Video Intervention to Raise Awareness of Common English Intonation Patterns

Mathew Porter and Azusa Okada

(Received on October 31, 2014)

Abstract

The researchers created three short videos introducing different types of sentence-level intonation patterns: fall (statement), rise (question), and high fall (enthusiasm) to help non-English majors improve their perception and production of L2 sounds. Learning-based and student-based evaluations of the materials were carried using two groups of students assigned to either a treatment group (n = 25) or contrast group (n = 26). The quasi-experimental design used for the learning-based evaluation consisted of production and perception pretests and posttests administered the following week to see if the materials were effective at improving the treatment group's perception and production of the targeted intonation pattern. Perception posttests for both group showed a significant improvement, but not as the result of the videos. The student-based evaluation was conducted using a survey and showed the intonation videos were well received because they helped to raise awareness and were easy to understand, interesting, and instructive.

Introduction

Unlike any of the four skills, pronunciation does not usually get its own course in a general English program at Japanese universities. As a result, any guidance about English pronunciation usually occurs at the discretion of the teacher, most likely incidentally during a speaking or communication course. However, with most classes only meeting 15 times in a semester, it is unlikely that students are getting enough guidance or practice in class to make significant improvements in their pronunciation abilities. In addition, required general English class sizes are often large making it difficult for teachers to address individual pronunciation problems when they are noticed. This is complicated by English teachers believing they lack the requisite knowledge of pronunciation and confidence to address pronunciation in their classes (McDonald, 2002).

One way to ease this situation is through the use of engaging and pedagogically-sound multimedia materials delivered online (Pennington, 1999). Porter (2012) designed

a supplemental, structured pronunciation course accessed via Moodle, which was used with 110 Japanese students taking a required general listening course. Using original materials, the course introduced segmental and suprasegmental features of English in Japanese and provided opportunities for focused listening and practice to raise students' awareness of these features. The 8-module course was completed by only 13% of the students and although there were positive comments indicating that the materials helped to raise some students' awareness of the features of English phonology, many students wanted to work with the materials in class. The course is being revised and the materials examined in this study were intended for addition.

The current study is a retrospective analysis using student-based and learning-based evaluations (Ellis, 1997) to examine the effectiveness of materials designed for supplemental, self-directed study of English intonation, focusing on three patterns: fall (statement), rise (question), and high fall (enthusiasm). The learning-based evaluation investigates whether a one-time intervention using videos to teach about intonation and provide structured practice are effective at helping improve both students' comprehension of English sentences with different intonation patterns and production of English utterances using those same targeted intonation patterns. The student-based evaluation measures students' attitudes towards studying intonation through our teacher-created videos. We answer the following research questions:

1. Does the treatment group outperform the contrast group on the perception posttest?
2. Does the treatment group outperform the contrast group on the production posttest?
3. Does the treatment group perform better on perception and production posttests after the intervention?
4. How do students feel about studying pronunciation in this way?

Background

It must first be noted that the goal of pronunciation training should not be the elimination of all traces of a foreign accent or the acquisition of "native-like" pronunciation, which according to Scovel (2000) may be impossible for most. More importantly, students need help reaching a level of intelligibility that will help them in contexts where English is being used as the Lingua Franca between non-native speakers of English, as is common in Japan. It is in those contexts that an English language learner's pronunciation abilities,

both their production and perception of English sounds, can stand in the way of communication. In a study of a spoken corpus of non-native speaker conversations, Jenkins (2000) found that pronunciation errors caused the majority of instances of communication breakdown, and these errors were the result of the transfer of L1 sounds.

We focus on intonation because it is crucial for effective communication since it can change the meaning of utterances. The use of different intonation delivers different information, which may be described as *attitudinal*, *grammatical*, *focusing*, *discourse*, *psychological*, *indexical* (Wells, 2006). For example, if the statement "That's my dad." is said with a falling intonation, the speaker is indicating to the hearer that the man is his father. The same sentence could also be said with a rising intonation, in which the speaker is having difficulty believing that the man is his father, perhaps because the man looks very different from what the speaker expected. If the sentence is said with an exclamation (which usually expresses surprise or excitement) as in "That's my dad!' with a falling intonation from very high pitch, it shows that the speaker has just realized that the man is actually his father. Wells mentions, in particular, that exclamations are "the simplest kind of utterance" in which ESL students could practice falling intonation, because exclamations are always said with a falling tone.

