Lexical Frequency and Neighbourhood Density effects on Vowel Production in Words and Nonwords

Lisa Stephenson

Macquarie Centre for Cognitive Science and the Speech, Hearing and Language Research Centre
Macquarie University, Sydney, Australia
lisa@maccs.mq.edu.au

Abstract

Lexical frequency and phonological neighbourhood density have been found to influence vowel production. This experiment investigated the effect of lexical frequency and phonological neighbourhood density on vowel space expansion in words and nonwords. Twenty speakers produced words varying in phonological neighbourhood density and word frequency and produced nonwords that varied in neighbourhood density. Words with low frequency and dense phonological neighbourhoods were produced with more expanded vowel spaces than frequent words with sparse phonological neighbourhoods. In contrast, nonwords from sparse neighbourhoods were produced with more vowel space expansion than those from dense neighbourhoods. The pattern for individual vowels was also investigated.

1 Introduction

Previous studies on the intelligibility of speech have found an association between increased intelligibility and expanded vowel spaces (e.g., Moon & Lindblom, 1994; Picheny, Durlach, & Braida, 1986). Bradlow, Toretta and Pisoni (1996) showed the mean Euclidean distance measure of vowel space expansion was a good predictor of intelligibility. In addition, Wright (2003) found vowel space expansion differed in words varying on lexical frequency and phonological neighbourhood density. Lexical frequency is a measure of how commonly a reader or a speaker will encounter a particular word. Phonological neighbourhood density is a measure of how many words differ by a single phoneme from the target word. For example, *"dog"* is from a dense phonological neighbourhood with many phonologically similar words, such as *"log"*, *"dot"*, and *"dig"*; whereas, *"zipper"* from a sparse neighbourhood like has no phonetically similar words. Wright used *Easy* and *Hard* words: *Easy* words came from sparse phonological neighbourhoods and relative to their phonological neighbours they had high lexical frequencies; *Hard* words came from densely populated phonological neighbourhoods and relative to their neighbours had low lexical frequencies. Wright found that *Hard* words showed larger vowel spaces, suggesting greater vowel dispersion when compared to the list of *Easy* words.

This study aimed to extend the study by Wright (2003) by increasing the number of tokens. In line with Wright's results, it is hypothesised that Easy (high frequency, sparse phonological neighbourhood) words will show a less vowel space expansion (or dispersion) in comparison to Hard (low frequency, dense phonological neighbourhood) words. This hypothesis is based on the hyper- and hypo-theory of speech as proposed by Lindblom (1990). Lindblom suggests speech production involves the balance between listener- and speaker-constraints: listener-constraints aim for speech that is clear and successful in delivering a message, whereas speaker-constraints attempt to produce speech with decreased articulatory effort. The hypothesis is that if listeners hear a word more frequently and this word has few similar sounding words, then these Easy words may be less constrained by listener-oriented forces and speakers may show more coarticulation and use less articulatory effort. Conversely, Hard words that are heard infrequently with many similar sounding words may require increased articulatory effort because they are more difficult for listeners to identify.

In addition to replicating Wright's experiment (2003), the present study will compare differences in the production of lexical (Word) and non-lexical (Nonword) items. Here, it is hypothesised that Words will show a reduction in vowel space in comparison to Nonwords. Low frequency words are produced with a

Proceedings of the 10th Australian International Conference on Speech Science & Technology

Accepted after full paper review

longer duration and with more vowel space dispersion than high frequency words (Umeda, 1975; Wright, 2003). Therefore, the nonwords without a lexical frequency will be produced with longer durations and greater vowel dispersion than low frequency words.

Finally, rather than just looking at vowel space expansion, this study will also investigate dispersion within vowel types. It is hypothesised that Easy items will show more variability and therefore more vowel dispersion within a vowel because these items will be less hyperarticulated. This may appear contrary to the predictions concerning vowel space expansion; however, items with more intra-vowel dispersion would show less vowel space expansion because they would have more items closer to the centre of the vowel space.

2 Method

2.1 Speakers

Twenty native speakers of Australian English (ten males and ten females) from the Macquarie University community participated in this experiment. Their average age was 29.3 years (S.D. 6.34 years and ranged between 22 to 40 years). The speakers were unaware of the experimental aims, did not have any known speech or hearing disorders and had normal or corrected to normal vision.

