Developing interpreter trainees' speech comprehensibility: does nativeness of the instructor matter?

Yenkimaleki, M.; Heuven, V.J.J.P. van

TESL-EJ Teaching English as a Second or Foreign Language
The Electronic Journal for English as a Second Language

Developing Interpreter Trainees' Speech Comprehensibility: Does Nativeness of the Instructor Matter?

February 2024 – Volume 27, Number 4
https://doi.org/10.55593/ej.27108a8

Mahmood Yenkimaleki
Nahavand Higher Education Complex, Bu-Ali Sina University, Iran
<m.yenkimaleki@basu.ac.ir>

Vincent J. van Heuven
Leiden University Centre for Linguistics, The Netherlands
University of Pannonia, Hungary
<v.j.j.p.van.heuven@hum.leidenuniv.nl>

Abstract

This study examines the effect of native *vs.* non-native prosody instruction on developing interpreter trainees' speech comprehensibility in English as a foreign language (EFL) using a pretest-posttest-delayed posttest design. Twenty-three groups of 28 interpreter trainees at a University in Iran (six different branches) took part in the study, all groups receiving the same amount of instruction (9 hours over 3 weeks). Three control groups listened to/viewed authentic audio recordings and movies in English, discussed their contents, and completed a variety of speaking tasks but received no specific prosody instruction. Twenty experimental groups spent part of the instruction time on theoretical explanation of, and practical exercises with, English prosody by thirteen nonnative instructors, and seven native instructors. Three experts evaluated the comprehensibility of the trainees in elicited speech samples collected during the pretest, immediate posttest and delayed posttest, and subsequently presented in random order. The findings revealed that the experimental groups gained between 1 and 2 points on the 0 to 10 comprehensibility scale, and lost little in the delayed posttest; however, hardly any changes were observed in the control groups. We conclude that native and non-native English instructors' prosody teaching were equally effective in enhancing EFL students' speech comprehensibility.

Keywords: Native instructor, Non-native instructor, Speech comprehensibility, Speaking skills

Despite its importance in speech perception and production, instructors in academic settings often do not address speech prosody appropriately (Yenkimaleki & van Heuven, 2021, 2022; Levis, 2015a, 2016). Non-native instructors may face a further hurdle in teaching prosody. They may have a feeling of insecurity in prosody teaching simply because they do not know

how to apply the nuances of English prosody. Therefore, they need both training and confidence that their prosody teaching can be effective. As both native and non-native language instructors do not have adequate training or are uncertain about the effectiveness of prosody instruction, they find prosody a difficult skill to teach (Yenkimaleki & van Heuven, 2022). The case can be more complicated for non-native language instructors since they may see themselves as an inadequate model to teach prosody (Golombek & Jordan, 2005; Yenkimaleki, 2019, 2021). Although studies show that prosody instruction is effective for interpreter trainees and/or learners of English as a foreign language (EFL) (Saito, 2012; Yenkimaleki & van Heuven, 2018, 2022), it has not been systematically elucidated if the enhancement relies on the instructors' native language, nor if students progress differently depending on whether their instructors are native or non-native speakers of the target language.

Researchers and practitioners confirm the strength of native and non-native instructors in EFL/ESL teaching programs (Alseweed & Daif-Allah, 2012; Braine, 2010; Levis et al., 2016). However, the effect of native *vs.* non-native prosody instructor for developing interpreter trainees' speech comprehensibility has not been investigated systematically. Therefore, the present study was set up to explore this aspect in depth so that the interpreter training experts will be on solid ground in their choice of instructors when aiming to enhance interpreter trainees' speech comprehensibility. Moreover, due to large number of intact classes of students and choice of nonnative English instructors, i.e., Persian and several varieties of Arabic, our results contribute to the generalizability of results reported in the literature on the relative merits of native and nonnative instructors in EFL teaching in general.

Previous related research

Pronunciation Teaching and Nativeness of Instructor

Instructors in foreign language/second language contexts do not have confidence to teach prosody because they do have no experience in teaching it or they have not received adequate training (Couper, 2017; Murphy, 2014; Yenkimaleki & van Heuven, 2019a, b). Some studies (e.g., Levis et al., 2016) propose that nonnative-speaking instructors and native-speaking instructors are equally effective in pronunciation instruction, however, the required training for non-native instructors has not been investigated (Buss, 2016; Couper, 2016). Buss investigated the beliefs and reported practices of EFL instructors in Brazil and stated that instructors addressed pronunciation by focusing on learners' errors generally on the word levels by repetition. Couper (2016) also pointed out that many instructors did not have confidence with their knowledge of phonetics and phonology in training. The results of these studies converged with the findings of native speaker instructors in contexts of English as a second language (Couper, 2017). The other issue is that EFL instructors in both studies stated their concerns about their suitability for pronunciation teaching because of their foreign accent (see Karakaş, 2019). Franceschi and Vettorel (2017) pointed out the general concern of EFL instructors regarding their language proficiency. Levis et al. (2016) studied the effectiveness of a native speaker instructor and non-native instructor in two pronunciation classes and reported no significant differences in the development of comprehensibility in second language learners from these two classes. This result is important since it showed that non-native instructors can be as efficient as native speaker instructors.

Levis et al. (2016) investigated the effectiveness of a Native-Speaker Teacher (NST and a Non-Native-Speaker Teacher (NNST) and reported no significant differences in the development of

comprehensibility in EFL students. The NST and NNST classes (N = 16 each) were taught by using almost the same procedures, materials, and classroom activities. The NNST was found to be as effective as the NST in pronunciation instruction. It was concluded that the effectiveness of pronunciation instruction depends largely on instructors' knowledge base of language teaching and not necessarily on their status as a native speaker instructor or non-native speaker instructor (see also Aslan & Thomson, 2017; Phillipson, 2018). However, it is important to analyze in detail the NNSTs' knowledge base of pronunciation instruction and how they use this knowledge in the classroom. The urgency of this research follows from the high number of NNSTs of English around the world and the corresponding need to reach a nuanced understanding of the different facets of their pronunciation instruction (Derwing & Munro, 2015; Levis et al., 2016). Furthermore, findings from such analyses could provide rich content that could be used for teacher training purposes in both ESL and EFL contexts (Murphy, 2018).

Ma (2012) investigated student perceptions of the advantage and disadvantages of learning English from NSTs and a special type of NNSTs, i.e., local English Teachers (LETs), who shared the non-native language with their students. Data were collected through semi-structured group interviews with 30 secondary-school students studying in three different schools in Hong Kong with both NSTs and LETs. LETs experienced difficulties when providing clarifications when the students had problems in grasping new points. In contrast to this, NSTs did not always understand the needs and problems of a particular group of students compared to LETs (Ma, 2012). NNSTs can have difficulties about their second language accent (Whitehead & Ryu, 2023) and lack of confidence (Ma, 2012) in pronunciation teaching. Therefore, NNSTs may not perceive themselves as good pronunciation teaching models (Ma, 2012), even though they might be confident about other areas of language.

Buss (2016) examined the beliefs and reported practices of NNSTs in Brazil, and found that these instructors approached pronunciation by focusing on learner errors (mostly at the word level) through repetition. Additionally, the NNSTs expressed a desire for more pronunciation-teaching training, which resembled what teachers expressed in Couper's (2016, 2017) studies of reported practices of NNSTs in Uruguay. Couper also found that many NNSTs did not feel confident with their knowledge of phonetics and phonology for teaching. The findings from these studies aligned with the findings of NSTs in contexts of English as a second language (ESL). The NNSTs from both studies also expressed concerns about their suitability for pronunciation teaching because of their foreign accent, which has also been documented previously in pronunciation teaching and is a general concern of NNSTs in terms of their language proficiency (Gordon, 2020).

Native English speakers are estimated to constitute a quarter of EFL/ESL instructors worldwide (Kochem, 2021). However, the encouragement of native speakers as ideal instructors may well be a fallacy (Selvi, 2014). Research on EFL/ESL learners' attitudes towards instructors reveal that using one's native language has a main impact on instructors' confidence, learners' perspectives about efficiency of the instructors, and administrative employment policies (Li & Zhang, 2016).

Li and Zhang (2016) studied the effect of NST *vs*. NNST (LET) for L2 pronunciation teaching. The Chinese participants' (subjectively rated) accentedness and comprehensibility improved both when were taught by NST and then again, in a second stage, by the NNST/LET but the gain was statistically significant only for the second stage. The students preferred being taught

by the NST even though they had gained less. Although the design of the experiment does not permit any firm conclusions, the authors tentatively consider these results "supporting evidence regarding NNSTs' suitability in L2 pronunciation teaching" (p. 95).

Speech Comprehensibility

Pronunciation instruction with an emphasis on comprehensibility can contribute to EFL/ESL students being understood when using a second language (Levis 2018; Yenkimaleki & van Heuven, 2021). We define the intelligibility of a speaker or of a speech utterance in the classical, rather narrow, sense as the degree to which a listener is able to recognize the linguistic units (e.g., morphemes, words) in the stream of sounds and to establish the order in which they were spoken (e.g., Denes & Pinson, 1963; Smith & Rafiqzad, 1979; Gooskens & van Heuven 2021). When a sufficient number of words are recognized in the correct order, the listener will be able to reconstruct the speaker's meaning and intention. This is what we call speech understanding or comprehension. In Applied Linguistics, however, speech intelligibility is more often defined as the degree to which a speaker or spoken utterance is understood in a functional test (e.g., writing down what a speaker said), while comprehensibility is the subjective impression or judgment of a speaker's intelligibility (Munro & Derwing, 1995). In the present study, we will use the terms intelligibility and comprehensibility in this latter ("Applied Linguistic") sense.

By prosody we mean the ensemble of phenomena in speech that cannot be predicted from the mere sequence of the vowels and consonants (the "segments") that make up the spoken utterance. We distinguish between word prosody and sentence prosody. Word prosodic phenomena are word stress and lexical tone. Correct word stress and tone contribute to word recognition, while stressing the wrong syllable or executing the wrong lexical tone reduce the chances of words being correctly recognized (e.g., Cutler, 2005, 2012; Field, 2005), especially when the segmental quality of the spoken utterance is poor or cannot be clearly heard (van Heuven, 2022). Sentence prosody comprises such phenomena as phrasing (by pausing, and local changes in speaking rate), sentence stress (making words communicatively important by adding a pitch change to a word stress) and intonation (or sentence melody). Incorrect prosody may negatively affect a speaker's comprehensibility.

