The CATESOL Journal

JOHN M. LEVIS
GRETA MULLER LEVIS
Iowa State University, Ames

Teaching High-Value Pronunciation Features: Contrastive Stress for Intermediate Learners

Pronunciation features are not equal in how they affect listeners' ability to understand. Some are low value, while others are high value. This study explores whether contrastive stress is high value. Previous research has shown that identification of contrastive stress is learnable (Pennington & Ellis, 2000), and that explicit teaching about contrastive stress patterns can improve production for advanced learners (Hahn, 2002; Muller Levis & Levis, 2012). To test whether instruction on contrastive stress improved comprehensibility and fluency in spontaneous speech, we developed a 3-week class for intermediate ESL learners, whose pre- and posttest productions were rated by native listeners. Ratings for fluency showed no improvement. Ratings for comprehensibility significantly improved for the experimental group while control participants showed no improvement. Improvement resulted both from better contrastive stress and greater comfort with producing grammatical frames to express the contrasts. The article concludes by discussing the importance of high-value pronunciation features for improved comprehensibility.

In a video of his younger than 2-year-old son, the linguist Deb Roy (personal communication, n.d.) shows a child whose contrastive stress is perfectly expressed. We hear the child saying, *"That's a BLUE car. That's a GREEN car. That's a ORange car. That's an OTHer orange car."* In these utterances, the child shows command of a basic prosodic pattern of English by calling attention to contrasting lexical items when everything else in the sentence pattern is the same. He also demonstrates anticipation of a pattern of contrasts by placing contrastive stress on *BLUE* before any other colors are mentioned.

Sentence stress, also called nuclear stress, primary phrase stress, sentence focus, prominence, and a variety of other names, uses prosody to draw listeners' attention to a certain word in a spoken phrase or sentence. In this article, we will use *sentence stress* as our term to refer to this feature. Sentence stress is typically indicated in careful speech by pitch movement, greater syllable length, and increased loudness on the last content word (e.g., noun, verb, adjective, adverb) of each sentence, and it is often represented by capital letters, as in the examples below. Except in uncommon cases (Allerton & Cruttenden, 1979), the most typical placement of sentence stress is on the last content word of a sentence, as the example in (1). This use of sentence stress, known as *unmarked sentence stress,* is so common in spoken English that Crystal (1969) found that up to 90% of spoken phrases/sentences stress the last content word. Contrastive stress, a second function of sentence stress, can be seen in (2) and (3). The example in (2) reflects a contrast that is explicitly present. In (3), we see an implied contrast, which although not present in the sentence, is retrievable or readily understood by a listener.

(1) *It's MONday. I have CLASS. So I can't GO.*

(2) *It's on the BOTTOM shelf, not the TOP shelf.*

(3) *You'll find it UNder the table.*

This study focuses on contrastive stress as shown in (2), in which the contrasting items are explicitly said.

Literature Review

Sentence Stress

Sentence stress is essential for intelligible communication in English no matter the teaching context (Jenkins, 2002). Besides unmarked stress, sentence stress serves other functions in spoken English. For example, it marks new information in discourse (Halliday, 1967), promotes listener comprehension (Hahn, 2004), and calls attention to contrasts (Bolinger, 1961). In general, every spoken phrase or sentence will have one word that is marked for sentence stress. In the unmarked case, sentence stress typically occurs toward the end of the phrase/sentence and is also the beginning of the pitch movement at the end phrase or sentence. For example, in (1), the pitch will jump on MON- and then fall to a low pitch. The fall is what is often referred to as falling intonation. In marking new information, sentence stress placement may be at the end of the sentence or somewhere else, and

a word may be marked with sentence stress because another word has been said already, that is, it is now given information. In (4), Person 1 starts by putting sentence stress on *TUESday*, but the response by Person 2 avoids putting sentence stress on *Tuesday* and instead moves it to *afterNOON*. Information structure is an important function of sentence stress, but we do not address it in this article.

(4) Person 1: So I made our appointment for after dinner on TUESday.

Person 2: Person 2: But I told you I could only do afterNOON on Tuesday.

The last major function of sentence stress is to call attention to contrasts. *Contrastive stress* (the term we will use to talk about this function of sentence stress) is not always easy to predict because of its dependence on the mind of the speaker (Bolinger, 1972). Contrastive stress placement does not match the expected patterns of unmarked sentence stress (i.e., stress falling on the final content word of the utterance). Contrastive stress may also show up on non-content words such as pronouns or prepositions because they are important for the contrastive meaning the speaker is conveying, as in (5) and (6).

(5) It's IN the dresser / not ON the dresser. (prepositions in contrast)

(6) I gave it to HIS mother, not HERS. (pronouns in contrast)

Contrastive Stress

Besides its pronunciation, contrastive stress is often signaled by lexical and grammatical features (Cowles, Walenski, & Kluender, 2007). For example, Theune (1999) argues that the existence of alternative items in speech (as in choice questions) and grammatical parallelism are strong predictors of contrastive stress. Contrastive stress is also associated with "focus-sensitive particles" (Zimmermann, 2008, p. 156) such as *even*, (*an*)*other,* and *only*. For example, the small child's use of "*That's a BLUE car. That's a GREEN car. That's a ORange car. That's an OTHer orange car"* at the beginning of this article has grammatically parallel sentences and a focus-sensitive particle (*another*).

