Lesnov, R., Wolhein, S. & Bogorevich, V. (2020). Can successful L2 pronunciation facilitate listening comprehension? The role of speech rate and pitch range. In O. Kang, S. Staples, K. Yaw, & K. Hirschi (Eds.), *Proceedings of the 11th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, Northern Arizona University, September 2019 (pp. 250–260). Ames, IA: Iowa State University.

CAN SUCCESSFUL L2 PRONUNCIATION FACILITATE LISTENING COMPREHENSION? THE ROLE OF SPEECH RATE AND PITCH RANGE

Roman Lesnov, Oakwood University
Sofia Wolhein Nava, Northern Arizona University
Valeriia Bogorevich, Northern Arizona University

The relationship between speech perception and production has been the topic of a continuous debate and investigation in second language (L2) contexts, with no conclusive outcome. Considering that perception is a fundamental part of listening comprehension, the question emerges: Can production of certain pronunciation features predict L2 learners' listening comprehension? Also, it is not yet clear if this relationship depends on students' speaking proficiency. This study acoustically analyzed speech samples from 43 English-as-a-second-language (ESL) students at three speaking proficiency levels to see whether and to what extent the production of the four suprasegmental features (i.e., speech rate, pauses, sentence stress, and pitch range) correlated with the learners' listening comprehension scores. The results showed that speech rate and pitch range were significant predictors of listening comprehension; the former having much more predictive power and the latter being especially prominent at the lower speaking proficiency. The findings suggest that ESL learners' success in listening may be associated with their speech rate and pitch production patterns. Therefore, teachers might want to consider teaching speech rate and intonation patterns to develop students' speaking and listening abilities.

INTRODUCTION

The link between pronunciation and other L2 skills, including listening, has been discussed by a number of researchers (e.g., Gilbert, 1995; Kennedy & Trofimovich, 2010; Trofimovich et al., 2009). Gilbert (1995) argued that pronunciation instruction can aid listening comprehension. On the other hand, Kennedy and Trofimovich (2010) discussed that in order to be successful in pronunciation, L2 learners might need to achieve a certain level of listening proficiency first. The-chicken-or-the-egg dilemma aside, these studies showed that speaking and listening skills are interdependent, as are speech production and speech perception. Motivated by this connection, the present study aimed to investigate if L2 suprasegmental production could be associated with successful L2 listening comprehension.

Pronunciation and Listening Comprehension

Listening comprehension

Listening has been widely considered a complex language skill consisting of different processes and stages. Drawing on Anderson's (1985) cognitive model of comprehension, Vandergrift and Goh (2012) have defined listening in terms of three concurrent stages: *perception*, *parsing*, and *utilization*. Perception involves recognizing sound units and syllable boundaries, and identifying words, while the parsing and utilization stages take charge of forming meaningful phrasal units from the listening input in short-term memory and matching these units with existing knowledge stored in long-term memory. In this definition of listening comprehension, speech perception plays a key role that triggers subsequent processes and involves both segmental (e.g., sound units) and suprasegmental (e.g., words, phrases, syntactic units) decoding and parsing. It should be noted that numerous other variables factor in listening comprehension, such as attention and processing speed, short-term memory capacity, strategic control, and processing styles (Rost, 2016; Vandergrift & Goh, 2012).

Suprasegmental production and listening comprehension.

Suprasegmentals were described by Ladefoged (2006) as speech features that involve more than single consonants and vowels. From this perspective, any features above the consonant or vowel level are considered suprasegmental and include, among others, stress, rhythm, and intonation, as well as fluency-based prosodic features, such as speech rate and pausing. Many of these features have been found important for L2 learners' oral proficiency as well as comprehensibility (e.g., Kang, 2010; 2013). For example, speech rate and pauses demonstrated a strong correlation with oral fluency (Derwing et al., 2004), which explains why these two temporal characteristics are used to measure fluency (Thomson, 2015). However, there remains a gap as to how the mastery of these features' production may affect L2 learners' listening comprehension.