Current Study

Prosody instruction is a fundamental element for the interpreter training programs in enhancing the perception and production skills of interpreter trainees for decoding and encoding the messages in interpreting performance (Yenkimaleki et al., 2021; Yenkimaleki & van Heuven, 2022). Levis et al. (2016) stated that native speaker instructors do not have any advantage over non-native instructors in teaching pronunciation for the students. There are different opposing and unresolved perspectives on the impact of being native speaker or not for efficient pronunciation instruction. This aspect needs to be investigated systematically to shed more light on the choice of native *vs.* non-native prosody instructor for developing interpreter trainees' speech comprehensibility. Therefore, the present study is set up to explore this issue in depth. The following research question is raised:

Compared to L1 native English instructors, to what extent do L2 (nonnative) English instructors enhance/compromise interpreter trainee comprehensibility through teaching prosody?

At this stage, we do not suggest specific hypotheses for the any advantages of the native instructors in teaching prosody for boosting speech comprehensibility of interpreter trainees, and of EFL learners in general. This will depend on the working languages, and the proficiency of instructors in faithfully observing the rules and guidelines of the specific method of instruction when teaching prosodic features.

Method

Participants

Six hundred forty-four Persian-speaking undergraduate students were chosen randomly at the University of Applied Sciences and Technology (UAST), different campuses, Iran to participate in the study. The campuses, number of groups and students involved and L1 of instructors are specified in Table 1. The native language of the instructors (identified by their initials) is specified. Each instructor taught one experimental group. Instructors identified in bold face also taught one control group.

Table 1. Number of groups and students (28 per group) for each of six UAST campuses enrolled in the experiment.

City/campus	Participants		L1 of instructor				
	Groups	N	Am. Eng	Persian	Iraq	Arabic Syr	Leb
Tehran West	4	112	LH, OS	**MY**			
Tehran East	4	112	AG, RS,WB		HRJ		
Isfahan	4	112	AR, **PG**		MHH		
Fars	4	112		RA, MG			MR, JH
Kermanshah	3	84		HK, HY, MA			
Alborz	4	112		MH		RM, **ZF**	
Total	23	644	224	224	56	84	56

Seven classes were assigned to prosody instruction by Iranian instructors, six to prosody instruction by Arabic EFL teachers (i.e., two Iraqi, two Syrian and two Lebanese instructors), and seven classes were taught by L1 American Native English speakers. Each of these experimental groups was taught by a different instructor. Three more classes served as the control groups with no prosody instruction at all. The instructors for the control groups also taught one experimental group each (see Table 1 for details).

None of participants had studied or lived abroad. They had learned English in secondary school for four years, with two hours of lessons per week. In university, they had studied translation and interpreting studies for three semesters before participating in the program at UAST. The age range of the students was between 18 and 27. They had passed the entrance exam for the translation and interpreting department. The students were randomly grouped into 23 classes of 28 students. Seven classes were assigned to prosody instruction by American native instructors.

Ethical Matters

Ethical approval to involve these participants in the experiments was obtained from University of Applied Sciences. All the participants agreed to take part in the research project by signing written informed consent.

Procedure

Twenty-three groups of 28 students were formed by random assignment. Three classes were control groups, which did not receive prosody instruction but instead did different types of exercises to develop their English speaking skills. They also watched authentic English materials (e.g., audios, videos), and discussed their contents in different sessions (nine sessions of one hour in three weeks). Three different instructors taught the control groups (see Table 1). The instructors were one Persian-speaking Iranian instructor, one American native instructor, and one Syrian Arabic instructor.

The experimental groups spent 20 minutes less time per session on the routine curriculum and instead received 20 minutes of explicit teaching of prosodic properties of English (see Table 2 for a summary of the activities and time spent by experimental and control groups).

Table 2. Summary of activities and time spent (minutes) by participants in 20 experimental and 3 control groups in the experiment.

Activity	Group	
	Control	Experimental
Listening to instructor explanations/guidance/comments	160	160
Audio tracks/movies	380	200
Prosodic theory, training, and practice	--	180
Total time spent	540	540

The materials for the treatment were American English speech fragments. The phonological phenomena to be demonstrated and explained were chosen on the basis of the authors' earlier experience and studies. The types of training also were based on the authors' previous studies (e.g., Yenkimaleki, 2017; Yenkimaleki & van Heuven, 2018, 2020). The authors knew that some features would be important in the speech comprehensibility of Persian EFL learners. The students already knew that they were going to receive prosody training for the program. The specific contents for each session, however, were not mentioned to the students beforehand. The reason was that the instructors did not want students prepare any materials outside of the classroom. The materials for prosody training were the same for all twenty experimental groups. Moreover, before the beginning of the training program, the twenty instructors consulted with each other to ensure that they would follow the same approach (see the treatment section below) and use the same materials in prosody teaching.

We will refer to the twenty experimental groups as Native instructor groups, and Non-native instructor groups. The Non-native instructors were native speakers of either Persian or Arabic. All of them had several years of experience in teaching pronunciation to EFL learners. Seven of them had defended a PhD in Applied linguistics. Three were doing their PhD in Applied linguistics, and were employed as pronunciation teachers for EFL students, while three more had obtained their MA in (Applied) Linguistics, and had on the average of ten years' experience in teaching pronunciation in EFL context. All Native instructors held an MA in (Applied) Linguistics, and had worked as EFL instructors for some years at the time the experiment was conducted.

Treatment

Control groups. The control groups received routine instruction in interpreting, i.e., the routine curriculum and the syllabus which has been used in the English Translating and

Interpreting Department at University of Applied Sciences. For these groups, the techniques of interpreting, different aspects of interpreting, and types of interpreting were normally instructed and practiced. Students practiced intensive listening tasks, which were followed by detailed comprehension questions, e.g., inferring the meaning of unknown words. The movies, that the students watched, were not subtitled. To help prepare students' expectations about what they are going to listen/watch and to aid in their comprehension, the instructor used warm-up questions, and brainstormed relevant vocabulary. The instructor introduced the topic, and got the students thinking about it. If felt necessary, the instructor presented a short list of keywords occurring in the movie/audio file that students might be unfamiliar with. The meaning of such keywords was illustrated by using them in disambiguating sentences before the audio file was started. The instructor played the movie clip/audio file first for general comprehension – to allow students to get the main idea. Then, he replayed it several times for the students to grasp more details. The pause button was used as needed to focus on sections students had difficulty in understanding. Next, students were asked to complete an exercise on the corresponding activity. Interesting, and in some cases, humoristic movies/audio files were chosen, covering a variety of topics such as politics, social issues, and scientific findings. Only good quality audio files with clear-voiced speakers were presented. The same procedure was followed at nine sessions of training program.

Experimental groups. The experimental groups spent 20 minutes less time per session on the routine curriculum for the control groups (see section 4.4.1), and instead received awareness training on prosodic features of English. The students received explicit teaching of prosodic features of English speech during each training session, in a six-stage instruction module:

Stage 1: Students acquired procedural knowledge about the prosodic features of English e, i.e., intuitive knowledge that cannot be verbalized (e.g.,Yenkimaleki & van Heuven, 2018) such as phonetic/phonological sensitivity. Prosodic sensitivity or awareness should be directed to three important elements: (i) lexical stress, i.e. the location of the syllable in a word that receives emphasis, (ii) intonation, i.e. the pattern of pitch rises and falls that is used to divide continuous speech into sentences and phrases and to mark specific units within those phrases as communicatively important, and (iii) temporal organization by which pauses (often reinforced by melodic means) are inserted between groups of words – roughly fulfilling the function of punctuation marks in a written text.

Stage 2: Students were involved in speech shadowing. Authentic audio files (i.e., materials not produced for teaching purposes in EFL classrooms but produced for movies, newscasts, and for other communication purposes in the media) were selected with a maximum duration of five minutes; they were based on the speech of a single native English speaker (e.g., zappenglish.com; see Appendix 1 for an example). A topic was chosen that students were already familiar with and which was interesting to them. The students listened to the audio files once quickly just to get used to the speaker. They paid attention to the speaker's rhythm, intonation, stress, and pace of speaking. The students were asked to shadow the input speech, i.e., to repeat the words immediately after the speaker, for about 30 seconds at a time. Students paused, tried again, and even recorded and listened back to their own versions (e.g., when somebody sings the words of a song s/he already knew well, s/he tries to imitate the speaker's pronunciation and pace as best as s/he could). This helped the students focus on how English speakers modulated speed, used intonation, and blended words together. By

mimicking the speaker, students could begin to improve their own intonation, connected speech, and overall fluency.

Stage 3: Students targeted problematic sounds. Throughout the lesson, the instructors listened to how students spoke, identified a number of words that reflected pronunciation challenges for the students. The instructors wrote the words on the board and explained the contrast that was missed, e.g., *steam/esteem, prayed/parade* (incorrect vowel epenthesis) or *foreign/for rain* (incorrect stress pattern). In this stage, the instructors also asked the students to transcribe and drill new vocabulary (e.g., by introducing new vocabulary, the instructors will be sure that the students will not conditionally answer the exercises from whatever which they already practiced). Here, the instructors focused on aspects of pronunciation such as word stress, sentence stress (e.g., the practice of sentence stress was based on listening and producing the same stress patterns that the native speakers/instructors had in uttering words (see Appendix 2 for an example). The new vocabulary items were in the new sentences, not in the already practiced sentences by the students), and intonation. New vocabulary that came up during the training program was written down.

Stage 4: The instructors asked the students to contextualize their tasks (e.g., repetition of key words in a listening passage; see Appendix 3 for an example).

Stage 5: Having done the required practical tasks on contextualization, the students performed meaningful, authentic tasks (e.g., choice of correct word in a sentence or a sentence in a paragraph).

Stage 6: Students were asked to do realistic tasks (e.g., a role-play of a situation similar to one that one may face in real life or a discussion of the students' real-life situation or concerns) For further details see Yenkimaleki, 2017: 50-88.