2.2 Stimuli

The complete stimuli set comprised the 68 words (34 Easy and 34 Hard) used by Wright (2003), along with 66 newly selected Words (32 Easy and 34 Hard) and 136 Nonwords (68 Easy and 68 Hard). The words were CVC in structure and were selected from CELEX (Baayen, Piepenbrock, & van Rijn, 1993). Nonword items were generated from the ARC Nonword database (Rastle, Harrington, & Coltheart, 2002) and were selected from sparse (Easy) and dense (Hard) phonological neighbourhoods. The average lexical frequency and phonological neighbourhood density for all the stimuli are given in Table 1.

Table 1: Average lexical frequency and phonological neighbourhood density across the Stimuli (Word/Nonword) and EH (Easy/Hard) conditions.

	Nonword		Word	
	Easy	Hard	Easy	Hard
Lexical Freq.	0	0	3069.9	308.75
Phon. N. Density	14.5	29.7	15.45	30.65

The “new” Easy and Hard words did not differ significantly from the Wright Easy and Hard words on lexical frequency or phonological neighbourhood density. The Easy and Hard Nonwords did not differ significantly from the Wright Easy and Hard words on phonological neighbourhood density.

As in the study by Wright (2003), the vowels were not distributed evenly through the lists. This was due to constraints regarding lexical frequency and neighbourhood density. Each speaker produced 270 tokens, giving a total of 5400 recorded tokens.

2.3 Procedure

Recordings took place in a sound-treated studio in the Speech, Hearing and Language Research Centre at Macquarie University. A beyerdynamic TG-X 45 microphone was used for the recording with a TASCAM DA P1 portable digital audio tape recorder at a sampling rate of 44.1 kHz. The digital audio tapes were then digitised onto a SUN workstation at a sampling rate of 20 kHz. Stimuli were presented on a 15-inch Apple emac monitor using PsyScript (Bates & D'Oliveiro, 2003). Items were presented in randomised orders and each item was presented once. The experiment was preceded by a practice block consisting of 12 items (both words and nonwords) not included in the main experiment. Speakers were asked to be accurate (rather than fast) in producing the items they were shown. The task was self-paced; participants pressed the space bar to continue to the next item.

2.4 Labelling and analysis

The recordings were segmented by hand from the acoustic waveform and spectrographic data using the EMU Speech Database System (Cassidy & Harrington, 2001). The recordings were then labelled automatically at the Bavarian Archive for Speech Signals, University of Munich, using a simple forced alignment technique based on Baum-Welch training (1-16 Gaussians per state depending on training material). These labels were then manually checked and corrected. Formants 1-4 were automatically tracked using the EMU Pitch and Formant tool and were manually checked and corrected.

The mean number of tokens analysed per participant was 151, and ranged from 144 to 154 tokens. This is less than the 270 recorded tokens because errors were removed and subsequently the decision was made to exclude diphthongs from the analysis. Although diphthongs were analysed by Wright (2003), it was decided to restrict the analysis to monophthongs as done by Munson and Solomon (in press). The number of vowels to be analysed in each condition is given in Table 2.

Table 2: Number of vowels analysed in each condition.

	Nonword		Word		
	Easy	Hard	Easy	Hard	Total
æ	60	182	40	80	362
ɪ	20	80	188	142	430
iː	273	129	80	198	680
ɔ	100	98	180	153	531
ʉː	93	59	80	20	252
ɐ	97	214	180	270	761
Total	643	762	748	863	3016

2.4.1 Vowel Formants – F1 & F2

Vowel targets were tracked automatically after excluding the initial 25% and final 25% of vowel duration to minimise the effects of the surrounding consonants. Targets were tracked according to vowel height and backness features: thus, the target in low and mid vowels including /æ ɐ ɔ/ were tracked at the F1 maximum. The target of the high vowels /ɪ/ and /i/ was marked at the F2 maximum. /ʉː/ is a central high vowel in Australian English and for most speakers its target could be identified at the F2-maximum.

