COMPREHENSIBILITY AND THE ACOUSTIC CONTRAST BETWEEN TENSE AND LAX VOWELS IN MANDARIN-ACCENTED ENGLISH SPEECH

Chien-Min Kuo, Purdue University
Olga Dmitrieva, Purdue University

This study investigates the relation between the comprehensibility of Mandarin-accented English speech and the degree of acoustic contrast in the vowel tenseness pairs [i] - [ɪ] and [u] - [ʊ]. We hypothesized that Mandarin speakers under-differentiate the tenseness vowel pairs and that variability in the acoustic contrast between tense and lax vowels correlates with variability in comprehensibility. Twenty Mandarin speakers and ten American English speakers produced sentences with the target vowels, and 26 English speakers rated the sentences' comprehensibility. We found that spectral difference between [u] - [ʊ] was relatively reduced in L2 speech and that L2 productions were overall more variable, compared to American English speech. We also found that the degree of spectral contrast correlated with comprehensibility rating, suggesting that the degree of acoustic differentiation between phonological categories contributes to the comprehensibility of L2 speech.

INTRODUCTION

Comprehensibility of speech, rather than its phonetic native likeness, has been gaining support as the primary goal of second language (L2) pronunciation teaching and learning (Derwing & Munro, 1997, 2005; Flege et al., 1995; Munro & Derwing, 1995). Comprehensibility refers to the degree of effort needed to understand an utterance, and in L2 pronunciation research it is typically measured using listeners' subjective scalar judgments (Derwing & Munro, 1997, 2005; Munro & Derwing, 1995).

Previous studies explored a number of factors that may influence the comprehensibility of L2 English speech, including: (1) grammatical accuracy and complexity (Derwing & Munro, 1997; Saito et al., 2017; Trofimovich & Isaacs, 2012); (2) lexical variety, accuracy, and complexity (Saito et al., 2017; Trofimovich & Isaacs, 2012); (3) articulation rate and other indices of fluency (Derwing et al., 1998; Derwing & Munro, 1997; Saito et al., 2017); (4) properties of rhythm, intonation and word stress (Derwing et al., 1998; Derwing & Munro, 1997; Isaacs & Thomson, 2020; Saito et al., 2017); and (5) segmental errors (Derwing et al., 1998; Derwing & Munro, 1997; Isaacs & Thomson, 2020; Saito et al., 2017). Despite correlations between impressionistically rated segmental errors and comprehensibility in previous studies (Derwing et al., 1998; Derwing & Munro, 1997; Isaacs & Thomson, 2020; Saito et al., 2017), the relationship between the instrumentally measured acoustic realization of segments and comprehensibility has not been widely investigated. While acoustic-phonetic native-likeness of L2 speech is not necessarily a strong predictor of comprehensibility, as accentedness and compressibility were shown to be relatively independent parameters (Derwing & Munro, 1997, 2005; Munro & Derwing, 1995), we propose that the degree of acoustic distinction between phonologically contrastive categories may play a role in determining comprehensibility. In the present study, we evaluate the degree of acoustic contrasts between English tense and lax vowels produced by L2 speakers and explore the

Kuo, C-M. & Dmitrieva, O. (2024). Comprehensibility and the Acoustic Contrast Between Tense and Lax Vowels in Mandarin-Accented English Speech. In D. J. Olson, J. L. Sturm, O. Dmitrieva, & J. M. Levis (eds), *Proceedings of 14th Pronunciation in Second Language Learning and Teaching Conference*, (pp. 1-13). Purdue University, September 2023. https://doi.org/10.31274/psllt.17569

relationship between acoustic contrast and comprehensibility.

More specifically, the study focuses on Mandarin speakers' production of the two tenseness pairs in English: [i] - [ɪ] and [u] - [ʊ]. These targets were chosen due to a mismatch between the vowel systems of the two languages: unlike English, Mandarin lacks the phonological tenseness contrast, as Mandarin has [i] and [u] but not [ɪ] and [ʊ] (Lee & Zee, 2003). Such mismatches between phonological inventories of the L1 and L2 are predicted to generate both perception and production difficulties for L2 learners. For example, the Perceptual Assimilation Model (PAM) (Best, 1995; Best & Tyler, 2007) proposes that pairs of L2 vowels that are acoustically similar to the same L1 vowel category are likely to both be perceptually assimilated to that L1 vowel - a single-category assimilation scenario. In our case, both English [i] and [ɪ] are predicted to assimilate to Mandarin /i/, while [u] and [ʊ] assimilate to Mandarin /u/. The Speech Learning Model (SLM) (Flege, 1995; Flege & Bohn, 2021) describes such cases as subject to equivalence classification, whereby English [i] and [ɪ] are categorized by Mandarin learners as equivalent to Mandarin /i/. Importantly, the SLM states that failure to notice acoustic differences between similar L1 and L2 vowels (such as, between Mandarin [i] and English [ɪ]) means that no such differences would be realized in production. Consequently, L2 speakers may not produce sufficient acoustic difference between English [i] and [ɪ]. In support of this prediction, previous studies have found that Mandarin speakers produced less acoustic contrast between [i] - [ɪ] and [u] - [ʊ], compared to English speakers (Chen, 2006; Li et al., 2020). Yet implementing a sufficient acoustic difference between phonologically contrastive vowels is important due to the functional load such these vowels carry in distinguishing words in minimal pairs, such as *beat* - *bit* and *pool* - *pull*. Building on previous findings, we predict that the variability in the degree of acoustic contrast produced by Mandarin speakers will correspond to variability in perceived comprehensibility of their utterances containing relevant vowels.

