



Universiteit
Leiden
The Netherlands

Effects of prosody awareness training on the intelligibility of Iranian interpreter trainees in English

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

Citation

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

Version: Publisher's Version
License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)
Downloaded from: <https://hdl.handle.net/1887/80379>

Note: To cite this publication please use the final published version (if applicable).

Effects of prosody awareness training on the intelligibility of Iranian interpreter trainees in English

Mahmood Yenkimaleki^{1,2} and Vincent J. van Heuven^{3,4}

¹ Free University | ² University of Nahavand | ³ University of Pannonia |

⁴ Leiden University

The present study investigates the effect of prosodic feature awareness training on the intelligibility of speech produced by Iranian interpreter trainees. Two groups of student interpreters were formed. All were native speakers of Farsi who studied English translation and interpreting at the BA level. Participants took a pretest of speaking skills before starting the program so that their speech intelligibility level was rated. The control group listened to authentic audio tracks in English and discussed their contents, watched authentic English movies, discussed issues in the movies in pairs in the classroom. The experimental group spent part of the time on theoretical explanation of, and practical exercises with, English prosody. Students then took a posttest in speaking skills so that the effect of treatment on the intelligibility of their speech could be assessed. The results show that the prosody awareness training significantly improved the students' speech intelligibility.

Keywords: prosody, speech intelligibility, awareness training, interpreter trainees, curriculum design

1. Introduction

Intelligibility has been recognized as an important goal for pronunciation teaching but little is known about the factors which make language learners' speech intelligible (Field, 2005). Jenkins (2000) tried to establish which aspects of pronunciation cause intelligibility problems by drawing up a pronunciation core from interactions among non-native speakers of English when English is spoken as an international language. Burns (2003) pointed out that it is very important for speakers to be able to achieve intelligibility (the sound patterns produced by

the speaker are recognizable as English). In examining the role of stress, i.e., the degree of force used in producing a syllable (Crystal, 2003), in intelligibility, Field (2005) asked trained listeners to transcribe recorded materials in which word stress and vowel quality were manipulated. When word stress was erroneously shifted to an unstressed syllable, even without a change in vowel quality, utterances were less intelligible than when only vowel quality was manipulated. Non-native pronunciation of segmentals and prosody in English as a foreign language (EFL) contribute to the perception of foreign accent, and may compromise intelligibility (Trofimovich & Baker, 2006; Munro & Derwing, 2008; Cutler, 2012).

Assuming that intelligibility is a primary goal of pronunciation teaching, we may ask what level of intelligibility we should aim for in the EFL curriculum. Abercrombie (1956) introduced the now classical concept of 'comfortable intelligibility', that is a level of intelligibility on the part of EFL speakers interacting with native English speakers such that the native listener need not make an effort to understand the EFL speaker. More concrete criteria for what constitutes comfortable intelligibility have been proposed by, e.g., Celce-Murcia, Brinton, & Goodwin (1996), Morley (1991) and Walker (2001).

This concept of comfortable intelligibility, however, has lost much of its relevance, since the number of EFL speakers around the world is now two to three times larger (over one billion) than the number of native English speakers (about 400 million) because of the advance of globalization through English (Kachru, 1985; Crystal, 2004; Graddol, 2006), and oral communication among EFL speakers from different first language (L1) backgrounds has increased substantially. As a result of this development, EFL learners are more often engaged in transactions with other EFL speakers than with native speakers of English. It is not entirely clear at this time whether the native English norm should be approximated more closely in the communication with EFL listeners than with native English listeners. Native listeners tolerate large deviations from the norm before they experience discomfort (e.g. Van Heuven, 2008) and will understand an EFL speaker who may be unintelligible to other EFL speakers – unless the EFL speaker and listener have the same native language (the so-called shared interlanguage speech intelligibility benefit, ISIB, e.g. Bent & Bradlow, 2003; Wang & Van Heuven, 2015).

Jenkins (1998) proposed a new concept of intelligibility, which she called mutual intelligibility. It is defined as the level of intelligibility which is required for EFL speakers to communicate successfully with other EFL speakers from different L1 backgrounds. This concept excludes the communication between EFL speakers from the same language background. Moreover, the concept does not involve any notion of comfort but is defined instead on the basis of successful communication through joint efforts by EFL speakers and listeners (Jenkins, 2000, 2002; Howlader, 2010). The element of speaker-hearer interaction, in our view, ren-

ders it rather difficult to lay down precise levels of pronunciation accuracy for EFL speakers: a poor speaker may yet be engaged in successful interaction if the interlocutor is an exceptionally gifted or experienced (or even native) listener.¹ Nevertheless, this type of intelligibility is now regarded as a legitimate goal of pronunciation teaching today.

Research shows that prosodic feature awareness training has positive effects both on speech perception and production in EFL contexts (e.g., Champagne-Muzar, Schneiderman, & Bourdages, 1993; Pennington, 1998; Ahrens, 2004; Derwing & Munro, 2005; Venkatagiri & Levis, 2007; Foote, Holtby, & Derwing, 2011; Derwing, Diepenbroek, & Foote, 2012; Robinson, Mackey, Gass, & Schmidt, 2012; Suwartono, 2014; Koike, 2014; Yenkimaleki, 2016, 2017; Yenkimaleki & Van Heuven, 2013, 2016a, b, c, d, e). Prosody makes an important contribution to EFL speech recognition, and especially to speech understanding (Munro & Derwing, 2008). Anderson-Hsieh, Johnson & Koehler (1992) compared the relative contribution of prosody versus segmentals to the (judged) intelligibility of L2 speech by groups of speakers belonging to eleven different language backgrounds. Irrespective of the participants' L1, prosodic errors (stress, rhythm, intonation) all affected native listeners' judgments more than segmental errors.