The research into stress, defined as a prominent pronunciation of a word syllable (Ladefoged, 2006), has mostly focused on the effects of stress production on oral proficiency (e.g., Wennerstrom, 2000; Kang, 2013). Few studies have investigated the relationship between L2 speakers' stress production and their listening comprehension. Huang (2009) examined how three suprasegmental features produced by Chinese English-as-a-foreign-language (EFL) learners -- stress, rhythm, and intonation -- related to their listening comprehension test scores. While it was found that overall suprasegmental production was highly correlated with listening comprehension, stress in particular did not significantly correlate with the listening scores.

Stress is closely related to pitch since a stressed syllable is normally characterized by higher pitch (Avery & Ehrlich, 1992). Pitch range, stress, and their relation to L2 learners' listening comprehension were investigated by Wang (1993), as cited in Huang (2009). Wang found that lower- and intermediate-level learners' pitch patterns were correlated with their listening comprehension while pre-advanced learners' pitch production did not show any relationship to listening comprehension. A significant correlation was also found between the number of stresses intermediate learners put in their speech and their listening comprehension.

Summary

Overall, the literature on the relationship between L2 learners' pronunciation and listening comprehension shows that the two are connected and may bear influence upon each other (e.g., Gilbert, 1995; Kennedy & Trofimovich, 2010). However, many aspects of this relationship remain unknown and call for further exploration. One of these aspects concerns the production of suprasegmental features, including speech rate, pausing, stress, and pitch, and its relationship to listening ability. In addition, it is largely unknown if higher- and lower-proficiency L2 learners would differ in how their pronunciation abilities relate to their listening.

Purpose and Research Questions

This study investigated the relationship between suprasegmental pronunciation features (i.e., speech rate, pauses, stress, and pitch range) produced by ESL learners and the learners' listening comprehension ability. The study also investigated this relationship at different levels of ESL speaking proficiency (i.e., lower, medium, and higher). Accordingly, the study sought to answer the following research questions.

1. Is there a relationship between ESL learners' suprasegmental production and their listening comprehension?
2. Does speaking proficiency play a role in this relationship?

METHODS

Data Sources

The study used archived de-identified test scores and speech samples obtained from forty-three students in an Intensive English Program (IEP) at a university in the Southwestern United States. The speech samples were recorded by 18-to-25-year-old Arabic, Chinese, and Brazilian students, who had been in the USA for less than six months. The forty-three speech samples were categorized by proficiency: higher (n = 14, including seven males and seven females), medium (n = 14, including nine males and five females), and lower (n = 15, including six males and nine females).

Speaking prompt

The students' speech samples were recorded in response to a speaking prompt on a placement test. The task displayed five pictures showing the process of photosynthesis in plants. The students had one minute to prepare an oral description of how plants produced energy and then one minute to record their answers using hand-held voice recorders. The instructions and a brief practice session on how to use a recorder were given by the proctor prior to the speaking test. To reduce classroom noise due to simultaneous speaking, a one-student-per-desk seating arrangement was used.

Speech samples

The students were divided into three proficiency groups based on their composite score for the speaking section of the placement test. They were categorized into three groups: lower (0-12 points), medium (13-17 points), and higher (18-30 points). There were 111 available student recordings, each of which was assigned a number (1-111). Out of the 111 available recordings, 15 were randomly selected from each proficiency group using the online random number service Random.com (Haahr, 2020), with a total of 45 speech samples. Later, two of the 45 speech samples were found to be defective (i.e., one in the medium proficiency group and the other in the higher proficiency group), necessitating their exclusion from the study.

Listening scores

Listening scores from the placement test were also included in the study. They were based on the students' answers to comprehension questions after listening to eight passages of several text types, including conversations, lectures, and news reports. Each passage was followed by several multiple-choice questions measuring students' abilities to understand explicit (i.e., details) and implicit (i.e., inferences) information. The total score for the listening section was out of 30 points. The reliability (Cronbach's alpha) was 0.72.

Acoustic Measures

The recordings were acoustically analyzed for suprasegmental pronunciation features including speech rate, pausing, stress, and pitch. Due to the fact that some recordings were only 30 seconds long, only the first 30 seconds of each recording were analyzed.

