Acoustic correlates of intelligibility in native and non-native speakers[1]

Jennifer Haywood

University of Washington

jhaywood@uw.edu

Introduction

Intelligibility, or "[the] extent to which a speaker's message is actually understood by a listener" (Munro and Derwing 1995), has long been a concern of researchers interested in non-native speech. The research in this area has largely been focused on deviations from native speaker norms or errors. Several studies (Anderson-Hsieh, Johnson, and Koehler 1992; Bent, Bradlow, and Smith 2007; Rogers 1997) have attempted to quantify the relative weight of different types of errors in determining intelligibility. Meanwhile, researchers focusing on English spoken by native speakers (both normal and disordered speech) have investigated several acoustic-phonetic factors that are thought to affect intelligibility in the speech of those talkers, such as slower speech rate, expanded pitch range, and expanded vowel space (Bond and Moore 1994; Bradow, Torretta, and Pisoni 1997; Neel 2008). Native speakers also exploit these factors when they are intentionally speaking clearly (Picheny, Durlach, and Braida 1986). The results of the intelligibility research have not been entirely consistent. This may be partly due to differences in methodology and measurements; for example, various measures of vowel space have been used. The present study represents an attempt to address these issues.

Research Questions and Experimental Design

The research questions under investigation in the present study are:

[1] Funded by NIH grant #10186254: Acoustic and Perceptual Effects of WDRC Compression.

1. Do acoustic factors such as pitch range and vowel space reliably correlate with intelligibility in non-native English speech?
2. Which measures of vowel space are most appropriate for measuring this factor in native and non-native speakers of English?

To answer these questions, the study was devised. First, a corpus was created of equivalent sentences spoken by native (Pacific Northwest English L1) and non-native (Mandarin L1) speakers of English. The corpus is fully-crossed for L1, gender, and sentence. Then, the stimuli were presented in noise to Pacific Northwest English listeners. The listeners repeated sentences as they heard them, and these repetitions were scored by accuracy of key words. Various acoustic measures were taken of the corpus, and these measures were statistically tested for correlation with the intelligibility scores.

Methods

Speaking Task

Talkers from two specific language backgrounds were recorded. Talkers in the MC group were native speakers of Mandarin from mainland China. They were all advanced speakers of English who use English on a daily basis in their work at the University of Washington, where they were researchers, administrative workers, and graduate students. Talkers in the NW group were native speakers of Pacific Northwest English. This dialect area was considered to include the states of Washington, Oregon, and Idaho, although all of the NW talkers used in the present study were from Washington. The NW talkers were all students at the University of Washington. Talkers were paid for their participation.

Each speaker was recorded reading a selection of the “Harvard sentences” (IEEE 1969). MC speakers read 124 sentences, each containing five key words, and NW speakers, who were recorded as part of a larger study, read 200 sentences, including the 124 read by the MC speakers. These sentences are designed to be phonetically balanced and include all phonemes of General American English. Most of these sentences are rather unnatural (e.g. *“A pot of tea helps to pass the evening.”).* This has benefits and drawbacks. They are less semantically predictable, so they are good for intelligibility testing. This may also cause them to be read less naturally, but this was seen as a necessary drawback.

Three randomizations of all sentences were presented via PsyScope (Cohen et al 1993) to each speaker. Speakers read the sentences and were recorded via a head-mounted microphone in a sound-attenuated booth.

Speakers in the MC group were given the sentences ahead of time so that they could practice them if they wished, and the researcher was available to them for consultation and correction while the recordings were being made.

Stimulus Preparation

From each group of speakers (MC and NW), the recordings of three males and three females were prepared for presentation to listeners. The most fluent, natural, error-free utterances of 120 of the sentences were selected. (See Appendix A for the list of these sentences.) The utterances were extracted by hand and equalized for RMS intensity. Then the utterances were mixed with corpus-based speech-shaped noise at 0dB and 6dB SNR.

Intelligibility Task

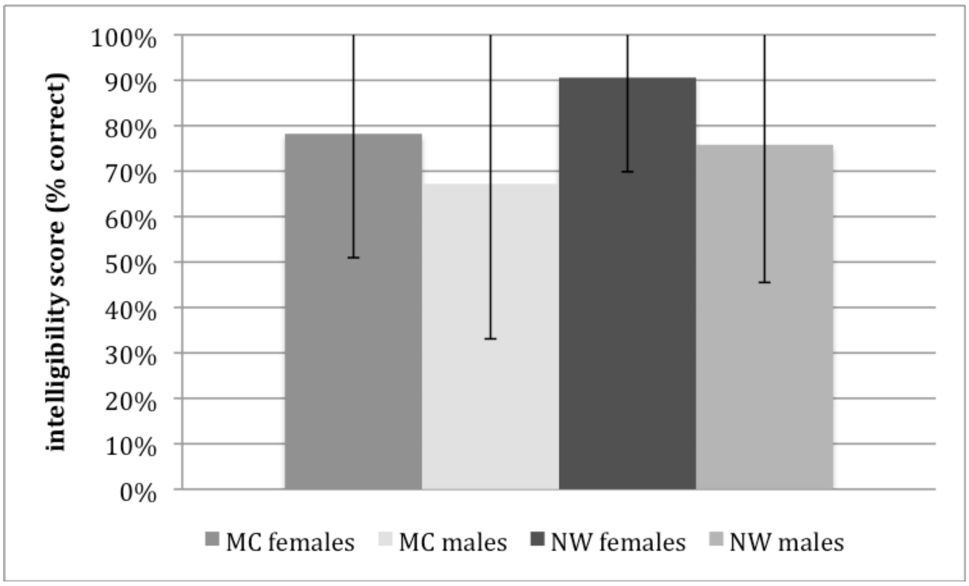
Listening

Stimuli were presented to twenty-four listeners, all university students in the Seattle area, and all native speakers of PNW English who had never studied or had significant exposure to a Chinese language. The listeners were paid for their participation in the study. The listeners all received a hearing screening prior to the study. After a training section, they heard a pseudo-randomized selection of the sentences, ten from each of the twelve speakers, half at 0dB SNR and half at 6dB SNR. Each listener heard one each of the entire set of 120 sentences, and each of the 2440 utterances in the stimulus set was heard by two listeners.