Assessment

The pretest, posttest, and delayed posttest consisted of interviews that were run systematically to evaluate the participants' speech comprehensibility. Three raters judged the participant's comprehensibility. Rater 1 (R1) completed his PhD in the United Kingdom in Applied Linguistics. Rater 2 (R2) specializes in pronunciation teaching, and defended his PhD dissertation in the Netherlands. Both nonnative raters R1 and R2 are native speakers of Persian. The third rater (R3) is an American native speaker, who obtained his MA in Linguistics from University of Arizona, USA. The native speaker was employed as a rater to add to the quality of the assessment since in some cases the nuances of pronunciation can be better recognized by native speakers. Judgements by native and non-native raters need to be consistent with each other (Yenkimaleki & van Heuven, 2021). Ekmekçi (2016) and Zhang and Elder (2014) concluded that native and non-native English language instructors display almost identical rating behavior in assessing EFL students' oral proficiency. The first step in our data analysis was to check whether this is also the case in our study.

The pretest, posttest, and delayed posttest consisted of six questions. The questions were descriptive in nature, so that world knowledge on the part of the students was not involved in answering the questions, e.g., *describe the historical sites of the city*. All the questions were open-ended, and students had one minute to answer each question. The same prerecorded questions were asked to all the participants, who were in different rooms and could not communicate with each other after answering the questions. The questions in the pretest,

posttest, and delayed posttest were different so that students would not remember any point from the pretest (to control the test effect). Raters, by consulting with each other, had written eighteen questions before the program started. Six of these were randomly chosen for the pretest, six for the posttest, and six for delayed posttest. Recordings were made directly onto a digital computer (44,100 Hz, 16 bit) through individual, table-mounted microphones. Near-native experts in the field of Applied Linguistics presented the (pre-recorded) questions, spoken in English by one female expert, to the participants. The experimenter had not been an instructor to any of the groups.

The pretest was administered in the week before the start of the treatment. The immediate posttest followed in the week after the last session of the treatment, while the delayed posttest was held one month after the immediate posttest.

The interview questions were spoken by one near-native female expert, and then the question files were sent to the local experimenters. One interviewer on each campus was engaged in interviewing the students. The students' responses were recorded as audio files by the interviewers on each campus. At the end of the semester, after the delayed posttest, all the recordings were sent to the first author. The 644 audio files were then identified by code numbers randomly assigned to locations (campuses), times of testing (pretest, immediate posttest or the delayed posttest) and speakers, and made available to all three raters. The code numbers did not reveal the campus, name of the student, or time of testing (e.g., pretest, immediate posttest, or delayed posttest).

Each rater was instructed to listen to at least 20 seconds of recording time before judging the speech comprehensibility of a student's recording. After listening to a recording, the rater entered his score in a spreadsheet that listed the code numbers of the recording, in the order in which the rater was asked to listen to the files. This amounted to around 20 hours of work per rater. The workload was distributed over five working days. The raters were asked to take a 20 minute break after one-hour of working time.

The raters listened to the recordings, in the same quasi-random order (excluding immediate succession of recordings by the same speaker), at different locations, independently of one another. They assessed the speaker's overall speech comprehensibility. They indicated the degree of comprehensibility on a 11-point scale from 0 (impossible to recognize even a single word) to 10 (all words can be effortlessly recognized in the intended order).

Statistical Analysis

When the number of categories along the rating scale are in excess of five, and when the categories at the low end of the scale are not used, rating scales such as the one employed here are generally treated as interval data (Apparent Equal Interval, AEI; Rietveld, 2020: 24-26, and references therein). Parametric inferential tests make the further requirement that the data are normally distributed. The first step in the data analysis would therefore be a check on the normality of the data distribution. The preliminary screening revealed that the ratings were strongly skewed so that only non-parametric inferential statistics could be applied.

In Stage 2, we checked whether the three raters employed were in sufficient agreement, using Cronbach's alpha as the relevant agreement measure. The results show that the raters were in excellent agreement so that we could perform subsequent analyses on the ratings averaged over the three raters.

In the main analysis, we tested the effect of teacher L1 on the pretest and two posttest scores, as well as on the gain from pretest to immediate posttest and subsequent loss from immediate to delayed posttest. We did this in separate non-parametric alternatives for each of the dependent variables. The Effects of Time of testing (pretest, immediate posttest, delayed posttest) were analyzed by the Friedman test as an alternative for the Repeated Measures Analysis of Variance, followed by post-hoc triplets of Wilcoxon tests with Bonferroni adjusted alphas to accommodate multiple testing. Effects of L1 of instructor and Group were tested by the Kruskal-Wallis alternative for the independent samples Analysis of Variance, with a post-hoc series of Mann-Whitney U-tests, again with Bonferroni correction for multiple comparisons.

Finally, we checked whether potential uncontrolled teacher variables, such as the difference between MA and PhD qualification and years of relevant teaching experience, might (partly) explain the results – and found that such was not the case.

Results

Preliminary Data Analysis

To explore the general distribution of the rating scores in the present experiment, we inspected the histograms of the scores (0, 1, … 10) for each rater separately as well as aggregated over all three raters. Since the effects of treatment potentially lead to a bimodal distribution, we only examined the distributions obtained for the pretest scores, i.e., the scores that were given when all participants were assumed to be sampled from the same underlying population. The four histograms (with normal curves superimposed) are shown in Figure 1.

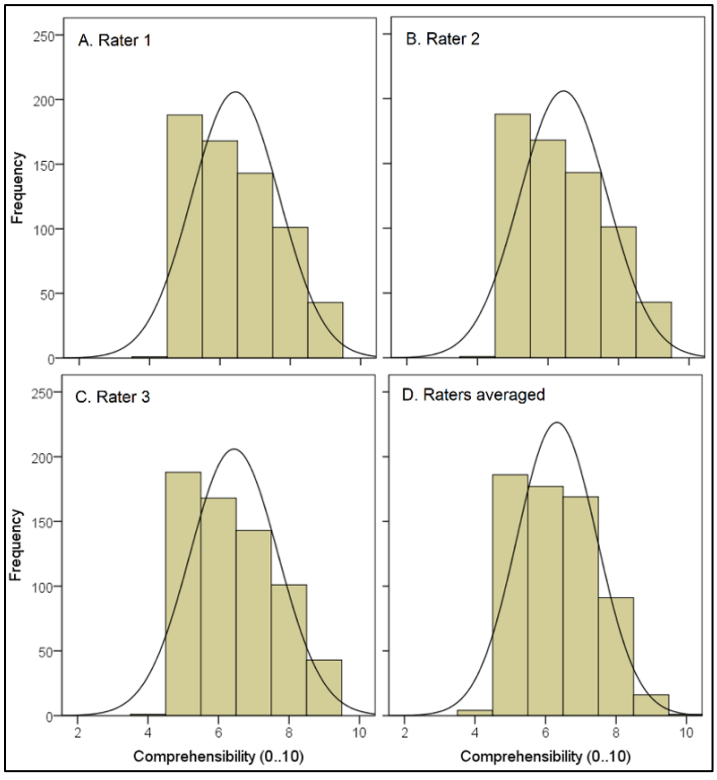

Figure 1. Distribution of comprehensibility scores on pretest given by raters R1, R2, R3, and all raters averaged. Normal curves are superimposed.

Summary statistics are given in Table 3. The Kolmogorov-Smirnov test (K-S, with Lilliefors significance correction) was used to determine deviation from normality.

Table 3. Mean, SD, quartiles, skew and kurtosis for pretest scores, for individual raters separately and averaged (*N* = 644).

Raters	Mean	SD	25%	50%	75%	Skew	Kurtosis	K-S	p
R1	6.42	1.23	5	6	7	.431	–.871	.192	< .001
R2	6.34	1.15	5	6	7	.590	–.461	.216	< .001
R3	6.34	1.15	5	6	7	.372	–.694	.182	< .001
Averaged	6.37	1.09	5	6	7	.305	–.959	.194	< .001

Inspection of Figure 1 shows that hardly any scores were given in the lower half of the comprehensibility scale. The reason for this must be due to the fact that all participants in the experiment had passed an entrance exam to enroll in the Departments of Translation and Interpreting. Candidates with below-average comprehensibility were effectively prevented from enrolling in the curriculum. As a result of this, only the upper half of the underlying distribution is seen, which is therefore strongly positively skewed. The Kolmogorov-Smirnov test (Table 3, right-most columns) bears out that the observed distributions deviate significantly from normality. Moreover, no transformations, e.g., logarithmic, reciprocal, square or cubic root, or even the Box-Cox method (Box & Cox, 1964), could be found that would sufficiently undo the skewedness. For this reason, we decided to perform subsequent inferential statistics on non-parametric alternatives.

Given 23 groups of 28 participants, whose performance was tested at three different points in time, there were 1,932 productions to be rated by each of three experts in terms of their comprehensibility on a rating scale between 0 and 10. Table 4 presents the mean and standard deviation of the ratings given by each of the three raters, overall and broken down by time of testing (i.e., pretest, immediate posttest, delayed posttest). Differences among the three raters were tested for statistical significance by the Friedman test for k related samples, and, when found significant, probed further by post-hoc tests, with Bonferroni correction for multiple comparisons. In Table 4, the column headed 'Post-hoc', the '>' sign indicates that the rater on the left-hand side of the sign was significantly more lenient than the rater on the right-hand side. Rater pairs that are not mentioned do not differ from one another. The effect of Rater was tested by the Friedman alternative for the RM ANOVA, followed by Wilcoxon pairwise tests. Thus 'x>y' means x is significantly larger than y ($p \leq .05$, Bonferroni corrected). The right-most column lists Cronbach's alpha, as an indication of agreement among the raters.

Table 4. Mean and SD of ratings specified per rater and by time of testing.

Time of testing	Nonnative				Native		Friedman test + Wilcoxon				Cronbach's alpha
	Rater 1		Rater 2		Rater 3						
	Mn	SD	Mn	SD	Mn	SD	N	$\chi^2(2)$	p	Post-hoc	
Pretest	6.42	1.23	6.34	1.15	6.34	1.15	644	2.1	.356		.903
Immediate posttest	7.86	.95	7.76	.89	7.84	.89	644	8.9	< **.011**	1>2, 1>3	.834
Delayed posttest	7.88	.94	7.64	.97	7.67	1.06	644	50.1	< **.001**	1>2, 1>3	.799
Tests combined	7.39	1.25	7.25	1.20	7.28	1.23	1932	40.3	< **.001**	1>2, 1>3	.906

The mean ratings given by the three raters were close together, with a maximum discrepancy between any two raters of .23 on the scale from 0 to 10. Nonnative Rater 1 was always more lenient than his colleagues, even though the difference was not significant in the pretest. The

analysis shows that there is always one nonnative rater who was equally lenient as the American L1 rater, i.e., Rater 3.