Speech rate

Syllables per second (SpS) and mean length of run (MLR) were used to measure speech rate. SpS was calculated as a ratio of the total number of syllables in a speech sample over the total length of the speech sample. MLR was a measure of the average number of syllables produced in utterances between pauses of 0.1 seconds and above, as suggested by Kang (2010). This was calculated as the total number of syllables divided by the total number of runs. To calculate the values for both measures, the speech samples were transcribed, and then acoustically analyzed for the length of the speech as well as the total number of runs in Praat (v. 5.4; Boersma & Weenink, 2020). Syllables were counted using Wordcalc.com.

Pauses

This study employed the number of silent pauses per 30-second speech file as a measure of pause use. Also, the mean length of silent pauses was calculated as the total length of silent pauses divided by the total number of silent pauses, a silent pause being of length 0.1 seconds and longer (Kang, 2010).

Stress

Pace (the average number of prominent words per run) and space (the proportion of prominent words to the total number of words) were used as measures of stress. First, prominent syllables were identified in the speech samples using both computer and human judgment. Praat spectrograms were analyzed in order to find syllables with higher pitch, intensity, and duration of a vowel (Kang, 2010). Next, pace and space were calculated.

Pitch range

The pitch range of prominent syllables was calculated as the difference between the lowest and the highest values of pitch in a speech sample (Kang, 2010). The highest and lowest pitch values were calculated manually using Praat spectrograms. The manual calculation precluded potential errors in pitch analysis due to Praat's tendency to assign unreasonably high pitch values to particular sounds, such as fricatives.

Analyses

The research questions, variables, and analyses are presented in Table 1 below. To answer the first research question, a multiple regression analysis was used. The second research question ran correlation analyses between listening proficiency scores and the seven independent variables within each of the three proficiency levels.

Table 1
Variables in the Study

	Variable	Function	Operationalization	Level of Measurement	Statistical Procedure
RQ1	LC	DV	Listening scores across all data sets	Interval	Multiple regression
	Rate	IV #1	Syllables per second (SpS)		
		IV #2	Mean length of run (MLR)		
	Pausing	IV #3*	Number of silent pauses (NSP)		
		IV #4	Mean length of silent pauses (MLSP)		
	Stress	IV #5*	Pace (aver. # of prominent words per run)		
		IV #6	Space (# prominent words / total # words)		
	Pitch	IV #7	Overall pitch range (OPR)		
RQ2	LC - High	DV #1	Listening scores within the higher speaking proficiency level dataset	Interval	Spearman correlation analyses
	LC - Mid	DV #2	Listening scores within the medium speaking proficiency level dataset		
	LC - Low	DV #3	Listening scores within the lower speaking proficiency level dataset		
	Rate	IV #1	Syllables per second (SpS)		
		IV #2	Mean length of run (MLR)		
	Pausing	IV #3	Number of silent pauses (NSP)		
		IV #4	Mean length of silent pauses (MLSP)		
	Stress	IV #5	Pace (aver. # of prominent words per run)		
		IV #6	Space (# prominent words / total # words)		
	Pitch	IV #7	Overall pitch range (OPR)		

Note: * = excluded from the analysis due to the linearity assumption violations (strong non-linear relationships with the dependent variable); LC = Listening Comprehension; Rate = Speech Rate; RQ = Research Question

RESULTS

Suprasegmental Production and Listening Comprehension

To answer the first research question -- "Is there a relationship between ESL learners' suprasegmental production and their listening comprehension?" -- a multiple regression analysis was performed, using a linear multiple regression analysis (forward stepwise method) in JASP (version 0.11.1; JASP Team, 2019). The forward stepwise method was selected based on its predictive accuracy and tendency to generate more parsimonious models (Roecker, 1991). The results are displayed in Table 2 below. Table 3 shows the descriptive statistics for the acoustic measures and listening comprehension.