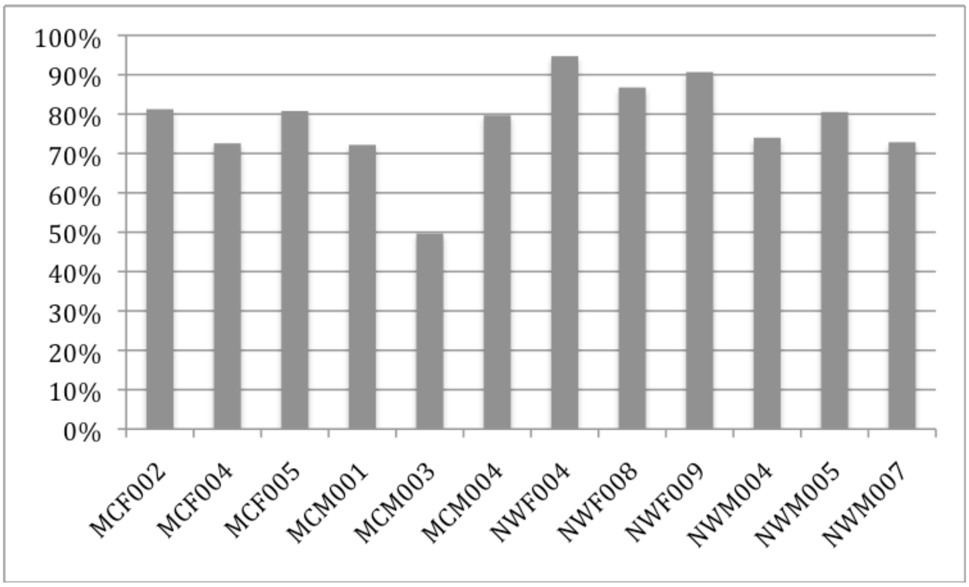
Stimuli were presented to the listeners in a randomized order via PsyScope, through closed supraaural headphones in a sound-attenuated booth. Listeners heard each sentence one time and then repeated it. Sentence repetitions were recorded and later scored for accuracy based on correctness of the first four key words in each sentence.

Scoring

Group intelligibility means, by language background and gender, are shown in Figure 1. As expected, paired t-tests show that the non-native English speakers (MC group) were significantly less intelligible than the native English speakers (NW group) (t=9.63, p<.0001). Also, as has been found in other studies (c.f. Bradlow *et al* 1997), female talkers were significantly more intelligible than male talkers (t=11.96, p<.0001).



What was less expected is that the non-native females were actually slightly more intelligible than the native male talkers, with mean group intelligibility scores of 78% correct for MC females and 76% for NW males. Figure 2 shows the mean intelligibility score for each talker. There is a great deal of variability within each talker group, but two of the MC females achieved higher mean intelligibility scores than two of the NW males.



Acoustic measures can perhaps provide a partial explanation for this unexpected result. Descriptions of the acoustic measures are presented in the next section, followed by results of statistical correlation tests in the section following.

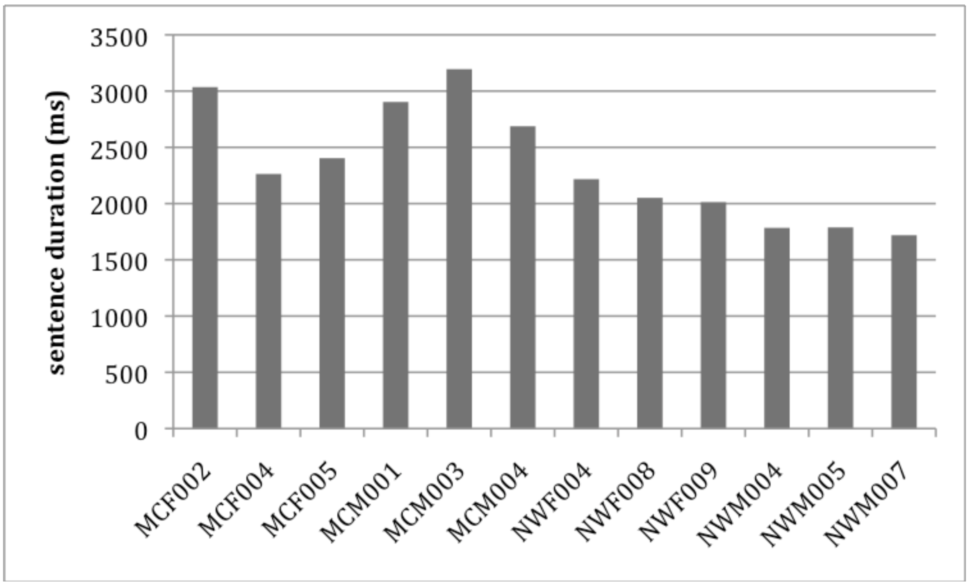
Acoustic Measurements

Speech rate

Two measures of speech rate were undertaken: sentence duration and vowel duration. Since all speakers produced the same sentences, total sentence duration was used as a rough estimate of speech rate, following Bradow, Torretta, and Pisoni (1996). The results are shown in Figure 3, as mean sentence duration for each individual speaker.

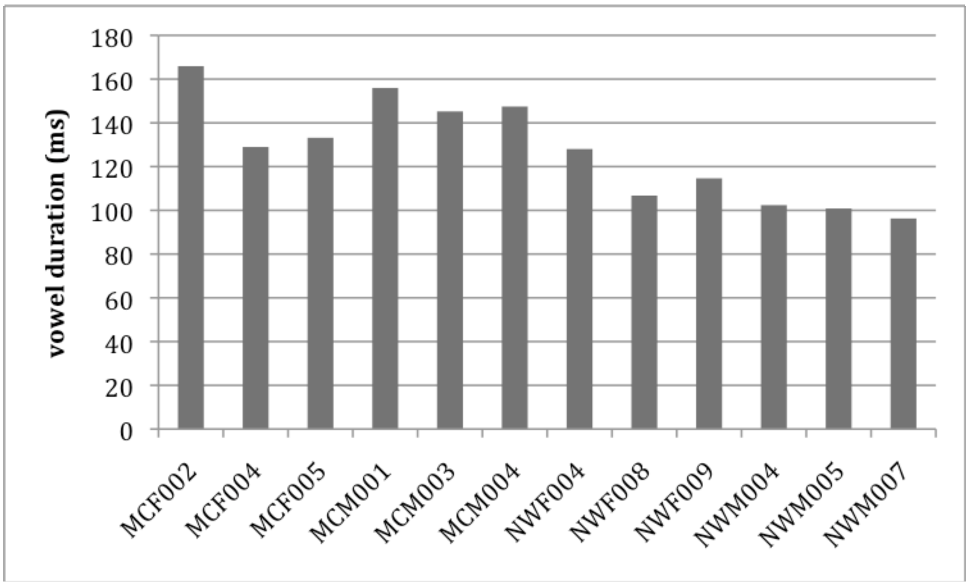
Previous studies of native speakers have found that a slower speech rate is correlated with intelligibility, and here we do find, for example, that the most intelligible NW speaker (and most intelligible speaker overall) has the greatest mean sentence duration of the NW group. Also, the NW female talkers were more intelligible than the NW males, and they also spoke more slowly.