Most EFL communication takes place between interactants who are both proficient in English, and who are able to use English as a lingua franca. All of the studies reviewed above address the intelligibility and comprehensibility of English with this type of interaction in mind. A more complicated situation arises when the primary interactants do not understand each other's language. In such situations an interpreter with command of both languages A and B may be called upon to mediate between the interactants. In the case of consecutive interpreting, the interpreter waits until speaker A has produced a chunk of speech in language A, typically of paragraph length, and then produces a semantic equivalent of the chunk in the language of listener B. Then speaker and listener will reverse roles and the interpreter will take B's response as input and produce the equivalent in language A. In our study the interpreters have no equal command of the languages A and B. They are native speakers of Farsi (Modern Persian) as spoken in Iran, with English as a foreign language. In the present study the interpreter's

1. It is, of course, possible to extend the concept of mutual intelligibility such that it includes successful communication between native and non-native speakers of English. Moedjito (2015) seems to imply that this 'global intelligibility' presupposes higher oral and aural skills than Jenkins' mutual intelligibility does – which, we would argue, is a fallacy. Native listeners typically understand non-native speakers better than non-native listeners do – even in the case of a shared interlanguage. The interlanguage benefit is only seen in relative and not in absolute terms (Van Heuven, 2015; Wang & Van Heuven, 2015).

task is to listen to passages spoken in the native language Farsi and to produce a spoken equivalent of the input in non-native English. This is often referred to as inverse (or *verso*) interpreting. *Verso* interpreting is held to be a more challenging task than direct (or *recto*) interpreting, i.e., converting input in the non-native language to the interpreter's native language. We assume that the non-native English output of the *verso* interpreting will be more intelligible (to native and non-native English listeners alike) as the prosody is closer to the English norm. It is our hypothesis that a better approximation of the prosody of English can be attained by the Farsi-English interpreters through explicit instruction in the prosodic features of English and the differences in word and sentence prosody between English and Farsi. The better use of English prosody will yield a more intelligible output in English.

In previous studies we tested the effect of a prosody-enhanced training program on the consecutive interpreting performance by groups of Iranian students of translation and interpreting. The effects were investigated both for groups of students concentrating on *recto* interpreting (Yenkimaleki & Van Heuven, 2013, 2018), and on *verso* interpreting (Yenkimaleki & Van Heuven, 2016a). In these studies, the students' interpreting skills were rated on ten different aspects, such as accuracy of content, use of technical vocabulary, grammatical correctness, and fluency, but none of the rating scales addressed the interpreter's intelligibility. The results showed that the prosody awareness training improved the student's interpreting performance, especially in terms of prosody-related rating scales, more so for *recto* interpreting than for *verso* interpreting. We also tested the effect of the program on the development of the students' English word recognition skills (Yenkimaleki & Van Heuven, 2016b) and on their overall English listening comprehension skills (Yenkimaleki & Van Heuven, 2016c). Again, the prosody-enriched training yielded better scores than the routine training program, which we consider evidence that awareness of the differences in word and sentence prosody between Farsi and English helps the Iranian listener to better decode the foreign English input. In one earlier study, finally, we aimed to test specifically whether the prosody awareness training improves the students' intelligibility, i.e. an aspect of the students' performance in *verso* interpreting (from native Farsi into non-native English). Before and after the training program the participants' speaking skills were assessed by three Iranian expert interpreters in an interview. The student's speech production was rated on four scales, i.e. comprehensibility, pronunciation, grammar and vocabulary. The results revealed a significant improvement of the student's pronunciation of the vowels and consonants (1.1 on the 5-point scale) and a smaller and only marginally significant improvement of comprehensibility (0.4 on the 5-point scale). Predictably, no effect of the prosody training was found for the grammar and vocabulary scales.

The present experiment aims to replicate the earlier study while targeting more specifically the effect of the prosody training on student's intelligibility. We did this by introducing several changes vis a vis the earlier experiment. First of all, we replaced one of the raters of the students' performance by a native speaker of English. This was done to ascertain whether the Iranian experts would judge the students' performance in the same way as native speakers of English would. Secondly, we instructed the raters to judge the students' performance by only one scale, i.e. intelligibility, technically defined as the ease with which the words in the speaker's vocal output can be recognized in the order in which they were spoken (e.g. Gooskens, Van Heuven, Van Bezooijen & Pacilly, 2010; Gooskens & Van Heuven, 2019). Instead of the 5-point rating scale we now asked the experts to rate the student's intelligibility on a scale between 0 and 10. The third change involved reducing the teaching time from 14 sessions of 90 minutes to 14 × 60 minutes. Since the amount of teaching time spent on matters of prosody (20 minutes per session) was unaffected by the reduction, we expect at least the same benefit as in the previous edition of the experiment.

Concretely, we asked the following research question:

Does explicit teaching of prosodic features enhance the speech intelligibility of Farsi-English interpreter trainees?

Our expectation is that the explicit teaching of prosodic features will enhance the speech intelligibility of Farsi-English interpreter trainees.

2. Methodology

2.1 Participants

Twenty-eight student interpreter trainees at the BA level who were majoring in interpreting and translation studies at University of Applied Sciences in Tehran, Iran, were chosen randomly to participate in this study. They were randomly divided into two classes of 14 students that each incorporated 7 male and 7 female students. The participants were native speakers of Farsi within an age range of 18 to 26 years. They participated in all sessions of the training program.

2.2 Procedure

At the beginning of the program all the participants took a pretest of general English proficiency. The test battery was the standard Longman's TOEFL English proficiency test, with separate modules testing the learner's (i) Listening

comprehension, (ii) Reading comprehension and (iii) Structure and writing skills. Then, all participants took a pretest of speaking skill so that intelligibility of participants' speech would be rated before starting the training program.