The agreement among the raters was quantified by means of Cronbach's alpha. Cronbach's alpha coefficient was .906 overall, which is generally considered as excellent agreement among raters (e.g., Nunnely, 1978). On the strength of these results, we decided to aggregate the scores over the three raters, and perform all further analyses on the means of the rating scores.

Main Analysis

Figure 2 presents a breakdown of the Mean Comprehensibility scores obtained by the 23 groups of participants, at each of the three times of testing. Ratings were first averaged over three independent raters. Error bars are 95% confidence intervals of the means (N = 28 per data point). Data are presented separately for groups taught by L1 American instructors (panel A), L1 Arabic EFL instructors (panel B) or L1 Persian EFL instructors (panel C). The solid marker (Group 0) represents the control groups with no dedicated prosody instruction module. There is no Group 7 in panel B.

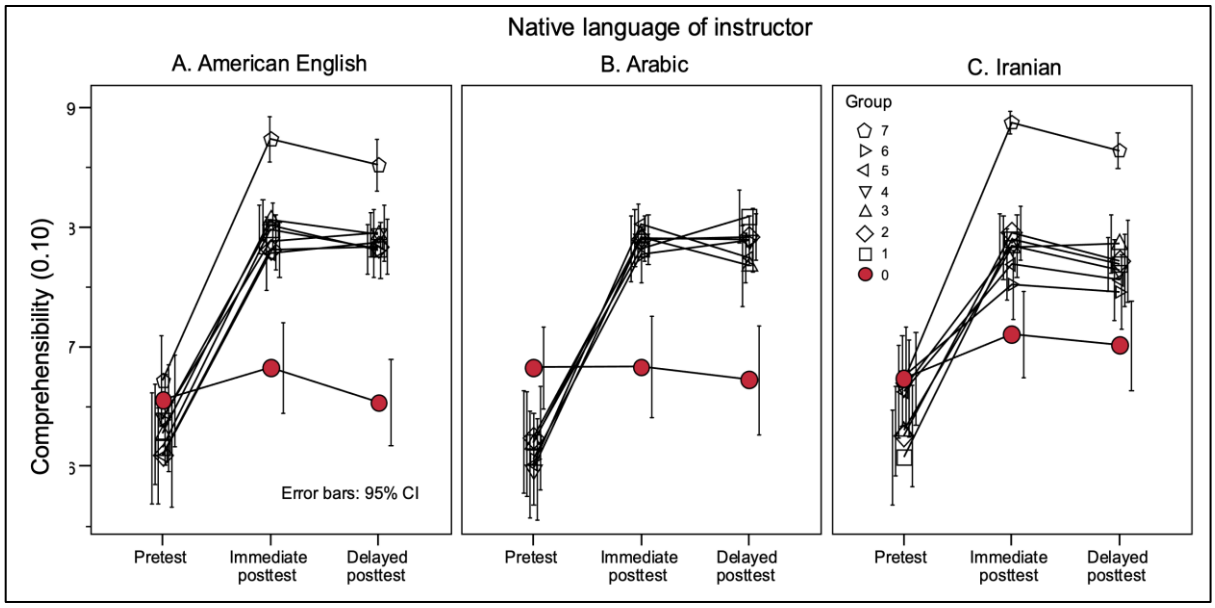

Figure 2. Comprehensibility (0 to 10) for up to seven groups of participants at three times of testing.

To test the significance of the differences seen in Figure 2, we analyzed the means in each of the three panels separately at each time of testing, using the Kruskal-Wallis non-parametric alternative for the one-way Analysis of Variance for independent samples. Post-hoc pairwise comparisons were made by Mann-Whitney U-tests, with Bonferroni-adjusted criteria. Adopting the basic .05 significance criterion, alpha has to be divided by the number of groups under comparison, so that adjusted alphas apply at .05/7 = .0024 for panel B (7 groups with Arabic L1 instructors), and at .05/8 = .0018 for panels A and C (American English and Persian L1 instructors, respectively). The results of these tests are summarized in Tables 5A, 5B and 5C below.

Inspection of Figure 2 shows that, generally, the 23 groups of students involved in the study are close together in terms of their comprehensibility, with mean scores between 6 and 7 on

the rating scale from 0 to 10. We performed three separate Kruskal-Wallis tests to establish whether there were significant differences among the 7 or 8 groups in each panel in Figure 2. For the American L1 instructors (panel A, Table 5A), and for the Persian instructors (panel C, Table 5C), no effect of group was found. For the Arabic L1 instructors the control group performed significantly better in the pretest than the other 6 groups (there was no group E7 here).

In the following three tables, scores were obtained in a pretest (P), immediate posttest (I), and delayed posttest (D). Gain (from P to I), and Loss (from I to D) are specified. $N = 28$ per cell. Significant overall effects are bolded. 'x>y' in the column headed "Post hoc" means "x is significantly larger than y" ($p \leq .05$, Bonferroni corrected). Superscripts denote significant post-hoc differences in columns.

Table 5A. Comprehensibility (mean and standard deviation on a scale from 0 to 10 = best) for 8 groups of Persian learners of English as a foreign language taught by American English L1 instructors.

Group		Pretest	Posttest		Δ		Friedman test + Wilcoxon post hoc		
			Imm.	Del.	Gain	Loss	$\chi^2(2)$	p	Post hoc
C0	Mean	6.55	6.83[b]	6.53[b]	.28	.29	8.6	**.014**	I>P, I>D
	SD	.99	.98	.93	.68	.56			
E1	Mean	6.15	7.98[ab]	7.82[ab]	1.83[a]	.16	38.8	**< .001**	I>P, D>P
	SD	1.20	.52	.53	.87	.36			
E2	Mean	6.10	7.81[ab]	7.84[ab]	1.72[a]	–.03	36.5	**< .001**	I>P, D>P
	SD	1.13	.59	.58	.71	.40			
E3	Mean	6.27	8.06[ab]	7.94[ab]	1.79[a]	.12	42.3	**< .001**	I>P, D>P
	SD	1.07	.44	.48	.79	.41			
E4	Mean	6.40	7.88[ab]	7.95[ab]	1.47[a]	–.07[a]	43.8	**< .001**	I>P, D>P
	SD	1.15	.59	.61	.80	.49			
E5	Mean	6.09	7.78[ab]	7.87[ab]	1.69[a]	–.09[a]	43.4	**< .001**	I>P, D>P
	SD	1.03	.79	.73	.56	.42			
E6	Mean	6.39	8.02[ab]	7.81[ab]	1.63[a]	.21	47.4	**< .001**	I>P, D>P
	SD	.97	.50	.60	.61	.31			
E7	Mean	6.72	8.73[a]	8.52[a]	2.02[a]	.22	50.2	**< .001**	I>D>P
	SD	.97	.49	.56	.82	.26			
	$\chi^2(7)$	12.0	71.3	70.6	60.9	25.6			
	p	.101	**< .001**	**< .001**	**< .001**	**.001**			
Post-hoc U-test			[a] ≠C0 [b] <E7	[a] ≠C0 [b] <E7	[a] >C0	[a] <E7			

Table 5B. With students taught by Arabic L1 instructors.

Group		Pretest	Posttest		Δ		Friedman test + Wilcoxon post hoc		
			Imm.	Del.	Gain	Loss	$\chi^2(2)$	p	Post hoc
C0	Mean	6.83	6.83	6.72	.00	.11	2.7	.263	
	SD	.88	1.09	1.18	.53	.48			
E1	Mean	6.21[a]	7.82[a]	8.09[a]	1.61[a]	–.27	43.1	**< .001**	I>P, D>P
	SD	1.11	.71	.57	.70	.58			
E2	Mean	6.24[a]	7.90[a]	7.92[a]	1.66[a]	–.02	41.3	**< .001**	I>P, D>P
	SD	1.11	.54	.50	.68	.39			
E3	Mean	6.19[a]	7.91[a]	7.68[a]	1.72	.23	40.1	**< .001**	I>P, D>P
	SD	1.13	.61	.86	.73	.69			
E4	Mean	5.98[a]	7.90[a]	7.89[a]	1.93[a]	.01	42.1	**< .001**	I>P, D>P
	SD	1.10	.49	.69	.76	.42			
E5	Mean	6.02[a]	8.03[a]	7.74[a]	2.01[a]	.28[a]	40.6	**< .001**	I>D>P
	SD	1.15	.44	.52	1.00	.54			
E6	Mean	6.06[a]	7.77[a]	7.89[a]	1.71[a]	–.12	39.5	**< .001**	D>I>P
	SD	1.00	.59	.53	.77	.57			
	$\chi^2(6)$	14.6	28.0	30.9	65.2	18.8			
	p	**.023**	**< .001**	**< .001**	**< .001**	**.001**			
Post-hoc U-test		[a] <C0	[a] >C0	[a] >C0	[a] >C0	[a] >E1			

Table 5C. With students taught by Persian L1 instructors.

Group		Pretest	Posttest		Δ		Friedman test + Wilcoxon post hoc		
			Imm.	Del.	Gain	Loss	$\chi^2(2)$	p	Post hoc
C0	Mean	6.74	7.10[b]	7.01[b]	.37[e]	.09	28.3	**< .001**	I>P, D>P
	SD	1.00	.93	.96	.24	.20			
E1	Mean	6.08	7.90[ab]	7.69[b]	1.82[a]	.21	45.4	**< .001**	I>D>P
	SD	1.02	.56	.58	.60	.35			
E2	Mean	6.26	7.95[ab]	7.72[b]	1.70[a]	.24	42.3	**< .001**	I>P, D>P
	SD	1.09	.58	.88	.65	.55			
E3	Mean	6.30	7.83[b]	7.86[ab]	1.53[a]	–.03	42.8	**< .001**	I>P, D>P
	SD	.97	.68	.62	.54	.34			
E4	Mean	6.66	7.84[ab]	7.64[b]	1.19[abe]	.20	37.0	**< .001**	I>P, D>P
	SD	1.05	.66	.77	.70	.45			
E5	Mean	6.63	7.69[b]	7.56[b]	1.06[abcde]	.13	43.8	**< .001**	I>P, D>P
	SD	1.00	.76	.86	.47	.40			
E6	Mean	6.75	7.52[b]	7.46[b]	.77[abcde]	.06	18.5	**< .001**	I>P, D>P
	SD	1.09	.74	.80	.53	.57			
E7	Mean	6.71	8.87[a]	8.64[a]	2.16[a]	.23	47.1	**< .001**	I>D>P
	SD	.98	.24	.39	.87	.38			
	$\chi^2(7)$	14.0	74.4	56.5	112.0	10.7			
	p	.052	**< .001**	**< .001**	**< .001**	.152			
Post-hoc U-test			[a] >C0 [b] <E7	[a] >C0 [b] <E7	[a] >C0 [b] <E1 [c] <E2 [d] <E3 [e] <E7				

In the immediate posttest, there is a highly significant main effect of Group in each panel of Figure 2. Post-hoc pairwise U-tests revealed that all experimental groups in each panel received better comprehensibility scores than the respective control groups. Also, group 7, when present (i.e., with American and Persian instructors) obtained significantly higher scores than all other groups in the immediate posttest. The same configuration of effects is found for the delayed posttest.

