson, P.K. Kuhl, R. Akahane-Yamada, E, Diesch, Y. Tohkura, A. Kettermann & C.Siebert (2003). "A perceptual interference account of acquisition difficulties for non-native phonemes" *Cognition* 87, B47-B57

[4] P. Iverson & B.G. Evans (2009). "Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers" *Journal of the Acoustical Society of America* 122(5), 2842-2854

[5] W.A. Van Dommelen & V. Hazan (2010). "Perception of English consonants in noise by native and Norwegian listeners". *Speech Communication* 52, 968-979

[6] A. Lengeris & V. Hazan (2010). "The effect of native vowel processing ability and frequency discrimination acuity on the phonetic training of English vowels for native speakers of Greek". *Journal of the Acoustical Society of America* 128(6), 3757-3768

[7] M. Hall (2007) "Phonological Characteristics of Farsi Speakers of English and L1 Australian English Speakers' Perceptions of Proficiency", M.A. Dissertation, Curtin University of Technology, Perth, Australia, available at http://www.asian-efl-journal.com/Thesis-M-Hall.pdf

[8] J.C. Wells (1962) "A Study of the Formants of the Pure Vowels of British English", M.A. Dissertation, University of London, U.K., available at http://www.phon.ucl.ac.uk/home/wells/formants/index.htm

[9] D. Deterding (1997). "The Formants of Monophthong Vowels in Standard Southern British English Pronunciation." *Journal of the International Phonetic Association*, 27, 47-55

[10] D. Deterding (2006). "The North Wind versus a Wolf: short texts for the description and measurement of English pronunciation" *Journal of the International Phonetic Association*, 36(2), 187-196.

[11] M.P. Lewis (ed.) (2009). "Ethnologue: Languages of the World", Sixteenth edition. Dallas, Tex.: SIL International. Online version: http://www.ethnologue.com/.

[12] D. Gharavian & S.M. Ahadil (2004) "Evaluation of the Effect of Stress on Formants in Farsi Vowels", *Proceedings of IEEE ICASSP 2004*

[13] A.A. Ansarin (2004) "An Acoustic Analysis of Modern Persian Vowels", *Proceedings of SPECOM 2004 : 9th ISCA Confererence on Speech and Computer, St. Petersburg, Russia*, September 2004

[14] Laryngograph Ltd., http://www.laryngograph.com/

[15] Speech Filing System, http://www.phon.ucl.ac.uk/resource/sfs/

[16] J.I. Hualde (2005) "The Sounds of Spanish", Cambridge University Press, Cambridge, U.K. pp 53-54, 120-128

[17] E. Zwicker & E. Terhardt (1980). "Analytical expression for critical-band rate and critical bandwidth as a function of frequency". *Journal of the Acoustical Society of America*, 68(5), 1523-1525.

[18] D. Deterding (2006) "Measuring and Plotting Vowels", http://videoweb.nie.edu.sg/phonetic/vowels/measurements.html