The participants were divided into control and experimental groups through the application of systematic random sampling. The control group received routine exercises, asking them to listen to authentic audio tracks in English and speaking about the issues brought up in the audio tracks. They also watched authentic movies and discussed the contents of the movie or talked about some proposed hot topic, in pairs in the classroom. The experimental group spent less time on these tasks and instead received awareness training of English prosody in the form of theoretical explanation by the instructor and practical exercises in prosody for 20 minutes during each training session. The participants took part in the program for 14 sessions (sixty minutes per session) in four weeks, i.e. 14 hours in all.

The control group spent 630 minutes in all doing speaking exercises and tasks in the classroom as explained above, while the instructor monitored the discussion and provided feedback whenever needed. Moreover, both the control group and the experimental group listened for 210 minutes to the Iranian instructor who explained how to do exercises and also provided feedback in pair discussions and in doing speaking tasks in the classroom. The experimental group altogether spent 350 minutes on speaking exercises and tasks which were the same as those of the control group. Additionally, the experimental group received 280 minutes of English prosody awareness training and did exercises based on the explanations of prosodic matters (for details of training program see Yenkimaleki, 2017, pp. 52–88). The activities covered by the two participant groups and the time (minutes) spent on them are summarized in Table 1. A more detailed example of one instruction session is given in Appendix 1 (for the experimental group) and Appendix 2 (for the control group).

Table 1. Summary of activities and time spent (minutes) by two groups of participants in the experiment

Activity	Group	
	Control	Experimental
Audio tracks/movies & discussion	630	350
Explanation and feedback by instructor	210	210
Explanation of prosodic theory and practice		280
Total time spent	840	840

In all the sessions, at different times, formative tests were administered to the participants in order to measure their progress and to diagnose problems on the part of the participants. All the participants, whether in the control group or in the experimental group, took a pre-test as well as a post-test on speaking skill so that the effect of treatment on their intelligibility could be assessed. Pre-test and post-test were interviews conducted systematically by two lecturers at the Interpreting and Translation department at the University of Applied Sciences and one native speaker from the United Kingdom. This native speaker had grown up in the south of England. He had obtained an MA degree in linguistics and had come to Iran as an exchange student of Eastern Studies for six months at the University of Applied Sciences in Tehran, Iran. The other two interviewers were native speakers of Farsi who had learned English as a foreign language, held MA degrees in English, worked in the department of English Translation and Interpreting, and were experienced professional consecutive interpreters between English and Farsi.

The interviews lasted eight minutes per participant. The students were asked to describe a scenery/a place to the interviewers that they were familiar with or had grown attached to. The same questions were asked from each participant. The students were given no time to prepare and were instructed to answer the questions on the spot. The questions were the same for the control and experimental groups. One of the interviewers (the first author) knew the students; the other two interviewers did not have any information about the participants' background. Different questions were asked in the post-test but an effort was made to make the pre-test and post-test questions equally difficult.

The students' intelligibility (as defined in the introduction) was rated by the three interviewers independently during or immediately after the interview. The interviewers noted down their marks on a sheet of paper, using an assessment scale running between 0 ('completely unintelligible') to 10 ('perfectly intelligible'). The ratings had to be expressed as whole numbers. The raters did not consult with each other at any time during or after the interview when evaluating the participants' speech intelligibility.

3. Results

Table 2 summarizes the raw component scores of the proficiency test of the control group (left-hand part of table) and of the experimental group (right-hand part).

There were no significant differences whatsoever between the control and experimental groups for any of the three components nor for the overall TOEFL scores (i.e., the unweighted mean of the three components multiplied by 10),

Table 2. Component and overall TOEFL scores obtained by participants in control (left) and experimental (right) groups

Control group						Experimental group					
		Struct.						Struct.			
		List.	&	Read.	Overall			List.	&	Reading	Overall
Nr. ID	Gend.	Comp	Writing	Comp	TOEFL	Nr. ID	Gend.	Comp	Writing	Comp	TOEFL
1.	HaN M	65	63	66	646.6	1.	MaT F	63	61	65	630.0
2.	MoR M	62	61	61	613.3	2.	FaB M	63	60	63	620.0
3.	DaS F	60	57	58	583.3	3.	NiZ M	61	58	56	583.3
4.	HoT M	59	54	56	563.3	4.	SaH F	58	56	60	580.0
5.	GoK F	55	52	55	540.0	5.	SiH F	54	54	55	543.3
6.	JaR M	55	53	53	536.6	6.	BiL M	53	54	54	536.6
7.	MaF M	54	53	52	530.0	7.	HaJ M	52	54	53	530.0
8.	KoB F	53	54	51	526.6	8.	MoH F	52	52	53	523.3
9.	ReL F	52	54	49	516.6	9.	ZiH M	51	53	52	520.0
10.	PeS M	51	52	49	506.6	10.	NeF F	50	52	52	513.3
11.	ZaK F	50	51	49	500.0	11.	NaK F	51	51	51	510.0
12.	KoD F	50	50	49	496.6	12.	HoJ M	49	50	50	496.6
13.	RoG F	49	50	49	493.3	13.	KaT M	48	50	50	493.3
14.	Rap M	48	50	49	490.0	14.	ReS F	47	49	49	483.3
Mean		54.50	53.87	53.28	538.77	Mean		53.71	53.85	54.50	540.21
SD		5.19	3.97	5.34	47.43	SD		5.39	3.71	4.92	46.13

according to t-tests for independent samples, $-.625 < t(26) < .392$ ($p \geq .537$ in all comparisons).

