Effects of Teaching Suprasegmentals on Second Language Learning and Motivation

Amy Joseph and Douglas Whitaker

College of Liberal Arts and Sciences, University of Florida

This study focuses on the effect of teaching one aspect of suprasegmental structure, sentence stress, on two factors: students’ motivation to learn the language and native speakers’ perception of the students’ speech. The experiment was carried out on a class of English language learners in a high school in Chile over the course of six weeks. The class was split into two groups, and each group was taught the same lessons, but one received explicit instruction about where sentence stress occurred when new phrases were taught. At the beginning and end of the study, students were given a survey measuring their motivation to learn English. Also, a panel of judges with varying exposure to Spanish evaluated recordings of two students from the test class and one from the control to determine how close the speakers sounded to native speakers. Statistical analysis showed that there was no significant difference between the motivation levels of the test class versus the control class. However, there was significantly more improvement between the pre and post recordings of the test class members than the recording of the control class student. Although it did not demonstrate a correlation between teaching sentence stress and motivation, this study reinforces the importance of teaching suprasegmental structure to second language learners.

Introduction

The purpose of this study is to discuss the effects of teaching suprasegmentals, specifically primary stress, on various aspects of language teaching. Traditionally, the focus of the language classroom was segmentals, individual sounds in a language. However, since the rise of communicative language teaching, many educators have put more emphasis on suprasegmentals. This term generally covers three features of connected speech: rhythm, intonation, and stress placement (Celce-Murcia, Brinton, & Goodwin, 1996). Studies have supported the push to concentrate on these elements. Kendrick (1997) found that while it is easiest for students to improve segmental production, improvement in suprasegmentals has the greater effect on intelligibility. Although McNerny and Mendelsohn (1992) believed that students can more easily improve in suprasegmentals, they agreed with Kendrick that suprasegmentals have a stronger impact on students’ comprehensibility. This impact can be seen in studies conducted on international teaching assistants (ITAs) at universities. In one study, Tyler, Jeffries, and Davies (1988) discovered that the ITAs who were viewed as unorganized marked too many syllables with primary stress, paused too much, and did not use appropriate intonation. Another study on ITAs postulated that unless segmental pronunciation makes an ITA’s utterance completely incomprehensible, suprasegmentals are more important for communication because they are most important in conveying meaning (Stevens, 1989).

As important as these features of speech are, English learners often have difficulty including them into their speaking. Various studies have found that this may have to do with transfer from students’ L1 (Adams, 1979; Anderson, 1993; Mochizuki-Sudo & Kiritani, 1991; Wenk, 1985). For example, Juffs (1990) found that Chinese speakers put primary stress on almost every word, whether it is a function or content word (Kendrick, 1997), perhaps due to the fact that their native language is tonal. The switch from a tonal to a non-tonal language often causes difficulties with imitating the intonation of the L2; the switch from a syllable-timed to a stress-timed language would presumably cause difficulties in L2 stress placement. In syllable-timed languages, like Spanish and Japanese, syllables occur after certain time intervals, whether or not they are stressed. In stress-timed languages, like English and Arabic, stress occurs after certain time intervals, regardless of how many syllables fall in this interval. Therefore, in these languages, each syllable takes about the same amount of time to produce and not all syllables have the same length (Abercrombie, 1967). It is perhaps for this reason that Chela-Flores (2001) identified syllable length as the most important phonetic difference between English and Spanish and the feature most significant for understanding non-native speakers’ speech. As syllable length is one of the defining characteristics of primary stress (Celce-Murcia et al., 1996), this study focuses on primary stress placement.

Further, the project focuses on the link between this aspect of pronunciation and motivation because motivation is seen as one of the key factors in predicting eventual success in language learning (Gardner & Lambert, 1959). Increasing this factor would increase the chances of the learner improving in the language after the learning...
experience is over, and a continued effect on the student even after they exit the classroom is a goal of all teaching. Furthermore, the project looked at overall improvement in language speaking ability to measure the general effectiveness of this teaching technique.

Methods

Treatment. This experiment was carried out in an English-as-Foreign-Language (EFL) setting of a high school English class in Chile. The researcher taught a conversational English class that complemented the regular academic class once a week for five weeks, and the first and last week of this period were used to explain and fill out the surveys. In conducting the research, the class was divided into halves of 22 students each. The halves learned identical lessons, but the first half, Group 1, was given explicit instruction and drilling of stress placement in sentences while the second half of the class, Group 2, did not. This explicit instruction consisted of the following: the lessons centered around a theme (e.g., vacations) and students were given appropriate phrases used to discuss these themes. The phrases contained new vocabulary and sometimes grammatical structures the students were not familiar with. However, there was no sustained emphasis on grammar. Attention was explicitly drawn to stress placement during the Group 1 class. In the first session with treatment, the importance of stressing the correct syllable was impressed on the students. Then, for the next two sessions after this, students were given the stress placement in new phrases. For example, if students were given the phrase, “Pancho goes to sleep at 5,” the researcher first elicited the vocabulary, wrote the sentence on the board, drew a box over the stressed syllable and underlined it, read the sentence aloud, asked students to repeat the sentence as a whole, and finally chose random individuals to repeat. The researcher would correct for incorrect stress placement. At the end of this sequence, the sentence was left on the board and looked like this:

\[
\text{Pancho goes to sleep at 5.}
\]

