Relative contribution of explicit teaching of segmentals vs. prosody to the quality of consecutive interpreting by Farsi-to-English interpreting trainees

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

Routledge
Taylor & Francis Group

Interactive Learning Environments

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/nile20

Relative contribution of explicit teaching of segmentals vs. prosody to the quality of consecutive interpreting by Farsi-to-English interpreting trainees

Mahmood Yenkimaleki & Vincent J. van Heuven

To cite this article: Mahmood Yenkimaleki & Vincent J. van Heuven (2020): Relative contribution of explicit teaching of segmentals vs. prosody to the quality of consecutive interpreting by Farsi-to-English interpreting trainees, Interactive Learning Environments, DOI: 10.1080/10494820.2020.1789673

To link to this article: https://doi.org/10.1080/10494820.2020.1789673

INTERACTIVE LEARNING ENVIRONMENTS
https://doi.org/10.1080/10494820.2020.1789673

Routledge
Taylor & Francis Group

Relative contribution of explicit teaching of segmentals vs. prosody to the quality of consecutive interpreting by Farsi-to-English interpreting trainees

Mahmood Yenkimaleki [a] and Vincent J. van Heuven [b,c]

[a]Department of English Language, University of Nahavand, Nahavand, Iran; [b]Hungarian and Applied Linguistics, University of Pannonia, Veszprém, Hungary; [c]Leiden University Centre for Linguistics, Leiden, The Netherlands

ABSTRACT

This study investigates the effect of explicit teaching of segmentals vs. prosody on the quality of interpreting by Farsi-to-English interpreter trainees. Participants were native speakers of Farsi and BA students of English translation and interpreting in Iran, who were assigned to one of three groups. No differences in English language skills were found between the groups prior to the experiment. The control group listened to and discussed English audio and video recordings and did exercises in consecutive interpreting. One experimental group instead spent part of the time on explanation of, and exercises with, English prosody. The second experimental group spent part of the time on explanation of, and exercises with, English segmentals. Total instruction time was 12 hours for each group. The students' performance in consecutive interpreting was then rated independently by three experts. Results showed that both experimental groups performed better than the control group. Moreover, teaching prosody had a larger positive effect on the overall quality of interpreting from Farsi into English than segmental instruction. We argue that the interpreting curriculum can be strengthened by devoting a small portion of the time to explicit instruction on segmental and (especially) prosodic differences between source and target language.

ARTICLE HISTORY
Received 24 May 2020
Accepted 25 June 2020

KEYWORDS
Segmentals; prosody; interpreter trainees; curriculum design; interpretation performance

1. Introduction

It has been observed in the context of foreign language teaching that most of the attention is given to the proper pronunciation of the vowels and consonants of the target language, whereas the prosody was traditionally dealt with briefly, and typically in the last stages of the course (e.g. Levis, 2018; Yenkimaleki & van Heuven, 2018). The last decade has witnessed a counter movement, however, which claims that prosody teaching should take precedence over the teaching of segmental properties of the target language (e.g. Levis, 2018; Yenkimaleki, 2019; Yenkimaleki & van Heuven, 2019a, 2019b, 2019c). The present study was set up to illuminate this issue of the relative importance of teaching of segmentals versus prosody in a specific teaching situation, where Iranian university students with Persian as the native language were trained to interpret messages into English in a conference-interpreting setting (also known as consecutive, or non-simultaneous, interpreting, see Section 3). In this particular study we used one specific teaching strategy, which is known as explicit instruction, which, we argue, is the optimal strategy for teaching foreign languages to adult learners.

CONTACT Mahmood Yenkimaleki myenkimaleki@gmail.com

In explicit instruction, the instructor clearly outlines what the learning goals are for the students, and offers clear, unambiguous explanations of the skills and information structures being presented. Explicit learning is an intentional process which requires learners to determine what will be learned such that the learners can express the acquired knowledge structure (Richards & Schmidt, 2002). Implicit teaching, in contradistinction, involves teaching a certain topic in a suggestive or implied manner; the objective is not plainly expressed. Its purposes are to introduce new concepts in a student-centered manner, to give students instruction with a variety of several examples and to allow students to create their own schemas for understanding rules instead of memorizing specific rules. Implicit teaching methods help students induce rules from examples given to them (Ellis, 1994).

In recent years, there has been considerable debate about the relative importance of segmental information versus prosody in the teaching of English as a foreign language (Yenkimaleki & van Heuven, 2018). In English writing, the larger units of discourse are explicitly coded prosodically. Paragraphs are separated by a new indented line. Sentence boundaries are coded by full stops and semicolons, and word boundaries are reflected by spaces. A question mark signals that the sentence is meant to elicit a response from the reader/listener but it does not automatically signal high pitch. Many *wh*-questions in English (and other languages) end with a low terminal boundary tone, but they are written with a question mark nevertheless. In same vein, an exclamation mark signals an exclamative but it is not clear if there are any linguistically relevant properties in the phonetics of an exclamative that are reliably coded by the exclamation mark (see van Heuven, 2017). The orthography of English does not mark stressed syllables in words or stressed words in sentences. Moreover, syllable boundaries are not explicitly marked in the spelling, and most of the spoken within-clause phrase boundaries remain implicit in the spelling, yielding ambiguous sentences such as *John hit the man with a stick* – which either means that John used a stick to hit the man, or that John hit the man who was carrying a stick (an audible phrase boundary before *with a stick* disambiguates the sentence and leaves only the second reading as a possible interpretation but this would not be visible in the orthography, Shattuck-Hufnagel & Turk, 1996, and references therein). It would appear to be the case, then, that to understand a sentence it is enough to recognize the words by their constituent vowels and consonants and that prosody is largely redundant. This is, in fact, easily demonstrated experimentally. If we synthesize sentences without any prosodic information, for instance by stringing together a sequence of syllables that were all spoken with the same amount of stress, with only their neutral durations and with no pitch inflections, the speech remains intelligible. However, if we leave all the prosodic information in place but obliterate the segmental information, for instance by low-pass filtering at 300 Hz, the melody, rhythm, phrasing and stress patterns remain audible but no words will be correctly recognized and no speech understanding will take place (van Heuven, 2017). The role of prosody would be confined to guiding the preliminary chunking of continuous speech into the larger units, such as paragraphs (for paragraph intonation see Lehiste, 1975; Sluijter & Terken, 1993), sentences and major phrases, and highlighting the communicatively most important elements within sentences through sentence stress. It has been shown that the word and sentence prosody of unintelligible speech constrains the native listeners' reconstruction of what might have been said to phrase structures and words that fit the prosody (e.g. Blesser, 1969; Lindblom & Svensson, 1973; Svensson, 1974).