Contrastive stress may also enhance the noticeability of a set of contrasts in the discourse (Cowles et al., 2007) by signaling the existence of a set of logically related contrasts (e.g., car colors in the previous paragraph), even if the set is not explicitly mentioned as in (7).

The sentence stress on *ONly* in Person 2's turn signals a contrast with a set of other people in addition to Jim that the speakers are likely able to identify, even though the others are not mentioned.

(7) Person 1: Who can pick her UP?
Person 2: Jim's the ONly one free.
Person 3: SEriously?

Contrastive stress is multifunctional in English (Boer, 1979). It can be used to signal corrections to previous information, such as "But I didn't say ENcourage, I said DIScourage" (Gökgöz, Bogomolets, Tieu, Palmer, & Lillo-Martin, 2016; Zimmermann, 2008). When the sentence stress signals a possible contrast occurring at the beginning of a speaker's contribution (e.g., *It's a BLUE car …*), it helps listeners anticipate the forthcoming contrast more effectively (Ito & Mester, 2012). Finally, contrastive stress has been reported to help resolve interpretations of ambiguity in speech (e.g., "Jim yelled at JOE, but HE didn't pay attention"), especially with relation to which noun phrase is referred to by a pronoun (Cowles et al., 2007).

These three primary uses of contrastive stress (correction, anticipation, and resolution of ambiguity) suggest that the feature is important for L2 learners both for perception and production. For perception, identification of contrastive stress is important in order for learners to interpret the speaker's intended meaning. For production, it is important for them to clearly mark what they are talking about so that listeners can attend to meaning more readily. Successful use of contrastive stress thus holds promise for improving the comprehensibility of L2 speech.

Comprehensibility

This study looks at how instruction on contrastive stress affects the comprehensibility of nonnative English speakers. By comprehensibility, we mean the amount of work that listeners do in understanding speech (Munro & Derwing, 1995). Comprehensibility should not be confused with concepts that are indirectly related to it, especially intelligibility (that is, whether listeners actually understand the words, meaning, or intentions of speech) or accentedness (the perceived difference in speech compared to a particular model).

Comprehensibility is not related to judgments involving only pronunciation. Isaacs and Trofimovich (2012) demonstrated that 18 of the 19 speech measures they examined correlated with comprehensibility ratings—including measures of discourse complexity, fluency,

grammatical accuracy, lexical richness, and pronunciation. In their study, they sought to create an oral testing rubric that would be usable by English as a second language (ESL) teachers teaching French students. Of the five features finally incorporated into the rubric, word stress was the only phonological feature included, and it was the only feature that distinguished the spoken language production of beginning, intermediate, and advanced learners. In other words, word stress was, for the French learners of English in the study, a high-value feature that correlated with differences in spoken proficiency.

High-Value Pronunciation Features and Contrastive Stress

This study examines whether teaching contrastive stress can lead to improvement in spontaneous speech, both in how easily the speech is understood (i.e., comprehensibility) and in how smoothly it is produced (i.e., fluency). In other words, we examined whether contrastive stress is a high-value feature. By *high value,* we mean that changes in the feature, even if everything else remains the same, are likely to make listeners rate speech more positively than if the feature is not correctly used. This hypothesis assumes that a particular feature can, by itself, help a speaker be better understood; that is, it can make the speaker more comprehensible. This is an underlying assumption of an intelligibility-based approach to pronunciation teaching: Some features will affect how listeners process foreign-accented speech more than others. Gilbert (2001) describes this approach to pronunciation teaching by analogy to the medical practice of triage, in which more important medical needs are addressed before less important ones.

We believe, from our previous research, that contrastive stress has promise as a high-value feature that can improve learners' comprehensibility. Previous research has shown that contrastive stress is learnable, both for perception (Pennington & Ellis, 2000) and production (Muller Levis & Levis, 2012). In the Pennington and Ellis study, Cantonese learners of English were asked to attend to the form and meaning of four intonational features (tag questions, contrastive stress, juncture, and compound nouns). The learners were asked to identify whether they had already heard a sentence such as "She's going, isn't she?" For the initial part of the study, learners received no information about intonational features. Their answers on the recognition task were judged as correct only if the learners identified the sentence both in its grammar and in its intonation. The researchers found that on this task, learners paid no attention to intonation cues, paying attention instead only to the lexical and grammatical features of the sentence. As a result, they performed well in remembering the lexical and grammatical content of sentences, but very poorly in re-

gard to remembering or identifying the prosodic differences. In the second part of the study, learners received instruction about each of the intonational features before being asked to perform the recognition task again. They showed improvement in recognition only for sentences with contrastive stress distinctions, suggesting that, for perception, contrastive stress was learnable but that the other features were not.