With one exception, all groups, control and experimental alike, were given significantly higher comprehensibility scores in the immediate posttest than in the pretest. The gain was small for the control groups ($\Delta \leq .37$ on the scale from 0 to 10) but substantial for all experimental groups ($.77 \leq \Delta \leq 2.16$). The single group that showed no gain is the same group that significantly outperformed the other six groups in the pretest stage in Figure 2B.

Although minor loss of comprehensibility was incurred after one month, the loss was significant in only 5 out of 23 groups, two of which involved group 7 (which will be discussed later). The relative positions and significance of the differences among the 7 or 8 groups remained the same between the immediate and the delayed posttest.

Since there is one control group that obtained better ratings in the pretest than the experimental groups, the best way to analyse the effect of the treatment is by computing the gain in ratings from pretest to immediate posttest, and the subsequent loss from immediate to delayed posttest. A separate Kruskal Wallis test was performed on the Gain from pretest to immediate posttest. Post-hoc U-tests bore out that the differences in Gain were significantly different for the 7 or 8 groups in each panel of Figure 2. The Gain observed for the control groups was always significantly smaller than for any experimental group, while the Gain for the two groups 7 was significantly larger than that in any other group. The Loss in comprehensibility incurred after one month was relatively minor, varying between –.27 and .29 (where a negative Loss represents a further gain in comprehensibility). Significant differences in Loss were observed in only 3 out of 23 groups, two of which were actually (insignificant) gains after the immediate posttest.

Table 6 lists the mean Gain and Loss scores obtained by the 23 groups in the study, broken down by L1 of the instructor. Means are presented with and without inclusion of Group 7.

Table 6. Mean Gain (from pretest to immediate posttest) and Loss (from immediate to delayed posttest) of Comprehensibility

		L1 of instructor			Kruskal-Wallis test	
		Am. English	Arabic	Persian	$\chi^2(2)$	p
Gain	all groups	1.55	1.52	1.32	1.07	.586
	without group 7	1.49	1.52	1.21	2.41	.299
Loss	all groups	.10	.03	.14	1.03	.598
	without group 7	.09	.03	.13	.78	.677

To test the main effect of L1 of instructors, we performed a Kruskal-Wallis test for independent samples on the mean Gain and Loss scores obtained by 23 groups in our study. The L1 of instructors has no significant effect on the Gain nor on the subsequent Loss of comprehensibility.

There was no group 7 with Arabic L1 instructors. The experimental group 7 with an American or an Iranian instructor performed significantly better than any other group in the study. Group

7 received the highest comprehensibility score in the immediate posttest, the only groups with scores > 8.0. They were also the two groups with the largest gain from pretest to posttest (> 2.0). The groups 7 were taught by the only instructors (the first author, and a former American colleague at the same university) who had done this type of prosody teaching before in similar experiments.

Potential Confounds

There were potentially important individual differences among the 20 instructors we employed in the present study. Although all instructors were experienced EFL teachers, their experience ranged between 7 and 15 years. Also, none of the American native English teachers had obtained a PhD degree, while three of the Arabic L1 and four of the Iranian (Persian L1) instructors had a PhD qualification in Applied Linguistics in addition to their MA in TEFL. Table 7 lists, for each of the 20 instructors employed in the study, the group(s) they taught, their nationality, highest academic degree obtained, and years of experience in teaching pronunciation.

Table 7. Instructors' L1, nationality, group(s) taught, academic degree, and years of experience in pronunciation teaching. Control groups are bolded.

#	Initial	L1	Nationality	Group taught	Degree	Experience (years)
1.	MY	**Persian**	Iranian	**0**, 7	PhD	12
2.	RA	Persian	Iranian	5	PhD	7
3.	MG	Persian	Iranian	1	MA	7
4.	HK	Persian	Iranian	3	MA	6
5.	HY	Persian	Iranian	2	MA	13
6.	MA	Persian	Iranian	4	PhD	11
7.	MH	Persian	Iranian	6	PhD	8
8.	LH	American	American	2	MA	15
9.	OS	American	American	4	MA	11
10.	AG	American	American	3	MA	9
11.	RS	American	American	1	MA	8
12.	WB	American	American	5	MA	11
13.	AR	American	American	6	MA	8
14.	PG	**American**	American	**0**, 7	MA	9
15.	HRJ	Arabic	Iraqi	2	PhD	10
16.	MHH	Arabic	Lebanese	5	MA	12
17.	RM	Arabic	Syrian	3	MA	11
18.	ZF	**Arabic**	Syrian	**0**, 4	PhD	9
19.	MR	Arabic	Lebanese	6	PhD	8
20.	JH	Arabic	Iraqi	1	MA	15

If it were true that more years of relevant teaching experience (within the range employed in our study) and an additional PhD diploma are conducive to superior pedagogic abilities, this should show up in larger gains from pretest to posttest, especially for the experimental groups.

Figure 3 shows the Gain (Y-axis) as a function of length of teaching experience (expressed in years, X-axis) for the three instructor groups in separate panels. MA (red/dark dots) and PhD-level (green/light squares) instructors are shown by different markers

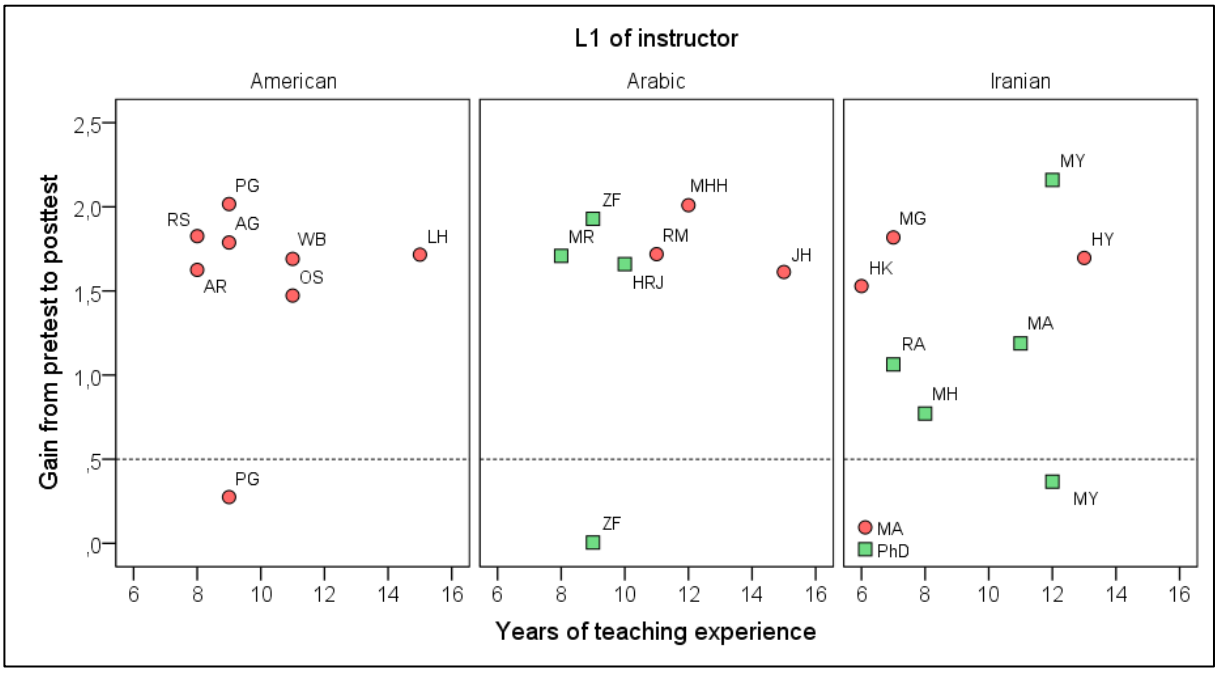

Figure 3. Gain scores of EFL students' comprehensibility from pretest to posttest as a function of teaching experience (in years), highest academic degree, and first language of instructor.

Gains below .5 on the 10-point rating scale (below the dotted line) were found for the control groups. It is easily seen in Figure 3 that there is hardly any correlation between Years of teaching experience and Gain (r = .116, N = 23, p = .298, one-tailed, or r = .186, N = 20, p = .216, one-tailed, when the control groups are excluded). The Gain found for groups taught by instructors with an MA degree (N = 14) was 1.63, while the Gain achieved by the PhD instructors (N = 9) was, in fact, smaller, i.e., 1.21. When the control groups are omitted from the comparison, the Gain for the MA instructors (N = 8) is 1.73, while the PhDs' Gain (N = 7) was 1.50. The effect of highest academic degree obtained was tested by Mann-Whitney U-tests. The effect of Degree was insignificant whether the control conditions were included, z = 1.323 (p = .201, n.s.) or excluded from the analysis, z = .911 (p = .393, n.s.).

Summary of Results

We may summarize the statistical analysis as follows. Before the start of the experiment all groups (except one) were equally comprehensible. The control groups gained only .4 on the comprehensibility scale from pretest to immediate posttest, and subsequently lost all the gain after one month. The experimental groups gained a substantial increase in judged comprehensibility as a result of the treatment, i.e., an improvement of close to 2 full points on the comprehensibility scale from 0 to 10, and they lost little in the four following weeks. The gain and subsequent loss in comprehensibility was roughly the same for all experimental groups, with discrepancies that never exceeded .25 point.

Crucially, the native language of the instructors, i.e., whether American English (target language), Persian L1 (shared with the students), or unrelated Arabic L1, did not significantly influence the effectivity of the prosody training for the development of the students' comprehensibility in the target language.