The present study begins by examining the effect of teaching segmentals vs. prosody on the quality of interpreting. Next, we review different studies about teaching segmentals and prosody and its merits and demerits for pedagogical purposes. We address the issue of relative contribution of explicit teaching of segmentals vs. prosody to the quality of consecutive interpreting by Farsi-to-English interpreting trainees. In doing so, we offer suggestions that policy makers and practitioners should observe and understand in the choice of effective methods in teaching segmentals and prosody for EFL learners and interpreter trainees.

2. Review of related literature

2.1. Segmental structure and prosody

Most linguistic theories assume that speech should be analyzed as a hierarchical structure of small units that are combined to higher-order units, which are gathered into yet higher units (e.g. Cutler & McQueen, 2014; van Heuven, 2017; Nespor & Vogel, 1986; Shattuck-Hufnagel & Turk, 1996). The smallest units are the speech sounds, i.e. the vowels and the consonants, which are combined into syllables, which are then united into words, phrases, clauses, sentences, paragraphs and texts. However, there are properties of spoken language which cannot be understood from the mere sequence of the vowels and consonants which make up a spoken sentence (van Heuven, 2008). The same string of segments, e.g. /trʌsti/, is an adjective meaning "trustworthy" when stressed on the first syllable but it refers to the member of the board of a foundation, i.e. a noun, with stress on the final syllable. The difference in stress position is a prosodic difference by definition: it is a property of the larger sound shape, i.e. the word, which does not follow from the sounds that are contained by it. At a higher level in the linguistic hierarchy the same sequence of words, e.g. *You think I am joking*, is perceived as a statement when it ends with low terminal pitch, but as a question, prompting confirmation on the part of the listener, when it ends with high pitch.

Traditionally, the pronunciation of vowels and consonants (also called "segmentals"), was considered having a primary role to play in speech communication, and they were the instructors' primary concern in foreign language teaching since segmental errors may have more detrimental effects than prosodic (also referred to as suprasegmental) inaccuracies for the intelligibility and comprehensibility of a foreign language (e.g. Derwing et al., 1998; Gilbert, 1984; Herry & Hirst, 2002; Hahn, 2004; Foote et al., 2011; Koike, 2014; Robinson et al., 2012; Yenkimaleki, 2016, 2017; Yenkimaleki & van Heuven, 2013, 2018).

It would seem to us that the relative importance of prosody versus segmentals will depend, first of all, on the structural differences that can be established between the learner's native language (L1) and the target language (L2) (Yenkimaleki & van Heuven, 2018, 2019a, 2019c). If the segmental structure differs a lot between the two languages but the prosodic systems are largely the same, as between English and Dutch, or English and German, the teaching of segmentals should take precedence (van Heuven, 2017). However, when both the segmental structure and the prosody differ between L1 and L2, the choice is less than straightforward. This is where a second aspect should be considered, which the directionality of the communicative process that will be targeted. We would argue that listening in the L2 does not make special demands on the part of the learner. Across languages the major division of speech into paragraphs, sentences and clauses is universally coded by the insertion of speech pauses, which will be longer as the prosodic boundary is deeper, i.e. separates units of higher hierarchical status (e.g. Swerts, 1997; Wightman et al., 1992). These breaks are found at moments when the speaker needs to inhale, and the acoustic effect of inhalation is a physical pause (van Heuven, 2017). Moreover, the deeper the break, the deeper the inhalation, and therefore the stronger the subglottal air pressure immediately after the inhalation. This in turn causes the vocal cords to vibrate faster (higher pitch) after deep breaks than after shallow breaks. The change in pitch before and after a break would then be a universal feature of language which need not be learnt anew for an L2 (van Heuven, 2008).

The positive effects of explicit instruction of phonological rules have been emphasized by different researchers (e.g. Ahrens, 2004; Champagne-Muzar et al., 1993; Derwing et al., 2012; Derwing & Munro, 2005; Foote et al., 2011; Koike, 2014; Leather, 1990; Pennington, 1998; Robinson et al., 2012; Suwartono, 2014; Venkatagiri & Levis, 2007; Yenkimaleki & van Heuven, 2016a, 2016b, 2016c, 2016d, 2018). Derwing et al. (1998) found that speakers who had received instruction emphasizing suprasegmental features, could transfer their learning to spontaneous speech production more effectively than those who received instruction with only segmental content i.e. vowels and consonants. Yenkimaleki and van Heuven (2016a) showed that the explicit teaching of prosodic (i.e. suprasegmental) features led

to a significant improvement of listening comprehension skills for interpreter trainees. Anderson-Hsieh et al. (1992) studied the effect of prosody and segmentals to the intelligibility of second language stress, rhythm, and intonation all affected native speakers' judgments more than that of segmental errors.

Capliez (2016) holds that in the comparison between the place of prosody and the place of segmentals in the learning of a foreign language, it has frequently been postulated that accuracy in second language prosody has a more positive impact on intelligibility, comprehensibility, and foreign-accentedness than an accurate realization of individual sounds. Therefore, segmental errors may not be as much of a hindrance to communication as prosodic errors. Firth (1992) made reference to the *zoom principle* that *global* aspects (e.g. prosody) should be taught before *local* aspects (e.g. segmentals).

Derwing et al. (2012) investigated three typical difficulties that learners encounter with English as a foreign language – including the perception of sentence stress, intonation, and *-teen/-ty* number distinctions (e.g. *eighteen – eighty*) – and they found a certain amount of effectiveness of the explicit instruction of prosodic features.

However, there are some studies which show that teaching segmentals could be more effective for interpreters in perception of message than teaching prosody (e.g. Yenkimaleki & van Heuven, 2016e). Some researchers hold that the first step to the teaching of oral English is through segmentals – as is the case in many academic curricula. Pennock and Vickers (2001) state that considering the quantity of rules and exceptions (e.g. in English lexical stress assignment) the teachability of prosody would not be effective and they recommend that instructors focus on segmentals. However, Cruttenden (2008) asserts that teaching English prosody (e.g. stress) is not so difficult to teach and learn.