Context can play a role in moderating the effect of acoustic realization due to the top-down effects in utterance processing (Davis & Johnsrude, 2007). To take this factor into consideration, our experiment incorporated two types of contexts: carrier phrases, which contained only one member of the minimal pair in a semantically neutral context, and contrastive sentences, which contained both members of the minimal pair in a semantically meaningful context. We hypothesized that vowel acoustics may play a lesser role in determining the comprehensibility of contrastive sentences due to the facilitative effect of context. In addition, it is possible that speakers may be subject to pragmatic pressures to realize greater acoustic differences between the vowels in contrastive sentences, given that the members of the minimal pair appear in a semantic juxtaposition to each other.

We evaluated vowel acoustics by collecting spectral and durational measurements, corresponding to the primary and secondary dimensions of distinction between English tense and lax vowels. Tense vowels in English are more peripheral in the vowel space and longer in duration (Hillenbrand et al., 1995; Hillenbrand et al., 2000). We used these measurements to calculate the degree of acoustic contrast and investigate the relationship between contrast and comprehensibility, measured as scalar judgments by human raters, to address the following research questions: (1) Does the tenseness contrast produced by Mandarin speakers differ acoustically from the contrast produced by native English speakers? Hypothesis: We expect to find less contrast in L2 speakers' production, especially in the spectral dimension and in semantically neutral contexts. (2) Does the degree of spectral and durational contrast correlate with perceived comprehensibility? Hypothesis: Greater durational and spectral contrast between the tense and lax vowel is expected to predict higher utterance

comprehensibility, especially in the absence of semantically meaningful context.

METHODS

Participants

Twenty (14 female, 6 male) native speakers of Mandarin (M_{age} = 24.7, *SD* = 3.6) with average US residency length of 34.5 months (*SD* = 30.5) and average self-rated English proficiency (participants reported their self-rated proficiency for each of the language they spoke on a scale from 0 to 10: 0 = none, 10 = perfect) of 7.0 (*SD* = 1.1) from Taiwan and Mainland China and 10 (9 female, 1 male) native speakers of American English (M_{age} = 28.0, *SD* = 14.5) participated in the sentence production task. Five of the Mandarin speakers reported intermediate proficiency in Taiwanese Southern Min (*M* = 4.5). Knowledge of additional languages in both groups was limited to low proficiency, with few exceptions (self-reported using the same 11-point scale).

Twenty-six (16 female, 10 male) native speakers of American English (M_{age} = 26.9, *SD* = 9.7), university students or employees, participated in the comprehensibility rating task. They were born, raised, and educated in the US. Self-reported proficiency in languages other than English was on average intermediate or below and did not include Mandarin: French (*N* = 1, *M* = 4.3), German (*N* = 1, *M* = 3.3), Hindi (*N* = 1, *M* = 1.7), Japanese (*N* = 1, *M* = 6.3), Korean (*N* = 1, *M* = 5.0), and Spanish (*N* = 9, *M* = 4.5).

Stimuli

Targets in the sentence production tasks were CVC words forming minimal pairs for vowel tenseness, e.g., *feet* [fi̲t], *fit* [fi̲t], *cooed* [kʰu̲d], and *could* [kʰʊ̲d]. Target words were incorporated into two sentence types: carrier sentences (*I like to say _____ some of the time)* and contrastive sentences. Each contrastive sentence had a clause containing a tense-vowel word and a clause containing a lax-vowel word, e.g., *The desks are eight feet long, so they can fit into the space.* There were four carrier sentences for each of the four vowels, and four contrastive sentences for each of the two vowel pairs, for a total of 24. For every sentence with a target word there was a corresponding distractor sentence with structurally comparable words without target vowels, for a total of 48 sentences. Participants read each sentence twice.

One of the two productions of each sentence was randomly selected for the comprehensibility rating task and normalized for intensity to 70 dB, avoiding sentences with errors. Around 3% of data was discarded due to mispronunciation unrelated to the tenseness of target vowels. Each rater was assigned a total of 48 sentences: 24 targets and 24 distractors. To ensure that each rater was exposed to a sufficiently representative set of productions, sentences were assigned following these criteria: (1) Each rated a unique combination of four different speakers. (2) Each speaker was assigned to four different raters. (3) Carrier sentences were assigned in pairs: a tense vowel sentence and its lax vowel counterpart produced by the same speaker. (4) Each rater evaluated both front vowel and back vowel sentences.