Discussion

The findings revealed that experimental groups' speech comprehensibility improved relative to the control groups. The results also showed that the instructor's nativeness or non-nativeness does not significantly impact the development of the interpreter trainees' speech comprehensibility by teaching prosody. These results are in line with Yenkimaleki and van Heuven (2019b, 2021), who found that prosody teaching enhances EFL students' speech comprehensibility. Our results also converge with Levis et al. (2016), who reported that nonnative-speaking instructors and native-speaking instructors are equally effective in pronunciation instruction.

Despite the well-attested importance of pronunciation skills for successful communication in the L2 (Derwing & Munro, 2015), recent studies have found that pronunciation instruction is still neglected in L2 classrooms (Foote et al., 2016). While reasons for these shortcomings are manifold, the two justifications most generally given by instructors are lack of time to address pronunciation in a very full curriculum, as well as a general sense of discomfort with teaching L2 pronunciation (Grim & Sturm, 2016). Taken together with findings that have shown that learners themselves see pronunciation training as very important (Couper, 2016) and that their pronunciation skills can improve through training (Lee et al., 2015), prosody instruction should be, and in fact has been, included in the standard interpreter training curriculum.

Research shows that EFL learners prefer native speaker instructors (Watson Todd & Pojanapunya, 2009) because they are assumed to be ideal models (Gurkan & Yuksel, 2012). The use of English with non-native instructors is likely to result in inauthentic L2 input (Flege & Liu, 2001). If the input comes from speakers of a student's own L1, it may reinforce the type of errors the students themselves might be apt to make. Learners' perceptions of native and nonnative instructors' teaching of prosody may vary depending on the accent of the instructor (Ma, 2012), the physical appearance of the instructor (Golombek & Jordan, 2005), the learners' proficiency level (Madrid & Canado, 2004), the learners' previous experience with nonnative instructors (Braine, 2005), perceptions of instructors' skill at teaching prosody (Madrid & Canado, 2004), and the instructors' strength in speaking confidently and fluently (Reves & Medgyes, 1994). Considering all these variables, the accent of the instructor, physical appearance, and the learners' proficiency levels may be especially crucial. Munro, Derwing and Morton (2006) pointed out that L2 students evaluate the accentedness of nonnative speech in much the same way native-English-speaking listeners do. This suggests that learners are likely to be attentive to the pronunciation of their instructors and use this information to gauge the instructor's appropriateness as a pronunciation instructor. In terms of L2 speech learning, to form new phonological categories, learners need to be exposed to as many different speakers as possible. Thomson and Derwing (2016) found that after more than a month of High Variability Phonetic Training (HVPT), English learners from various L1 backgrounds showed different amounts of improvement in production depending on both the type of training materials and the way the gains were assessed.

The findings in the present study can be related to the awareness of non-native instructors about the international variety of English. For instance, Cook (2005) indicated that NNESTs provide models of proficient second language users in action in the classroom, and also examples of people who have become successful second language users. Moreover, Modiano's (2005) study showed that NNESTs would be more aware of learning an international variety of English and would be in a better position to encourage diversity since they did not belong to a

specific variety of English. As a result, students would learn more about how English operates in a diverse number of nation states so that they can gain better understanding of the wide range of English language usage (Tsou & Chen, 2017).

Some practitioners believe that employing a native instructor would somehow result in greater enhancement by catching pronunciation in much the same way one catches a corona virus, through exposure alone (see Levis, 2015b). The corona virus perspective also may elucidate their stated uncertainty about non-native instructors, who could potentially infect them with bad pronunciation. Their positive mentality to state a preference for a native speaker instructor regarding pronunciation teaching mirrors the power of the nativeness principle over the intelligibility principle (Levis, 2005b; Yenkimaleki & van Heuven, 2022). However, the findings reflects that their perspectives are wrong and that even though a native speaker bias for pronunciation teaching is solid, learning pronunciation is seemingly dependent upon variables other than whether the instructor is or is not a native speaker of the language being taught.

Therefore, to move forward, and in line with the assertions of Levis (2020) and Jenkins (2007), a major paradigm shift is required. Understanding and confidence in NNESTs' own pronunciation must be fostered in all stakeholders to support the teaching of English pronunciation by non-native speaking instructors. Moreover, non-native English instructors need to become comfortable and confident with their own pronunciation and pronunciation teaching abilities. This is not an easy task as it challenges a history of long held beliefs in EFL settings. However, it can start by educating stakeholders concerning recent developments in the pronunciation of English and providing clear descriptions, demonstrations, and models of what intelligible pronunciation sounds like. It might be possible to help these instructors move forward through focused attention on cognition development and identity construction (e.g., Burri et al., 2017). It could be pointed out that challenging their current beliefs on good pronunciation and how they view themselves in relation to that, might help to shift their focus to considering things from an intelligibility standpoint (rather than a nativelike standpoint). Such cognition and identity focus can result in a different outlook toward their own pronunciation and pronunciation instruction resulting in increased acceptance and confidence in their own intelligible pronunciation model. Through changes in instructors' cognition and identity related to English pronunciation and teaching, they may be able to overcome self-held pronunciation biases stemming from the long history of following the nativeness principle (Yenkimaleki & van Heuven, 2022). Thus, in line with Burri (2015), instructors must be provided with the opportunity to attend pronunciation-focused teacher education programs or courses which focus on fostering the understanding and acceptance of the intelligibility principle and intelligible models in English pronunciation teaching instruction and building up their ability to implement intelligibility focused pronunciation instruction in their classroom. Fostering confidence in non-native instructors' pronunciation and teaching pronunciation despite the impact of the existing native model prejudice in the EFL settings would allow EFL instructors to facilitate a learning environment in which their students are more likely to embrace their pronunciation teaching (Whitehead & Ryu, 2023).

This study has some limitations. We did not have access to a large number of instructors with the same degree, and years of experience in pronunciation teaching in an EFL context. For instance, the comparison groups for native English instructors and non-native English instructors were uneven (thirteen L2 instructors against seven L1 instructors). The non-native

English instructor group consisted of three PhDs while the native English instructor group was Masters-level.

The non-native instructors in the present study all speak fluent English, albeit with an identifiable Arabic or Persian accent. Their English can be easily understood by listeners who share the L1 with the instructor (or have grown accustomed to their accent in some other way). Our study cannot answer the question of what the effect would be on the development of comprehensibility of the students if the NNESTs' accent had been stronger. The present findings should not be generalized beyond the instruction of highly qualified NNESTs at the university level to less well trained instructors, for instance, in secondary school settings. On the one hand, the use of authentic, native English audio and video materials may provide an adequate pronunciation model for the EFL learners. On the other hand, students may not be challenged to develop pronunciation that would sound more native-like than that of their instructors. A follow-up study would be required to determine the effect of systematically varied strength of nonnative accent of the instructors.

Finally, the expert raters employed in the present study were highly familiar with the type of non-native accent produced by the interpreter trainees. Two of the raters shared the L1 (i.e., Persian) with the students, while the third rater, a native speaker of American English, had been employed as an EFL teacher in several consecutive appointments in universities in Iran. It has been shown in the literature that a speaker's intelligibility and comprehensibility is better when the listener shares the L1 with the speaker (the shared interlanguage intelligibility benefit, e.g., Bent & Bradlow, 2003; Wang & van Heuven, 2015, and references therein). More research is needed to determine the Persian EFL speakers' comprehensibility for listeners with other native languages, and – especially – if the positive effects of dedicated prosody training persist when the listeners do not share the L1 of the speakers.

Conclusion

Overall, the results showed that speech comprehensibility of the experimental groups improved compared with the control groups. Furthermore, the performance of the interpreter trainees in the experimental groups (with either native or non-native instructor groups) did not differ significantly. It could be concluded that the effectiveness of prosody instruction in developing students' speech comprehensibility depends mainly on instructors' knowledge base of pronunciation teaching and not necessarily on their status as a native or non-native speaker instructor.

Pronunciation is an essential component for intelligible speech, an aim for almost all language students. Development of speech comprehensibility does not require native speaker instructors, as it is unlikely that learners will catch good pronunciation skills from native speakers and equally unlikely that they will catch bad pronunciation from nonnative speakers (Levis, 2015). Nonnative instructors, as in all language skills, can bring great advantages to the teaching of prosody. Non-native instructors know what it is like to learn the target language, and they perceive the tricks and pitfalls their learners may fall into. Moreover, countries require nonnative instructors since native speakers are a minority of language instructors in the world, and if pronunciation is to be taught, it must be taught by all kinds of qualified teachers (Levis, 2015).

Acknowledgement

We would like to thank all the students for their invaluable help in accomplishing this study. A special word of thanks goes to the two colleagues at UAST who generously donated their time and effort to assist in the rating of the 644 recordings.

About the Authors

Mahmood Yenkimaleki (PhD, Leiden University) is an assistant professor of Interpreting and Translation Studies at Bu-Ali Sina University, Iran. He is also an affiliated researcher at Leiden University, The Netherlands. His area of interest is Interpreting Studies and Applied Linguistics. ORCID ID: 0000-0002-8712-9004

Vincent van Heuven (PhD, Utrecht University) is emeritus professor of Experimental Linguistics and Phonetics at Leiden University and the University of Pannonia (Veszprém, Hungary) and former director of the Leiden University Centre for Linguistics. He served on the editorial boards (and as associate editor) of *Journal of Phonetics* (Elsevier) and *Phonetica* (Karger) and was editor of the series *Speech Research* (Mouton de Gruyter). He is a life member of the Royal Netherlands Academy of Arts and Sciences. ORCID ID: 0000-0003-3631-5699

To Cite this Article

Yenkimaleki, M. & van Heuven, V. (2024). Developing interpreter trainees' speech comprehensibility: Does nativeness of the instructor matter? *Teaching English as a Second Language Electronic Journal (TESL-EJ), 27* (4). https://doi.org/10.55593/ej.27108a8

References

Alseweed, M., & Daif-Allah, A. (2012). University students' perceptions of the teaching effectiveness of native and nonnative teachers of English in the Saudi context. *Language in India, 12* (7), 35–60. http://www.languageinindia.com/july2012/alseweedaymannonnativefinal.pdf

Aslan, E., & Thompson, A.S. (2017). Are they really 'two different species'? Implicitly elicited student perceptions about NESTs and NNESTs. *TESOL Journal, 8* (2), 277–294. https://doi.org/10.1002/tesj.268

Bent, T., & Bradlow, A. R. (2003). The Interlanguage speech intelligibility Benefit. *Journal of the Acoustical Society of America, 114* (3), 1600–1610. https://doi.org/10.1121/1.1603234

Borg, S. (2006). The distinctive characteristics of foreign language teachers. *Language Teaching Research, 10* (1), 3–31. https://doi.org/10.1191/1362168806lr182oa

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological), 26* (2), 211–252.