Contrastive stress also seems to be learnable for production. In a study of sentence stress learning, Hahn (2002) looked at 36 participants' pretest (Time1), posttest (Time2), and delayed posttest (Time3) production accuracy for nine sentence stress patterns, three of which involved contrastive stress (specifically, contrasts in choice questions, either/or statements, and shifting stress on "you" pronouns as in *How ARE you? Fine, How are YOU?*). The pronunciation instruction that the participants received during the course included the nine patterns and many other pronunciation features, including word stress, intonation, and segmentals. Instruction took place during a 15-week semester at the university level. All sentence stress patterns showed significant improvement at Time2. For five of the patterns (including the three contrastive stress patterns), participants showed levels of learning above pretest performance at Time3. However, at Time3, the other four patterns retreated from Time2 improvement levels and went back to Time1 levels or worse. Since all tests used controlled language production (reading aloud), the results suggest that only some patterns, including those with contrastive stress, are likely to maintain some level of improvement beyond classroom instruction.

In another study involving production, Muller Levis and Levis (2012) studied whether a group of advanced L2 learners consisting of international teaching assistants (ITAs) could improve their use of contrastive stress in controlled contexts using language appropriate to a professional academic context. Using a pretest-posttest design, the subjects took both a recognition test and completed an oral reading of sentences containing contrastive stress. They were then taught about contrastive stress and practiced it during four hours of class time (three 80-minute class periods). Instruction included production, perception, and prediction practice, as well as instruction on grammatical and lexical features related to using contrastive stress. After instruction, the ITAs took an identical posttest. They had high recognition scores at pretest and at posttest, indicating that, unlike the subjects in Pennington and Ellis (2000), they were able to identify the placement of contrastive stress. For production, their performance improved significantly, and the number of accurate contrastive stress placements went from below 50% to nearly 75% accurate. In sum,

with a modest amount of instruction, ITAs were able to improve their controlled reading of sentences with contrastive stress. The study did not test whether listeners heard the productions as being more comprehensible, nor was improvement measured for spontaneous speech, both of which are important for our present study.

In a follow-up study, Muller Levis, Levis, and Benner (2014) showed that intermediate learners using nonacademic language also seemed to improve their command of contrastive stress after instruction. Instead of using academic language (which could be overly complicated for students at this level), students compared pictures, described other contrasting illustrations such as illustrated weather forecasts (http://www.weather.gov/), and read a variety of arithmetic equations and conversational sentences.

The Effect of Instruction

Pronunciation instruction is successful in improving L2 learners' pronunciation. This is the inescapable conclusion of several recent analyses of instructional studies. Lee, Jang, and Plonsky (2015) conducted a meta-analysis of 86 widely varied studies of instruction on pronunciation improvement, 83 of which had English as the target language. In almost all cases, the studies reported that learners became more accurate in their pronunciation as a result of being taught. Sometimes they were taught single features, such as the initial consonants in *rock-lock*, and sometimes they were taught multiple features. Overall, there was a large statistical effect for instruction, and the effects of instruction were particularly strong when instruction took place over an extended period of time, when feedback was given on pronunciation, and when production tasks were carefully controlled. In other words, longer instructional treatments are better, especially if learners receive appropriate corrective feedback (Saito & Lyster, 2012). In addition, learners are especially likely to show improvement when the tasks allow them to concentrate on their accuracy.

This last feature is a limitation of many of the studies in the meta-analysis. Ultimately, learners need to improve their pronunciation not only in controlled speech tasks (e.g., repeat-after-me or reading aloud), but also in their spontaneous speech. Transfer from controlled to free speech is particularly challenging in pronunciation learning, and frameworks for communicative pronunciation instruction distinguish between activities that are primarily form based and those that are meaning based (Celce-Murcia, Brinton, & Goodwin, 2010). Pronunciation improvement is relatively easy in controlled contexts. Improvement that can be heard in free speech is a far more challenging goal. This is the message of a narrative analysis by Thomson and

Derwing (2015), which looks at many of the same studies examined by Lee et al. (2015). Rather than examining the statistical evidence for improvement, their narrative analysis looked at how instruction occurred and at which types of improvement were demonstrated. The narrative analysis points out that very few instructional studies have examined improvements in comprehensibility or intelligibility; most simply looked at accuracy. The authors argue that accuracy alone is an insufficient goal. Instead, improvements in comprehensibility or intelligibility should be considered the gold standard for instructional studies because learner improvement is important only if it is something that average listeners notice. The analysis also points out that many studies demonstrate improvement for features that are less likely to affect comprehensibility and intelligibility. Improvements in some features simply do not make much difference in how listeners evaluate speech. For example, high functional load segmentals (contrasting sounds that have many minimal pairs, such as the initial sound in *lead-read*) are more valuable for improved intelligibility and comprehensibility than low functional load ones (those with few minimal pairs, such as the vowels in *Luke-look*; see Brown, 1988).