The explicit teaching of phonological features of English in this study is targeted at improving students' cognitive capacities as well as their auditory and motor skills. Fraser (2008) introduced a framework for the teaching of English pronunciation to English language learners based on cognitive phonology. She explained that when students imitate sounds, those sounds aren't just simply heard and repeated, but "heard, subconsciously broken down or processed, and then recreated" (p. 82). It is in that intermediary stage that concepts, or ideas about what constitutes meaningful sounds, are subconsciously applied to what is being heard. For adults, these concepts were formed with the acquisition of the L1. She suggests that both perception and production activities are essential in aiding students to change their concepts.

Wrembel (2006) proposed a model for teaching pronunciation emphasizing the cognitive and metacognitive aspects of pronunciation learning. Similar to Fraser's framework, the metacompetence-oriented model of acquisition of second language phonology also considers that students process sounds through their L1. Wrembel explains the need for a kind of procedural knowledge, or know-how, "which is understood as conscious knowledge of and about the grammar of the language and which may be developed by making

the learner metalinguistically aware of L2 phonetics and phonology" (p. 2). This phonological metacompetence is thought to aid in relaxing the effect of the L1 on the perception and representation of L2 sounds as well as support the development of the ability to self-monitor and self-correct.

Few studies have tested the effectiveness of targeting cognition in pronunciation teaching through explicit instruction, but two studies by Abe (2007, 2011) have shown some success. In Abe (2007), students learning about connected speech were divided into two groups and exposed to two different teaching methods. Both methods relied on input enhancement, with one group listening to a sample of careful and connected speech and then receiving teacher explanation and the other group listening to the samples and then exploring connected speech in pairs. After pair-work was completed, the pairs shared their findings with the group. The group that worked in pairs significantly outperformed the group receiving explanation on both perception and production tests. Abe (2011) once again looked at connected speech, this time over a 4-week period using form focused instruction. Similar to his earlier study, the treatment group listened to careful and connected speech recordings and worked in pairs to notice aspects of connected speech while the contrast group received explanations of connected speech and carried out listen and repeat exercises. The treatment group, again, outperformed the contrast group on both perception and production tests, and these results were still true in a second posttest conducted one month later. Our study attempts to find similar positive results from raising awareness in the learning of pronunciation through online, self-directed study.

Method

Participants

All participants (n = 51) were non-English majors from a private university in western Japan, taking a required English listening or general English course. Most participants were in their first (n = 29) or second year (n = 21) of university and the proportion of first- and second-year students were nearly equal in both groups. A group of 25 randomly chosen students (18 Male/7 Female) served as the treatment group. A comparable group of 26 students (22 Male/4 Female) participated in an untrained contrast group. Finally, two lecturers of English at universities in Japan were recruited to evaluate pre- and posttest utterances. Both lecturers were native speakers of American English. Rater 1 had taught

for 3 years in the US and 11 years in Japan. Rater 2 had taught for 2 years in the US and 7 years in Japan.

Procedure

A pre-test was given to both groups before the video intervention consisting of both a production and perception test. The production test consisted of a single sentence which used the three intonation patterns: fall (F), rise (R), and high fall (HF). The intonation patterns were represented by a single sentence which used different punctuation, and each sentence was shown individually on students' computer monitors. Students were not given any auditory prompts. The sentences appear in Table 1.