Among the MC talkers, the most intelligible talker does have one of the slowest speech rates, but the longest mean sentence duration of any individual talker is that of the least intelligible talker, MCM003. This can probably be explained by the fact that he may be less fluent, as well as less intelligible, than the other non-native speakers.



Duration of the stressed vowels used for vowel space measures (explained below) was also measured via Praat script, as an estimate of the speech rate of the key words in the sentences. The results are shown in Figure 4, as mean stressed vowel duration for each individual speaker.

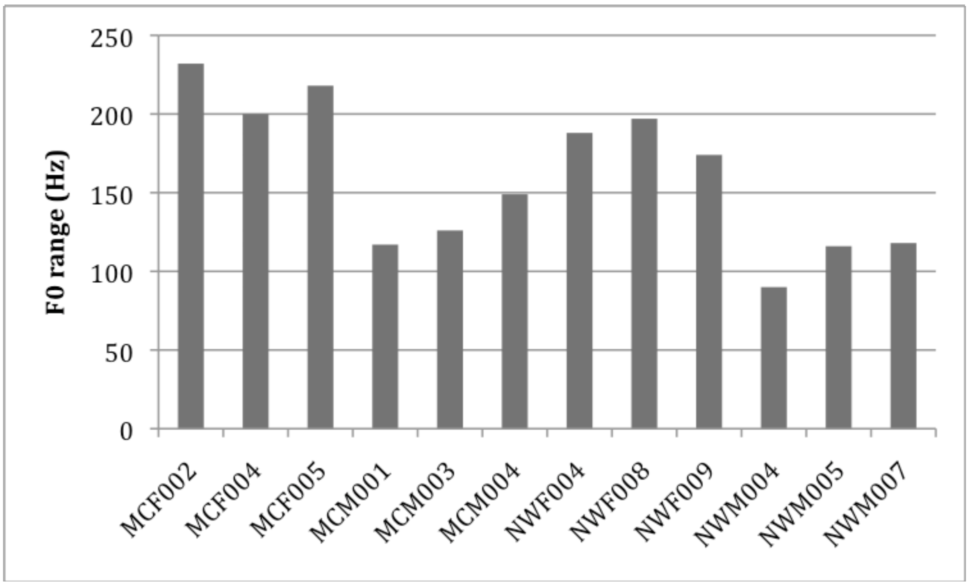
Longer vowel durations were expected to correlate with higher intelligibility. NW females used longer vowel durations than males, and were more intelligible. Also, the most and least intelligible females in the NW and MC groups had the longest and shortest vowel durations, respectively, in each of those groups. However, the results for the males are not clear from speaker means alone.



Pitch range

The pitch range for each sentence in the corpus was measured via a Praat script. F0 maxima and minima were extracted for each sentence, and the range recorded as the difference between the two in Hz. The results are shown in Figure 5, as mean F0 range for each individual speaker.

It was expected that a larger pitch range would correlate with higher intelligibility. Female talkers from each L1 group used larger F0 range than their male counterparts and were also more intelligible, as expected. And in the MC female group, the most and least intelligible talkers used the largest and smallest pitch range, respectively. In the other groups, the results are not clear from speaker means.



Vowel space measures

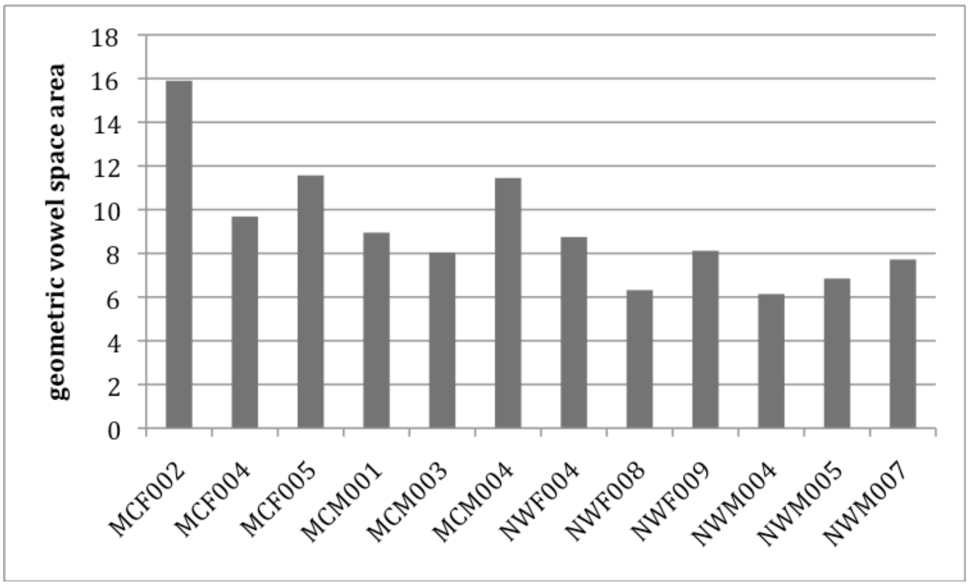
A sample of vowels was measured for each talker. Vowels were marked by hand on the waveform, with reference to spectrogram, using Praat (Boersma 2001). A Praat script was used to extract first and second formant values at the midpoint of each vowel, and these values were hand-corrected. F1 and F2 values were converted to Bark, which is perceptually motivated, but were not normalized in any other way, as other normalization methods result in undesirable alterations of the vowel space.