Gold-standard instructional studies typically show improvement in comprehensibility for spontaneous speech only when instruction takes place over an extended period and involves suprasegmentals (i.e., stress, rhythm, and intonation). In a study of different effects of segmentals or suprasegmentals on comprehensibility, Derwing, Munro, and Wiebe (1998) showed that instruction on suprasegmentals and segmentals both led to improvements in comprehensibility for reading aloud, but that only instruction on suprasegmentals led to an improvement in the comprehensibility of learners' spontaneous speech.

Research Questions

The goals of this study were to determine whether intermediate-level ESL learners could, after a short instructional period, learn to produce contrastive stress more successfully, such that their post-instruction production would be heard as having improved in comprehensibility and in fluency. This leads to two research questions.

1. Does the fluency of spontaneous speech improve after instruction on contrastive stress for intermediate-level learners?
2. Does the comprehensibility of spontaneous speech improve after instruction on contrastive stress for intermediate-level learners?

Methods

To address whether the comprehensibility of learners' free speech improved after instruction on contrastive stress, we developed a three-week class to teach contrastive stress (six sessions at 25 minutes/session) to intermediate-level students studying English at an Intensive English Program at a large Midwestern university.

Pretest-Posttest

The pretest and posttest tasks were identical. They included oral reading of sentences (32 sentences), math equations (e.g., 10-2=8; 10+2=12), and a short passage. In addition, students completed tasks that were designed to promote spontaneous speech. They compared daily weather forecasts using visual illustrations and six sets of two pictures in which differences were limited and noticeable. In the first stage of the study, we examined the oral reading tasks (Muller Levis et al., 2014). For this study, we asked listeners to rate three of the picture comparisons for comprehensibility. We used the ratings to examine whether there was improvement for participants' spontaneous speech based on the comprehensibility ratings. The pictures used to elicit the spontaneous speech are shown in Figures 1, 2, and 3. Participants were simply asked to compare the two pictures. Their comparisons were unscripted and were subsequently used for the rating task reported in this study.

Participants

In this study, two groups of participants took part in classroom instruction, an experimental group and a control group. All participants were studying English in the intermediate-level listening-speaking classes of an Intensive English Program at a US Midwestern university during the spring semester (January-May) of 2014. The control group (n=9), composed of students from Oman, Korea, and China, did not receive instruction on contrastive stress but completed other language tasks during their class. The experimental group (n=18) received instruction on contrastive stress. Participants in the experimental group came from Oman, Korea, Thailand, Colombia, and Japan. Of the 18 experimental group participants, 10 finished the posttest. Of the 9 in the control group, only 3 finished the posttest. The drop-off in participation was related to unexpected problems related to spring semester scheduling. Participants were enrolled in an Intensive English Program and took the institutional TOEFL test two weeks before spring break, around the beginning of March. After this, attendance became unexpectedly sporadic, and many did not come to do the posttest. Participants were not paid.

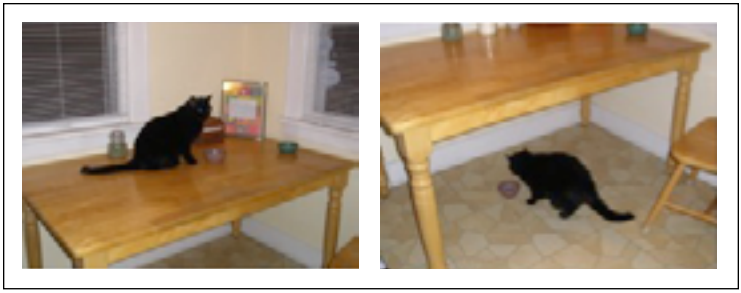

Figure 1. Two cats picture comparison.

Figure 2. Upstairs and downstairs picture comparison.

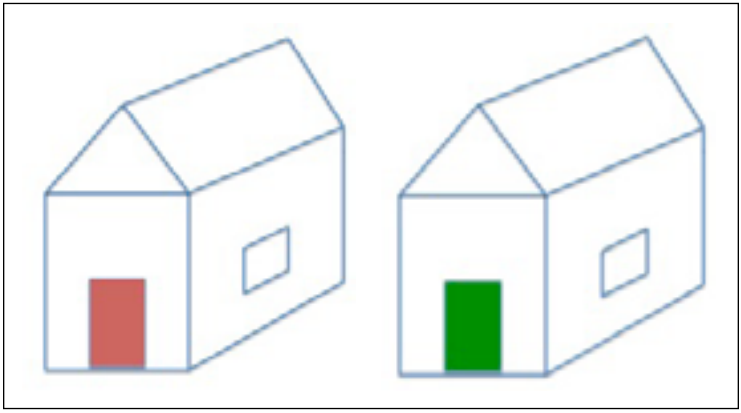

Figure 3. Two houses picture comparison.