The three raters differed in their overall assessment of the interviewees such that the native English rater found the speakers less intelligible (5.64 on the scale from 0 to 10) than the Iranian raters did, of whom one (the first author) was more critical (6.09) than his colleague (6.96). The effect of rater was highly significant by a Repeated Measures Analysis of Variance with rater as a within-items factor, $F(2, 110) = 61.3$ ($p < .001$, $p\eta^2 = .527$). Since the Iranian raters shared the interviewees’ mother tongue, we would explain the higher scores as an instance of the shared interlanguage intelligibility effect (Bent & Bradlow, 2003; Wang & Van Heuven, 2015). All three raters differed significantly ($\alpha = .050$) from each other by post-hoc analyses with Bonferroni correction for multiple comparisons. In spite of the overall difference between the raters they were in good agreement with respect to the relative ratings across participants, with a Cronbach’s alpha of .898. We therefore decided to analyse the effects of the treatment on the mean of the three ratings given to each participant.

Table 3 presents the mean intelligibility ratings of the participants in the control and experimental groups obtained in the pre-test and the post-test, as well as the gain, i.e., the difference between the post-test and the pre-test. Since the participants in the two groups are matched on their TOEFL scores, the differences between the groups can be evaluated in a within-subjects design. Probabilities will be reported on the basis of two-tailed testing; partial eta squared ($p\eta^2$) will be used as the measure of effect size.

The small difference between the control (5.57) and experimental (5.79) groups in the pre-test is not statistically significant, $t(13)=1.4$ ($p=.189$, $p\eta^2=.054$). This means that we are justified in considering the two groups equivalent at the start of the experiment, as was evidenced earlier by the absence of any difference on the TOEFL test.

Table 3. Pre-test and post-test scores and gain (difference) in speech intelligibility for control (left) and experimental (right) groups. The bottom two rows contain the mean and standard deviation of the scores. Participants are ordered as in Table 2

Control group						Experimental group					
Nr.	ID	Gender	Pretest	Posttest	Gain	Nr.	ID	Gender	Pretest	Posttest	Gain
1.	HaN	M	7.33	7.33	0.00	1.	MaT	F	6.33	7.67	1.34
2.	MoR	M	6.67	8.67	2.00	2.	FaB	M	6.67	7.67	1.00
3.	DaS	F	5.67	7.33	1.66	3.	NiZ	M	6.67	9.33	2.66
4.	HoT	M	5.67	6.67	1.00	4.	SaH	F	7.00	9.33	2.33
5.	GoK	F	6.33	6.00	−0.33	5.	SiH	F	6.33	8.33	2.00
6.	JaR	M	4.67	5.33	0.66	6.	BiL	M	5.67	7.67	2.00
7.	MaF	M	5.33	6.33	1.00	7.	HaJ	M	4.67	6.67	2.00
8.	KoB	F	5.33	5.00	−0.33	8.	MoH	F	4.67	6.67	2.00
9.	ReL	F	6.33	6.67	0.34	9.	ZiH	M	6.33	6.67	0.34
10.	PeS	M	5.33	5.00	−0.33	10.	NeF	F	6.00	7.33	1.33
11.	ZaK	F	4.67	5.67	1.00	11.	NaK	F	5.33	7.33	2.00
12.	KoD	F	4.67	6.33	1.66	12.	HoJ	M	5.33	5.33	0.00
13.	RoG	F	5.33	6.00	0.67	13.	KaT	M	5.00	6.67	1.67
14.	Rap	M	5.00	5.33	0.33	14.	ReS	F	4.67	5.67	1.00
Mean			5.60	6.26	0.67	Mean			5.76	7.31	1.55
SD			0.81	1.03	0.77	SD			0.82	1.17	0.76

After the treatment, the intelligibility ratings improved significantly for both groups, with a gain of .67 points for the control group, $t(13)=3.2$ ($p=.008$, $p\eta^2=.429$), and of 1.55 points for the experimental group, $t(13)=7.6$ ($p<.001$). The

larger gain for the experimental group is significant as well, $t(13)=2.9$ ($p=.012$, $p\eta^2=.399$).

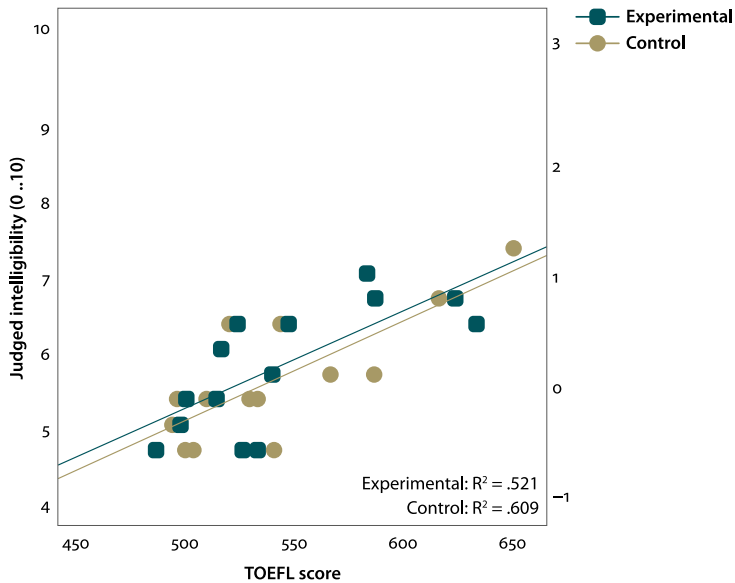
Figure 1 plots the pre-test score (panel A), the post-test score (panel B) and the gain after the treatment (panel C) for each individual participant with separate markers for members of the experimental and control groups.