As in previous studies of lexical competition and vowel space expansion (Munson & Solomon, in press; Wright, 2003), formant values in Hertz were converted to Bark values. The Bark scale attempts to represent an auditory scale, rather than the Hertz linear measure of frequency. The Bark scale represents critical bands of sensitivity to frequency differences and these bands differ in size depending on the frequency. Listeners behave in a similar way, being more sensitive to frequency changes at lower frequency in Hertz than the same magnitude of change at higher frequencies.

2.4.2 Vowel Space Expansion

Two methods were used to investigate vowel space expansion: the average Euclidean distance between every vowel to every other vowel (Inter-vowel D) and the average Euclidean distance between all tokens of the same vowel (Intra-vowel D).

2.4.2.1 Euclidean Distance between all vowels (Inter-vowel D)

Vowel expansion in the F1/F2 space was measured as the average Euclidean distance between each individual vowel token to every other vowel token (without repetition). For example, to calculate inter-vowel Euclidean distances let us consider Figure 1, there are 9 vowel tokens with three tokens from each of /i/, /a/ and /u/. This measure of Euclidean distance calculates the distances between all the vowels without measuring the same distance twice. Therefore, in Figure 1, the Euclidean distances to be calculated are all the black or dashed lines. The dashed lines represent the distances to be calculated for a single vowel token for /i/.

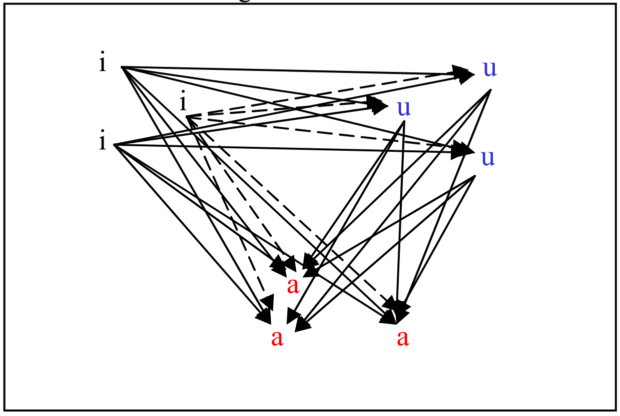

Figure 1: Inter-Euclidean distances are calculated between each vowel token to every other token. The Euclidean distances to be calculated are represented by the dashed and black lines. Dashed lines represent the Euclidean distances to be calculated from a single vowel token.

2.4.2.2 Intra-vowel Euclidean distance (Intra-D)

The distance between tokens of a vowel type also give an indication of vowel space expansion albeit for single vowels. This measure represents the variability within a vowel type. Intra-vowel distances are represented by the black lines in Figure 2.

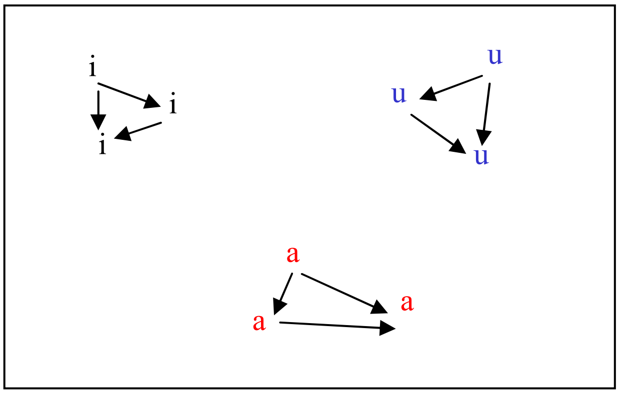

Figure 2: Intra-Euclidean distances are calculated between each vowel token to every other token within the same vowel category. The Euclidean distances to be calculated are represented by the black lines.

Accepted after full paper review

2.4.3 Statistical Models

Welch two sample t-tests will be reported for the analysis of inter-vowel and intra-vowel Euclidean distances. Due to the large number of tokens in the database, the possibility of Type I errors is increased and therefore a conservative significance level of $p<.01$ will be adopted (p-values between .01 and .05 will be interpreted as a trend).