Instructors

The two instructors were the first researcher and a senior lecturer. Both were experienced ESL teachers with at least 25 years of experience in teaching pronunciation, and they coordinated their teaching lessons, using the same materials in the same order on the same days.

Instruction

The experimental group was taught in six 25-minute lessons, two times a week over three weeks, for a total of 150 minutes of instruction. The instruction was part of their normal listening/speaking classes. The rest of the time in class, they were taught by their regular instructor. Pronunciation instruction involved traditional exercises such as reading aloud and calling attention to prosodic prominence patterns, but it also included practice using pictures with limited differences to promote better spontaneous speech, as in Figure 4. These pictures prompted comparisons that naturally elicited contrastive stress and could be used with multiple students. For example, comparisons included "the first house is small but the second house is big" and "the red house is on the right but the gray house is on the left."

Instructions: Give a sentence comparing the two houses.

Figure 4. Sample picture task from instruction.

Instruction also included explicit attention to lexicogrammatical patterns that are used to compare and contrast (e.g., *other*, *first/second*, *on the left/right*), and grammatical parallelism. In Muller Levis and Levis (2012) and Muller Levis et al. (2014), we found that to express comparisons efficiently in speech, contrastive stress needed to be taught in conjunction with these patterns. These lexicogrammatical

patterns took the form of grammatical frames provided during class and helped make productions of contrastive stress more noticeable. Some grammatical frames for comparison that we used are shown in Table 1.

Table 1
Grammatical Frames That Co-Occur With Contrastive Stress

Comparison/ contrast words	***Example***
"but"	The first house is gray, **but** the second house is red. The house in the first picture has _____, but the house in the second picture has _____.
"while"	One house is _____, while the **other** house is _____.

The frames were also used to bridge between sentence reading and free speech by pairing them with pictures, as in Figure 5. In these tasks, we showed the visual on PowerPoint along with the grammatical frame and asked students to compare the two bears. When a student compared the bears, we provided feedback on his or her performance and encouraged a correct form. We then asked another student to give us another sentence comparing the two bears, and so forth until several students had a chance to answer.

Figure 5. Activity using a grammatical frame to practice contrastive stress. Students were provided with the frame "the bear on the right … while …" and asked to compare the bears.

Rating Procedures

Raters were 20 undergraduate linguistics students, all native speakers of American English. All were paid for their participation. They were shown the three picture comparisons, one at a time, before rating randomly ordered pretest and posttest descriptions for each picture. The first picture description involved two cats (on and under the table, Figure 1), followed by two doors (Figure 3) and the upstairs-downstairs pictures (Figure 2).

Each set of ratings followed the same procedure: Raters were shown the two pictures that the students had described, and then as a group they rated three examples from students who had not completed the posttests, after which we discussed the ratings and they asked questions. They then rated 29 picture descriptions. These included the 13 pre- and posttest descriptions for the experimental and control participants, two nonparticipants, and one description from a native English speaker. Pre- and posttest descriptions for each picture were presented in a random format to the whole group at one time. Each section took about 6-7 minutes to complete. After the comprehensibility ratings, raters again listened and rated the same picture descriptions in the same order, but they listened for fluency. Altogether, there were 1,560 total ratings (20 raters x 3 pictures x 2 tasks x 13 speakers), 300 pretest and 300 posttest experimental ratings, 90 control pretest and 90 control posttest ratings for comprehensibility, and the same number for fluency) for the experimental and control subjects.

Ratings for comprehensibility used a 9-point Likert scale, from 1 *(very difficult to understand)* to 9 *(very easy to understand).* This is the opposite orientation from Munro and Derwing (1995) because we tested the task with nonraters and found they did better when 9 was associated with higher ratings, meaning improved comprehensibility. Fluency ratings used a similar 9-point scale from 1 *(extremely dysfluent)* to 9 *(extremely fluent),* again based on the fluency rating scale of Derwing, Munro, and Thomson (2007). Although they used a 7-point Likert scale, we used a 9-point scale to be consistent with the comprehensibility scale.

Analysis

Because the number of participants was small and because there were only 20 raters, we used the Wilcoxon 2-tailed nonparametric signed-rank test to analyze the ratings. This test was appropriate because we could not assume that the sample was normally distributed.

Results

The first research question explored whether the fluency ratings for the spontaneous speech of intermediate-level learners improved as a result of instruction on contrastive stress. This question was originally suggested when we played the recordings of pre- and posttest picture comparisons during a conference presentation. One of the audience members suggested that the posttest recordings sounded more fluent, leading us to ask this research question. Table 2 shows the means and standard deviations for the pretest and posttest, but we did not find a significant difference in fluency ratings for the experimental group (z=.889, p=.374). Descriptively, two of the three picture descriptions, comparing cats and comparing stairs (Figures 1 and 2), had a higher posttest rating while one, comparing doors (Figure 3), had a lower average. When the findings for all three picture comparisons were combined, the posttest average was very close to the pretest.