The following measures of the vowel space were calculated for each talker: vowel space area, vowel dispersion, repulsive force, F1 range, F2 range, /i/ F2-F1, and /ɑ/ F2-F1.

Vowel space is generally measured using only three or four point vowels; for example, Bradlow, Torretta, and Pisoni (1996) used the point vowels /i, a, o/, and Neel (2008) used the vowels /i, æ, ɑ,u/. However, point vowels are problematic in that they vary depending on dialect. For example, in the case of the Neel (2008) study, the Hillenbrand (1995) corpus on which the study was based included mostly speakers having a raised /æ/ relative to other North American English speakers (Wright and Souza 2012), making this a less appropriate point vowel for those speakers, less representative of their overall working vowel space.

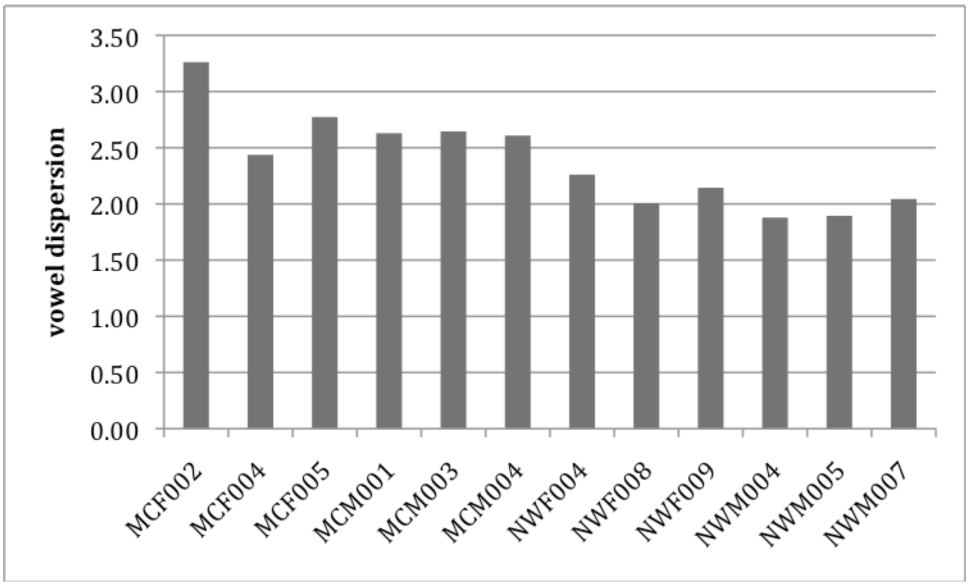
Because of these considerations, for my calculations of vowel space I used /i, ɪ, e, ɛ, æ, ɑ, o, ʊ, u/--all the monophthongal vowels measured with the exception of /ʌ/. Using these vowels to sketch the perimeter, the geometric area of the resulting polygon was calculating, giving a more accurate measure of overall working vowel space. The results are shown in Figure 6, as geometric vowel space area for each individual speaker.

It was expected that a larger vowel space would correlate with higher intelligibility. This looks true from these results for gender within each L1, and for individual speakers within gender-L1 groups with the exception of the NW males. Of course, what was not expected is that the non-native speakers would have larger geometric vowel space areas than the native speakers, but this may start to explain the relatively greater intelligibility of the MC females compared to the NW males.



Dispersion of vowels within the vowel space was arrived at, following Neel (2008), by calculating the Euclidean distance between each pair of the ten-vowel category means for a given speaker. The results are shown in Figure 7, as vowel dispersion for each individual speaker.

It was expected that greater dispersion would correlate with higher intelligibility. The results are similar to those for geometric vowel space area, except that dispersion appears to be slightly less representative (based on mean intelligibility scores alone) for the MC male speakers than the geometric vowel space measure.



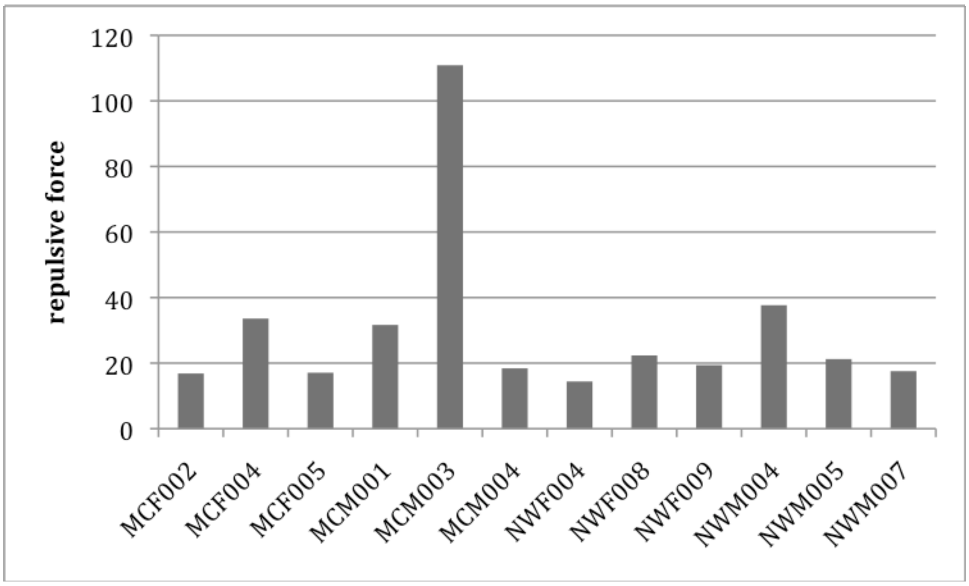
The alternative measure of repulsive force was also used to measure the relative compression of each speaker's vowel space. Liljencrants and Lindblom (1972) developed the calculation of repulsive force, which can be used to measure the extent to which maximal perceptual contrast, and presumably intelligibility, are present in a vowel system. This formula has also been used by Wright (1998) as an alternate measure of "compactness of the vowel space". In this formula, below in (1), *i* and *j* represent all pairs of vowels in the system.