Figures 1A–B show that there are substantial and highly significant ($p\leq.005$) correlations between the TOEFL scores and the intelligibility ratings, both in the pre-test and in the post-test, which explain between 45 and 61 percent of the variance in the intelligibility scores even when the control and experimental groups are kept separate (as indicated by the r^2 -values in the figure). Moreover, the relationship between the TOEFL scores and the criterion, i.e. pre-test or post-test score, can well be captured by linear functions. Clearly, students are rated as being more intelligible – whether before or after the treatment – as they are more proficient in English at the beginning of the experiment. There is, however, no such linear relationship between the student's proficiency level at the beginning of the treatment and the gain that s/he will obtain by the treatment (Figure 1C). The linear correlation coefficients are low ($r^2\leq.080$), and fail to reach significance ($p=.388$ for the experimental group and $p=.523$ for the control group). Closer inspection of Figure 1C shows that the relationship is U-shaped (quadratic) rather than linear. The quadratic r^2 -value for the control group is insignificant but the correlation for the experimental group is substantial ($r^2=.375$) and significant ($p=.013$). We will come back to this non-linear relationship in the discussion section.

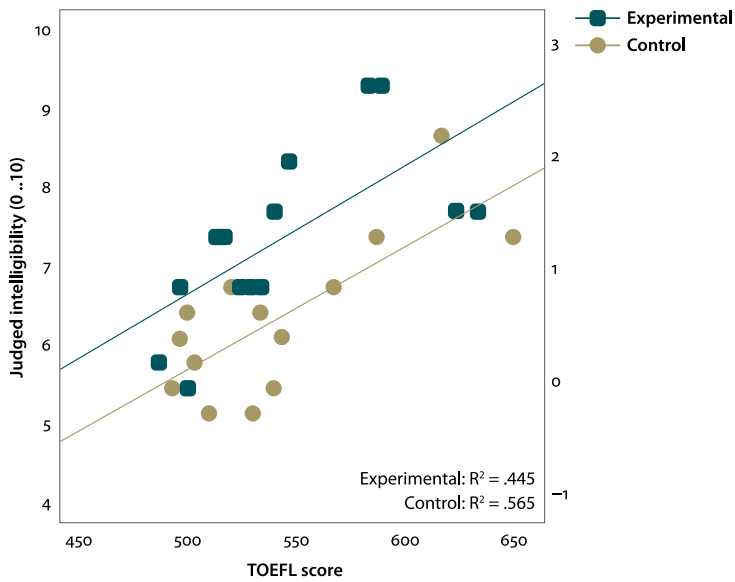
The relative contribution of the overall proficiency in English at the start of the experiment and that of the treatment can be quantified through multiple linear regression. In such an analysis, run in stepwise mode, the TOEFL scores account for 40.9% of the variance in the intelligibility rating in the post-test. The treatment adds another 18.6% so that the total percentage of the variance accounted for equals 59.5. The contribution of the TOEFL score is stronger ($\beta=.632$, $t(26)=5.0$, $p<.001$) than that of the treatment ($\beta=.432$, $t(25)=3.4$, $p=.002$) but the difference is small.

4. Conclusion and discussion

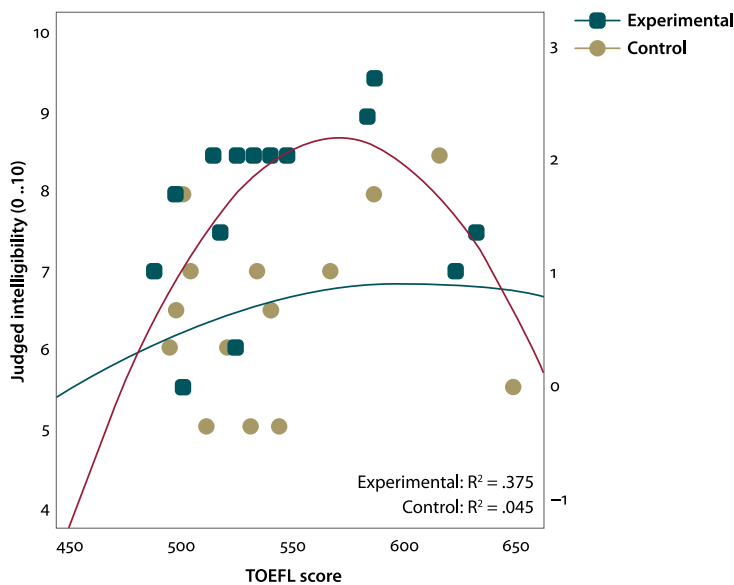
This study investigated the effect of prosodic feature awareness training on the intelligibility of interpreter trainees. The results showed that prosodic feature awareness training significantly contributed to speech intelligibility of interpreter trainees. This perspective is supported by Tsurutani and Ishihara (2012) who



a. Pre-test



b. Post-test



c. Gain (post-test – pre-test)

Figure 1. Intelligibility ratings obtained by experimental and control groups in the pre-test (panel A), post-test (panel B), and the gain after the treatment (panel C), as a function of the participant’s TOEFL score at the start of the experiment. R² values are for linear regression for panels A and B, and for quadratic regression for panel C

stated that prosodic features have significant impact on the intelligibility of L2 learners’ pronunciation. Our finding converges with language researchers’ claim (e.g. Anderson-Hsieh et al., 1992; Gilbert, 1995; Celce-Murcia et al., 1996; Munro & Derwing, 1997; Mouri, Hirose, & Minematsu, 2003) that prosodic errors seriously compromise speech intelligibility and impair the intelligibility of L2 learners’ speech.

In earlier studies we investigated the effects of our prosody training program on the quality of the interpreter trainees’ output directly. In these earlier studies (see introduction) we ran experiments on groups of students who interpreted English into Farsi (*recto*) and of students who interpreted from Farsi into English (*verso*). We found that, indeed, the prosody training yielded interpreting into (non-native) English that was rated more favorably than the performance by the control group which had not received the prosody training. The prosody training was shown to be beneficial in terms of improved pace of delivery, better accentuation as judged by (non-native) Farsi-English interpreting instructors. It was assumed at the time that these improvements could be summarized as compo-

nents of better intelligibility of the non-native output. The results of the present experiment indicate that this is indeed the case.