3 Results

3.1 Inter-vowel D (distance between all vowel tokens)

Welch Two Sample t-tests were calculated on the Euclidean distance between all vowel tokens. To control for an inflated Type I Error rate, Bonferroni adjusted planned t-tests will use a .01 familywise error rate. Therefore, p-values less than .0017 will indicate significant mean differences. Mean and standard deviations of the Inter-vowel distances are given in Table 3. Vowel plots are given in Figure 3 and Figure 4.

Table 3: Mean (S.D.) Inter-vowel D by Stimuli (Word/Nonword) and EH (Easy/Hard) type.

	Easy	Hard
Word	4.35 (1.99)	4.58 (2.14)
Nonword	4.56 (2.08)	4.22 (1.86)

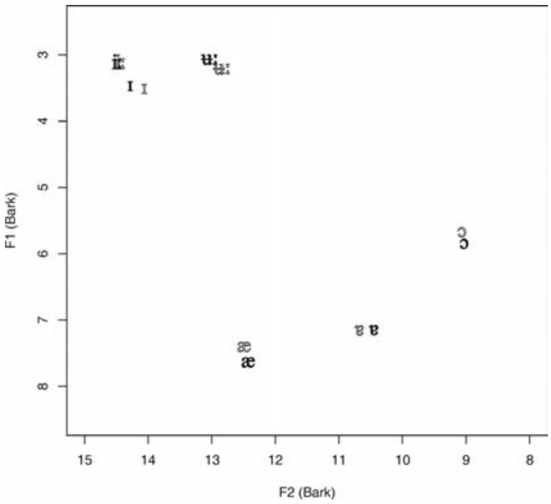

Figure 3: A F1 by F2 vowel plot of Hard Words (dark symbols) and Easy Words (light symbols) with F1 on the vertical axis and F2 on the horizontal axis.

A t-test of the main effect of Stimuli type on the Inter-vowel D revealed the vowel space of Easy items was significantly less dispersed than the space of Hard items (t(1637310) = -20.96, p < .0017). A t-test of the main effect of EH type on the Inter-vowel D showed the vowel space of Words was significantly more dispersed than the F1/F2 space of Nonwords (t(1751384) = 20.95, p < .0017).

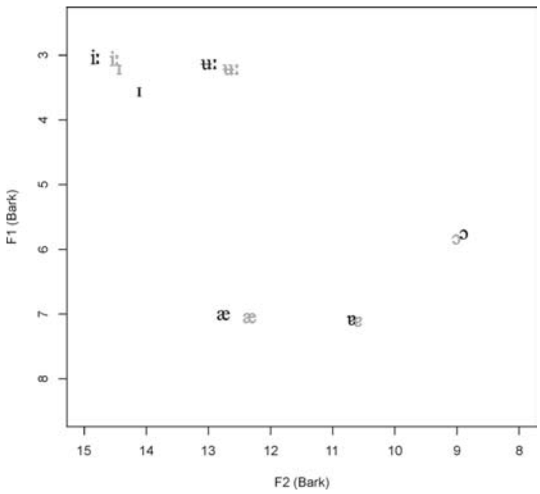

Figure 4: A F1 by F2 vowel plot of Hard Nonwords (dark symbols) and Easy Nonwords (light symbols) with F1 on the vertical axis and F2 on the horizontal axis.

The t-test between Easy and Hard Words indicated the vowel space of Hard Words was significantly larger or more dispersed than for Easy Words (t(494601.2) = -40.15, p < .0017). The t-test between Easy and Hard Nonwords revealed Easy Nonwords showed more vowel space expansion than Hard Nonwords (t(302605.1) = 51.08, p < .0017). This is opposite to the effect found for Easy and Hard Words.

Testing for differences between Easy Words and Easy Nonwords showed the vowel space of Easy Nonwords was more dispersed than for Easy Words (t(320716.9) = -30.73, p < .0017). A comparison between Hard Words and Hard Nonwords found more dispersion in Hard Words (t(520513.5) = 64.92, p < .0017).

3.2 Intra-vowel D (distance between tokens for each vowel type)

The results for the intra-vowel distances do not reveal a consistent pattern of variability across the vowel types. Means and standard deviations for each vowel (/æ, ɪ, iː, ɔ, ʉː, ɐ/) are given in Tables 4, 5, 6, 7, 8 and 9. The Intra-vowel D results are summarised in this section and significant differences are reported if p < .0025.