Braine, G. (2010). *Nonnative speaker English teachers: Research, pedagogy, and professional growth*. Routledge.

Brinton, D. (2004). Nonnative English-speaking student teachers: Insights from dialogue journals. In L. D. Kamhi-Stein (Ed.), *Learning and teaching from experience: Perspectives on nonnative English-speaking professionals* (pp. 190–205). University of Michigan Press.

Burri, M. (2015). “My perspective changed dramatically”: A case for preparing L2 instructors to teach pronunciation. *English Australia Journal, 31* (1), 19–37. Retrieved from https://ro.uow.edu.au/sspapers/1945/

Burri, M., Chen, H., & Baker, A. (2017). Joint development of teacher cognition and identity through learning to teach L2 pronunciation. *The Modern Language Journal, 101* (1), 128–142. https://doi.org/10.1111/modl.12388

Buss, L. (2016). Beliefs and practices of Brazilian EFL teachers regarding pronunciation. *Language Teaching Research, 20* (5), 619–637. https://doi.org/10.1177/1362168815574145

Butler, Y. (2007). How are nonnative-English-speaking teachers perceived by young learners? *TESOL Quarterly, 41* (4), 731–755. https://doi.org/10.2307/40264404

Canagarajah, S. (2005). *Reclaiming the local in language policy and practice.* Lawrence Erlbaum.

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

Clark, E., & Paran, A. (2007). The employability of non-native speaker teachers of EFL: A UK survey. *System, 35* (4), 407–430. https://doi.org/10.1016/j.system.2007.05.002

Cook, V. (2005). Materials for adult beginners from an L2 user perspective. In B. Tomlinson (Ed.), *Developing materials for language teaching* (pp. 275-290). Continuum.

Couper, G. (2016). Teacher cognition of pronunciation teaching amongst English language teachers in Uruguay. *Journal of Second Language Pronunciation, 2* (1), 29–55. https://doi.org/10.1075/jslp.2.1.02cou

Couper, G. (2017). Teacher cognition of pronunciation teaching: Teachers’ concerns and issues. *TESOL Quarterly, 51* (4), 820–843. https://doi.org/10.1002/tesq.354

Cutler, A. (2005). Lexical stress. In D. B. Pisoni, & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 264–289). Blackwell. Retrieved from https://repository.ubn.ru.nl/handle/2066/56545

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. MIT Press.

Denes, P. B., & Pinson, E. N. (1963). *The Speech Chain. The physics and biology of spoken language*. Bell Telephone Laboratories. https://www.academia.edu/32963495/

Derwing, T. M., & Munro, M. J. (2005). Pragmatic perspectives on the preparation of teachers of English as a second language: Putting the NS/NNS debate in context. In E. Llurda (Ed.), *Non-native language teachers: Perceptions, challenges and contributions to the profession* (pp. 179–191). Springer. https://doi.org/10.1007/0-387-24565-0_10

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Benjamins. https://doi.org/10.1075/lllt.42

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39* (3), 399–423. https://doi.org/10.2307/3588487

Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition, 23* (4), 527–552. https://doi.org/10.1017/S0272263101004041

Ekmekçi, E. (2016). Comparison of native and non-native English language teachers' evaluation of EFL learners' speaking skills: Conflicting or identical rating behaviour? *English Language Teaching, 9* (5), 98–105. https://doi.org/10.5539/elt.v9n5p98

Foote, J., Holtby, A., & Derwing, T. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal, 29* (1), 1–22. https://doi.org/10.18806/tesl.v29i1.1086

Foote, J. A., Trofimovich, P., Collins, L., & Soler Urzúa, F. (2016). Pronunciation teaching practices in communicative second language classes. *The Language Learning Journal, 44* (2), 181–196. https://doi.org/10.1080/09571736.2013.784345

Franceschi, V., & Vettorel, P. (2017). ELF users' perceptions of their non-nativeness in digital communication through English: Falling short of the target? *Altre Modernità*, 4, 133–148. https://doi.org/10.13130/2035-7680/8307

Gardner, R. (2007). Motivation and second language acquisition. *Porta Linguarum*, 8, 9–20. Retrieved from http://hdl.handle.net/10481/31616

Golombek, P., & Jordan, S. (2005). Becoming 'black lambs' not 'parrots': A post-structuralist orientation to intelligibility and identity. *TESOL Quarterly, 39* (3), 513–533. https://doi.org/10.2307/3588492

Gooskens, C., & Heuven, V. J. van (2021). Mutual intelligibility. In M. Zampieri, & P. Nakov (Eds.), *Similar languages, varieties, and dialects: A computational perspective* (pp. 51–95). Cambridge University Press. https://doi.org/10.1017/9781108565080

Grim, F., & Sturm, J. L. (2016). Where does pronunciation stand in the 21st century foreign language classroom? Views of educators and learners. In J. Levis, H. Le, I. Lucic, E. Simpson, & S. Vo (Eds.), *Proceedings of the 7th pronunciation in second language learning and teaching conference* (pp. 51–64). Iowa State University.

Gurkan, S., & Yuksel, D. (2012). Evaluating the contributions of native and nonnative teachers to an English Language Teaching program. *Procedia — Social and Behavioral Sciences*, 46, 2951–2958. https://doi.org/10.1016/j.sbspro.2012.05.596

Jenkins, J. (2007). *English as a lingua franca: Attitude and identity*. Oxford University Press.

Hayati, A., & Mohmedi, F. (2011). The effect of films with and without subtitles on listening comprehension of EFL learners. *British Journal of Educational Technology, 42* (1), 181–192. https://doi.org/10.1111/j.1467-8535.2009.01004.x

Heuven, V. J. van (1986). Some acoustic characteristics and perceptual consequences of foreign accent in Dutch spoken by Turkish immigrant workers. In J. van Oosten, & J. F. Snapper, (Eds.), *Dutch Linguistics at Berkeley* (pp. 67–84). The Dutch Studies Program, UC Berkeley. Retrieved from https://scholarlypublications.universiteitleiden.nl/handle/1887/2831

Heuven, V. J. van (2022). Resolving the prosody paradox. In P. Arantes, & A. Post da Silveira, (Eds.), *Prosody and Bilingualism* (168–204). Letraria. https://www.researchgate.net/publication/359025608

Heuven, V. J. van, & Yenkimaleki, M. (2024). Crash modules to help Persians speak more intelligible and comprehensible English, emphasizing either production or perception of either sounds or melodies. *Journal of Second Language Pronunciation.* https://benjamins.com/catalog/jslp.23053.van

Karakaş, A. (2019). Preferred English accent and pronunciation of trainee teachers and its relation to language ideologies. *PASAA, Journal of Language Teaching and Learning in Thailand, 58* (2), 264–294. Retrieved from http://files.eric.ed.gov/fulltext/EJ1227154.pdf

Kochem, T. (2022). Exploring the connection between teacher training and teacher cognitions related to L2 pronunciation instruction. *TESOL Quarterly, 56* (4), 1136–1162. https://doi.org/10.1002/tesq.3095

Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics, 36* (3), 345–366. https://doi.org/10.1093/applin/amu040

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39* (3), 369–377. https://doi.org/10.2307/3588485

Levis, J. M. (2015a). Pronunciation for English as an international language: from research to practice. *Asian Englishes, 17* (2), 174–177. https://doi.org/10.1080/13488678.2015.1021034

Levis, J. (2015b). Learners' views of social issues in pronunciation learning. *Journal of Academic Language and Learning, 9* (1), A-42–A-55. https://Journal.aall.org.au/index.php/jall/article/view/367

Levis, J. M. (2016). Accent in second language pronunciation research and teaching. *Journal of Second Language Pronunciation, 2* (2), 153–159. https://doi.org/10.1075/jslp.2.2.01lev

Levis, J. (2018). *Intelligibility, oral communication and the teaching of pronunciation*. Cambridge University Press.

Levis, J. (2020). Revisiting the intelligibility and nativeness principles. *Journal of Second Language Pronunciation, 6* (3), 310–328. https://doi.org/10.1075/jslp.20050.lev

Levis, J. M., Sonsaat, S., Link, S. & Barriuso, T. (2016). Native and nonnative teachers of L2 pronunciation: Effects on learner performance. *TESOL Quarterly, 50* (4), 894–931. https://doi.org/10.1002/tesq.272

Li, Y. & Zhang, G. (2016). Native or non-native-speaking teaching for L2 pronunciation teaching? — An investigation on their teaching effect and students' preferences. *English Language Teaching,* 9 (12), 89–97. https://doi.org/10.5539/elt.v9n12p89

Lippi-Green, R. (2011). *English with an accent: Language, ideology, and discrimination in the United States*. Routledge.

Ma, L. (2012). Advantages and disadvantages of native- and nonnative-English speaking teachers: Student perceptions in Hong Kong. *TESOL Quarterly, 46* (2), 280–305. https://doi.org/10.1002/tesq.21

Madrid, D., & Canado, M. L. P. (2004). Teacher and student preferences of native and nonnative foreign language teachers. *Porta Linguarum, 2,* 125–138. https://dialnet.unirioja.es/descarga/articulo/1129932.pdf

Mahboob, A. (2010). The NNEST lens. In A. Mahboob (Ed.), *The NNEST lens: Non-native English speakers in TESOL* (pp. 1–12). Cambridge Scholars.

Mahboob, A., & Golden, R. (2013). Looking for native speakers of English: Discrimination in English language teaching job advertisements. *Voices in Asia Journal, 1* (1), 72–81. Retrieved from https://www.academia.edu/4517679/

Modiano, M. (2005). Cultural studies, foreign language teaching and learning practices, and the NNS practitioner. In E. Llurda (Ed.), *Non-native language teachers: Perceptions, challenges, and contributions to the profession* (pp. 25–43). Springer.