The effect obtained in the present experiment is a gain in intelligibility of 1.55 points on a scale from 0 to 10 (= 14%) for the prosody group versus .67 points (6%) for the control group. In the earlier experiment we found a gain in comprehensibility (rather than intelligibility) of .4 points on a scale from 1 to 5 (= 8%) for the experimental group against zero gain (0%) for the control group. The quality of the segmental pronunciation did not improve significantly for the control group (.1 point = 2%) but it did for the prosody group (1.2 point = 24%).

Prosody impacts differently on speech intelligibility than on speech understanding. If we define a speaker's intelligibility as the ease with which a listener may recognize the speaker's words in the order as produced by the speaker, then sentence stress and intonation will impact only marginally on a speaker's intelligibility. The words will be recognized irrespective of the sentence melody and phrasing. A misplaced sentence stress will not prevent the listener from recognizing a word as long as the pitch change that signals the sentence stress is on the correct syllable in the word. Sentence prosody is relevant to speech understanding rather than speech recognition – and only the latter is the process that defines intelligibility. Intelligibility is affected by incorrect word stress – especially when the word stress is realized as a sentence stress (i.e. occurs in a communicatively important word). However, whether incorrect word stress creates a word recognition problem depends on the language background of the listener. If the listener's L1 is not a language with contrastive stress, the person will be largely stress deaf, so that for such a listener the stress error is not a major problem (e.g. Peperkamp & Dupoux, 2002; Dupoux, Peperkamp, & Sebastián-Gallés, 2010). We predict from this that French listeners are not harmed very much by incorrect English or Spanish word stress but Dutch listeners would suffer seriously. In fact, Cutler showed that Dutch listeners are (even) more susceptible to incorrect word stresses in English than native English listeners are (for details see Cutler & McQueen, 2014).

The benefits of our prosody training for Iranian-to-English interpreters should therefore be differentiated in terms of sentence prosody (intonation pattern, sentence stress marking focus and phrasing) and word prosody (stress placement). Improved sentence prosody will make the interpreter's output more comprehensible, while improved word prosody will yield better intelligibility, i.e., the words spoken by the interpreter will be recognized more easily. In the earlier experiment (Yenkimaleki & Van Heuven, 2016d) we showed that our prosody awareness training yielded a small (and marginally significant) improvement in comprehensibility, together with a large increase in the judged quality of the pronunciation. The present experiment provides the additional insight that the prosody training increases the intelligibility of the Iranian learners of English. It

would seem reasonable to assume that the increased intelligibility is due to the improved pronunciation on the part of the learner, and that the better intelligibility in turn boosts the learner's comprehensibility.

That the success of speech communication depends on the quality of the speaker's pronunciation can be argued as a matter of logic – rather than as the results of experimental studies. If a listener cannot recognize the sounds, word recognition fails, and communication breaks down. Incorrect choice of words and flawed word order can only compromise intelligibility if the incorrectly used or placed words are recognized in the first place (Van Heuven & De Vries, 1981; Van Heuven, 1986; Wang, 2014). So, the real question is not whether pronunciation is important to speech intelligibility but rather: can we predict how far the sounds in a word may deviate from the listener's norm before word recognition fails? When the listener is a non-native speaker of English the answer depends on the interaction between the phonologies of the speaker's language and the listener's language. The closer the phonologies match, the better the chances of successful sound identification and word recognition (Wang & Van Heuven, 2015; Van Heuven, 2016; Van Heuven & Gooskens, 2017).

The pedagogical implications of our study would be that instructors in EFL settings should consider, and then include, prosody teaching in the curriculum. This will help EFL learners (including interpreter trainees) to increase their second language proficiency and become more intelligible and comprehensible, which is a precondition for successful communication.

References

- Abercrombie, D. (1956). *Problems and principles in language study*. London: Longman.
- Ahrens, B. (2004). *Prosodie beim Simultandolmetschen*. Frankfurt am Main: Peter Lang.
- Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speakers' judgements of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure, *Language Learning*, 42, 529–555.
<https://doi.org/10.1111/j.1467-1770.1992.tb01043.x>
- Bent, T., & Bradlow, A.R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*, 114, 1600–1610. <https://doi.org/10.1121/1.1603234>
- Burns, A. (2003). *Clearly speaking: Pronunciation in action for teachers*. Sydney: National Center for English Language Teaching and Research, Macquarie University.
- Celce-Murcia, M., Brinton, D.M., & Goodwin, J.M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.
- Champagne-Muzar, C., Schneiderman, E.I., & Bourdages, J.S. (1993). Second language accent: The role of the pedagogical environment. *International Review of Applied Linguistics*, 31(2), 143–160.
- Crystal, D. (2003). *A dictionary of linguistics and phonetics*. Oxford: Blackwell.