Table 2
Fluency Ratings for Experimental Group

	Pretest ***Mean / Standard deviation***	***Posttest*** ***Mean / Standard deviation***
Cats	5.665 /.75757	5.96 /1.069
Doors	5.66 /1.368	5.395 /1.217
Stairs	5.08 /1.898	5.60 /1.295
Combined	5.467 /.625	5.653 /.789

Note. Rating based on a Likert scale from 1-9, with 9 being highest.

The second research question explored whether comprehensibility improved as a result of instruction. Table 3 shows the means and standard deviations for comprehensibility ratings; there was a significant improvement at posttest using the Wilcoxon signed-rank test (z=2.09, p=.037). All picture comparisons were rated higher on average at the posttest, indicating that the listeners found the speakers' posttest comparisons easier to understand. The gain in the ratings matched what we noticed in listening to the picture comparisons, that is, that the learners often more skillfully expressed the contrasts after instruction.

Finally, the three control group speakers who finished the posttest seemed to show no difference in ratings. Their overall comprehensibility ratings at pretest were 4.7 and at posttest 4.87, while fluency ratings were 5.23 at pretest and 5.53 at posttest. Because of the small number of subjects, we did not compute significance levels.

Table 3
Comprehensibility Ratings for Experimental Group

	Pretest *Mean / Standard deviation*	*Posttest* *Mean / Standard deviation*
Cats	5.659 / .943	6.035 /.996
Doors	4.845 / 1.241	5.295 / 1.262
Stairs	4.30 / 1.399	5.05 /1.447
Combined	4.991 / .771	5.459 /.996

Note. Rating based on a Likert scale from 1-9, with 9 being highest.

Discussion

This study examined whether instruction on contrastive stress would result in improved comprehensibility and fluency for intermediate learners of English. The results show that the instruction resulted in improved comprehensibility but not fluency. This suggests that instruction on certain pronunciation features can lead to L2 speech that listeners find easier to listen to. Previous research (Derwing et al., 1998) has shown that global instruction on suprasegmentals over a longer period of time led to improved comprehensibility in spontaneous speech. The results of this study suggest that a shorter period of instruction on particular suprasegmentals can also result in improved comprehensibility. We make no claim that the spontaneous speech of these learners would show the same improvements in comprehensibility if the speaking task did not narrowly target contrastive stress, but the results are encouraging because they indicate that appropriate unscripted speech tasks for intermediate learners can show comprehensibility improvements as a result of instruction.

Confirming previous research, our results indicate that contrastive stress was learnable in a classroom context in a relatively short period. Learnability has been put forth as a criterion for why pronunciation features should or should not be prioritized (Jenkins, 2000, 2002). In a field such as pronunciation teaching, however, such assertions have little evidence to support them beyond a teacher's experience. Oftentimes, teachers notice that L2 learners may improve in controlled contexts within the classroom, but the improvement typically does not transfer to speech in which attention is primarily given to meaning. Such lack of transfer from controlled speech to free communication says little about whether features are learnable, but only that they are not learnable within a particular period of time or with a particular teaching approach. In general, we simply do not know to what extent features are learnable because of the limited research that

looks at improvement over a longer period. Hahn (2002) looked at whether nine sentence stress patterns were learnable. Learners in her study were otherwise advanced learners at an American university. Their pronunciation was poor enough that they were required to take a semester-long pronunciation class (3 classes/week, 15 weeks, a total of 45 hours of instruction) in which sentence stress was an important topic. Pronunciation practice was cognitively oriented (through explicit teaching of predictive rules) and used controlled practice (reading aloud, repetition). All patterns showed large improvements in production in controlled contexts at the end of the class, but in delayed posttests, four of the patterns showed no improvement over pretest production. Four showed significant increases in performance over the pretest (two of which involved contrastive stress patterns), but only one pattern (again, involving contrastive stress) showed continued improvement over the posttest.

These results suggest the difficulty of making learnability arguments at our current level of knowledge. If we look simply at the pretest-posttest results of Hahn's study, students learned the patterns very well. However, we have no information whether this learning transferred to free speech since it was not tested. We also see that improvement at the delayed posttest can be interpreted in various ways. Eight of the nine patterns decreased from posttest performance, but five patterns still showed improvements over pretest performance. These results can be used to make different claims about learning, but the results can be used only to suggest answers about whether teaching predictive rules can result in learners' producing correct patterns of pronunciation while reading aloud. Hahn's results also may indicate that focusing only on controlled practice is unlikely to influence free speech. Darcy, Ewert, and Lidster (2012) argued that while controlled practice is needed for improvement and for automaticity of production, it is not sufficient for transfer to spontaneous speech. For transfer, we also need practice that focuses on communicating meaning when pronunciation is essential to the communicative task (Darcy, 2018 [this issue]).

