

Investigating the Relationship between Suprasegmental Production
and ESL Students' Listening Comprehension

Roman Lesnov

Northern Arizona University

Abstract

The relationship between speech perception and production has been the topic of continuous debate and investigation in L2 contexts with no conclusive outcome. Some studies found that perception precedes production (e.g., Flege, 1995), others confirmed that the opposite is possible (e.g., Trofimovich & Baker, 2006). In the light of this, the question emerges: Can production of certain pronunciation features predict L2 learners' listening comprehension? Comparatively few studies (e.g., Romanini, 2008) investigated how the production of ESL learners' suprasegmental features affect listening comprehension. Also, little to no research has been conducted into how the relationship between suprasegmental production and listening comprehension may depend on students' ESL speaking proficiency level. This study sought to reduce the scarcity of this knowledge. It acoustically analyzed speech samples produced by 43 ESL students of different speaking proficiency levels to see whether and how the production of the four overarching suprasegmental features – speech rate, pausing, sentence stress, and pitch – correlate with the learners' listening comprehension scores. It was found that the productions of speech rate and pitch range were important predictors of L2 listening ability, the former having much more predictive power and the latter especially significant for lower level students. These findings shed more light on the relationship between ESL speech perception and production, and brought important insights into ESL classroom practices.

Background

Pronunciation is an essential part of both second language listening and speaking skills. To put it differently, L2 learners' speech perception and production would be impossible without their knowledge of correct pronunciation of the target language sounds. This importance raises the question about how the two constructs are connected with each other, the answer to which can provide additional insights into L2 pronunciation instruction and its effects on other L2 skills. In this realm, investigating whether speech perception leads production, or vice versa, abundant research has led to the understanding that the two are strongly linked to each other. However, few studies went beyond the perception of particular pronunciation features and looked at the overall listening comprehension and its relationship to the production of those features.

Regrettably, research into how production of suprasegmental features (i.e., features above the consonant/vowel level of pronunciation) can affect ESL learners' listening comprehension is even scarcer (e.g., Romanini, 2008), particularly in relation to fluency production. Unfortunately, there is no research known to the author of this study that would investigate the relationship between L2 speakers' fluency production and their listening comprehension. Still, there are reasons to hypothesize the presence of such a relationship due to the strong interconnection between speech perception and production as described above. Dissimilar to that, it is harder to hypothesize that L2 learners' speech rate and pauses contribute to their listening comprehension selectively depending on their speaking proficiency levels due to the absence of any theoretical or empirical support.

Few studies investigated the relationship between L2 speakers' stress production and their listening comprehension leaving the gap in this knowledge barely filled. Xiaoyu (2009)

examined how three suprasegmental features produced by Chinese EFL learners – stress, rhythm, and intonation, related to their scores for the listening comprehension of a TOEFL lecture. It was found that overall suprasegmental production was highly correlated with listening comprehension, which, according to the author, implied that pronunciation teaching could bolster listening comprehension. However, stress in particular did not significantly correlate with listening scores.

It is evident that many aspects of the relationship between pronunciation and listening comprehension remain unknown and call for further exploration. First, the production of fluency-based suprasegmental features, namely speech rate and pausing, have not been explored in terms of its influence on L2 learners' listening comprehension. Second, the research into the effects of stress and pitch production on listening comprehension has been scarce in amount. Finally, L2 proficiency level has not been taken into consideration as a variable that may intervene and affect the hypothesized relationship.

Research questions

In an attempt to fill the gaps in the literature, the study was guided by the following research questions.

1. To what extent does English learners' suprasegmental production (speech rate, pausing, stress, and pitch) contribute to their listening comprehension?
2. Does speaking proficiency affect contributions of English learners' suprasegmental production (speech rate, pausing, stress, and pitch) to their listening comprehension?

Methods

There were no human participants in the study per se. Rather, archived speech samples obtained from the students in an Intensive English Program (IEP) in a southwestern American

university were used. So, the primary source of data in the study was the students' speech samples recorded for a speaking task in the IEP Fall 2015 placement test. Forty-three out of 111 speech samples were randomly chosen for analysis. The 43 samples represented three speaking proficiency categories. They consisted of 15, 14, and 14 samples from lower, middle, and higher proficiency categories respectively. The three proficiency categories were formed based on the placement speaking scores. Speaking scores in the range of 0-12 were used to designate the lower proficiency pool. The ranges of 13-17 and 18-30 were used to form the middle and high proficiency categories.

The speech samples were students' responses to one of the speaking tasks of the placement test. The prompt of the speaking task displayed five pictures showing the process of photosynthesis in plants. Based on these pictures, the students had 1 minute to prepare a description of how plants produced energy and then 1 minute to record their answers. The students' recordings were first reduced in length to around 30 seconds each, and then acoustically analyzed for suprasegmental pronunciation features including speech rate, pausing, stress, and pitch.