- Crystal, D. (2004). The past, present, and future of World English. In A. Gardt, & B. Hüppauf (Eds.), *Globalization and the future of German* (pp. 27–46). Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110197297.27>
- Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9012.001.0001>
- Cutler, A., & McQueen, J. (2014). How prosody is both mandatory and optional. In J. Caspers, Y. Chen, W. F. L. Heeren, J. Pacilly, N. O. Schiller, & E. van Zanten (Eds.), *Above and beyond the segments. Experimental linguistics and phonetics* (pp. 71–82). Amsterdam: John Benjamins. <https://doi.org/10.1075/z.189.06cut>
- Derwing, T. M., Diepenbroek, L. G., & Foote, J. A. (2012). How well do general-skills ESL textbooks address pronunciation? *TESL Canada Journal*, 30, 22–44. <https://doi.org/10.18806/tesl.v30i1.1124>
- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39, 379–397. <https://doi.org/10.2307/3588486>
- Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2010). Limits on bilingualism revisited: Stress ‘deafness’ in simultaneous French-Spanish bilinguals. *Cognition*, 114, 266–275. <https://doi.org/10.1016/j.cognition.2009.10.001>
- Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly*, 39, 399–423. <https://doi.org/10.2307/3588487>
- Foote, J. A., Holtby, A. K., & Derwing, T. M. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal*, 29, 1–22. <https://doi.org/10.18806/tesl.v29i1.1086>
- Gilbert, J. B. (1995). Pronunciation practice as an aid to listening comprehension. In D. J. Mendelsohn, & J. Rubin (Eds.), *A guide for the teaching of second language listening* (pp. 97–112). San Diego, CA: Dominie Press, San Diego, CA.
- Gooskens, C., & Van Heuven, V. J. (2019, to appear). Mutual intelligibility. In M. Zampieri, & P. Nakov (Eds.), *Similar languages, varieties, and dialects: A computational perspective*. Cambridge University Press.
- Gooskens, C., Van Heuven, V. J., Van Bezooijen, R., & Pacilly, J. J. A. (2010). Is spoken Danish less intelligible than Swedish? *Speech Communication*, 52, 1020–1037. <https://doi.org/10.1016/j.specom.2010.06.005>
- Graddol, D. (2006). *English next*. London: The British Council.
- Haug Hilton, N., Gooskens, C., & Schüppert, A. (2013). The influence of non-native morphosyntax on the intelligibility of a closely related language. *Lingua*, 137, 1–18. <https://doi.org/10.1016/j.lingua.2013.07.007>
- Howlader, M. R. (2010). Teaching English pronunciation in countries where English is a second language: Bangladesh perspective. *ASA University Review*, 4, 233–244.
- Jenkins, J. (1998). Which pronunciation norms and models for English as an international language? *ELT Journal*, 52, 119–126. <https://doi.org/10.1093/elt/52.2.119>
- Jenkins, J. (2000). *The phonology of English as an international language: New models, new norms, new goals*. Oxford: Oxford University Press.
- Jenkins, J. (2002). A sociolinguistically based, empirical researched pronunciation syllabus for English as an international language. *Applied Linguistics*, 23(1), 83–103. <https://doi.org/10.1093/applin/23.1.83>
- Kachru, B. B. (1985). Standards, codification, and sociolinguistic realism: The English language in the outer circle. In R. Quirk, & H. G. Widdowson (Eds.), *English in the world* (pp. 11–34). Cambridge: Cambridge University Press.

- Koike, Y. (2014). Explicit pronunciation instruction: Teaching suprasegmentals to Japanese learners of English. In N. Sonda, & A. Krause (Eds.), *JALT 2013 Conference Proceedings* (pp. 361–374). Tokyo.
- Moedjito (2015). Factors determining global intelligibility of EFL learners in Asian contexts. *Research on Humanities and Social Sciences*, 5, 155–164.
- Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25(3), 481–521. <https://doi.org/10.2307/3586981>
- Mouri, T., Hirose, K., & Minematsu, N. (2003). Consideration on vowel durational modification for Japanese CALL system. *Proceedings of Eurospeech 2003*, 3153–3156.
- Munro, M. J., & Derwing, T. (1997). Accent, intelligibility and comprehensibility. *Studies in Second Language Acquisition*, 19, 1–16. <https://doi.org/10.1017/S0272263197001010>
- Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, 34, 520–531. <https://doi.org/10.1016/j.system.2006.09.004>
- Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning*, 58, 479–502.
- Pennington, M. C. (1998). The teachability of phonology in adulthood: A re-examination. *International Review of Applied Linguistics*, 36, 323–341. <https://doi.org/10.1515/iral.1998.36.4.323>
- Peperkamp, S., & Dupoux, E. (2002). A typological study of stress ‘deafness’. In C. Gussenhoven, & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 203–240). Berlin: Mouton-De Gruyter. <https://doi.org/10.1515/9783110197105.203>
- Robinson, P., Mackey, A., Gass, S. M., & Schmidt, R. (2012). Attention and awareness in second language acquisition. In S. Gass, & A. Mackey (Eds.), *The Routledge handbook of Second Language Acquisition* (pp. 247–267). London: Routledge.
- Suwartono, S. (2014). Enhancing the pronunciation of English suprasegmental features through reflective learning method. *TEFLIN Journal*, 25, 80–93.
- Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28, 1–30. <https://doi.org/10.1017/S0272263106060013>
- Tsurutani, C., & Ishihara, S. (2012). Naturalness judgement of prosodic variation of Japanese utterances with prosody modified stimuli. *Proceedings of Interspeech 2012*, 646–649.
- Van Heuven, V. J. (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). Berkeley: The Dutch Studies Program, U. C. Berkeley.
- Van Heuven, V. J. (2008). Making sense of strange sounds: (Mutual) intelligibility of related language varieties. A review. *International Journal of Humanities and Arts Computing*, 2, 39–62. <https://doi.org/10.3366/E1753854809000305>
- Van Heuven, V. J. (2015). A relative measure of the interlanguage speech intelligibility benefit: A meta-analytic exercise. In S. Bátyi, & M. Vigh-Szabó (Eds.), *Language – System, Usage, Application* (pp. 31–52). Budapest: Tinta Könyvkiadó.
- Van Heuven, V. J. (2016). An acoustic characterisation of English vowels produced by American, Dutch, Chinese and Hungarian speakers. *Hungarian Journal of Applied Linguistics*, 16(2), 1–20. <https://doi.org/10.18460/ANY.2016.2.004>