Table 2
Multiple Regression Results

Source	Sum of Squares	*df*	Mean Square	*F*	*β*	SE	*β* *standardized*	*t*	*p*
Model	319.39	2	159.70	12.83					.000**
SpS		1			3.46	0.86	0.51	4.04	.000**
OPR		1			0.01	0.01	0.28	2.26	.032*
Residual	497.72	40	12.44						
Total	817.11	42							

Note: R = 0.63; R^2 = 0.39; Adjusted R^2 = 0.36; ** = significant at the 0.01 level; * = significant at the 0.05 level; SpS = Syllables per Second; OPR = Overall Pitch Range.

Table 3
Descriptive Statistics

Variable	*N*	Mean	*SD*
ListScore	43	13.29	4.41
SpS	43	2.10	0.65
MLR	43	4.01	1.57
MLSP	43	0.78	0.50
Space	43	0.68	0.13
OPR	43	138.59	95.55

ListScore = Listening Comprehension Score; SpS = Syllables per Second; MLR = Mean Length of Run; MLSP = Mean Length of Silent Pauses; OPR = Overall Pitch Range

As shown in Table 2, the regression model was reached in two steps – by including SpS as a predictor first, and OPR next. The final model F-test is significant, which shows that there was a strong linear relationship between each of these two variables and listening comprehension. In other words, SpS and OPR had considerable importance in explaining the variability of listening comprehension whereas MLR, MLSP, and Space did not display such contributions. Yet, it can be seen that SpS was a stronger contributor than OPR (β = 3.46 versus 0.01). Both SpS and OPR had positive Beta values. This means that the higher SpS or OPR the participants spoke with, the better their listening comprehension tended to be. Again, SpS seemed to predict listening comprehension more than 3 times as well as OPR.

The Effect of Speaking Proficiency

To investigate the relationship between the seven independent variables and listening comprehension at each of the three proficiency levels, a Spearman correlation analysis was used. A Spearman correlation was preferred to a Pearson correlation analysis because the distributions of the independent variables were not normal. The results are presented in Table 4. There were no significant correlations between any variables and listening comprehension for higher and medium proficiency groups. For the lower proficiency group, only the correlation between listening comprehension and OPR was significant.

Table 4
Spearman Correlation Coefficients by Level

	SpS	MLR	NSP	MLSP	Pace	Space	OPR
ListScore High (n=14)	.467	.493	-.387	-.126	.191	-.377	.193
ListScore Mid (n=14)	.267	.253	-.137	.234	.078	-.130	-.171
ListScore Low (n=15)	.237	.106	.133	-.099	-.110	-.072	.645**

Note: **. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed). ListScore = Listening Comprehension Score; SpS = Syllables per Second; MLR = Mean Length of Run; NSP = Number of Silent Pauses; MLSP = Mean Length of Silent Pauses; OPR = Overall Pitch Range

DISCUSSION

As we have seen, there is an association between the L2 learners' suprasegmental production and listening comprehension. This association, however, is not necessarily a cause-and-effect relationship. For instance, we cannot claim that the faster students' speech rate is, the better their listening comprehension will be. The study simply suggests that this causal relationship is likely and should be an object of future, more methodologically rigorous research studies (e.g., a pedagogical intervention study). The following discussion lies within these boundaries.

Suprasegmental Production and Listening Comprehension

Based on the results for the first research question, it may be argued that certain aspects of suprasegmental production by ESL learners predict their listening comprehension. More specifically, higher speech rate and broader pitch range were associated with higher listening comprehension scores. These associations are in line with the established link between speech perception and speech production and with previous research, and they may imply that the higher L2 learners' pronunciation competence, the more perceptive learners are towards connected native speech (Gilbert, 1995; Huang, 2009; Kennedy & Trofimovich, 2010; Wang, 1993). The possibility exists that, as also alleged in Huang (2009), students with better suprasegmental pronunciation mastery tend to recognize suprasegmental patterns framed by speech rate and pitch production in a listening stimulus more skillfully, thereby improving the perceptual and parsing stages within the listening comprehension process (Anderson, 1985; Vandergrift & Goh, 2012). Future research could shed more light on the veracity of this supposition.