Table 1. ***Sentences Used in the Production Test***

Sentence	Sentence Type/Function	Tone used at the End
It stopped raining.	Statement	Fall (F)
It stopped raining?	Question	Rise (R)
It stopped raining!	Exclamation	High fall (HF)

Students read each sentence aloud once and recorded it by themselves. A second production test, asking students to say the same three sentences in Japanese, was conducted afterward. All recordings were made using Victor Software Recorder. This was followed by a 24-item perception test using 11 sentences spoken with the three intonation patterns conducted in Moodle and automatically marked. These sentences and intonation patterns are shown in Table 2. Students were then asked to write their impressions of the

Table 2. ***Sentences and Intonation Patterns Used in the Perception Test***

Sentence	Intonation
I got a car.	R, F, HF
I got an A on my English homework.	R, F, HF
You're getting married.	R, F, HF
Our neighbors moved.	R, F, HF
They bought him a present.	R, F, HF
You've lost weight.	R, HF
You found a job.	R, HF
We sold the house.	R, F
Mom cooked dinner.	F
You passed the exam.	F
We won the contest!	HF

three intonation patterns in Japanese. (Data gathered in Japanese will be examined in another paper.) Finally, students in both groups were shown videos.

The stimuli for the intervention were three, two-minute videos created around the sentence, "It stopped raining," using the three different intonation patterns, which we called statement (fall), question (rise), and enthusiasm (high fall). Each video included a sketch to contextualize the sentences, visual representation of the appropriate intonation pattern, and opportunities for listen and repeat practice. The sketch was created and performed by two American lecturers of English. Figure 1 shows a visual representation of the high fall tone used in the enthusiasm video.

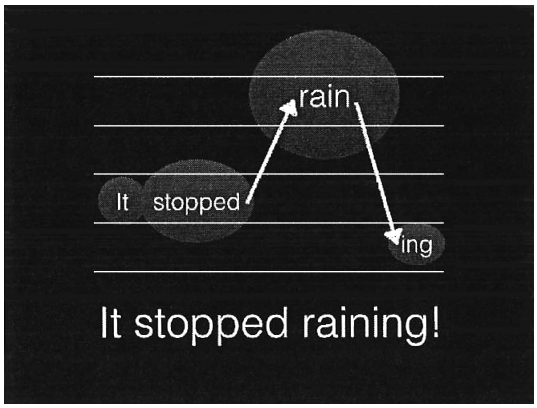

Figure 1. It Stopped Raining: High Fall (Enthusiasm)

Treatment group participants watched the intervention videos at individual computer workstations in a CALL classroom while participants in the contrast group watched three, two-minute videos with explanations and activities for practicing an English phoneme known to be difficult for Japanese speakers such as the /w/ in "woman" and /ŋ/ in sing (Rogerson-Revell, 2011). One week later, participants in both groups were given a posttest consisting of the same production and perception tests, in addition to a questionnaire. Then both groups were shown the other videos.

Perceptual Evaluations

Production data files were cleaned of identifying data and numbered sequentially and then assigned an identification number using a random number generated by Random.org, a True Random Number Generator that uses atmospheric noise, rather than a computer program, to generate randomness. The production data underwent a single perceptual

evaluation carried out independently by the two raters, who were asked to identify the intonation used in each of the 312 recordings gathered in the pre- and posttests. Each utterance was assigned to one of four categories: statement (fall), question (rise), enthusiasm (high fall), or none of these. Rater evaluations were compared to the intended utterances and assigned a value of 1 if the intended utterance and evaluation matched and 0 if it did not. The raters had substantial agreement, Cohen's Kappa = .70.

Results

Perception Tests

The treatment and contrast groups' pretest scores were analyzed using a two-tailed independent *t* test to confirm that there was no significant differences between the groups and the groups could be considered homogenous at the start of the treatment. The average score for both treatment and contrast groups were 22.92 with SD of 1.41 for treatment group and 1.23 for contrast group and the differences in mean score were not significant at the .05 alpha level, $t(49) = 0.008$, $p = .993$, revealing that the two groups are in fact comparable. These results are summarized in Table 3.