Procedure

Participants read sentences aloud from a screen in randomized order and were audio-recorded

in a sound-attenuated booth using Shure KSM32 microphone. Stimuli were presented using the experiment builder Gorilla (Anwyl-Irvine et al., 2020). Figure 1 shows the trial structures. Each participant was compensated with an $8 Amazon eGift Card.

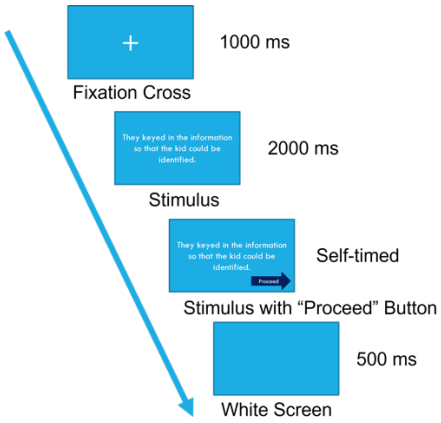

Figure 1. Diagram of a trial in the sentence production task

The comprehensibility task was conducted online through Gorilla. The participants listened to sentences in randomized order, typed what they heard, and answered the question 'How easy was it to understand the sentence?' using a mouse-controlled slider labelled *'very easy'* and *'very hard'* on the opposite ends. A continuous slider was designed to be divisible into 100 segments and was used instead of a discrete numeric scale to capture a greater variability in ratings and to obtain a more continuous measurement for analysis (see also Saito et al., 2017). The trial structure is shown in Figure 2. The order of the transcription task and comprehensibility rating task was counterbalanced across participants. The transcription task was included as a measure of intelligibility, which is not reported in this paper. Each participant was compensated with a $10 Amazon eGift Card.

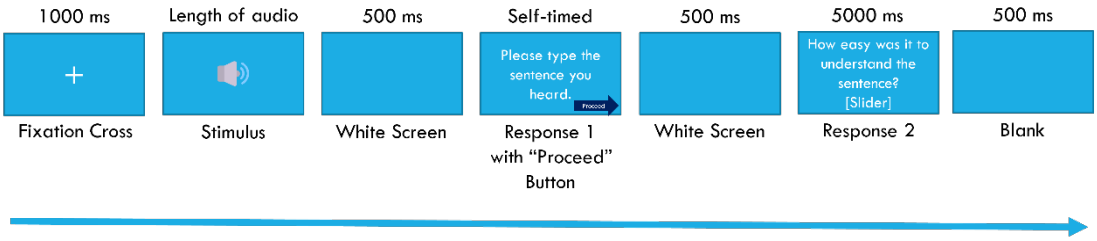

Figure 2. Diagram of a trial in the comprehensibility rating task

Analysis

One repetition of each sentence was selected for acoustic analysis. Vowels were segmented manually using Praat 6.3.18 (Boersma & Weenink, 2023). Vowel duration and formant values (F1 and F2) at vowel midpoint were measured using Praat scripts. Formant measurements were normalized using the R package *phonTools* (Barreda, 2023). Spectral contrast was calculated for each sentence pair as the Euclidean distance between the tense vowel and the lax vowel based on the formant values. Durational contrast was calculated for each sentence pair as tense minus lax vowel duration. A linear-mixed effect model was conducted in R, using the package *lmerTest* (Kuznetsova et al., 2017), with Spectral Contrast or Durational Contrast as the dependent variables, L1 (English and Mandarin), Sentence Type (carrier and contrastive), Vowel Pair (front and back), L1 × Sentence Type, and L1 × Vowel Pair

interactions as independent variables, and Participant as the random effect. The two models were conducted to determine whether the degree of tenseness contrast differed between Mandarin speakers and English speakers.

Comprehensibility ratings were converted to a score from 0 to 100, based on Saito et al. (2017). For contrastive sentences, the rating each sentence received was correlated with the tenseness contrast between the two vowels produced in each sentence. For the carrier sentences, which contained one vowel per sentence, the rating each tense-vowel sentence received was averaged with the rating of its lax-vowel counterpart and the resulting number was correlated with the tenseness contrast between the vowels in the two sentences. Since comprehensibility ratings and acoustic contrast data were not normally distributed according to the Shapiro-Wilk, nonparametric correlation tests using Kendall's tau were conducted in R (R Core Team, 2023).

RESULTS

Production of Spectral Contrast

As a group, Mandarin speakers produced less F1 (height) difference between the back vowels and less F2 (backness) difference between the front vowels, than American speakers (Figure 3). However, the Euclidean distance was visibly affected only for back vowels (Figure 4). Euclidean distance was considerably more variable in Mandarin than in American speakers' productions (Figure 4).