Students were then asked to use these phrases in controlled and finally freer practice. During the last class, students were asked to identify where the stress was based on the researcher’s production of the sentence.

Group 2, on the other hand, was not told of the importance of stress placement and was not told where the stressed syllable was. However, they followed the same sequence otherwise: the researcher elicited the vocabulary, wrote the sentence on the board, read it aloud, asked the students to repeat the sentence as a whole and then individually, and finally prompted them to use the phrases in controlled and freer practice.

Measuring Motivation. Both groups were given surveys measuring their motivation to learn English and their assessment of their own English, especially when compared to that of native speakers of English (see Appendix 1). This survey was handed out during the first week of class, before treatment was given to either group, and it was expected to be handed in by the end of that week. Unfortunately, students did not return this survey, so a few weeks later, the same survey was given in class. The students were explicitly told to answer the survey based on how they felt at the beginning of the semester, before the researcher had interacted with them. This same survey was given to both classes during the last class of the program.

Motivation was measured from the surveys in the following way. The first twelve questions were given with a set pattern. First, students were asked how often or what percentage of the time they did a certain activity. Then they were asked how often or how much they carried out this activity in English. For these questions, the second question in the series (how often the activity was done in English) was divided by the first question (how often the activity was done in general) to give a number. The six numbers generated from this portion of the survey were added to the final score.

The next three questions asked students about their willingness to use English in three situations: with people who knew almost no English, with those who were bilingual, and those who knew almost no Spanish. Students were given a full point if they tried to speak English with those who did not know much English, 0.75 points if they tried to speak English with bilinguals, and 0.5 points if they tried to speak it with people who knew almost no Spanish. Related to these questions was the next question: did students look for opportunities to speak English? If they answered “yes,” they were given 0.5 points and 0 if they said “no.” In addition, another question on the survey asked if students would go abroad to study English if given the opportunity. This question was not used to calculate the overall score because the choice to study abroad might very well have more to do with the novelty of going to a new place and meeting new people instead of a desire to improve language skills.

The next three questions asked students to rate how much they agreed with statements related to trying hard to learn English and wanting to sound like native speakers. The rating scale was from one to five (five showing the most motivation), so the score for each question was calculated by dividing the rating students chose by five.

All of the scores for the separate sections were added together to generate an overall motivation score. The Wilcox Rank-Sum test was then used to determine if there was a significant difference between the motivation scores at the beginning of the program and the scores at the end (see Appendix 3).
Evaluating Recorded Speech. Apart from motivation, this study also examined recorded speech from different students. Two speakers from Group 1 and one speaker from Group 2 were recorded before and after the treatment. On both occasions, the students were asked to describe a picture in as much detail as possible and then were asked freer questions about their lives and town. These recordings were assessed by a panel of six undergraduate students: two who had very little exposure to Spanish, two who had studied Spanish for several semesters at the college level, and two who had lived in Spanish-speaking countries for about the first ten years of their lives and learned English as a second language. The judges were asked to rate the students in the following categories: fluency, ease of speech, content, consonant and vowel pronunciation, prosody and intelligibility. In each category, they gave the students a rating between one (nothing like a native speaker) and 5 (exactly like a native speaker). The prosody ratings were compared using the Wilcoxon Rank-Sum test (see Appendix 4). Furthermore, an overall score was assigned to each recording, which was computed by adding together all the numerical ratings given by the judges. These overall scores were also compared using the Wilcoxon Rank-Sum test (see Appendix 4).

Measuring Differences between Judges. The scores given by the three different groups of judges (those with no, little, and large amounts of exposure to Spanish) were compared using the Kruskal-Wallis test to see if judges with more exposure to Spanish would rate the recordings as closer to native speakers than those with less exposure to Spanish (see Appendix 4).

Hypotheses

The researcher posed the following hypotheses concerning EFL classes for Spanish-speaking high school students. Sustained emphasis on primary stress would increase language learning motivation and improve prosody and overall speaking ability. Also, judges who have more exposure to Spanish will judge English learners from Spanish-speaking countries as having better speaking abilities in English than judges who have less exposure to Spanish.