Table 3. ***Results of t test Comparing Pretests for Homogeneity***

Test	*n*	*M*	*SD*	*Df*	*t*	*p*
Contrast	26	22.92	1.41	49	0.0083	0.00934
Treatment	25	22.92	1.23	49		

A two-way ANOVA was conducted to investigate the effects of condition (treatment/contrast group), time (pretest/posttest), and the mutual effect of condition and time. The main effect of time was significant at the .01 alpha level, $f(1,49) = 10.4895$, $p = .002$, but the main effect of condition as well as the interaction effect were not significant at the .05 alpha level, $f(1,49) = 00018$, *ns*; $f(1,49) = 0.0032$, *ns*. These results are summarized in Table 4.

Table 4. ***Summary of Two-Way ANOVA for Condition, Time, and Interaction***

Source	*df*	*SS*	*MS*	*F*	*p*
Condition	1	0.0039	0.00039	0.0018	0.9666
	49	106.9569	2.1828		
Time	1	7.1394	7.1394	10.4895	0.0022**
Interaction	1	0.0022	0.0022	0.0032	0.9552
	49	147.4608	0.6806		

Looking at the incorrect answers from both groups on both the pre- and posttests, the number of incorrect answers in both groups decreased on the posttests for fall and rise, but not for high fall. These numbers are shown in Table 5.

Table 5. ***Changes in Missed Answers on Perception Tests for Both Groups***

	Treatment		Contrast	
	Pretest	posttest	Pretest	posttest
Statement (F)	9	2	10	1
Question (R)	10	5	13	6
Enthusiasm (HR)	8	7	5	7

Production Tests

Rater evaluations of both groups' posttest production data were examined using Fisher's Exact Test to see if there was any statistically significant relationship between condition and performance evaluations. Fisher's Exact Test was used instead of a standard chi test because the expected frequency of incorrect answers was lower than five. None of the observed results were significant at the .05 alpha level, fall, $p = 0.63$; rise $p = 0.53$; and high fall, $p = 0.57$. The 2x2 contingency tables for the three intonation patterns can be found in Table 6.

Table 6. ***Results of Fisher's Exact Test Comparing Both Group's Posttest Production Data***

	Fall		Rise		High Fall	
	Yes	No	Yes	No	Yes	No
Contrast	21	4	19	6	10	15
Treatment	22	4	19	7	10	16

Rater evaluations of the treatment group's production data from the pre- and posttests were analyzed using McNemar's Test for Paired Samples to compare differences in the number of students in the treatment group whose utterances were rated as accurate. The tests revealed that the number of students rated accurately did not change significantly for any of the intonation patterns: rise, $\chi^2(1, N = 26) = 0.14$, $p = 1$; fall, $\chi^2(1, N = 26) = 0.50$, $p = .72$; and high fall, $\chi^2(1, N = 26) = 0.14$, $p = 1$. The observed frequencies for each of the intonation patterns can be found in Table 7.

Table 7. 2 *x* 2 *Contingency Table of Treatment Group's Production Data*

		Posttest					
		Fall		Rise		High fall	
		Yes	No	Yes	No	Yes	No
Pretest	Yes	17	3	15	3	7	4
	No	5	1	4	4	3	12

Student Attitudes

Students were given an 11-item questionnaire after the posttest. (See Appendix A.) The questionnaire had ten items addressing three areas: attention during perception and production activities in English, past English pronunciation learning experience, and impression of the videos. The items used a 4-point Likert scale with no midpoint. Here we focus only on impressions. The descriptive statistics are shown in Table 8.

Table 8. *Impressions about Studying Pronunciation through Teacher-Created Videos*

	Enjoyed watching videos	*Want to watch again*
Treatment	2.96 (0.93)	2.96 (0.89)
Contrast	2.65 (0.8)	2.58 (0.9)

The final item on the questionnaire was open-ended and asked students to freely write what they felt about the videos. The comments (Appendix B) were translated and key words identified, categorized, and counted. A large majority of responses (n = 19) contained positive key words, which were organized under 4 themes: interest, difficulty, instructiveness, and awareness. Responses (n = 6) with negative key words were organized under two themes: interest and difficult. Respondents often had more than one key word. The counts for these key words are shown in Table 9. The comments can be seen in Appendix B.