Speech rate was measured by syllable per second (SpS) and mean length of run (MLR). The former is calculated as a ratio of total number of syllables over total length of speech. The latter is a measure of the average number of syllables produced in utterances between pauses 0.1 second and above, as suggested by Kang (2010). To measure pausing, the number of silent pauses (NSP) per 30-seconds speech and mean length of silent pauses (MLSP), which was calculated as the total length of silent pauses divided by the total number of silent pauses, were used. The approach to measuring stress in this study was adopted from Vanderplank (1993) by measuring pace and space. Pace was the average number of prominent words per run. Space was

computed as the proportion of prominent words to the total number of words. Finally, the overall pitch range (OPR) was calculated as the subtraction between the lowest and the highest values of pitch in a speech sample (Kang, 2010). Pitch was obtained only for prominent syllables. All the parameters were measured in Praat.

Lastly, for each of the 43 speech samples, the corresponding students' scores for the listening section were obtained. The listening scores were then analyzed to see if they associated with the pronunciation features measures both across the three proficiency sets of data and within each of them.

Results

After checking the necessary assumptions, the multiple regression analysis was performed to answer RQ1. The dependent variable was listening proficiency, operationalized by the students' scores on the listening section of the IEP Fall 2015 placement test. On the basis of the assumption checks, the decision was made to exclude the two variables from the regression analysis - NSP and Pace due to their strong non-linear relationships with the dependent variable and markedly low degrees of those relationships. Thus, 5 predictors were to be used in the regression analysis – SpS, MLR, MLSP, Space, and OPR.

For the regression analysis, the Automatic Linear Modeling procedure in SPSS (Version 22) was used. As displayed in Table 3, the final (i.e., corrected) model was reached in two steps – by including SpS as a predictor first, and OPR next. 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 contribution. Yet, it can be seen that SpS was a much stronger contributor than OPR (0.77 and 0.23 respectively), with SpS explaining 77% of variance in listening comprehension scores and OPR 23%. Both SpS and OPR had positive Beta values,

which leads to the conclusion that the higher SpS or OPR the participants spoke with, the better their own listening comprehension of L2 speech tended to be. Again, SpS seems to predict listening comprehension more than 3 times as good as OPR.

Table 1

Regression Model Output: Effects and Coefficients

Source	Sum of Squares	df	Mean Square	F	Beta coefficient	Sig.	Importance
Corrected Model	320.69	2	160.34	12.92		.000**	
SpS_transformed	204.17	1	204.17	16.45	3.54	.000**	0.77
OPR_transformed	61.84	1	61.84	4.98	.01	.031*	0.23
Residual	496.43	40	12.41				
Corrected Total	817.11	42					

Note: **. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).

Acronyms: 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

To gauge the degree of the relationship between the seven independent variables and listening comprehension for each of the three proficiency levels (RQ2), it was unreasonable to use the regression analysis due to the small numbers of the groups' sample sizes. Therefore, a less sophisticated procedure was used, namely Spearman correlation. Spearman correlation was preferred to Pearson correlation because some of the independent variables were not normally distributed. The results of the Spearman correlation analysis are presented in Table 2.

Table 2

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).

There were no significant correlations between any variables and listening comprehension for higher and middle proficiency groups. For the lower proficiency group, only

the correlation between listening comprehension and OPR was significant. These results may suggest that there is no relationship between the speech measures and listening comprehension except for OPR in the lower proficiency level. This, to a certain degree, reflects the finding for RQ1.

Relevance to PIE and Second Language Learning

From a practical standpoint, the findings of this study may shed more light on the role of pronunciation teaching in the development of ESL students' listening comprehension. First, it seems worthwhile for teachers to work on improving students' speech rate, which may eventually benefit both their speaking and listening. In particular, teachers can allocate adequate time for fluency activities in their speaking classes at all proficiency levels and expect improvements on, at least, two fronts. Similarly, the teaching focus on pitch production and intonation patterns may be used to improve L2 learners' listening ability. As the results implied, it can be expected to turn most effective for lower-level students. This finding adds to the growing awareness of the effectiveness of the instruction on intonation in discourse. According to Ford & Thompson (1996), pitch contours are closely related to grammatical units 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 multiple skills (e.g. pronunciation, speaking, grammar, listening).

Besides actual teaching, L2 learners can be informed of the benefit faster speech rate and more accurate pitch production, and, possibly, putting stress into words more selectively can add to their progress in listening. This awareness, in turn, may increase their motivation to work on the improvement of their L2 pronunciation in a number of specific areas.

References

- Baker, W., & Trofimovich, P. (2006). Perceptual paths to accurate production of L2 vowels: The role of individual differences. *IRAL*, 44, 231-250.
- 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*. Timonium, MD: York Press.
- Ford, C. E., & Thompson, S. A. (1996). International units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and Grammar* (pp. 134-184). New York: Cambridge University Press.
- Kang, O., Rubin, D., Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal*, 94, 554-566.
- Romanini, A. (2008). The Influence of production accuracy on suprasegmental listening comprehension. *Master's thesis*. Brigham Young University BYU ScholarsArchive. Retrieved from <http://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=2585&context=etd>
- Rost (1991). *Listening in Action*. Prentice Hall International (UK) Ltd.
- Vanderplank, R. (1993). Pacing and spacing as predictors of difficulty in speaking and understanding English. *English Language Teaching Journal*, 47, 117-125.
- Xiaoyu, H. (2009). The relationship between Chinese EFL learners proficiency in suprasegmental features of pronunciation and their listening comprehension. *CELEA Journal*, 32, 31-39.