Results

Learner Motivation. The first independent variable examined was the motivation of the learners using the Wilcoxon-Rank-Sum Test. As is shown in Figure 1, it was found that there was no significant difference (p-value = 0.881).

Prosody Improvement. The next independent variable was the ratings judges gave the recordings of students from the different classes. The prosody ratings on their own were compared using the Wilcoxon Rank-Sum test. As shown in Figure 2, it was found that Group 1 students did not improve significantly more between the first and second recording with regards to this category than the Group 2 student (p=0.1006).

Overall Speaking Improvement. An overall score was assigned to each recording as well, which was computed by adding together all the numerical ratings given by the judges. These scores were compared using a one-sided test, which was statistically significant at the 5% level (p-value = 0.045). Therefore, it is reasonable to conclude that C1 students had improved significantly more between the two recordings than the Group 2 student. Figure 3 shows this difference between Group 1 and Group 2.
Figure 3: Difference between the overall scores given by the judges.

Rating Differences Based on Exposure to Spanish. The last independent variable examined was the difference between ratings given by the judges. Judges were chosen so two had basically no exposure to Spanish (Judge Group 1), two studied Spanish at the college level (Judge Group 2), and two spoke Spanish natively as well as English (Judge Group 3). When the scores assigned by these three groups were compared using the Kruskal-Wallis Test ($\chi^2 = 0.9405$, df = 2, p-value = 0.6249), it was found that there was no significant difference between them. Figure 4 shows this lack of difference between Judge Groups 1, 2, and 3.

Figure 4: The difference in the overall ratings given by each group of judges.

Discussion

After finishing the experiment, three conclusions were reached regarding the explicit instruction of sentence stress. First, it did not increase language learning motivation. Second, it did make students’ speech sound closer to that of native speakers. Third, it did not make students’ prosody closer to that of native speakers. Furthermore, it was posited as judges’ exposure to Spanish increased, their ratings of students’ English speech as being closer to that of native speakers would also increase. It was found, however, that there was no significant difference between the groups of judges.

This study was useful for several reasons. Apart from reinforcing the belief that students’ overall speech benefits from explicit suprasegmental instruction, it also showed that varying degrees of exposure to Spanish among academically oriented people does not affect perception of how close non-native English is to native English.

It would be useful to continue investigation in several ways. First, because sample sizes were a limiting factor in this study, an expanded study on the effect of explicit suprasegmental instruction on overall speech would make results more reliable. Equally, an expanded study on the effect of suprasegmental instruction on motivation would verify or refute that there is no relationship between these variables. Second, the native Spanish speakers on the panel of judges were both interested in languages and language teaching. This may have caused them to try to be more “objective” about their ratings than a native Spanish speaker who had no background in this area. Further study that includes judges who are not involved with second language acquisition or language study in general would be useful. Third, the judges who had basically no knowledge of Spanish and had very limited exposure to Spanish speakers did have significant exposure to non-native speakers of English. This may have made them more sympathetic to the efforts of the students in the recordings. It would be beneficial to repeat this study using non-native speakers with very limited exposure to non-native English.
References


Appendix 1

Por favor, lee cada pregunta con cuidado y contesta.

1. En promedio, por día y fuera de la clase, ¿cuántas horas estudias? _ horas.
2. En promedio, por día y fuera de la clase, ¿cuántas horas estudias inglés? _ horas.
3. En promedio, ¿cuántas horas escuchas música diariamente? _ horas.
4. En promedio, ¿cuántas horas escuchas música en inglés diariamente? _ horas.
5. En promedio, ¿cuántas horas ves televisión diariamente? _ horas.
7. Más o menos, ¿cuántas películas ves por semana? _ películas.
8. Más o menos, ¿cuántas películas en inglés ves por semana? _ películas.
9. En promedio, ¿cuántas horas lees por semana fuera del material que tus profesores te dan? _ horas.
10. En promedio, ¿cuántas horas lees por semana en inglés fuera del material que tus profesores te dan? _ horas.
11. En promedio, ¿qué porcentaje de tiempo que pasas hablando en la clase de inglés hablas en inglés (y no en español)? _ %
12. En promedio, ¿qué porcentaje del tiempo que pasas hablando afuera de la clase de inglés hablas en inglés? _ %.
13. Imagina que conocieras a alguien que sabía español perfectamente y un poco de inglés, ¿hablarías con esta persona en español o inglés? (Marca con un círculo.)
   español _ inglés _
14. Imagina que conocieras a alguien que sabía tanto español como inglés, ¿hablarías con esa persona en español o en inglés? (Marca con un círculo.)
   español _ inglés _
15. Imagina que conocieras a alguien que solamente sabía inglés, ¿intentarías hablar con esta persona en inglés? Sí _ No _
16. ¿Intentas encontrar oportunidades de hablar en inglés? (Marca con un círculo.) Sí _ No _
17. Si tuvieras la oportunidad de estudiar inglés en un país anglohablante, ¿la tomarías? (Marca con un círculo) Sí _ No _
18. ¿Es más importante para ti mejorar tu gramática o tu pronunciación (marca 1 para totalmente gramática, 5 para totalmente pronunciación)?
   1. Solamente me importa la gramática. 2. Ambos son igualmente importante 3. 4. 5. Solamente me importa la pronunciación. No me importa la gramática para nada.