Table 9. Positive and Negative Comments from Students in the Treatment Group

	n	Interest	Difficulty	Instructive-ness	Awareness-raising
Positive	20	5	7	4	4
Negative	5	3	3		

Discussion

Comparisons of the treatment and contrast group's scores on the perception posttests did not differ significantly, and differences between groups in improvement on the posttest were also not significant. Scores in both groups were high, with a mean score nearly perfect on both tests, implying a ceiling effect and suggesting our 24-item test was too easy for participants in both groups. Examining wrong answers on the pre- and posttests revealed that the number of wrong answers in both groups decreased for rising and falling tones. This was untrue in the case of the high-falling intonation used with enthusiastic exclamations.

It is possible that students in both groups were primed to notice intonation due to the earlier description task. In addition, rising and falling tones are very common both in English and in Japanese and although pitch contours for Japanese have a narrower range, six years of required English in secondary school plus exposure to spoken English through various media could have made these two patterns rather easy to identify. The high-falling pattern could have been successfully identified by students through a process of elimination, but its high pitch contour could have been a little troubling without the nonverbal cues that often occur with it. With the help of those cues, students should not have a problem identifying the pattern outside of the classroom.

The production of appropriate intonation patterns is a more challenging problem. Results from the treatment group's production showed that a one-time intervention with no further activities or guidance is not enough to influence changes in student production. Although there was not a statistically significant difference between the two groups or within the treatment group, high-fall performance was slightly different from the other two sentence types. The number of the utterances judged as having an accurate tone was only half for both groups and the figures show that half the students remain inaccurate even after intervention. In addition, inaccurate patterns were almost double the numbers of fall or rise. This shows students certainly have difficulty in producing high fall. As with perception, students are most likely already familiar with creating rising and falling intonation patterns, but the high-falling pattern is not commonly taught in English classrooms. Furthermore, if high fall is hard to perceive, it only follows that it would also be hard to produce.

Additionally, according to (Rogerson-Revell) "conveying attitude through intonation involves a lot more than tone choice, for example, voice quality, speed, loudness, pitch range, [and] non-verbal features." (p. 193) Our materials focused solely on intonation patterns, or as Celce-Mucia, Brinton, & Goodwin (1996) call them, "the entire melodic line" of an utterance (p. 184). Different results might be possible if the videos were preceded by other activities to raise students' awareness and experience with pitch on single words as well as word- and sentence-level stress. This again shows the need for a structured curriculum to address the English pronunciation of remedial English learners.

One favorable result from this study was found in the data related to student attitudes about the videos. Encouragingly, the majority of the students thought the videos were interesting and wanted to use videos to study pronunciation again. Perhaps more importantly, students mentioned in the comments that the videos helped introduce them to an aspect of the English language that they hadn't considered before. If this awareness is maintained, students might begin to notice these intonation patterns in their other classes or in bits of spoken English that they have encountered, gradually leading to improvements in perception and production.

Conclusion

Although no statistically significant changes in perception and production were found in this study, the results suggest some concrete changes in our methodology. Rather than focusing on the most common intonation types, additional materials should be made that focus on other examples of intonation being used to express other attitudinal functions, such as friendliness, sarcasm, or disinterest. It may not be useful for English learners to spend time on trying to learn to produce these intonation patterns, but failure to notice the intended meaning of an utterance can lead to communication breakdown. Moreover raising awareness of English intonation may encourage students to pay more attention to supra-segmental features of speech, preparing students to answer questions about speakers' feelings on major listening tests. Finally, since syllable and stress play a role in intonation as well, materials aimed at raising students' awareness of these features and contextualization intonation patterns within a larger framework of English phonology could lead to changes in perception and production of sentence-level intonation. However, further classroom research is needed.

Acknowledgements

The authors would like to express their gratitude to Professor Daisuke Nakanishi for providing timely advice and assistance with the statistics in this paper. The authors would also like to thank Jacob Lawrence at Hiroshima Bunkyo University for his help in creating the intonation videos.