Cutler and Clifton (1984) measured the reaction time of native English listeners instructed to recognize disyllabic words in which the lexical stress had been deliberately misplaced (e.g. *can'teen* was pronounced *'canteen*). They found no effect on word recognition if the stress had been shifted from right to left, but word recognition suffered when the stress had been shifted from left to right, and was even more compromised in the case of a change of vowel quality (e.g. *wallet* ['wɔlɪt] pronounced as [wɔ'let]). They concluded that stress errors significantly impact intelligibility, especially when stress affects segmental features.

Considering the different results in recent studies on the effectiveness of teaching prosody and of segmentals by some scholars, this domain needs to be investigated systematically in wider contexts with different participants. Given that interpreter training programs and foreign-language curricula have to make strict choices as to how much time should be spent on teaching particular skills, it is important to know whether teaching time is better spent on segmentals or on prosody. The results of such studies can be incorporated in interpreter training programs for training the next generation of interpreters.

2.2. English and Farsi sound systems compared

The vowel inventory of English is considerably richer than that of Persian. American English is generally analyzed as a system of eleven monophthongs (pure vowels), three diphthongs, and a non-stressable reduction vowel schwa /ə/ (Yavaş, 2011). The monophthongs are subdivided into a group of seven tense/long vowels against a smaller group of four lax vowels; the diphtongs behave as tense and long vowels. Farsi has only six vowels with three short and three long vowels that are not implicated in a tenseness contrast. The Contrastive Analysis Hypothesis (e.g. Lado, 1957; Odlin, 1989) predicts problems in perception and production of the English vowels, which is underdifferentiated in the native system of the Persian EFL learners. The consonant inventories are more similar in the two languages, be it that Persian lacks the sounds /ð/ as in *the*, θ as in *think* and /w/ as in *we*. These sounds are predictable learning problems for Persian EFL learners (Yavaş, 2011, pp. 203–204).

Languages can be classified as stress-timed or syllable-timed (Aquil, 2012). In stressed-timed languages like English, words can be reduced. Languages take different measures so that stress would occur at equal time intervals (Aquil, 2012). As a result the duration of syllables varies depending on the number of unstressed syllables that intervene between two stressed syllables. Also, the stressed syllable is more complex (contains more sounds) and is pronounced more carefully (yielding longer duration), all else being equal, than unstressed syllables (van Heuven, 2018; van Heuven & Turk, 2020). Farsi, in contrast to English, is held to be a syllable-timed language (Hall, 2007; Windfuhr, 1979). In a syllable-timed language every syllable, whether stressed or not, takes up roughly the same amount of time so that the number of syllables in a sentence yields a perfect prediction of the total sentence duration. Syllable-timed languages, such as Farsi, tend to have simple syllable structures, no difference between short and long vowels, no diphthongs, and no vowel reduction in unstressed positions (Dauer, 1983; Nespor et al., 2011), while stressed-timed languages, such as English, allow complex syllable structures with up to three consonants (C) preceding the vowel (V) and up to four following V, and typicaly contrast short and long vowels (V vs. VV) and have diphthongs. The syllable structure of Farsi can be presented as CV(C)(C), i.e. syllables must begin with exactly one C and contain only one (short) vowel, followed by maximally two Cs. This schema allows three legal syllable types in Farsi, i.e. CV, CVC and CVCC, whereas at least 18 different types of syllable are permissible in English. Shademan (2002) observes that an initial consonant cluster in an English word is broken up by vowel epenthesis by Farsi learners of English. The Contrastive Analysis Hypothesis predicts that this area of differences can cause problems in perception, namely, parsing and segmenting an English auditory input, and in the production of speech for Farsi second language learners of English (see Samareh (1986) and Yenkimaleki (2017) for a more detailed comparison of English and Farsi sound system).

Persian word stress is fixed prefinal, i.e. the second but last syllable of Persian words is stressed with no exceptions. The stress system of English is very different, and allows stresses to occur in different locations in words, depending on the segmental composition of the syllable (''quantity sensitive'', see Kager, 1989), and the morphological composition of the word. On top of that there many exceptions and subregularities, which renders English word stress a lexical property, i.e. a property that has to be learned by heart (Cutler, 1984). As a result of the mismatch in word stress locations Persian EFL learners often mis-stress English words, which compromises their intelligibility.

Persian and English are both Indogermanic languages and share a number of phonological features. Both languages mark important words in the sentence by sentence stresses, i.e. by executing a perceptually prominent change in vocal pitch on the stressed syllable of that word (e.g. van Heuven, 2018 and references therein). Nevertheless, there are small but important discrepancies between the sentence stressing between English and Persian (e.g. Abolhasanizadeh et al., 2012) which may lead to communicative problems for Persian EFL speakers. It has also been observed that the melodies of English and Persian, especially in the case of *wh*-questions (Shiamizadeh, 2018), differ in ways that may compromise the comprehensibility of Persian-accented English.

3. Motivation for the present study

The task of a professional interpreter is to understand a spoken message in one language and produce a semantically equivalent spoken message in another language, either immediately (simultaneously with the original speaker) or after the original speaker has finished his turn (consecutive interpreting). Typically, one language is the interpreter's native language (the A language), the other language was learned at a later stage in life as a foreign language (the B language). The interpreter's perceptual and productive command of the foreign language has to be excellent. In the field of interpreting studies much research has been done on improving the oral skills in the B language of the trainee's choice. There is a widespread belief that the explicit teaching of contrastive phonetics of the A and B language should be beneficial to interpreting trainees (Ahrens, 2004;

Yenkimaleki & van Heuven, 2018, 2019a, 2019b). However, the amount of teaching time available in the interpreter training curriculum is limited so that choices may have to be made as to what priority should be given to which aspects of the complex process of interpreting. One such choice is whether raising the student's awareness and explicit knowledge of differences between his native (A) language and the foreign (B) language yields enough benefits for the trainee's interpreting skills to include this in the interpreting curriculum, rather than, say, spending a similar amount of time and effort on expanding the trainee's vocabulary or on cultural and socio-economic background knowledge of the country where the B language is used. And if it should be decided that explicit phonetic training deserves a place in the interpreting curriculum, then the question is how much attention should be given to the teaching of segmentals versus that of prosody. The literature reviewed above suggests that attention to (contrastive) phonetics improves the trainee's interpreting performance, both in the *recto* and in the *verso* direction. However, the relative importance of teaching segmentals versus prosody is undecided.

The present study, therefore, addresses the importance of explicit teaching of English segmentals vs. prosody for Iranian interpreter trainees. In the experiment, Farsi (Modern Persian) is the native language while English is the non-native language. This is the inverse direction of interpreting (also called *verso*). We concretely asked the following research question:

> Which of two areas of explicit teaching yields better consecutive interpreting performance for Farsi-English interpreter trainees given the same of amount of training time: segmental training or suprasegmental training?