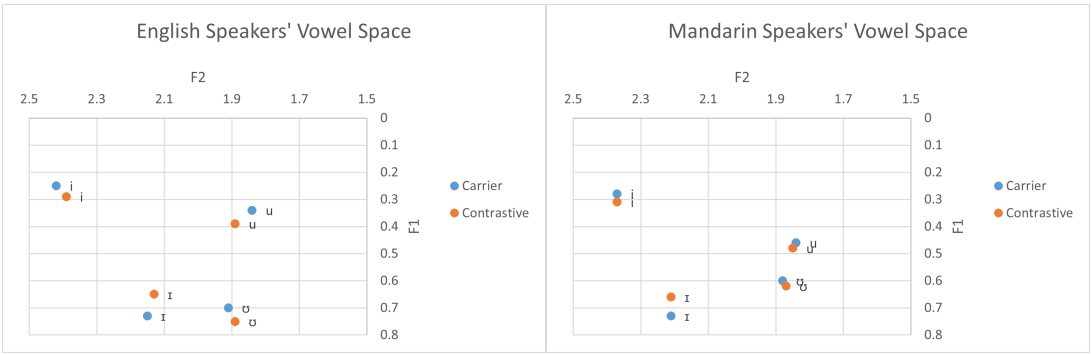

Figure 3. Average vowel space by Speaker Group

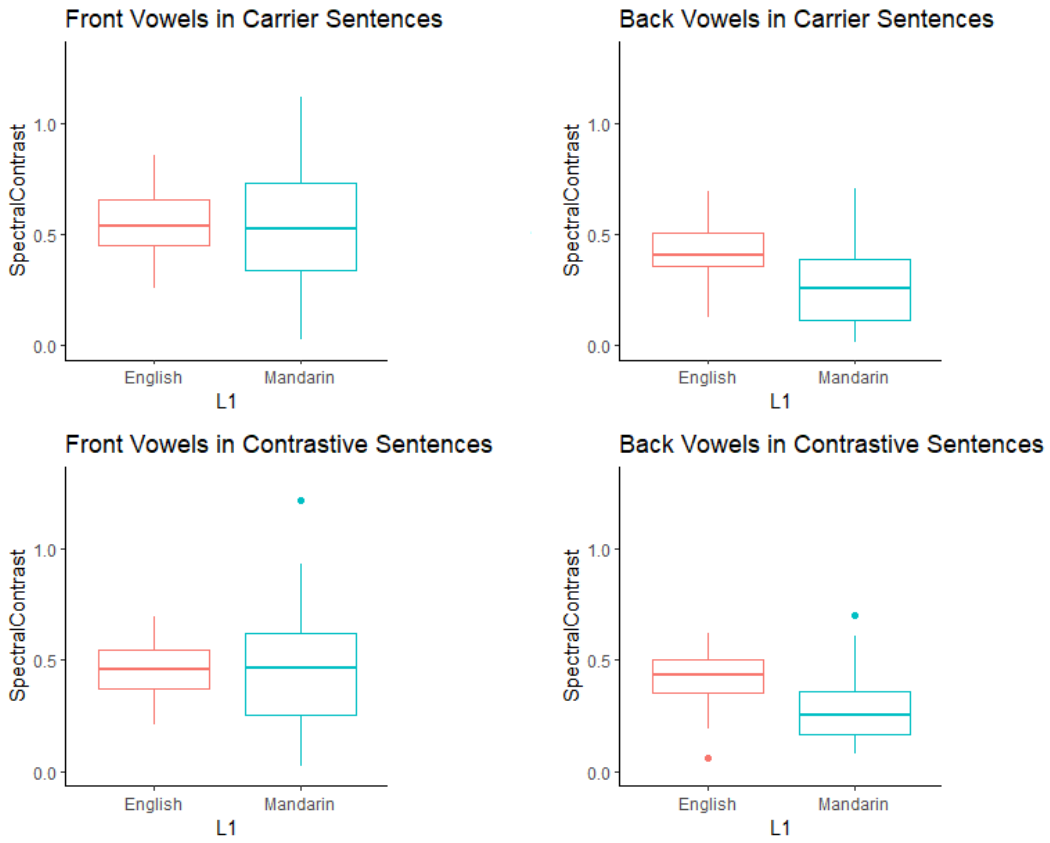

Figure 4. Spectral Contrast by Speaker Group, Sentence Type and Vowel Pair. Each bar represents the average Euclidean distance between the two contrasting vowels produced by the speaker group.

The linear mixed-effect model (conditional $R^2 = 0.46$, marginal $R^2 = 0.18$) indicated English speakers' front vowel spectral contrasts in carrier sentences differed significantly from their back vowel spectral contrasts in carrier sentences, and English speakers' front vowel spectral contrasts in carrier sentences differed significantly from Mandarin speakers' back vowel spectral contrasts in carrier sentences (Table 1). Post-hoc tests using the Tukey method showed Mandarin speakers produced a significantly smaller spectral distance between back vowels than English speakers ($\beta = 0.14$, $SE = 0.05$, $p < 0.05$).