Role of Speaking Proficiency

The analysis for the second research question shows that pitch range predicts listening comprehension differentially, depending on learners' speaking proficiency. Pitch range positively correlated with listening comprehension of lower-level learners only, which was also the case in Wang (1993). One possible explanation for this finding relates to lower-level L2 learners' tendency to focus on bottom-up skills. The ability to use wider pitch may help the learner to better recognize pitch patterns in other people's speech, including native and non-native speakers. It may be especially helpful for lower-level students by causing them to shift focus from bottom-up to top-down strategies, of which intonation is a part. This tentative explanation is in need of empirical verification and thus serves as a preliminary call for future research.

In contrast, the predictive power of speech rate does not seem to depend on students' speaking proficiency. While a significant predictor of the students' listening comprehension in general, speech rate did not display significant relationships with listening comprehension within any of the three proficiency groups. This can be partly explained by small sample sizes of the data subsets (i.e., $n = 15$, 14, and 14 for the lower, mid, and higher proficiency groups respectively), which may have precluded statistical significance. The correlation coefficients between speech rate and listening comprehension were close to moderate within each of the proficiency levels and, presumably, would have been significant if more data were collected.

Practical Implications

The findings of this study have implications for L2 teaching. Assuming that English learners can control their rate when speaking (see Munro & Derwing, 2001), it seems worthwhile for the teacher to work on improving students' speech rate because this work may eventually benefit both their speaking *and* listening. Teachers could allocate some time for fluency activities in their speaking classes at all proficiency levels. If the focus on speech rate in a particular L2 program seems weak, reconsidering its role may increase the effectiveness of L2 instruction in the development of the overall L2 communicative competence.

In addition, a teaching focus on pitch production and intonation patterns may have a positive effect on lower-level L2 learners' listening ability. This supposition adds to the growing awareness of intonation instruction effectiveness for linguistic development. According to Ford and Thompson (1996), pitch contours are closely related to grammatical unit boundaries in a spoken sentence and thus must help L2 learners decode 'pragmatic units' (Rost, 1991) from listening. Coupled with the findings of the present research, this may suggest that teaching pitch contours productively is a desirable addition to teaching them receptively as it may result in the development of integrated skills (e.g. pronunciation, speaking, grammar, listening).

Besides actual teaching, L2 teachers can inform students that faster speech rate and more modulations in pitch production may be associated with students' progress in listening. This awareness will likely increase students' motivation to autonomously work on their speech rate and intonation. Yet, it might be advisable for students to avoid speeding up their utterances too much. According to Munro & Derwing (2001), if non-native English speakers over-accelerate their speech, they can "introduce new segmental and prosodic errors that might increase perceived accentedness and reduce comprehensibility" (p. 467).

LIMITATIONS AND FUTURE RESEARCH

The study was limited in two ways. First, only a limited set of independent variables was used in the study. This may have precluded the use of other parameters of pronunciation production that can possibly affect listening perception (e.g., tone units, intonation, and pronunciation of segmentals). It is suggested that future research endeavors apply existing theory and hypothesis testing to uncover other variables that may bear influence upon listening comprehension. Second, the method of determining the speakers' proficiency level was problematic due to an arguably insufficient gap between adjacent proficiency categories (i.e., 0-12; 13-17; 18-30). Setting a larger gap (e.g., 0-10; 12-16; 18-30) would be one way to avoid an unwanted overlap between proficiency levels.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Okim Kang and the anonymous reviewers for useful comments on earlier drafts of this manuscript, as well as Jacqueline Church and the Program in Intensive English at Northern Arizona University for assistance with data access.

ABOUT THE AUTHORS

Roman Lesnov (rlesnov@oakwood.edu) is an Associate Professor of English at Oakwood University, Alabama. He teaches TESOL, Linguistics, and ESL. He holds a PhD in Applied Linguistics from Northern Arizona University (2018) and has several years of experience teaching English and linguistics in the US and Russia. His research interests include second language listening assessment and research methods for applied linguistics.