Our results also suggest that short periods of instruction can be effective not only in learning to pronounce particular features but in affecting comprehensibility in limited free speech contexts. In a similar study to ours, Gordon, Darcy, and Ewert (2013) taught classes in an Intensive English Program over three weeks (with three 25-minute pronunciation lessons per week). Their study compared three groups of 10 students receiving segmental-based, suprasegmental-based, or no explicit pronunciation instruction, with results showing that comprehensibility ratings improved only for the suprasegmental in-

struction group. It may be that suprasegmental instruction, because it applies over stretches of speech, is more likely to affect comprehensibility than segmental instruction, which addresses more localized errors. Thus in a short-term instructional setting, comprehensibility may be more quickly changed by attention to general speech habits and suprasegmentals (for more on this argument, see Firth, 1992, and McNerney & Mendelsohn, 1992). In the long run, however, instruction on segmentals, especially those more likely to lead to misunderstanding, is likely to be crucial as well (Derwing, Munro, & Wiebe, 1997, 1998).

The results also suggest that it is possible to identify pronunciation features as high value by controlling both what is taught and the types of spoken tasks used to elicit the feature in spontaneous speech. Using a general spontaneous speech task (such as a picture narrative or description of a favorite holiday) may not have revealed comprehensibility improvement because such a general task would not easily target the use of contrastive stress. While sentence stress in its other manifestations cannot be avoided (specifically the unmarked placement in which sentence stress goes on the last content word), contrastive stress is required in much more limited contexts and may be stylistically appropriate in others. In this study, the use of contrastive stress was required in the answers given by the L2 learners, and this allowed us to isolate the contribution of contrastive stress to comprehensibility.

Why did comprehensibility improve when fluency did not? According to previous research (Isaacs & Trofimovich, 2012; Kang, 2010; Trofimovich & Isaacs, 2012), comprehensibility ratings correlate with fluency-related features, pronunciation features, lexical/grammatical features, and the way speakers construct discourse (a story's cohesion, breadth, and depth). In this study, discourse features may not have contributed to ratings because the task did not especially promote breadth and depth. The relative brevity of the comparisons may also have been insufficient to contribute to fluency ratings. However, the use of contrastive stress (a pronunciation feature) and the lexical and grammatical strategies to present contrasts may have contributed to increases in comprehensibility even when fluency measures were not rated as improving. From this study, we do not have a way to tell whether pronunciation, grammar, or both contributed to improved comprehensibility. To elaborate, we noticed both of these features in different participants' productions that were rated higher at posttest. The first, a young man from Oman, compared the pictures in a rambling manner at pretest (8) but in a grammatically efficient way at posttest (9). His comparison sounded far better at posttest.

(8) *Here is uh two houses that have the same sizes but they are the different is just the door colors. The first one is uh … the first … the first one uh the door color is red and the second one the door color is green.* (Pretest)

(9) *There are two houses but the first one has a red door and the and the second one has a a green door.* (Posttest)

In the second example, a young woman, also from Oman, showed little difference in her grammar but enormous changes in her contrastive stress, shown in (11), which was marked in multiple locations with large pitch changes and duration changes on the prominent syllables. At pretest, contrasts were not noticeable, shown in (10).

(10) *The cat the cat /in in the second pictures /it's under / the TABLE / but in the first picture / it's on the TAble.* (Pretest)

(11) *The cat is ON the table, / the, the, but in / in the SEcond picture /the cat is UNder the table.* (Posttest)

Limitations

The study has several limitations. First, the ratings were done for only one of the tasks, the picture comparison. This task involved spontaneous speech, but the type of speech was limited. More extended free speech that required the use of contrastive stress may have given a more complete view of how much of an effect improved contrastive stress has on longer stretches of spoken discourse. A second limitation was, of course, the number of students. We chose these students because, according to the Intensive English Program placement test, they were similar in spoken proficiency. Initially, we had 27 students, but we did not anticipate the dramatic drop-off after the midpoint of the semester. This was especially problematic for the control group, in which six of the nine students did not complete the posttest despite our repeated attempts to record them. Third, the first languages of the students may have played a role in the results. A number of our students were from Oman and spoke Arabic as a first language. Arabic also has lexical stress and this may have facilitated learning of contrastive stress for them. We noticed during instruction that several of these students seemed to have "Aha!" reactions and quickly began to produce contrastive stress accurately and quite expressively. Finally, a future study should compare instruction on contrastive stress with and without instruction on lexical and grammatical features. This may help to distinguish the contributions of each feature to improved comprehensibility.

Authors

John Levis is founding editor of the Journal of Second Language Pronunciation, *the founder of the Pronunciation in Second Language Learning and Teaching Conference, and the co-developer of pronunciationforteachers.com. He is co-editor of several books, including the* Handbook of English Pronunciation, Social Dynamics in Second Language Accent, *and* Critical Concepts in Linguistics: Pronunciation. *He is also the author of the forthcoming Cambridge University Press book,* Intelligibility, Oral Communication and the Teaching of Pronunciation.

Greta Muller Levis is senior lecturer in TESL and Applied Linguistics at Iowa State University, where she also is adviser for undergraduate linguistics students. She teaches Introductory Linguistics and Second Language Acquisition. Her interests include pronunciation teaching. She has presented in the US, Czech Republic, Denmark, France, Canada, and Taiwan.