Table 1

Summary of the Linear-Mixed Effect Model with the Dependent Variable Spectral Contrast

	Estimate	***SE***	***df***	***t* value**	***p***	
(Intercept)	0.53	0.04	41.48	12.14	<0.001	***
L1 - Mandarin	-0.03	0.05	41.64	-0.57	0.574	
Sentence Type - Contrastive	-0.05	0.03	422.99	-1.80	0.073	
Vowel Pair - [u] - [ʊ]	-0.09	0.03	422.98	-3.43	<0.001	***
L1 - Mandarin : Sentence Type - Contrastive	0.02	0.03	423.06	0.56	0.576	
L1 - Mandarin : Vowel Pair - [u] - [ʊ]	-0.11	0.03	423.11	-3.51	<0.001	***

Note. The intercept refers to: English (L1), Carrier (Sentence Type), [i] - [ɪ] (Vowel Pair).

Production of Durational Contrast

With a possible exception of back vowels in carrier sentences, Mandarin speakers and American speakers produced a visually comparable degree of durational contrasts between tense and lax vowels, as shown in Figure 5. The figure also suggests a tendency for greater variability in durational contrast produced by Mandarin speakers in comparison to American speakers, although not as consistently as for the spectral contrast.

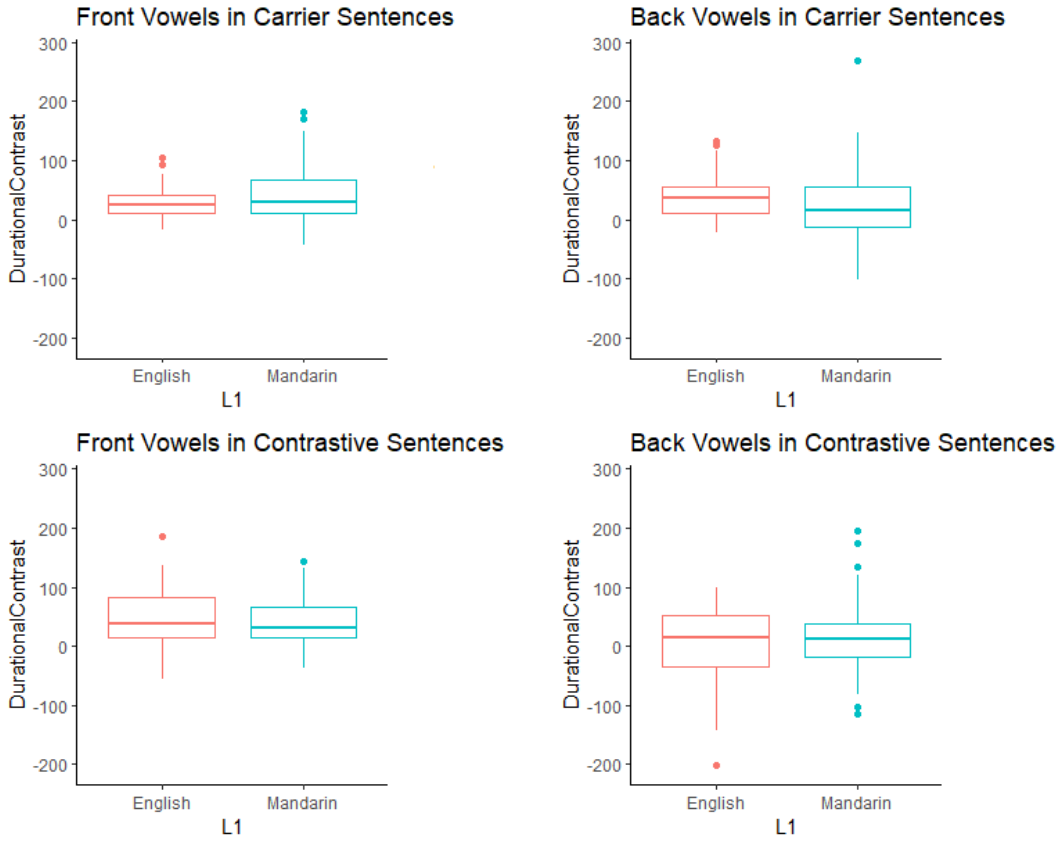

Figure 5. Durational contrast by Speaker Group, Sentence Type and Vowel Pair

The linear mixed-effect model (conditional $R^2 = 0.07$, marginal $R^2 = 0.05$) indicated English speakers' front vowel durational contrasts in carrier sentences differed significantly from their back vowel durational contrasts in carrier sentences (Table 2). Post-hoc tests using the Tukey method showed all speakers produced a significantly smaller durational contrast for back vowels than front vowels ($\beta = 20.5$, $SE = 5.03$, $p < 0.001$).