Our basic hypothesis is that this depends on the similarities and differences between the phonological systems (and their phonetic implementation) of L1 and L2. In the case of Persian as the L1 and English as the target language L2, there are many differences, both in the segmental and in the prosodic domain (see Section 2.2). In such a situation the importance of prosody increases, since native English listeners re-interpret sub-optimal segmental information so as to fit the perceived prosodic structure (Cutler, 1984; Cutler & McQueen, 2014). We predict, accordingly, that in the present case, explicit instruction in prosody should be more beneficial than segmental instruction.

4. Methodology

4.1. Participants

Forty-five student interpreter trainees at the BA level who were majoring in interpreting and translation studies at the University of interpreting studies, and applied Sciences (number four) in Tehran, Iran, were chosen randomly to participate in this study. They were randomly divided into three classes of 15 students with the restriction that each group should comprise 8 male and 7 female students. The participants were native speakers of Farsi within an age range of 21–26 years. They participated in all sessions of the training program.

4.2. Ethical issues

We received approval from the ethics committee of the Dept. of Modern Languages for the present study. All the participants agreed to take part in the research project on the basis of informed consent and received a small amount of money for their services.

4.3. 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.

Participants were then randomly assigned to one control group and two experimental groups. The control group received interpretation exercises, asking them to listen to authentic audio tracks in Farsi and interpret them into English. They also watched authentic Iranian movies and interpreted parts of the movie. The first 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 (stress at word and sentence level) for 20 min during each training session. The second experimental group received explicit teaching of segmentals (vowels, consonants, syllable structure) for 20 min during each training session. The participants took part in the program for 12 sessions (sixty minutes per session) during four weeks, i.e. 12 h in all.

Altogether, the control group listened to 540 min of authentic audio tracks and did the interpretation exercises based on them. The control group and the two experimental groups listened during 180 min to the Iranian instructor who explained how to do exercises in interpreting. The experimental groups listened for 300 min in total to authentic audio tracks and did the interpretation exercises based on them. The first experimental group listened for 120 min to the theoretical explanation of English prosody that was provided by the Iranian instructor and spent 120 min in all doing practical exercises in English prosody (see Yenkimaleki, 2017, pp. 55–81 for a detailed description of the training program). The second experimental group listened for 120 min to the theoretical explanation of English segmentals (vowels, consonants and syllable structure) that was provided by first author and 120 min in all doing practical exercises in English segmentals. For segmental practice, minimal pair drills (e.g. beat vs. bit; leave vs. live) were presented to the trainees to learn to perceive the differences in vowels. Moreover, consonants and syllable structures of English were introduced and practiced according the explanations and exercises of Celce-Murcia et al. (2010). The activities covered by the three participant groups and the time (minutes) spent on them are summarized in Table 1.

In all the sessions, at different times, interpreting tests were administered to the participants in order to measure their progress and to diagnose problems on the part of the participants. At the end of the training program the same post-test was administered to control and experimental groups alike to measure the quality of (consecutive) interpreting at the end of the treatment. The test included five 30-second extracts to be interpreted.

The post-test took place in a language laboratory in the presence of a classroom instructor. Students were seated in sound-proofed half-open cubicles. Source texts were presented over loudspeakers at a comfortable listening level; note taking was allowed. After every fragment, participants were given one minute to consult their notes and to record an interpretation in English. Recordings were made directly onto a computer through individual, table-mounted microphones.

The same three out of the five recorded texts per participant were evaluated independently by three experts, who were senior colleagues at the Department of English of University of interpreting studies, and applied Sciences in Tehran, Iran (age range of judges 36–46 years). Two of them had studied abroad (received their PhD out of Iran, one in UK and one in the Netherlands). Selection of the judges was based on their first language, professional and academic background, as well as their willingness and availability to participate. The judges were experienced instructors

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

	Group		
Activity	Control	Prosody	Segments
Audio tracks/ movies & interpretation	540	300	300
Listening to instructor explanations	180	180	180
Prosodic theory		120	
Prosodic practice		120	
Segmental theory			120
Segmental practice			120
Total time spent	720	720	720

in interpreting between Farsi and English, and did not know the students they judged. They did their ratings at different places and at different times. Evaluation criteria (based on Sawyer, 2004, see Table 2) were explained beforehand.[1]

5. Results

There were no significant differences between the control group (548) and the two experimental groups in terms of their TOEFL scores, with mean scores of 546 (segmental group) and 545 (prosody group), $F(2, 28) < 1$, $p\eta^2 = .033$. At the start of the intervention, therefore, we assume that the three groups were equal in their overall command of English. At the end of the intervention, however, the scores on the posttest differed substantially, depending on the group, i.e. on the type of instruction and training the participants had received.

We then examined the reliability of the three judges, R1, R2 and R3. Cronbach's alpha was between .864 (for the Grammar scale) and .962 (Accuracy). The overall judgements were most reliable with alpha = .983. We conclude that the raters' reliability (agreement) was very good to excellent for every single rating scale, and excellent for the overall assessment of the participants' performance. Independently of this, R1's ratings were somewhat lower on all component scales as well as on the overall performance index, than those of the other two raters. The effect of rater was significant for all scales and the overall index by separate Repeated Measures Analyses of Variance (RM-ANOVA) with rater as a within-subjects factor with $F(2, 88) > 4.5$ ($p < .016$). Rater R1 differed from the other two on all component scales as well as on the overall rating (Bonferroni test). On the basis of this analysis it was decided to run all further data analyses after averaging the scores over the three raters.

The mean scores, averaged over the three judges, for each of the eight rating scales, as well as the overall performance index (the sum score as explained in Table 2) are presented in Table 3, broken down by the three groups of participants. For a survey of individual scores see Appendix 1.

The differences between the three groups of participants were tested by a series of oneway repeated measures analyses of variance. Degrees of freedom were Greenhouse-Geiser corrected in case the assumption of sphericity was violated (but the nominal, uncorrected degrees of freedom are mentioned in the table banner).[2] *P*-values, however, are based on the corrected degrees of freedom. The effect size (partial eta squared) is indicated in the table. Post hoc analyses, with Bonferroni correction for multiple comparisons ($p \leq .05$) were performed to see which groups did or did not differ from each other. These are specified in the rightmost column of Table 3.