References

Abe, H. (2007). The Effect of Interactive Input Enhancement on the Acquisition of English Connected Speech by Japanese College Students. In M. Ashby & J. Maidment (Eds.) *Phonetics Teaching and Learning Proceedings*. PTLC 2007 London. Retrieved from https://www.ucl.ac.uk/pals/study/cpd/cpd-courses/ptlc/proceedings_2007

Abe, H. (2011). Effects of Form Focused Instruction on the Acquisition of Weak Forms by Japanese EFL Learners. In W. S. Lee & E. Zee (Eds.) *Proceedings of the 17th International Congress of Phonetic Sciences*. ICPhS-11 Hong Kong, 184–187.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages*. Cambridge: New York.

Ellis, R. (1997). The empirical evaluation of language teaching materials. *ELT Journal, 51*(1), 36–42.

Fraser, H. (2006). Helping Teacher Help Students with pronunciation: A cognitive approach. *Prospect, 21*(1), 80–96.

Jenkins, J. (2000). *The Phonology of English as an International Language*. Oxford: Oxford University Press.

MacDonald, S. (2002). Pronunciation–Views and Practices of Reluctant Teachers. *Prospect, 17*(3), 3–18.

Pennington, M. (1999). Computer-aided pronunciation pedagogy: Promise, limitations, and directions. *Computer Assisted Language Learning, 12*(5), 427–440.

Porter, M. (2012). Developing an online environment to enable the independent learning of English pronunciation. *Learning Learning, 19*(1), 21–34.

Rogerson-Revell, P. (2011). *English Phonology and Pronunciation Teaching*. Continuum: London.

Scovel, T. (2000). A critical review of the critical period research. *Annual Review of Applied Linguistics 20*, 213–223.

Wells, J. C. (2006). *English Intonation: An Introduction*. Cambridge University Press, New York.

Wrembel, M. (2005). Metacompetence-oriented Model of Phonological Acquisition: Implications for the Teaching and Learning of Second Language Pronunciation. *Phonetics Teaching and Learning Proceedings*. PTLC 2005 London, 1–5. Retrieved from http://www.phon.ucl.ac.uk/home/johnm/ptlc2005/ptlc2005_proc_index.htm

Appendix A: Post-test Survey (originally in Japanese)

Past Pronunciation Learning

1. I have learned about English segmental sounds before.
2. I have learned about English intonation before.

Attention

3. I pay attention to segmental sounds in listening to English.
4. I pay attention to intonation in listening to English.
5. I pay attention to segmental sounds in speaking English.
6. I pay attention to intonation in speaking English.
7. From now on I would like to pay attention to segmental sounds in listening to English.
8. From now on I would like to pay attention to intonation in listening to English.

Impression

9. I enjoyed watching the videos.
10. I want to watch the videos again.

Appendix B: Students' Comments (originally in Japanese)

Positive	-Interesting. -Easy to understand! -Interesting. -They were easy to understand. -Instructive. -Not interesting, but instructive. -The pronunciation was good and easy -I understood that differences are based on region. -They were exaggerated, but easy to understand and interesting. -It was an easy sentence, so it was easy to understand. -I thought that impressions change because of intonation and loudness. It would have been hard to understand had it been audio only, but I could understand facial expressions and mannerisms, so the differences were interesting. -The foreigners in the video had many facial expressions, so it was interesting. I'd like to see it again. -The pronunciation was clear and careful. -I was taught something I had never thought about before, so I enjoyed it. -It was good that it was easy to understand. -It was instructive because I never paid attention to it when -I listened in high school or in regular classes. -I was able to become aware of the differences between English and Japanese rather than study English. -This was never explained in detail when I was a high school student, so it was instructive. -I really realized that I don't put enough effort into my own pronunciation and intonation. I want to watch the video many times and overcome this.
Negative	-A little hard to understand. -It was interesting because there was a video, but I felt the contents weren't that great. -I didn't really understand. Not interesting. -I didn't really understand.