Table 2

Summary of the Linear-Mixed Effect Model with the Dependent Variable Durational Contrast

	Estimate	*SE*	*df*	*t* value	*p*	
(Intercept)	43.06	7.56	137.28	5.70	<0.001	***
L1 - Mandarin	2.07	9.29	138.10	0.22	0.824	
Sentence Type - Contrastive	-10.03	8.14	423.35	-1.23	0.218	
Vowel Pair - [u] - [ʊ]	-17.18	8.14	423.26	-2.11	<0.05	*
L1 - Mandarin : Sentence Type - Contrastive	2.24	10.06	423.68	0.22	0.824	
L1 - Mandarin : Vowel Pair - [u] - [ʊ]	-6.68	10.06	424.01	-0.66	0.507	

Note. The intercept refers to: English (L1), Carrier (Sentence Type), [i] - [ɪ] (Vowel Pair).

Correlation between Comprehensibility Rating and Spectral Contrast

The correlation test demonstrated a weak positive relationship between the Comprehensibility Rating and Spectral Contrast for the vowel pairs in both carrier sentences and contrastive sentences produced by Mandarin and American speakers (Table 3). The relationship is illustrated in Figure 6, which also shows a greater range of comprehensibility ratings in contrastive sentences (mean = 83.5, IQR = 54.5) than in carrier sentences (mean = 81, IQR = 29).

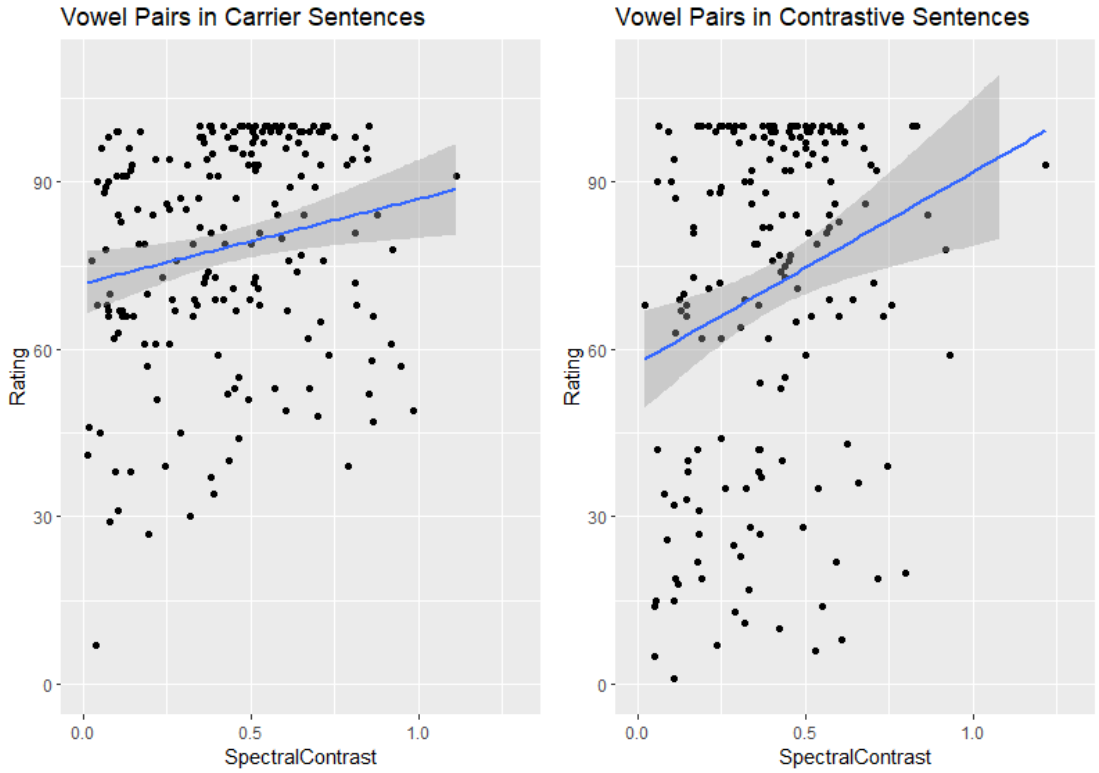

Figure 6. Comprehensibility Rating against Spectral Contrast by Sentence Type

Table 3

Summary of the Correlational Analysis for Comprehensibility Rating and Spectral Contrast

Model	Variables	*z*	*p*	*tau*	Effect size
Vowel pairs in carrier sentences	Comprehensibility Rating & Spectral Contrast	3.13	< 0.01**	0.153	0.023
Vowel pairs in contrastive sentences	Comprehensibility Rating & Spectral Contrast	3.11	<0.05*	0.157	0.025

Correlation between Comprehensibility Rating and Durational Contrast

No significant correlation was found between Comprehensibility Rating and Durational Contrast in either of the two contexts (Figure 7, Table 5).