In terms of accuracy of interpreting, the prosody group did significantly better than the segmental group, which in turn outperformed the control group. On the component scales Omissions and Additions, both experimental groups did better than the control group but the advantage of the prosody group over the segmental group is not significant. The segmental group scored significantly better than the other two groups when it comes to the grammatical quality of the interpreting performance. Exactly the reverse order is seen for the appropriateness of expression (formality). A very small effect, bordering on insignificant, favoring the control group over the experimental groups, is observed for the choice of terminology. By far the largest effects, however, are obtained for the two rating scales that pertain to prosody. The prosody group did significantly better than the other two groups (which do not differ significantly from each other) on both pace (fluency of delivery) and – even more so – on accentuation.

Table 2. Eight evaluation criteria and weights subdivided into three domains used in the quality judgment of interpreting performance. Weights add up to 100. After Sawyer (2004).

Meaning		Language use		Presentation	
Accuracy	20	Grammar	10	Pace	10
Omissions	15	Expression	10	Accentuation	10
Additions	15	Terminology	10		

Table 3. Results of posttest. Mean scores for eight rating scales and overall interpreting performance broken down by three groups of participants (control, experimental segments, experimental prosody).

Rating scale	Contr	Segm	Pros	$F(2, 28)$	p	$p\eta^2$	Bonferroni
Accuracy	13.1	14.0	15.0	14.7	< .001	.513	C < S < P
Omissions	11.4	12.0	12.1	6.4	.007	.315	C < { S, P}
Additions	11.4	12.0	12.1	6.4	.005	.315	C < { S, P}
Grammar	7.8	8.1	7.7	4.4	.025	.237	{C, P} < S
Expression	8.1	7.4	7.7	8.1	.002	.367	S < {P, C}
Terminology	7.8	7.3	7.5	3.3	.051	.191	{S, P} < C
Pace	6.9	7.0	8.6	50.9	< .001	.784	{C, S} < P
Accentuation	6.8	6.7	8.6	81.1	< .001	.853	{S, C} < P
Total	73.3	74.5	79.3	26.4	< .001	.654	{C, S} < P

Notes: Conditions (C for Control, S for Segments, P for Prosody) included in curly brackets are not significantly different from each other; (groups of) conditions separated by "<" differ significantly (with the lower score to the left of the "<" symbol and the better condition to the right of it).

The weighted sum of the component scores reflects the above configuration of effects, such that the prosody group outperforms the segmental group as well as the control group, which do not differ significantly from each other.

In the second part of this section we will examine the contribution of the individual participants' command of English, as established by the TOEFL score at the beginning of the intervention, to the quality of their interpreting performance in the posttest. The effects of TOEFL score and type of training are easily observed in Figure 1, which plots the overall interpreting quality index (on a scale between 0 and 100, see Table 2) based on the eight rating scales as a function of the individual participant's TOEFL score, in three separate panels, for the control group (panel A) and for the two experimental groups, i.e. the segmental group (panel B) and the prosody group (panel C).

We notice, first of all, that the distribution of the pretest TOEFL scores is roughly the same for the three groups of participants. Second, we may observe that the overall interpreting quality index (plotted vertically) asymptotes at ca. 85 points. For all students, irrespective of group membership, the asymptote begins at TOEFL-scores at or above 575. The relationship between the pretest and the posttest scores is captured better by a logarithmic than a linear curve-fit.[3] The results show that there is a very strong correlation between the TOEFL-scores and the participant's later performance on the interpreting task. Depending on the group, the R^2-values are between .75 and .92.

The contribution of the treatment and that of the TOEFL score can be decomposed by first computing the overall effect of the TOEFL score, again assuming a logarithmic relationship between the TOEFL score and the posttest interpreting performance. The R^2 of this correlation is .753. The net effect of the treatment is computed by first adjusting the scores obtained by the individual

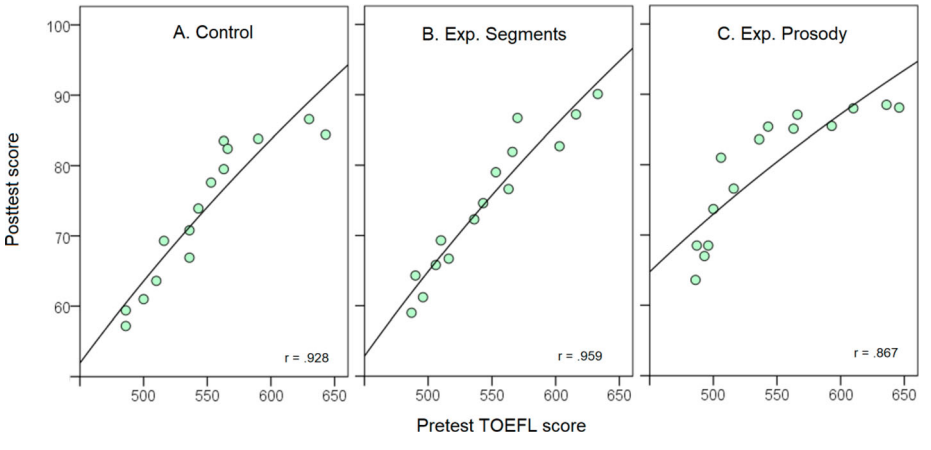

Figure 1. Posttest score as a function of pretest TOEFL score for three groups of participants. Correlation coefficients r are based on logarithmic curve-fits.

participants so as to eliminate the main effect of the treatment, i.e. by subtracting the group mean from the individual participant's posttest score (after log-transformation). R^2 then rises to .805, so that the net effect of the treatment amounts to 5 percent of the variance. The conclusion follows that the effect of the TOEFL score is 25 times larger than that of the treatment.

Visual inspection of Figure 1 reveals that the beneficial effect of the prosody training treatment tends to be concentrated on the poorer students, i.e. those with low pre-treatment TOEFL scores. In order to quantify and statistically test this impression, we compute a relative gain measure by comparing the change in relative position in the distribution of scores before and after the treatment. The gain measure is defined as the z-score of the participant's overall interpreting performance minus the person's z-transformed TOEFL score. Figure 2 plots the gain as a function of the participant's TOEFL score, in separate panels for the three groups.

The correlation between the gain and the student's TOEFL score is close to zero for the segmental group ($r = .055$, ins.) and for the controls ($r = -.078$, ins.). The correlation is better and highly significant for the prosody group ($r = -.605$, $p = .009$, one-tailed). This result suggests that the prosody training is especially beneficial for students with relatively poor overall proficiency in the foreign language.