No difference in Intra-vowel distances were seen between Easy and Hard Words for /æ, ʉː/. For /ɪ, ɔ, ɐ/, Easy Words had greater intra-vowel distances than Hard Words, whereas for /iː/ Hard words were more dispersed.

Accepted after full paper review

Easy and Hard Nonwords were not significantly different in intra-vowel dispersion for /ɪ, ɔ/. Easy Nonwords were more dispersed than Hard Nonwords for /ʉː, ɐ/, whereas Hard Nonwords were significantly more dispersed than Easy Nonwords for /æ, iː/.

Easy Words and Easy Nonwords did not differ in Intra-vowel distances for /æ, ɐ/. Easy Words were significantly more dispersed than Easy Nonwords for /ɪ, iː, ɔ/, whereas Easy Nonwords were more dispersed than Easy Words for /ʉː/. For /iː/, Hard Nonwords and Hard Words were not differently dispersed. Hard Nonwords were more dispersed for /æ, ɔ, ʉː, ɐ/ than Hard Words, whereas Hard Words were significantly more dispersed than Hard Nonwords for /ɪ/.

Table 4: Mean (S.D.) of Intra-vowel D by Stimuli (Word/Nonword) and EH (Easy/Hard) for /æ/.

	Easy	Hard
Word	0.82 (0.44)	0.79 (0.43)
Nonword	0.83 (0.47)	1.17 (0.68)

Table 5: Mean (S.D.) of Intra-vowel D by Stimuli (Word/Nonword) and EH (Easy/Hard) for /ɪ/

	Easy	Hard
Word	1.04 (0.91)	0.98 (0.86)
Nonword	0.69 (0.41)	0.70 (0.40)

Table 6: Mean (S.D.) of Intra-vowel D by Stimuli (Word/Nonword) and EH (Easy/Hard) for /iː/

	Easy	Hard
Word	0.73 (0.46)	0.77 (0.59)
Nonword	0.66 (0.37)	0.79 (0.53)

Table 7: Mean (S.D.) of Intra-vowel D by Stimuli (Word/Nonword) and EH (Easy/Hard) for /ɔ/

	Easy	Hard
Word	1.34 (0.80)	1.06 (0.57)
Nonword	1.23 (0.71)	1.24 (0.77)

Table 8: Mean (S.D.) of Intra-vowel D by Stimuli (Word/Nonword) and EH (Easy/Hard) for /uː/

	Easy	Hard
Word	0.72 (0.39)	0.71 (0.36)
Nonword	1.46 (1.20)	0.87 (0.56)

Table 9: Mean (S.D.) of Intra-vowel D by Stimuli (Word/Nonword) and EH (Easy/Hard) for /ɐ/

	Easy	Hard
Word	1.18 (0.72)	1.01 (0.58)
Nonword	1.15 (0.63)	1.06 (0.70)

4 Discussion

This experiment found vowel production was influenced by lexical frequency and phonological neighbourhood density. Hard words, low frequency and dense phonological neighbourhoods, were produced with more expanded vowel spaces than Easy words that are high frequency with sparse phonological neighbourhoods. This finding in conjunction with other research (Munson & Solomon, in press; Wright, 2003) suggest speakers adjust the degree of hyper-articulation to compensate for factors that reduce the intelligibility of their speech. This result is consistent with Lindblom's (1990) hyper- and hypo-speech theory which proposes that speech is a constant balance between listener and speaker constraints. It appears speakers may actively modify their articulations in different tasks or environments to maintain an adequate level of intelligibility. Additionally, this evidence lends support to exemplar or usage based models of phonology (Bybee, 2001; Pierrehumbert, 2002), where individual words may each have a unique entry in the lexicon which contains specific phonetic characteristics. These unique phonetic specifications may lead to productions which maximise their comprehension.

The vowel space of nonwords was less dispersed than for words. This result might fit within a listener-based hypothesis because nonwords do not carry sensible linguistic information. On the other hand, it could be suggested that nonwords might need extra articulatory effect to distinguish the segments because they are unknown items. An exemplar framework may be able to account for this result as nonwords would not have existing exemplars for reference. Therefore, the vowels used in nonwords might be an average of many other exemplars and nonwords may be produced with a reduced vowel.