(1)
$$E = \sum_{i=1}^{n-1} \sum_{j=0}^{i-1} 1 / r^2 ij$$

where $r^2 ij$ is $\sqrt{(\mathrm{F1_i} - \mathrm{F1_j})^2 + (\mathrm{F2_i} - \mathrm{F2_j})^2}$

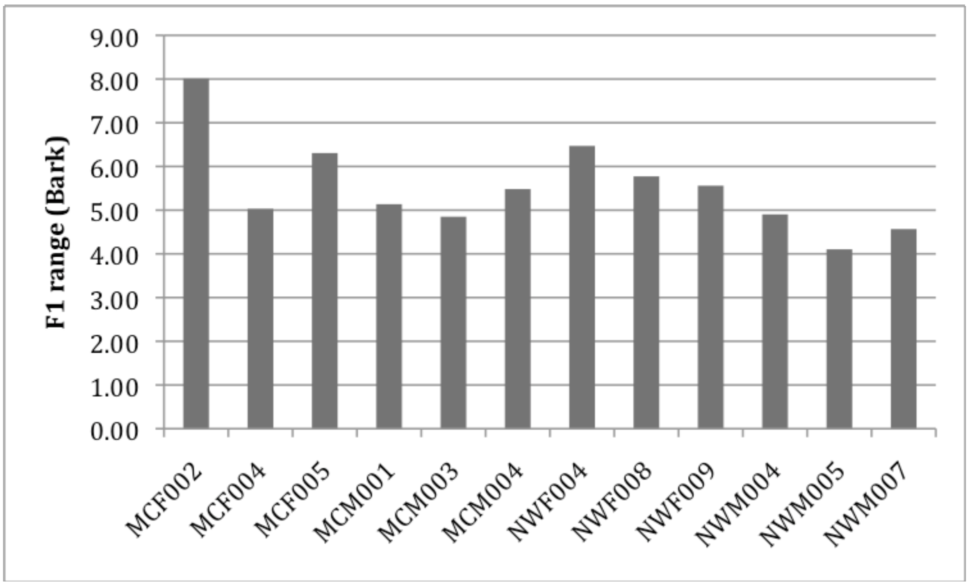
With this repulsive force measure, closer items produce greater repulsive force, so a system in which the vowels are grouped closely together will have a greater force than a system in which the vowels are more spread out. Thus it was expected that a lower repulsive force score would correlate with higher intelligibility.

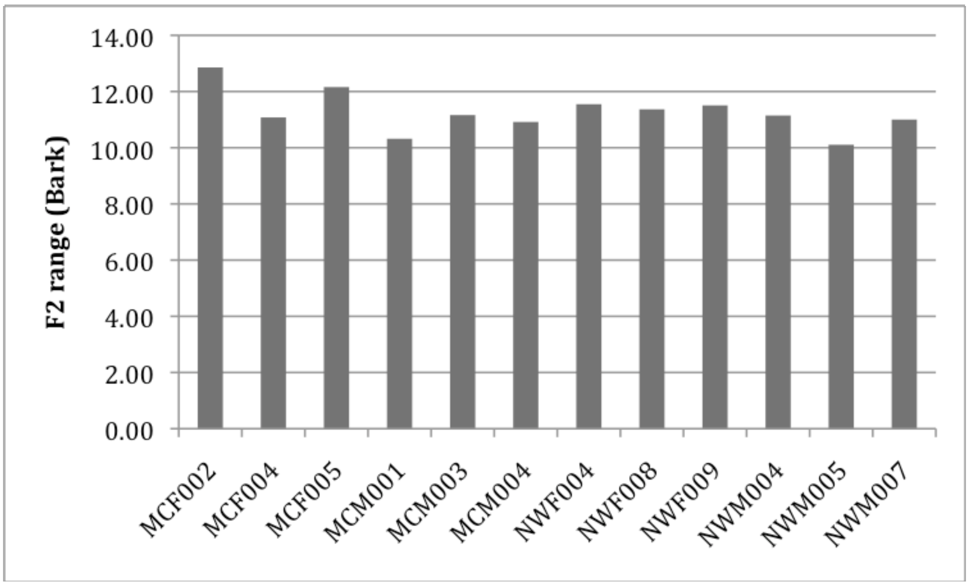
The results are shown in Figure 8, as repulsive force based on the vowel category means for each individual speaker. What is most striking is the very high repulsive force score for speaker MCM003, who was also much less intelligible than any of the other speakers.



F1 range and F2 range were also calculated for each speaker, to give an idea of the importance of the height dimension and front-back vowel dimension in intelligibility. It was expected that both greater F1 range and greater F2 range would correlate with increased intelligibility.

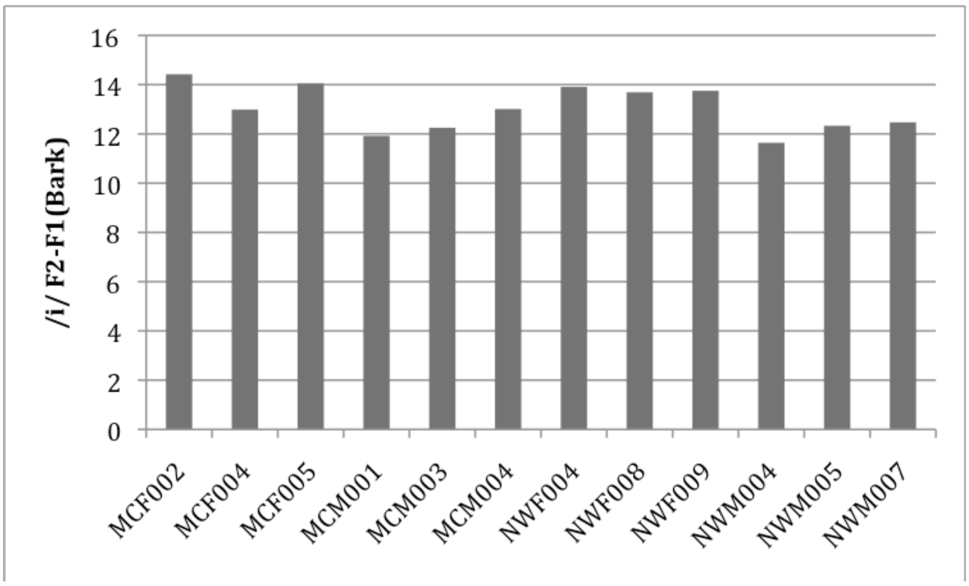
The results are shown in Figures 9 and 10, as F1 range and F2 range (in Hz) based on the vowel category means for each individual speaker.