Department of English and Foreign Languages
Moran Hall, Room 104
7000 Adventist Blvd NW
Oakwood University
Huntsville, AL 35896
Phone: 929.225.9330

Sofía Wolhein Nava (sofia.wolhein@nau.edu) is a Lecturer of Spanish at the Global Languages and Cultures Department at Northern Arizona University. Sofía teaches intermediate and advanced Spanish language courses for L2 Learners and Heritage Speakers. Her research interests are in second language curriculum development.

Department of Global Languages and Cultures
Babbitt Academic Annex, Room 206
S. San Francisco Street
Northern Arizona University
Flagstaff, AZ 86011
Phone: 928-523-9011

Valeriia Bogorevich (vb283@nau.edu) is a Lecturer at the Program in Intensive English at Northern Arizona University. Valeria has extensive EFL/ESL teaching experience. Her main research interests lie in the area of L2 assessment, curriculum development, and the use of technology in the classroom.

Program in Intensive English
506 S Beaver St,
Flagstaff, AZ 86011

Northern Arizona University
Phone: 928-523-6726

REFERENCES

Anderson, J. R. (1985). *Cognitive psychology and its implications* (2nd ed.). New York, NY: W. H. Freeman and Company.

Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation.* Oxford, UK: Oxford University Press.

Boersma, P. & Weenink, D. (2020). Praat: doing phonetics by computer [Computer program]. Version 5.4, retrieved from http://www.praat.org/

Derwing, T., Rossiter, M., Munro, M., & Thomson, R. (2004). Second language fluency: Judgments on different tasks. *Language Learning, 54*, 655-679.

Ford, C., & Thompson, S. (1996). International units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns. In E. Ochs, E. Schegloff, & S. Thompson (Eds.), *Interaction and grammar* (pp. 134-184). New York, NY: Cambridge University Press.

Gilbert, J. (1995). Pronunciation practice as an aid to listening comprehension. In D. Mendelson & J. Rubin (Eds.), *A guide for the teaching of second language listening* (pp. 97-112). San Diego, CA: Dominic Press.

Haahr, M. (2020, March 26). Random.org: True Random Number Service. Retrieved from https://www.random.org

Huang, X. (2009). The relationship between Chinese EFL learners' proficiency in suprasegmental features of pronunciation and their listening comprehension. *CELEA Journal, 32,* 31-39.

JASP Team (2019). JASP (Version 0.11.1) [Computer software]. Retrieved from https://jasp-stats.org/download/

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, 38, 301-315.

Kang, O. (2013). Relative impact of pronunciation features on ratings of non-native speakers' oral proficiency. In J. Levis & K. LeVelle (Eds.), *Proceedings of the fourth pronunciation in second language learning and teaching conference.* (pp. 10-15). Ames, IA: Iowa State University.

Kennedy, S., & Trofimovich, P. (2010). Language awareness and second language pronunciation: a classroom study. *Language Awareness*, *19,* 171-185.

Ladefoged, P. (2006). *A course in phonetics* (5th ed.). Boston, MA: Thomson.

Munro, M., & Derwing, T. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech: the role of speaking rate. *Studies in Second Language Acquisition*, *23,* 451-468.

Roecker, E. (2012). Prediction error and its estimation for subset-selected models. *Technometrics, 33, 459-468.*

Rost, M. (1991). *Listening in action.* New York, NY: Prentice Hall.

Rost, M. (2016). *Teaching and researching listening* (3rd ed.). New York, NY: Routledge.

Thomson, R. (2015). *Fluency.* In M. Reed & J. Levis, (Eds.), *The handbook of English pronunciation* (pp. 209-226). Malden, MA: John Wiley & Sons.

Trofimovich, P., Lightbown, P., Halter, R., & Song, H. (2009). Comprehension-based practice: The development of L2 pronunciation in a listening and reading program. *Studies in Second Language Acquisition*, *31*, 609-639.

Vandergrift, L. & Goh, C. (2012). *Teaching and learning second language listening: Metacognition in action.* New York, NY: Routledge.

Wang, G. (1993). Pitch movements of Chinese and English utterances and the Chinese EFL learner. Paper presented at Pac SLRF, University of Sydney, Australia.

Wennerstrom, A. (2000). The role of intonation in second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 102-127). Ann Arbor, MI: University of Michigan Press.