6. Discussion

The present study investigated the effect of the explicit teaching of segmentals vs. prosody on the quality of interpreting by Farsi-to-English interpreter trainees. The results showed that the teaching of prosody had a significant positive effect on the overall quality of *verso* interpreting compared with the traditional interpreting curriculum, which spends all available time on practical interpreting exercises. The results also revealed that explicit instruction in the use of prosody improves interpreting quality more than the explicit teaching of segmentals, and that the gain yielded by prosody instruction was especially beneficial as the trainee was less proficient in English at the start of the training program. These findings are in line with Yenkimaleki and van Heuven (2016a, 2016b, 2016c, 2016d), who found that the explicit teaching of prosodic features had a significant positive effect on the interpreter trainees' perception skills in the L2 (in *recto* interpreting) and/or the quality of the interpreter's delivery in the L2 (in *verso* interpreting). Our findings replicate the greater effect of the explicit prosody training for students whose command of English is less well developed at the start of the program, that was found earlier by Ueno (1998).

In our research we studied the effects of the teaching of prosody and segmentals on the use of English by consecutive interpreters between English and Farsi. The participants in the research,

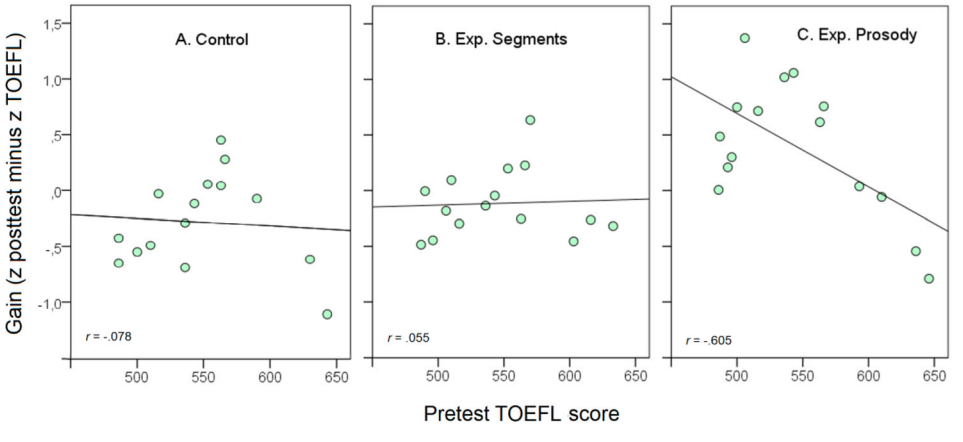

Figure 2. Gain (Z-postestscore minus Z-TOEFL score) as a function of TOEFL score, for three groups of participants. Correlation coefficients r are based on linear fits.

students of interpreting and translation, were native speakers of Farsi (New Persian) who have learnt English as a foreign language. We studied the performance of these students both when they translated from the foreign language (English) into their native language – a process often referred to as straight or *recto* interpreting – and when they interpreted spoken input in their native language into English, which is referred to as inverse or *verso* interpreting. In an earlier study we found that the explicit teaching of prosody (a cognitive method which explained the structural differences between Farsi and English in word and sentence stress and their functions, followed by short periods of practice, see Yenkimaleki, 2017, pp. 52–81) had a small but significant positive effect on the students' interpreting performance relative to the performance of a control group which received the routine interpreting curriculum with practice only. Crucially, however, a second experimental group who spent a similar amount of time on explicit instruction in the segmental differences between English and Farsi obtained a significantly larger gain than the prosody group. The awareness of the segmental differences between English and Farsi is conducive to better processing of the input in the foreign language by the non-native listener. The listener knows what specific sounds to expect and what contrasts are relevant for word recognition in English. In such situations the prosody is largely redundant and will only occasionally yield a measurable benefit (most likely in the area of word stress, which is indeed rather different between the two languages, since Farsi has fixed penultimate stress and English has a highly variable word stress pattern, Yenkimaleki, 2016).

In the present study we examined the relative merits of the teaching of segmentals versus that of prosody in the process of *verso* interpreting. Here, explicit knowledge and awareness of the peculiarities of English segmentals and prosody helped the Farsi learner of English to correctly realize the word and sentence prosody of his/her English speech output. The crucial question was whether getting the sounds right (or at least approximately right) would be a more desirable goal than getting the prosody (approximately) right.

The literature on this issue contains numerous claims to the effect that the teaching of prosody should be given priority over instruction targeting segmental correctness. Nevertheless, the evidence that is quoted in favor of this claim is circumstantial at best and no direct comparisons can be found in the literature. It is our understanding that native listeners who are confronted with speech with a foreign accent (or with speech in a dialect of their own language or in a closely related different language), rely on strategies they have developed over the lifetime to cope with deviant speech in general (van Heuven, 2008). Speech perception in quiet is the exception rather than the rule. In everyday communication, we always perceive speech against a noisy background (ambient noise, reverberation due to room acoustics, multiple speakers talking in the background, degraded speech quality due to bandwidth limitations in telephone communication). However, native speech perception is robust (Cutler, 2012).

An important insight here is that prosody is more robust against noise than segmental properties. Prosody is resistant to noise, filtering, distortion and even to competing voices (van Heuven, 2017). Giving priority to prosody in speech perception is an excellent strategy as long as the input speech is native. In that case, hearing a specific sequence of stressed and unstressed syllables will prompt the listener to initiate a search in his mental lexicon for words that fit the rhythmic pattern of the input. This will enable the native listener to limit the lexical search space so that only a small number of recognition candidates remain even if the segmental information is poorly defined. But this strategy backfires if the input speech is prosodically flawed, for instance when the segments are unreliable because they are spoken with a foreign accent. The problem would remain manageable if the prosody were correct. But if the prosody is wrong, the listener will rely on the only strategy he has available, i.e. trust the prosody and reinterpret the incoming segments so as to match the perceived prosody. As a case in point, Bansal (1976) showed that native listeners of British English reinterpret segments to fit the incorrectly perceived rhythmic pattern in the English as pronounced by Indian speakers. To the British-English ear the stress in the words *character* and *written* is on the second syllable when pronounced by an Indian speaker. Given that the segmental pronunciation of Indian English strongly deviates from the British norm, it makes sense for the British listener to reinterpret the information as *director* and *retain*, respectively (errors reported by Cutler, 1984).