The pattern of intra-vowel variability was not consistent across the vowel types. Some vowels showed no difference in the intra-vowel distances between Easy and Hard Words, others showed greater distances for Easy words and one vowel showed greater distances in Hard words. A similar range of effects was seen for Easy and Hard nonwords. This variability on intra-vowel dispersion suggests the pattern of vowel space

Accepted after full paper review

expansion may need in future to be investigated separately for each vowel type.

Considering that the effects of vowel space expansion on speech intelligibility have been well established (e.g., Bradlow et al., 1996), one would predict that the expanded vowel spaces found in Hard words would facilitate their perception. In contrast, research has showed that these words are harder to perceive than Easy words (e.g., Luce & Pisoni, 1998). Future research could consider whether the increase in vowel space expansion seen for Hard words actually produces a difference in perception accuracy. Moreover, future research should examine whether the influence of neighbourhood density on speech extends to other aspects of production, such as consonant articulation.

Another possibility is that these lexical competition effects are due to the influence of phonetic context. This stimuli set and those used previously to investigate vowel space expansion (Wright, 2003) have found effects that are just an artefact of phonetic context effects. Unfortunately, this remains a limitation of this type of study because it is extremely difficult (perhaps impossible) to match stimuli lists on lexical frequency, phonological neighbourhood density and phonetic context. Care must be taken regarding the generalisation of these results because they are susceptible to contextual effects rather than lexical competition effects. Future research must use carefully matched stimuli because of the many effects that phonetic context produces in speech, for example vowel lengthening before voiced obstruents than voiceless ones (Chen, 1970). Additionally, syllable-initial /l/'s are longer and darker when the syllable-final consonant is voiced (Hawkins & Nguyen, 2003). Therefore, future research must investigate lexical competition effects with stimuli that aim to be maximally matched for phonetic context.

Finally, as called for by Munson and Solomon (in press) future research should use a range of experimental paradigms to further explore whether these lexical frequency and neighbourhood density effects are caused by active attempts to maximize speech clarity or are the consequence of how different exemplars are encoded.

5 Acknowledgements

The author thanks Jonathan Harrington, Marija Tabain and Sallyanne Palethorpe for their comments on drafts of this paper, and Chris Callaghan for technical support.

6 References

Baayen, R., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database (CD-ROM). Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Bates, T. C., & D'Oliveiro, L. (2003). PsyScript: A Macintosh application for scripting experiments. *Behavior Research Methods, Instruments, and Computers, 35*(4), 565-576.

Bradlow, A. R., Toretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication, 20*, 255-272.

Bybee, J. (2001). *Phonology and language use.* Cambridge: Cambridge University Press.

Cassidy, S., & Harrington, J. (2001). Multi-level annotation in the EMU speech database management system. *Speech Communication, 33*, 61-77.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica, 22*, 129-159.

Hawkins, S., & Nguyen, N. (2003). Effects on word recognition of syllable-onset cues to syllable-coda voicing. In J. Local, R. Ogden & R. Temple (Eds.), *Phonetic Interpretation: Papers in Laboratory Phonology VI* (pp. 38-57). Cambridge: Cambridge University Press.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modelling* (pp. 403-439). Dordrecht: Kluwer Academic Publishers.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: Neighborhood activation model. *Ear and Hearing, 19*, 1-36.

Moon, S.-J., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America, 96*, 40-55.

Munson, B., & Solomon, N. P. (in press). The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language and Hearing Research*.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing: II. Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research, 29*, 434-446.

Pierrehumbert, J. (2002). Word-specific phonetics. In N. Warner & C. Gussenhoven (Eds.), *Laboratory Phonology VII* (pp. 101-139). New York: Walter De Gruyter.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology, 55A*, 1339-1362.

Umeda, N. (1975). Vowel duration in American English. *Journal of the Acoustical Society of America, 58*(2), 434-445.

Wright, R. (2003). Factors of lexical competition in vowel articulation. In J. Local, R. Ogden & R. Temple (Eds.), *Phonetic Interpretation: Papers in Laboratory Phonology VI* (pp. 75-87). Cambridge: Cambridge University Press.