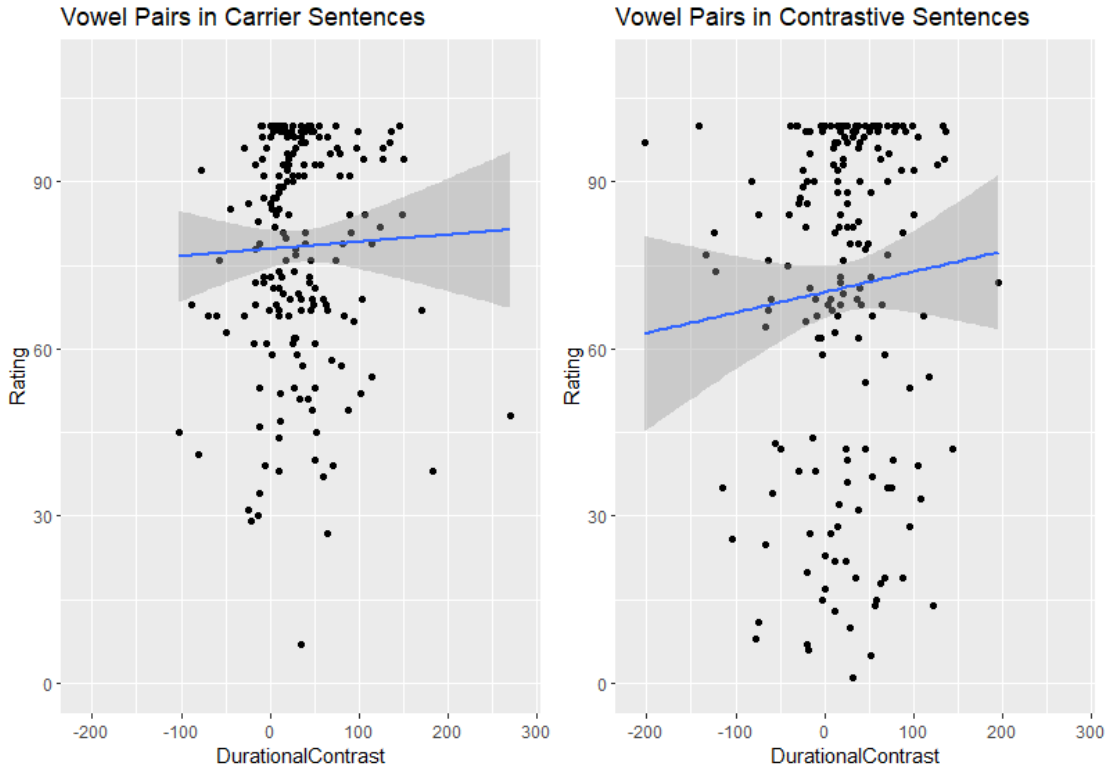

Figure 7. Comprehensibility Rating against Durational Contrast by Sentence Type

Table 4

Summary of the Correlational Analysis for Comprehensibility Rating and Durational Contrast

Model	Variable	*z*	*p*	tau	Effect size
Vowel pairs in carrier sentences	Comprehensibility Rating & Durational Contrast	0.48	0.634	0.023	0.0005
Vowel pairs in contrastive sentences	Comprehensibility Rating & Durational Contrast	1.59	0.111	0.081	0.007

DISCUSSION

The acoustic results indicated that, across the two groups of participants, front tense and lax vowels were more different from each other spectrally and in terms of duration, than back vowels. This effect was stronger for the spectral difference produced by Mandarin speakers – the only significant indication in the present data that the two groups differed from each other in terms of the degree of the acoustic tenseness contrast. Beyond this difference, speakers' L1 background did not significantly affect tenseness distinctions at a group level. Nevertheless, visualizations of the tenseness contrast suggested that there was considerably more variability in realizing the contrast in the Mandarin than in the American group.

The effect of sentence type on the degree of contrast was also not significant: tense and lax vowels were not differentiated more strongly in contrastive sentences than in carrier sentences. Thus, our first hypothesis was only partially supported: The tenseness contrast reduction in L2 speech was observed only for back vowels and was limited to the spectral dimension, while semantic context played no role.

These results seemingly go against the predictions of SLM (Flege, 1995; Flege & Bohn, 2021) and PAM-L2 (Best & Tyler, 2007) that second language speakers would under-differentiate certain L2 contrasts not present in L1 inventories. However, both theories also stipulate that greater L2 exposure and proficiency counteracts the effects of assimilation among L1 and L2 sounds, leading to more successful differentiation among L2 sounds both in production and perception. Mandarin speakers recorded for the present experiment were all residents of the USA with relatively high self-rated English proficiency (7.0 out of 10 on average), which could help explain the acoustic results. Future work should consider populations with a greater range of L2 experience and proficiency.