7. Conclusion

Overall, the results showed that the teaching of prosody had a significant positive effect on the overall quality of *verso* interpreting compared with the traditional interpreting curriculum, which spends all available time on practical interpreting exercises. The results also revealed that explicit instruction in the use of prosody improves interpreting quality more than the explicit teaching of segmentals. The effects of the prosody training on prosody-related evaluation criteria (accentedness and pace) were strong – and significantly stronger than the effects on the other seven criteria.

In our study the interpreter trainees had no equal command of the languages A and B. They were native speakers of Modern Persian (Farsi) as spoken in Iran, with English as a foreign language. In the present study the interpreter's task was to listen to passages spoken in the native language, 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 (Gile, 2005, 2009). Our results suggest that the non-native English output of the *verso* interpreting will be more intelligible (to native and non-native English listeners alike) when student interpreters have a conscious knowledge of prosodic features.

The pedagogical implications of the present study could be interpreting studies, and applied to interpreting programs. The learners' first language should be taken into account in the curriculum of interpreter training by policy makers. The current practitioners in different academic settings of interpreter training should have in-service training so that they could be updated on the role of prosodic features in message perception and production.

Notes

1. The design of this study is quasi experimental since there is no pretest that has the same structure as the posttest. We assume that three participant groups would have obtained equal scores on a proper pretest, and that, therefore, any differences that may be found in the posttest must be due to the different treatments. We know from earlier research (Yenkmeleki & van Heuven 2017, 2018) that the scores on pretests identical to our posttest correlate with the TOEFL scores of Iranian student-interpreters at $r = .940$ or better, which means that the TOEFL test almost perfectly predicts the interpreting pretest scores. This strengthens our assumption that the three groups would have obtained equal scores on an interpreting pretest.
2. The Greenhouse-Geiser correction is the most conservative method available when the requirement of sphericity is violated.
3. An (even) better fit was obtained by a quadratic (parabolic) curve but there is no theoretical ground for assuming such a relationship, since there is no reason why posttest interpreting performance should go down for students with very high TOEFL-scores.
4. In Appendices 2, 3 and 4 "I." stands for "Instructor".

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Mahmood Yenkimaleki (PhD) is a researcher at VU University Amsterdam, the Netherlands, and assistant professor of applied linguistics at University of Nahavand, Iran. His area of interest is interpreting studies and applied linguistics. Email: myenkimaleki@gmail.com.

Vincent J. van Heuven (BA, MA, and PhD from Utrecht University) is emeritus professor of Experimental Linguistics and Phonetics and former director of the Leiden University Centre for Linguistics. He is now a professor at the University of Pannonia in Veszprém, Hungary. He served on the editorial boards (and as associate editor) of *Journal of Phonetics* (Academic Press/Elsevier) and *Phonetica* (Karger) and was editor of the series *Speech Research* (Mouton de Gruyter). Forty-five doctoral dissertations were written under his (co-)supervision. He is a life member of the Royal Netherlands Academy of Arts and Sciences. Email: V.J.J.P.van.Heuven@hum.leidenuniv.nl.

ORCID

Mahmood Yenkimaleki http://orcid.org/0000-0002-8712-9004
Vincent J. van Heuven http://orcid.org/0000-0003-3631-5699

References

Abolhasanizadeh, V., Bijankhan, M., & Gussenhoven, C. (2012). The Persian pitch accent and its retention after the focus. *Lingua. International Review of General Linguistics. Revue internationale De Linguistique Generale*, *122*(13), 1380–1394. https://doi.org/10.1016/j.lingua.2012.06.002

Ahrens, B. (2004). *Prosody beim Simultandolmetschen* [Prosody in simultaneous interpreting]. Peter Lang.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgements of non-native pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, *42*(4), 529–555. https://doi.org/10.1111/j.1467-1770.1992.tb01043.x

Aquil, R. (2012). Listening to English connected speech: A problem and solutions. *Arab World English Journal*, *3*, 329–364. https://doi.org/10.1037/0033-295X.89.4.369

Bansal, R. K. (1976). *The intelligibility of Indian English: Measurements of the intelligibility of connected speech, and sentence and word materials presented to listeners of different nationalities*. Central Institute of English and Foreign Languages.

Blesser, B. A. (1969). *The perception of spectrally rotated speech* [Doctoral dissertation]. Massachusetts Institute of Technology. www.blesser.net/downloads/Perception%20of%20Spectrally%20Rotated%20Speech.pdf

Capliez, M. (2016). *Acquisition and learning of English phonology by French speakers on the roles of segments and suprasegments* [Doctoral dissertation]. Université Lille III – Charles De Gaulle.

Capliez, M. (2016). *Acquisition and learning of English phonology by French speakers on the roles of segments and suprasegments* [PhD thesis]. Université Lille III – Charles De Gaulle. https://www.researchgate.net/publication/311431450

Celce-Murcia, M., Brinton, D., & Goodwin, J. (2010). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. 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*, 143–160.

Cruttenden, A. (2008). *Gimson's pronunciation of English* (7th ed.). Hodder Education.

Cutler, A. (1984). Stress and accent in language production and understanding. In D. Gibbon & H. Richter (Eds.), *Intonation, accent and rhythm: Studies in discourse phonology* (pp. 77–90). de Gruyter.

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

Cutler, A., & Clifton, C. (1984). The use of prosodic information in word recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X* (pp. 183–196). Lawrence Erlbaum.

Cutler, A., & McQueen, J. M. (2014). How prosody is both mandatory and optional. In J. Caspers, Y. Chen, W. Heeren, J. Pacilly, N. O. Schiller, & E. van Zanten (Eds.), *Above and beyond the segments: Experimental linguistics and phonetics* (pp. 71–82). John Benjamins.

Dauer, R. (1983). Stress timing and syllable timing re-analyzed. *Journal of Phonetics*, *11*(1), 51–62. https://doi.org/10.1016/S0095-4470(19)30776-4

Derwing, T. M., Diepenbroek, L. G., & Foote, J. A. (2012). How well do general-skills ESL textbooks address pronunciation? *TESL Canada Journal*, *30*(1), 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*(3), 379–397. https://doi.org/10.2307/3588486

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, *48*(3), 393–410. https://doi.org/10.1111/0023-8333.00047

Ellis, R. (1994). *The study of second language acquisition*. OUP.

Firth, S. (1992). Pronunciation syllabus design: A question of focus. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 173–183). Oxford University Press.