The final measures of vowel space, found by Bradlow, Torretta, and Pisoni (1996) to be reliably correlated with intelligibility in their corpus, are designed to estimate the extremity of the edges of the vowel space. The first, /i/ F2-F1, indicates a more peripheral location for /i/ the greater the distance between the two formant values, and the other, /ɑ/ F2-F1, indicates a more peripheral location for /ɑ/ the smaller the distance between the two formant values. Thus, it was expected that greater /i/ F2-F1 values and smaller /ɑ/ F2-F1 values would correlate with higher intelligibility.

The results are shown in Figures 11 and 12, as /i/ F2-F1 values and /ɑ/ F2-F1 values for each individual speaker.



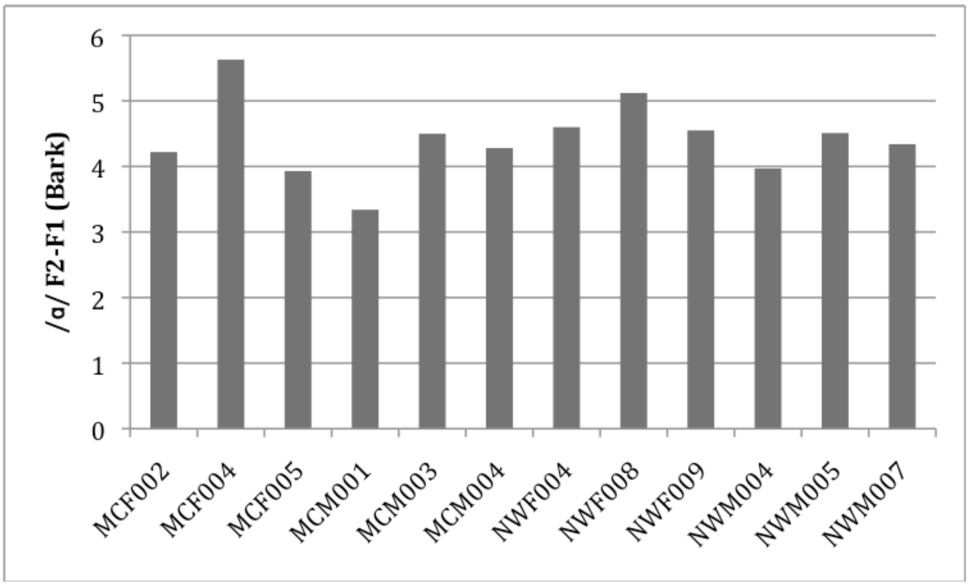


Table 1 presents a summary of the acoustic measures for each of the speakers in the corpus.

Table 1. Summary of acoustic measures for each of the 12 talkers. The acoustic measures are mean sentence duration, mean vowel duration, mean F0 range, geometric vowel space area, vowel dispersion, repulsive force, F1 range, F2 range, /i/ F2-F1, and /ɑ/ F2-F1.

	mean sentence duration (ms)	mean vowel duration (ms)	mean F0 range (Hz)	geometric vowel space area (Bark^2)	vowel disper-sion (Bark)	repulsive force	F1 range (Bark)	F2 range (Bark)	/i/
MCF002	3035	166	232	15.9	3.26	17	8.01	12.86	
MCF004	2263	129	200	9.69	2.44	34	5.03	11.08	
MCF005	2404	133	218	11.57	2.77	17	6.31	12.16	
MCM001	2904	156	117	8.95	2.63	32	5.14	10.32	
MCM003	3195	145	126	8.04	2.65	111	4.85	11.16	
MCM004	2688	147	149	11.45	2.61	18	5.48	10.92	
NWF004	2218	128	188	8.75	2.26	14	6.47	11.55	
NWF008	2051	107	197	6.32	2.01	22	5.77	11.37	
NWF009	2013	115	174	8.12	2.14	19	5.56	11.51	
NWM004	1784	102	90	6.14	1.88	38	4.90	11.15	
NWM005	1788	101	116	6.85	1.89	21	4.10	10.11	
NWM007	1719	96	118	7.72	2.04	18	4.57	11.00	

Results of statistical tests

Spearman’s correlations were run for the acoustic measures against sentence intelligibility scores for all speakers, for just the non-native speakers (MC L1 group), and for just the non-native speakers. Correlations shown to be significant at p<.0001 are shown with Spearman’s rho values in Table 1.

Table 1. Rho Values for Spearman's correlation tests between acoustic measures and intelligibility scores

	Spearman's Rho Values, p<.0001		
	all speakers	**MC speakers**	**NW speakers**
sentence duration	-.136	-.170	.307
vowel duration	-.078	not significant	.301
F0 range	.140	.204	.258
vowel space	not significant	.287	.183
vowel dispersion	-.105	not significant	.210
repulsive force	-.236	-.288	.110
F1 range	.223	.288	.254
F2 range	.130	not significant	.266
/i/ F2-F1	.228	.270	.225
/ɑ/ F2-F1	-.078	.264	-.139

Ranked by the strength of the correlation with intelligibility, the results for each group are as follows: for all speakers, force > /i/ F2-F1 > F1 range > F0 range > sentence duration > F2 range > /ɑ/ F2-F1; for Mandarin L1 speakers, force and F1 range > vowel space > /i/ F2-F1 > /ɑ/ F2-F1 > F0 range > sentence duration; for NW English speakers, sentence duration > vowel duration > F2 range > F0 range > F1 range > /i/ F2-F1 > dispersion > vowel space > force.

Discussion

As expected, all the acoustic measures tested correlated reliably with the intelligibility scores for the native English speakers. However, the repulsive force measure showed a positive, rather than the expected negative, correlation. This may be because repulsive force was calculated using vowel category means, rather than individual vowel tokens. Recalculation using individual vowel tokens may result in the expected negative correlation. The acoustic factors with the strongest correlation to intelligibility in this group related to speech rate, with faster talkers being less intelligible. Of the vowel space measures, F2 range, followed by F1 range, showed the strongest correlation with intelligibility.