Para las preguntas 19-22, utilizando una escala del 1 al 5 (1 es lo menos, 5 es lo más) decide más o menos cuanto estás de acuerdo con las oraciones dadas.

19. Me esfuerzo en mis clases de inglés. 1 _ 2 _ 3 _ 4 _ 5 _

20. Tengo muchas ganas de aprender inglés. 1 _ 2 _ 3 _ 4 _ 5 _

21. Quiero mucho que mi inglés suene como el inglés de los hablantes nativos de inglés. 1 _ 2 _ 3 _ 4 _ 5 _

22. Basado en mi habla, podría pasar por un hablante nativo de inglés. 1 _ 2 _ 3 _ 4 _ 5 _
Appendix 2

Judging Guidelines

Name: ___________________________  Recording #: ___________________________

I. As you listen to the following recordings, please rate how close the speaker is to a native speaker of English in each of the following categories.

a. Fluency (rate of speech, hesitation, stumbling over phrases, self-repetition, self-correction, false starts)
   
   1. Nothing like a native speaker
   2. Somewhat like a native speaker
   3. Exactly like a native speaker

b. Ease of speech (how much effort is put into producing speech, how naturally the phrases come to mind, confidence in speaking)
   
   1. Nothing like a native speaker
   2. Somewhat like a native speaker
   3. Exactly like a native speaker


c. Content (variety in vocabulary, correctness and complexity of grammar and sentence structures)
   
   1. Nothing like a native speaker
   2. Somewhat like a native speaker
   3. Exactly like a native speaker


d. Consonant and vowel pronunciation
   
   1. Nothing like a native speaker
   2. Somewhat like a native speaker
   3. Exactly like a native speaker


e. Prosody (rhythm, intonation, smoothness of speech)
   
   1. Nothing like a native speaker
   2. Somewhat like a native speaker
   3. Exactly like a native speaker


f. Intelligibility
   
   1. Nothing like a native speaker
   2. Somewhat like a native speaker
   3. Exactly like a native speaker

II. Please also answer the following questions about each speaker:

1. Do you think this person could hold a conversation with a native English speaker who is not accustomed to speaking with Chileans? (Circle one)
   
   Yes  No

   Why or why not?

2. Give three personality adjectives describing the person who is speaking.
Appendix 3

[1] "H0: Group1 = Group2 vs. HA: Group1 < Group2"

Wilcoxon Rank-Sum test

data: half1 and half2

W = 68, p-value = 0.881

alternative hypothesis: true location shift is less than 0

[1] "H0: Group1 = Group2 vs. HA: Group1 != Group2"

Wilcoxon Rank-Sum test

data: half1 and half2

W = 68, p-value = 0.2685

alternative hypothesis: true location shift is not equal to 0
Appendix 4

[1] "Differences between groups TOTAL score"
[1] "H0: Group1 = Group2 vs. HA: Group1 > Group2"

Wilcoxon rank sum test with continuity correction

data: group1tot and group2tot
W = 54.5, p-value = 0.04449
alternative hypothesis: true location shift is greater than 0

[1] "H0: Group1 = Group2 vs. HA: Group1 != Group2"

Wilcoxon rank sum test with continuity correction

data: group1tot and group2tot
W = 54.5, p-value = 0.08897
alternative hypothesis: true location shift is not equal to 0

[1] "Differences between groups PROSODY score"
[1] "H0: Group1 = Group2 vs. HA: Group1 > Group2"

Wilcoxon rank sum test with continuity correction

data: group1pros and group2pros
W = 49, p-value = 0.1006
alternative hypothesis: true location shift is greater than 0

[1] "H0: Group1 = Group2 vs. HA: Group1 != Group2"
Wilcoxon rank sum test with continuity correction

data: group1pros and group2pros

\( W = 49, \text{ p-value } = 0.2012 \)

alternative hypothesis: true location shift is not equal to 0

[1] "H0: All distributions equal vs. HA: At least one pair of differences"

Kruskal-Wallis rank sum test

data: c(judges1, judges2, judges3) and ind

Kruskal-Wallis chi-squared = 0.9405, df = 2, p-value = 0.6249