References

Allerton, D. J., & Cruttenden, A. (1979). Three reasons for accenting a definite subject. *Journal of Linguistics*, *15*(1), 49-53.

Boer, S. E. (1979). Meaning and contrastive stress. *The Philosophical Review*, *88*(2), 263-298.

Bolinger, D. L. (1961). Contrastive accent and contrastive stress. *Language*, *37*(1), 83-96.

Bolinger, D. (1972). Accent is predictable (if you're a mind-reader). *Language*, *48*(3), 633-644.

Brown, A. (1988). Functional load and the teaching of pronunciation. *TESOL Quarterly*, *22*(4), 593-606.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (with Griner, B.). (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). New York, NY: Cambridge University Press.

Cowles, H. W., Walenski, M., & Kluender, R. (2007). Linguistic and cognitive prominence in anaphor resolution: Topic, contrastive focus and pronouns. *Topoi*, *26*(1), 3-18.

Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge, England: Cambridge University Press.

Darcy, I. (2018). Powerful and effective pronunciation instruction: How can we achieve it? *The CATESOL Journal, 30*(1), 13-46.

Darcy, I., Ewert, D., & Lidster, R. (2012). Bringing pronunciation instruction back into the classroom: An ESL teachers' pronunciation "toolbox." In J. Levis & K. LeVelle (Eds.), *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching*

Conference (pp. 93-108). Ames: Iowa State University.
Derwing, T. M., Munro, M. J., & Thomson, R. I. (2007). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, *29*(3), 359-380.
Derwing, T. M., Munro, M. J., & Wiebe, G. (1997). Pronunciation instruction for fossilized learners. Can it help? *Applied Language Learning*, *8*(2), 217-235.
Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, *48*(3), 393-410.
Firth, S. (1992). Pronunciation syllabus design: A question of focus. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 173-184). New York, NY: Oxford University Press.
Gilbert, J. (2001). Six pronunciation priorities for the beginning student. *The CATESOL Journal*, *13*(1), 173-182.
Gökgöz, K., Bogomolets, K., Tieu, L., Palmer, J. L., & Lillo-Martin, D. (2016). Contrastive focus in children acquiring English and ASL: Cues of prominence. In L. Perkins, R. Dudley, J. Gerard, & K. Hitczenko (Eds.), *Proceedings of the 6th Conference on Generative Approaches to Language Acquisition–North America* (pp. 13-23). Somerville, MA: Cascadilla Proceedings Project.
Gordon, J., Darcy, I., & Ewert, D. (2013). Pronunciation teaching and learning: Effects of explicit phonetic instruction in the L2 classroom. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference* (pp. 194-206). Ames: Iowa State University.
Hahn, M. K. (2002). *The persistence of learned primary phrase stress patterns among learners of English* (Unpublished doctoral dissertation). University of Illinois at Urbana-Champaign, Urbana-Champaign, IL.
Hahn, L. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, *38*(2), 201-223.
Halliday, M. A. K (1967). *The intonation and grammar of British English*. The Hague, The Netherlands: Mouton.
Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition*, *34*(3), 475-505.
Ito, J., & Mester, A. (2012). Recursive prosodic phrasing in Japanese. In T. Borowsky, S. Kawahara, T. Shinha, & M. Sugahara (Eds.), *Prosody matters: Essays in honor of Elisabeth Selkirk* (pp. 280-303). Bristol, CT: Equinox.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford, England: Oxford University Press.

Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics*, *23*(1), 83-103.

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, *38*(2), 301-315.

Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, *36*(3), 345-366.

McNerney, M., & Mendelsohn, D. (1992). Suprasegmentals in the pronunciation class: Setting priorities. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 185-196). Oxford, England: Oxford University Press.

Muller Levis, G., & Levis, J. (2012). Learning to produce contrastive focus: A study of advanced learners of English. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference* (pp. 124-133). Ames: Iowa State University.

Muller Levis, G., Levis, J., & Benner, S. (2014, September) Contrastive stress can be learned – But can it be taught at lower levels? Paper presented at the Pronunciation in Second Language Learning and Teaching Conference, Santa Barbara, CA.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *45*(1), 73-97.

Pennington, M., & Ellis, N. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *The Modern Language Journal*, *84*(3), 372-389.

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Language Learning*, *62*(2), 595-633.

Theune, M. (1999). Parallelism, coherence, and contrastive accent. In *Proceedings of the Sixth European Conference on Speech Communication and Technology*. Baixas, France: International Speech Communication Association. Retrieved from https://pdfs.semanticscholar.org/33d9/e31d0895ce1a594c874944f1a98ec0649247.pdf

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, *36*(3), 326-344.

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from com-

prehensibility. *Bilingualism: Language and Cognition*, *15*(4), 905-916.
Zimmermann, M. (2008). Contrastive focus and emphasis. *Acta Linguistica Hungarica*, *55*(3-4), 347-360.