For the non-native speakers, most of the acoustic measures correlated reliably, in the expected direction, with intelligibility. The exceptions are F2 range, vowel dispersion, and vowel duration, which showed no reliable correlation with intelligibility, and sentence duration, which showed a negative correlation rather than the expected positive one. This last result is probably because the most intelligible non-native speakers were probably also the most fluent ones, thus the fastest talkers. Whereas native speakers who spoke more slowly were also speaking more clearly, this was not the case for the non-native speakers. This may also have been the reason for a lack of correlation with vowel duration. As for F2 range, Mandarin has a relatively rich set of front-back vowel distinctions, whereas English has more height distinctions, so it may be that only F1 range makes a difference in English intelligibility for native speakers of Mandarin.

Future research plans for this corpus include:

- refining speech rate measures by using syllables per second rather than sentence duration;
- combining stressed vowel and sentence duration measures for a stressed vowel duration ratio measure and also determining stressed to unstressed vowel duration ratios;
- refining pitch range measures by using semi-tones; refining the repulsive force measure by calculating it based on individual vowel tokens rather than vowel category means;
- implementing measures of vowel-inherent spectral change, which were found by Neel (2008) to be relevant to vowel intelligibility.

Statistical methods also will continue to be refined.

Conclusions

Acoustic factors including expanded vowel space and pitch range did reliably correlate with intelligibility in these non-native (MC L1) speakers of English.

Although a slower rate of speech reliably correlates with *increased* intelligibility in the NW L1 speakers, it correlates with *decreased* intelligibility in this group of MC L1 speakers.

The measures of vowel space that correlate most reliably with intelligibility are different for the NW L1 speakers and MC L1 speakers: F2 range and F1 range for NW speakers; repulsive force, F1 range, and vowel space area for MC speakers.

Non-native speakers may need to use expanded vowel space compared to native speakers to compensate for other issues in their speech. It is clear that intelligibility results from a complex group of acoustic factors, and these can even override native vs. non-native speaker status in certain cases.

Fig. 1. Group intelligibility means, by language background and gender.

Fig. 2. Mean intelligibility score for each talker.

Fig. 3. Mean sentence duration for each individual speaker

Fig. 4. Mean stressed vowel duration for each individual speaker.

Fig. 5. Mean F0 range for each individual speaker.

Fig. 6. Geometric vowel space area for each individual speaker.

Fig. 7. Vowel dispersion for each individual speaker.

Fig. 8. Repulsive force based on the vowel category means for each individual speaker.

Fig. 9. F1 range (in Hz) based on the vowel category means for each individual speaker.

Fig. 10. F2 range (in Hz) based on the vowel category means for each individual speaker.

Fig. 11. /i/ F2-F1 values for each individual speaker.

Fig. 12. /ɑ/ F2-F1 values for each individual speaker.

Appendix A: Sentences in corpus

The corpus used in the study consists of the following sentences (from IEEE 1969) spoken by each speaker. Key words are shown in bold.

LIST	NUMBER	SENTENCE
1	7	The **box** was **thrown beside** the **parked truck**.
2	6	A **pot** of **tea helps** to **pass** the **evening**.
3	2	The **fish twisted** and **turned** on the **bent hook**.
4	2	**Take** the **winding path** to **reach** the **lake**.
5	2	The **ship** was **torn apart** on the **sharp reef**.
5	5	The **lazy cow lay** in the **cool grass**.
5	7	The **rope** will **bind** the **seven books** at **once**.
6	1	The **frosty air passed through** the **coat**.
6	5	A **saw** is a **tool used** for **making boards**.
6	9	**Place** a **rosebush near** the **porch steps**.
6	10	**Both lost** their **lives** in the **raging storm**.
7	2	**Use** a **pencil** to **write** the **first draft**.
7	8	**This** is a **grand season** for **hikes** on the **road**.
7	10	**Those words** were the **cue** for the **actor** to **leave**.
8	4	The **walled town** was **seized without** a **fight**.
8	7	The **horn** of the **car woke** the **sleeping cop**.

8	8	The **heart beat strongly** and with **firm strokes**.
10	10	The **bill** was **paid every third week**.
11	5	**Add** the **sum** to the **product** of **these three**.
11	7	The **ripe taste** of **cheese improves** with **age**.
12	2	**Leaves turn brown** and **yellow** in the **fall**.
13	3	The **boss ran** the **show** with a **watchful eye**.
13	7	It **caught** its **hind paw** in a **rusty trap**.
13	9	**Feel** the **heat** of the **weak dying flame**.
14	4	**Two plus seven** is **less** than **ten**.
15	7	The **tree top waved** in a **graceful way**.
15	9	**Mud** was **spattered** on the **front** of his **white shirt**.
15	10	The **cigar burned** a **hole** in the **desk top**.
16	1	The **empty flask stood** on the **tin tray**.
16	2	A **speedy man** can **beat** this **track mark**.
16	10	The **sofa cushion** is **red** and of **light weight**.
17	5	An **abrupt start** does **not win** the **prize**.
17	6	**Wood** is **best** for **making toys** and **blocks**.
18	2	The **child almost hurt** the **small dog**.
18	4	The **sky** that **morning** was **clear** and **bright blue**.
18	5	**Torn scraps littered** the **stone floor**.
19	6	**Add** the **column** and **put** the **sum here**.
20	4	The **paper box** is **full** of **thumb tacks**.
22	2	The **loss** of the **second ship** was **hard** to **take**.
22	3	The **fly made** its **way along** the **wall**.
22	6	The **large house** had **hot water taps**.
22	7	It is **hard** to **erase blue** or **red ink**.
22	9	The **doorknob** was **made** of **bright clean brass**.
23	1	A **pencil** with **black lead writes best**.
23	7	The **shaky barn fell** with a **loud crash**.
24	2	They are **pushed back each time** they **attack**.
24	7	**Some ads serve** to **cheat buyers**.
24	9	A **waxed floor makes** us **lose balance**.
25	1	On the **islands** the **sea breeze** is **soft** and **mild**.
25	2	The **play began** as **soon** as we **sat down**.
25	4	**Add salt before** you **fry** the **egg**.
27	1	The **dark pot hung** in the **front closet**.
27	6	The **rude laugh filled** the **empty room**.
27	7	**High seats** are **best** for **football fans**.
27	9	A **dash** of **pepper spoils beef stew**.
30	4	**Watch** the **log float** in the **wide river**.
30	6	The **heap** of **fallen leaves** was **set** on **fire**.
31	10	The **fight** will **end** in **just six minutes**.
32	9	The **purple tie** was **ten years old**.
33	1	**Fill** the **ink jar** with **sticky glue**.
33	5	The **crunch** of **feet** in the **snow** was the **only sound**.
33	8	The **plush chair leaned against** the **wall**.
34	2	The **beach** is **dry** and **shallow** at **low tide**.
34	5	**Pages bound** in **cloth make** a **book**.
34	6	**Try** to **trace** the **fine lines** of the **painting**.
34	7	**Women form less** than **half** of the **group**.
34	8	The **zones merge** in the **central part** of **town**.
36	10	It **takes** a **good trap** to **capture** a **bear**.
37	1	**Feed** the **white mouse** some **flower seeds**.
37	5	**Plead** to the **council** to **free** the **poor thief**.