- Van Heuven, V.J., & Gooskens, C. (2017). An acoustic analysis of English vowels produced by speakers of seven different native-language backgrounds. In M. Wieling, G. Bouma, & G. van Noort (Eds.), *From semantics to dialectometry. Festschrift in honour of John Nerbonne* (pp. 129–139). Groningen: University of Groningen.
- Van Heuven, V.J., & De Vries, J.W. (1981). Begrijpelijkheid van buitenlanders: de rol van fonische versus niet fonische factoren [Intelligibility of foreigners: the role of phonic versus non-phonic factors]. *Forum der Letteren*, 22, 309–320.
- Venkatagiri, H.S., & Levis, J.M. (2007). Phonological awareness and speech comprehensibility: An exploratory study. *Language Awareness*, 16, 263–277. <https://doi.org/10.2167/la417.o>
- Walker, R. (2001). Pronunciation for international intelligibility. *English Teaching Professional*, 21, 1–7.
- Wang, Z. (2014). Developing accuracy and fluency in spoken English of Chinese EFL learners. *English Language Teaching*, 7, 110–118. <https://doi.org/10.5539/elt.v7n2p110>
- Wang, H., & Van Heuven, V.J. (2015). The interlanguage speech intelligibility benefit as bias toward native-language phonology. *i-Perception*, 6(6), 1–13. <https://doi.org/10.1177/2041669515613661>
- Yenkimaleki, M. (2016). Why prosody awareness training is necessary for training future interpreters. *Journal of Education and Human Development*, 5, 256–261. <https://doi.org/10.15640/jehd.v5n1a26>
- Yenkimaleki, M. (2017). Effect of prosody awareness training on the quality of consecutive interpreting between English and Farsi (LOT dissertation series 459). Utrecht: LOT.
- Yenkimaleki, M., & Van Heuven, V.J. (2013). Prosodic feature awareness training in interpreting: An experimental study. In L. Gómez Chova, A. López Martínez, & I. Candel Torres (Eds.), *Proceedings of the 6th International Conference on Education, Research and Innovation, November 18th-20th, 2013, Seville, Spain*, 4179–4188.
- Yenkimaleki, M., & Van Heuven, V.J. (2016a). Effect of explicit teaching of prosodic features on the development of listening comprehension by Farsi-English interpreter trainees: An experimental study. *International Journal of English Language Teaching*, 4, 32–41.
- Yenkimaleki, M., & Van Heuven, V.J. (2016b). The effect of prosody teaching on developing word recognition skills for interpreter trainees: An experimental study. *Journal of Advances in Linguistics*, 7, 1101–1107. <https://doi.org/10.24297/jal.v7i1.5158>
- Yenkimaleki, M., & Van Heuven, V.J. (2016c). Prosody teaching matters in developing speaking skills for Farsi-English interpreter trainees: An experimental study. *International Journal of English Language and Linguistics Research*, 4(5), 82–91.
- Yenkimaleki, M., & Van Heuven, V.J. (2016d). Effect of prosody awareness training on the performance of consecutive interpretation from Farsi into English: An experimental study. *Asia Pacific Translation and Intercultural Studies*, 3(3), 235–251. <https://doi.org/10.1080/23306343.2016.1233930>
- Yenkimaleki, M., & Van Heuven, V.J. (2016e). Explicit teaching of segmentals versus suprasegmentals: which would yield better listening comprehension skills for interpreter trainees? An experimental study. *British Journal of English Linguistics*, 4(6), 11–22.
- Yenkimaleki, M. & Van Heuven, V.J. (2018). The effect of teaching prosody awareness on interpreting performance: an experimental study of consecutive interpreting from English into Farsi. *Perspectives, Studies in Translation Theory and Practice*, 26(1), 84–99. <https://doi.org/10.1080/0907676X.2017.1315824>

Appendix 1. Sample of training program for prosody group

Time	Opening	Monitor/Feedback
5 mins.	Greeting students. Asking questions about previous discussions to get feedback on how things are going with the program.	
	Activities	
30 mins.	Students listened to authentic audio tracks in English and talked about the issues brought up in the audio tracks. They also watched authentic movies and discussed their contents. They talked about some proposed hot topic, in pairs in the classroom.	I. moved around class and helped students when needed.
	Prosody training	
20 mins.	Prosodic theory: I. explained to the students that change of stress in English may result in different interpretations. Prosodic practice: Marking syllables: I. played a list of words/ sentences and had learners count syllables and mark stressed syllables. Examples: Words: <i>deport, demarcation, campsite, cardiologist, carnival, catastrophe, cavalry, champion, charger, cheery, chowder.</i> Sentences: <i>The increased pressure within the muscle compresses nerves and blood vessels. The players had swelling in their triceps. I was just kind of shocked this was happening to us. The students said they did not take any body building supplements. We believe it was a strenuous workout, but we don't believe it was excessive. That's used so commonly by athletes of all ages.</i>	I. asked students to mark the syllables on work sheet and hand in to me to assess their work. I. asked some students to come in front of class to read the words/sentences aloud again to see how much they were able in practice to produce the correct stress patterns of words and sentences.
	Homework	
5 mins.	Students were asked to record Voice of America for five minutes, then listen to it and talk about its content for their classmates in the next session.	I. instructed students to do this task outside the classroom and present it in class the next session.

Note. I. = Instructor

Appendix 2. Sample training program for control group

Time	Opening	Monitor/Feedback
5 mins.	Same as prosody group	
	Activities	
50 mins.	Same as prosody group (but 30 minutes longer)	Same as prosody group
	Homework	
5 mins.	Same as prosody group	Same as prosody group

Address for correspondence

Mahmood Yenkimaleki
 Boerhaavalaan 84
 2334 ES Leiden
 The Netherlands
 myenkimaleki@gmail.com

Co-author information

Vincent J. van Heuven
 Emeritus professor of Experimental Linguistics and Phonetics
 Leiden University Centre for Linguistics LUCL
 Leiden Institute for Brain and Cognition LIBC
 Professor of Phonetics
 University of Pannonia, Veszprém, Hungary
 v.j.j.p.van.heuven@hum.leidenuniv.nl

Publication history

Date received: 26 June 2017
 Date accepted: 1 December 2018
 Published online: 30 September 2019