Moussu, L. (2006). Native and non-native English-speaking as a second language teachers: Student attitudes, teacher self-perceptions, and intensive English program administrator beliefs and practices. Unpublished doctoral dissertation, Purdue University. Retrieved from https://eric.ed.gov/?id=ED492599

Moussu, L., & Llurda, E. (2008). Non-native English-speaking English language teachers: History and research. *Language Teaching, 41* (3), 315–348. https://doi.org/10.1017/S0261444808005028

Munro, M. & Derwing, T. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language earners. *Language Learning, 45* (1), 73–97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28* (1), 111–131. https://doi.org/10.1017/S0272263106060049

Murphy, J. (2014). Intelligible, comprehensible, non-native models in ESL/EFL pronunciation teaching. *System, 42,* 258–269. https://doi.org/10.1016/j.system.2013.12.007

Nunnally, J. (1978). *Psychometric Theory (Vol. 2)*. New York: McGraw-Hill.

O'Brien-Wilson, E., & Spaulding, T. J. (2010). Effects of noise and speech intelligibility on listener comprehension and processing time of Korean-accented English. *Journal of Speech, Language, and Hearing Research,* 53 (6), 1543–1554. https://doi.org/10.1044/1092-4388(2010/09-0100)

Phillipson, R. (1992). *Linguistic imperialism*. Oxford University Press.

Phillipson, R. (2018). Linguistic imperialism and NNESTs. In J. I. Liontas, & M. DelliCarpini (Eds.), *The TESOL Encyclopedia of English Language Teaching*. Wiley-Blackwell.

Reves, T., & Medgyes, P. (1994). The non-native English speaking EFL/ESL teacher's self-image: An international survey. *System, 22* (3), 353–367. https://doi.org/10.1016/0346-251x(94)90021-3

Rietveld, T. (2020). *Human measurement techniques in speech and language pathology: Methods for research and clinical practice*. Routledge.

Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly, 46* (4), 842– 854. https://doi.org/10.1002/tesq.67

Selvi, A. (2014). Myths and misconceptions about nonnative English speakers in the TESOL (NNEST) movement. *TESOL Journal*, 5(3), 573–611. https://doi.org/10.1002/tesj.158

Smith, L. E., & Rafiqzad, K. (1979). English for cross-cultural communication: The question of intelligibility. *TESOL Quarterly, 13* (3), 371–380. https://doi.org/10.2307/3585884

Tsou, S. Y. & Chen, Y. L. (2017). EFL college students' perceptions toward native and non-native English speaking teachers. *American Journal of Educational Research, 5* (12), 1182–1190. https://doi.org/10.12691/education-5-12-2

Vančová, H. (2019). Current issues in pronunciation teaching to non-native learners of English. *Journal of Language and Cultural Education, 7* (2), 140–155. https://doi.org/10.2478/jolace-2019-0015

Walkinshaw, I., & Duong, O. (2012). Native-and non-native speaking English teachers in Vietnam: Weighing the benefits. *TESL-EJ, 16* (3), 1–17. Retrieved from http://www.tesl-ej.org/wordpress/issues/volume16/ej63/ej63a1/

Wang, H. & Heuven, V. J. van (2015). The interlanguage speech intelligibility benefit as bias toward native-language phonology. *i-Perception, 6* (6), 1–13. https://doi.org/10.1177/2041669515613661

Wang, Z. (2014). Developing accuracy and fluency in spoken English of Chinese EFL learners. *English Language Teaching, 7* (2), 110–118. Retrieved from https://files.eric.ed.gov/fulltext/EJ1075642.pdf

Watson Todd, R., & Pojanapunya, P. (2009). Implicit attitudes towards native and non-native speaker teachers. *System, 37* (1), 23–33. https://doi.org/10.1016/j.system.2008.08.002

Whitehead, G. E. K. & Ryu, Y. (2023). "I am not a native speaker…": Exploring English pronunciation teaching difficulties faced by Korean public elementary school English teachers under the lens of the nativeness principle. *System, 115*, 103056. https://doi.org/10.1016/j.system.2023.103056

Yenkimaleki, M. (2017). *Effect of prosody awareness training on the quality of consecutive interpreting between English and Farsi*. LOT. Retrieved from https://www.lotpublications.nl/Documents/459_fulletext.pdf

Yenkimaleki, M. (2019). Explicit teaching of segmentals versus suprasegmentals in developing speaking skills by interpreter trainees: an experimental study. *FORUM: International Journal of Interpretation and Translation, 17* (2), 250–268. https://doi.org/10.1075/forum.18007.yen

Yenkimaleki, M. (2021). Prosody training benefits in perception vs. production skills in simultaneous interpreting: an experimental study. *Dutch Journal of Applied Linguistics*, 10, 1–15. https://doi.org/10.51751/dujal9888

Yenkimaleki, M., & Heuven, V. J. van (2018). The effect of teaching prosody teaching on interpreting performance: An experimental study of consecutive interpreting from English

into Farsi. *Perspectives: Studies in Translatology*, *26* (1), 84–99. https://doi.org/10.1080/0907676X.2017.1315824

Yenkimaleki, M., & Heuven, V. J. van (2019a). The relative contribution of computer assisted prosody training vs. instructor based prosody teaching in developing speaking skills by interpreter trainees: an experimental study. *Speech Communication, 107*, 48–57. https://doi.org/10.1016/j.specom.2019.01.006

Yenkimaleki, M., & Heuven, V. J. van (2019b). Effects of prosody awareness training on the intelligibility of Iranian interpreter trainees in English. *Dutch Journal of Applied Linguistics, 8* (2), 291–309. https://doi.org/10.1075/dujal.17023.yen

Yenkimaleki, M., & Heuven, V. J. van (2019c). Prosody instruction for interpreter trainees: does methodology make a difference? An experimental study. *Across Languages and Cultures: A Multidisciplinary Journal for Translation and Interpreting Studies, 20* (1), 117–133. https://doi.org/10.1556/084.2019.20.1.6

Yenkimaleki, M., & Heuven, V. J. van (2021). Effects of attention to segmental *vs.* suprasegmental features on the speech intelligibility and comprehensibility of the EFL learners targeting the perception or production-focused practice. *System, 100*, 1–12. https://doi.org/10.1016/j.system.2021.102557

Yenkimaleki, M., & Heuven, V. J. van (2022). Comparing the nativeness vs. intelligibility approach in prosody instruction for developing speaking skills by interpreter trainees: An experimental study. *Speech Communication, 137*, 92–102. https://doi.org/10.1016/j.specom.2022.01.007

Yenkimaleki, M., & Heuven, V. J. van (2023). Relative contribution of explicit teaching of segmentals vs. prosody to the quality of consecutive interpreting by Farsi-to-English interpreting trainees. *Interactive Learning Environments, 31* (1), 451–467. https://doi.org/10.1080/10494820.2020.1789673

Yenkimaleki, M., Heuven, V. J. van, & Moradimokhes, H. (2021). The effect of prosody instruction in developing listening comprehension skills by interpreter trainees: does methodology matter? *Computer Assisted Language Learning, 36* (5–6), 968–1004. https://doi.org/10.1080/09588221.2021.1957942

Zhang, Y., & Elder, C. (2014). Investigating native and non-native English-speaking teacher raters' judgements of oral proficiency in the College English Test-Spoken English Test (CET-SET). *Assessment in Education: Principles, Policy & Practice, 21* (3), 306–325. https://doi.org/10.1080/0969594X.2013.845547

Appendix 1. Example from Zappenglish.com: Talking about hobbies and interests

A: Hi everyone, welcome to ZappEnglish.com. In this lesson, we're going to talk about hobbies and interests.
B: That's right. Hobbies are activities we do for fun or pleasure, and interests are things we like or enjoy.
A: So, let's start with hobbies. What are some of your hobbies?
B: Well, I love playing football and watching movies.
A: Ah, that's great. I'm more into reading and listening to music.
B: What about interests? Do you have any particular interests?
A: Yes, I'm really interested in learning new languages and cultures.
B: That's fascinating. I'm interested in history and politics.
A: So, how about you guys? What are your hobbies and interests?
C: Hi, I'm John. My hobby is playing guitar and my interest is photography.
D: Hi, I'm Sarah. My hobby is hiking and my interest is environmental issues.
A: Excellent. Now, let's practice some vocabulary related to hobbies and interests. Repeat after me: gardening, cooking, painting, dancing, playing chess...
B: And some common interests could be: science, fashion, sports, travel, music...
A: Great job everyone. Remember to use these words and expressions when talking about your hobbies and interests in English.

Appendix 2. Problematic stress patterns

- Students were divided into pairs or small groups. Each group was given a set of sentences or short dialogues in which target words needed to be stressed correctly. The students took turns reading the sentences aloud, paying attention to their own pronunciation and prosody.
- After each student read a sentence, group members provided feedback and discussed whether or not the stress was placed correctly. This feedback and discussion allowed the students to reflect on their own pronunciation and prosody, identifying areas for improvement.
- Throughout the activity, the students were actively thinking about and monitoring their own pronunciation and prosody, making adjustments as necessary to improve their communication skills. They experimented with different stress patterns, listened to their own recordings, and compared them to native speakers' pronunciation (see Yenkimaleki & van Heuven, 2019) to further refine their skills.
- By doing so, students not only practiced and improved their procedural knowledge of English stress patterns but also developed their metacognitive awareness. They became more self-directed and independent in their learning, as they were able to identify their own problem areas and find ways to address them.

An example of a short dialogue

Instructor: Alright, everyone, let's begin our activity on stress patterns in English. I will give you a short dialogue, and your task is to correctly place stress on certain words. Here's the dialogue:
Student A: Hey, have you seen my new car?
Student B: No, I haven't. Where did you park it?
Student A: I left it in the parking lot near the mall.

Appendix 3. Repetition of key words

Instructor: Let's move on to our next activity. This time, we will focus on repetition of key words in a listening passage to practice prosody. I will play a short audio clip for you. Listen carefully and identify the key words that are stressed in the passage. Afterwards, we will discuss together which stresses were placed correctly. Here's the audio clip:

Narrator: In today's fast-paced world, it's important to find ways to relax and unwind. Many people turn to activities such as yoga, meditation, or simply taking a walk in nature. These practices can help reduce stress and improve overall well-being. So, next time you feel overwhelmed, remember to prioritize self-care and find what works best for you.

Instructor: Now, let's listen to the audio clip again. This time, pay close attention to the emphasized words. Afterward, we will have a discussion.

More and more detailed information on the segmental and prosodic pronunciation instruction modules can be found in van Heuven & Yenkimaleki (2024).