38	2	**Mark** the **spot** with a **sign painted red**.
40	9	**Raise** the **sail** and **steer** the **ship northward**.
40	10	A **cone costs five cents** on **Mondays**.
41	8	The **sense** of **smell** is **better** than **that** of **touch**.
42	9	The **point** of the **steel pen** was **bent** and **twisted**.
43	2	**Draw** the **chart** with **heavy black lines**.
45	2	The **bloom** of the **rose lasts** a **few days**.
45	5	**Bottles hold four kinds** of **rum**.
45	8	**Drop** the **ashes** on the **worn old rug**.
46	2	The **couch cover** and **hall drapes** were **blue**.
46	5	The **clothes dried** on a **thin wooden rack**.
46	10	The **price** is **fair** for a **good antique clock**.
47	4	He **sent** the **figs**, but **kept** the **ripe cherries**.
48	1	The **kite flew wildly** in the **high wind**.
48	3	The **tin box held priceless stones**.
49	5	He **offered proof** in the **form** of a **large chart**.
50	1	A **man** in a **blue sweater sat** at the **desk**.
50	9	The **work** of the **tailor** is **seen** on **each side**.
51	4	The **dusty bench stood** by the **stone wall**.
52	10	The **beetle droned** in the **hot June sun**.
53	9	His **wide grin earned many friends**.
55	1	**Those last words** were a **strong statement**.
55	2	He **wrote** his **name boldly** at the **top** of the **sheet**.
56	2	**Clams** are **small**, **round**, **soft**, and **tasty**.
56	7	A **brown leather bag hung** from its **strap**.
56	8	A **toad** and a **frog** are **hard** to **tell apart**.
56	10	A **break** in the **dam almost caused** a **flood**.
58	3	It was **hidden** from **sight** by a **mass** of **leaves** and **shrubs**.
58	8	The **lobes** of her **ears** were **pierced** to **hold rings**.
59	5	**Keep** the **hatch tight** and the **watch constant**.
60	3	**Slide** the **tray across** the **glass top**.
60	9	**Get** the **trust fund** to the **bank early**.
62	6	The **early phase** of **life moves fast**.
64	5	**Crouch before** you **jump** or **miss** the **mark**.
64	6	**Pack** the **kits** and **don't forget** the **salt**.
65	3	They **sang** the **same tunes** at **each party**.
65	8	**Pile** the **coal high** in the **shed corner**.
65	9	A **gold vase** is **both rare** and **costly**.
66	1	The **rarest spice comes** from the **far East**.
66	5	The **aim** of the **contest** is to **raise** a **great fund**.
67	5	**Pick** a **card** and **slip** it **under** the **pack**.
67	6	A **round mat** will **cover** the **dull spot**.
67	9	The **mail comes** in **three batches** per **day**.
67	10	You **cannot brew tea** in a **cold pot**.
68	2	**Put** the **chart** on the **mantel** and **tack** it **down**.
68	8	We **don't like** to **admit** our **small faults**.
70	1	The **store** was **jammed before** the **sale** could **start**.
70	2	It was a **bad error** on the **part** of the **new judge**.
70	4	**Take** the **match** and **strike** it **against** your **shoe**.
70	6	The **baby puts** his **right foot** in his **mouth**.
70	9	The **streets** are **narrow** and **full** of **sharp turns**.
71	5	The **big red apple fell** to the **ground**.
71	10	The **corner store** was **robbed last night**.
72	10	**When** you **hear** the **bell**, **come quickly**.

References

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The Relationship Between Native Speaker Judgments of Nonnative Pronunciation and Deviance in Segmentals, Prosody, and Syllable Structure. *Language learning.*, *42*(4), 529–555.

Bent, T., Bradlow, A.R., & Smith, B. (2007). Phonemic errors in different word positions and their effects on intelligibility of non-native speech: All's well that begins well. In O.-S. Bohn & M. J. Murray J. (Eds.), Language experience in second language speech learning: In honor of James Emil Flege (pp. 331-348). Amsterdam: John Benjamins.

Boersma, Paul (2001). Praat, a system for doing phonetics by computer. *Glot International* **5:9/10**, 341-345.

Bond, Z. S., & Moore, T. J. (1994). A note on the acoustic-phonetic characteristics of inadvertently clear speech. *Speech Communication*, *14*(4), 325–337. doi:10.1016/0167-6393(94)90026-4

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

Cohen J.D., MacWhinney B., Flatt M., and Provost J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. Behavioral Research Methods, Instruments, and Computers, 25(2), 257-271.

Hillenbrand J, Getty LA, Clark MJ, & Wheeler K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, *97*(5), 3099–111.

IEEE. (1969) IEEE Recommended practice for speech quality measurements, IEEE Report No. 297.

Liljencrants, J. & Lindblom, B. (1972) Numerical simulation of vowel quality systems: the role of perceptual contrast, *Language*, **48** (4), 839-862.

Munro, M. J. & T. M. Derwing (1995). Foreign accent, intelligibility and comprehensibility in the speech of second language learners. *Language Learning* **45**, 73–97**.**

Neel, A. (2008) Vowel Space Characteristics and Vowel Identification Accuracy, *Journal of Speech, Language, and Hearing Research*, **51**, 574-585.

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** (4), 434–446.

Rogers, C.L. (1997). Intelligibility of Chinese-accented English. Unpublished doctoral dissertation, Indiana University, Bloomington.

Wright, R. (1998) *Factors of lexical competition in vowel articulation*, Labphon 6.

Wright R, & Souza P. (2012). Comparing identification of standardized and regionally valid vowels. *J. Speech Lang. Hear. Res. Journal of Speech, Language, and Hearing Research*, *55*(1), 182–193.