The asymmetry between front and back vowels is also not without explanation. The [i] – [ɪ] contrast in English has a much higher functional load that the [u] – [ʊ] contrast: front vowels distinguish a greater number of highly frequent words, while minimal pairs for the back contrast are scarce. The overall frequency of these phonemes in English confirms this observation: the frequency of [i] and [ɪ] is 3.69% and 3.64% respectively, while the frequency of [u] and [ʊ] is 1.13% and 0.76% respectively (Mines et al., 1978). The reduced degree of contrast between [u] and [ʊ] in L2 speech can stem both from lower exposure to these sounds, as well as to the lack of functional pressures to differentiate the two.

Our results also showed that greater spectral contrast between tense and lax vowels correlated with higher comprehensibility rating, supporting the second hypothesis that the degree of

acoustic differentiation between phonological categories should covary with the variability in comprehensibility. The correlation was present for both types of sentences, with a similar, albeit small effect size, suggesting that spectral difference was not more predictive of comprehensibility in the absence of meaningful semantic context.

While our study indicates that acoustic contrast is a potentially relevant factor in speech comprehensibility, it is clear that comprehensibility ratings are a result of multiple contributors, including speaking rate, intonation, word stress and more (Derwing et al., 1998; Derwing & Munro, 1997; Isaacs & Thomson, 2020; Saito et al., 2017), which explains a relatively low predictive power of acoustic contrast alone. This view is also supported by the finding that ratings were more variable for contrastive than carrier sentences, as their relative complexity provided more opportunity for estimating multiple dimensions of comprehensibility.

Finally, the lack of correlation between durational contrast and comprehensibility ratings agrees with the fact that the durational cue is of lesser importance than spectral cue in differentiating tense and lax vowels in English (Hillenbrand et al., 2000).

To conclude, our study demonstrates the potential significance of the acoustic properties of L2 speech in determining comprehensibility, not as a measure of accentedness or distance to the native speakers' model but as a measure of differentiation between contrastive phonological categories.

ACKNOWLEDEMENTS

We would like to thank our experimental participants for their time, the Purdue College of Liberal Arts PROMISE award for funding this research, and lab members Dr. Ye-Jee Jung and Dr. Yuhyeon Seo for assistance in data processing and analysis.

ABOUT THE AUTHORS

Chien-Min Kuo is an MA student of Linguistics at Purdue University and a Purdue Experimental Phonetics and Phonology Lab member. His research interests are phonetics and second language acquisition.
Contact Information:
Purdue University
640 Oval Dr.
West Lafayette, IN 47907
Email: kuo79@purdue.edu

Olga Dmitrieva is Associate Professor of Russian and Linguistics at Purdue University. Her research interests include acoustic phonetics, speech perception, language acquisition, and bilingualism. Her current research focuses on crosslinguistic interaction in bilingual speech production and perception and factors affecting acoustic detail and perceptual processing in non-native and bilingual speakers.
Contact Information:
Purdue University
640 Oval Dr.
West Lafayette, IN 47907
Email: odmitrie@purdue.edu

REFERENCES

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407.

Barreda, S. (2023). *phonTools: Functions for phonetics in R*. (R package version 0.2-2.2).

Best, C. T. (1995). A direct realistic view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross language research* (pp. 167–200). York Press.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13–34). John Benjamins,

Boersma, P. & Weenink, D. (2023). *Praat: Doing phonetics by computer* (Version 6.3.18).

Chen, Y. (2006). Production of tense-lax contrast by Mandarin speakers of English. *Folia Phoniatrica et Logopaedica*, *58*(4), 240–249.

Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hearing Research*, *229*(1–2), 132–147.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility. *Studies in Second Language Acquisition, 19*(1), 1–16.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, *39*(3), 379–397.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in Favor of a Broad Framework for Pronunciation Instruction. *Language Learning, 48*(3), 393–410.

Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 229–273). Timonium, MD: York Press.

Flege, J. E., & Bohn, O. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3-83). Cambridge: Cambridge University Press.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America, 97*(5), 3125–3134.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical society of America*, *97*(5), 3099–3111.

Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. *The Journal of the Acoustical society of America*, *108*, 3013–3022.

Isaacs, T., & Thomson, R. I. (2020). Reactions to second language speech. *Journal of Second Language Pronunciation, 6*(3), 402–429.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13).

Lee, W. S., & Zee, E. (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association*, *33*(1), 109–112.

Li, M., Chen, Y., & Cui, J. (2020). The effect of input on the production of English tense and lax vowels by Chinese learners: Evidence from an elementary school in China. *Interspeech 2020*. 2372–2376.

Mines, M. A., Hanson, B. F., & Shoup, J. E. (1978). Frequency of occurrence of phonemes in conversational English. *Language and Speech*, *21*(3), 221–241.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *45*(1), 73–97. Portico.
R Core Team. (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing.
Saito, K., Trofimovich, P., & Isaacs, T. (2017). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics, 38*(4), 439–462.
Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, *15*(4), 905–916.