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), 1–22. https://doi.org/10.18806/tesl.v29i1.1086

Gilbert, J. (1984). *Clear speech. Pronunciation and listening comprehension in American English. Student's Book*. Cambridge University Press.

Gile, D. (2005). Directionality in conference interpreting: A cognitive view. In R. Godijns & M. Hindedael (Eds.), *Directionality in interpreting. The 'retour' or the native?* (pp. 9–26). Communication and Cognition.

Gile, D. (2009). *Basic concepts and models for interpreter and translator training* (revised edition). John Benjamins.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, *38*(2), 201–223. https://doi.org/10.2307/3588378

Hall, M. (2007). *Phonological characteristics of Farsi speakers of English and L1 Australian English speakers' perception of proficiency* [Unpublished M.A. thesis]. Department of Linguistics, Curtin University.

Herry, N., & Hirst, D. (2002). Subjective and objective evaluation of the prosody of English spoken by French speakers: The contribution of computer assisted learning. In B. Bel & I. Marlien (Eds.), *Proceedings of speech prosody 2002* (pp. 383–386). Laboratoire Parole et Langage. https://www.researchgate.net/publication/2925313

Kager, R. (1989). *A metrical theory of stress and destressing in English and Dutch*. Foris.

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). Japan Association for Language Teaching. https://jalt-publications.org/sites/default/files/pdf-article/jalt2013_036.pdf

Lado, R. (1957). *Languages across cultures*. University of Michigan Press.

Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather & A. James (Eds.), *New sounds. Proceedings of the 1990 Amsterdam Symposium on the Acquisition of second language speech* (pp. 72–97). University of Amsterdam.

Lehiste, I. (1975). The role of temporal factors in the establishment of linguistic units and boundaries. In W. U. Dressler & F.-V. Mares (Eds.), *Phonologica* (pp. 115–122). Wilhelm Fink Verlag.

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

Lindblom, B. E. F., & Svensson, S.-G. (1973). Interaction between segmental and nonsegmental factors in speech recognition. *IEEE Transactions on Audio & Electroacoustics*, *AU-21*(6), 536–545. https://doi.org/10.1109/TAU.1973.1162527

Nespor, M., Shukla, M., & Mehler, J. (2011). Stress-timed vs. Syllable-timed languages. In M. van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell Companion to Phonology* (pp. 1147–1159). Blackwell.

Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Foris.

Odlin, T. (1989). *Language transfer*. Cambridge University Press.

Pennington, M. C. (1998). The teachability of phonology in adulthood: A re-examination. *IRAL - International Review of Applied Linguistics in Language Teaching*, *36*(4), 323–341. https://doi.org/10.1515/iral.1998.36.4.323

Pennock, B., & Vickers, P. (2001). Some notes on the teaching of English pronunciation. In H. Ferrer Mora, B. Pennock Speck, P. Bou Franch, C. Gregorii Signes, & M. Martí Viaño (Eds.), *Teaching English in a Spanish setting* (pp. 231–241). Departamiento de Filologia Inglesa y Alemana, Universidad de Valencia.

Richards, J. C., & Schmidt, R. (2002). *Longman dictionary of language teaching & applied linguistics* (3rd ed). Pearson Education, Ltd.

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). Routledge.

Samareh, Y. (1986). *Phonology of Farsi language*. Markaz e Nashre Daneshgahi.

Sawyer, B. (2004). *Fundamental aspects of interpreter education. Curriculum and assessment*. John Benjamins.

Shademan, S. (2002). *Epenthetic vowel harmony in Farsi* [Unpublished M.A. thesis]. University of California.

Shattuck-Hufnagel, S., & Turk, A. (1996). A prosodic tutorial for investigators of auditory speech processing. *Journal of Psychological Research*, *25*, 193–247.

Shiamizadeh, Z. (2018). *Prosody and processing of wh-in-situ questions in standard Persian*. LOT.

Sluijter, A. M. C., & Terken, J. M. B. (1993). Beyond sentence prosody: Paragraph intonation in Dutch. *Phonetica*, *50*, 180–188.

Suwartono, S. (2014). Enhancing the pronunciation of English suprasegmental features through reflective learning method. *TEFLIN Journal*, *25*, 80–93. https://doi.org/10.15639/teflinjournal.v25i1/80-93

Svensson, S.-G. (1974). *Prosody and grammar in speech perception*. Monograph 2, Dept. of Linguistics, University of Stockholm.

Swerts, M. (1997). Prosodic features at discourse boundaries of different strength. *Journal of the Acoustical Society of America*, *101*(1), 514–521. https://doi.org/10.1121/1.418114

Ueno, N. (1998). Teaching English pronunciation to Japanese English majors: A comparison of a suprasegmental-oriented and a segmental-oriented teaching approach. *JACET Bulletin*, *29*, 207–225.

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*(1-2), 39–62. https://doi.org/10.3366/E1753854809000305

van Heuven, V. J. (2017). Prosody and sentence type in Dutch. *Nederlands Taalkunde*, *22*(3-29), 41–43. https://doi.org/10.5117/NEDTAA2017.1.HEUV

van Heuven, V. J. (2018). Notes on the phonetics of word and sentence stress: A cross-linguistic (re-)view. In H. van der Hulst, J. Heinz, & R. Goedemans (Eds.), *The study of word stress and accent: Theories, methods and data* (pp. 13–59). Cambridge University Press. https://doi.org/10.1017/9781316683101.002

van Heuven, V. J., & Turk, A. (2020). Phonetic correlates of word and sentence stress. In C. Gussenhoven & A. Chen (Eds.), *The Oxford handbook of linguistic prosody*. Oxford University Press (in press).

Venkatagiri, H. S., & Levis, J. M. (2007). Phonological awareness and speech comprehensibility: An exploratory study. *Language Awareness*, *16*(4), 263–277. https://doi.org/10.2167/la417.0

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, *91*(3), 1707–1717. https://doi.org/10.1121/1.402450

Windfuhr, G. (1979). *Persian grammar: History and state of its study*. Mouton.

Yavaş, M. (2011). *Applied English Phonology*. Wiley-Blackwell.

Yenkimaleki, M. (2016). Why prosody awareness training is necessary for training future interpreters. *Journal of Education and Human Development*, *5*(1), 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.

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*, *2*, 250–268. https://doi.org/10.1075/forum.18007.yen

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 18-20, 2013, Seville* (pp. 4179–4188). IATED Academy. https://library.iated.org/view/YENKIMALEKI2013PRO

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. http://hdl.handle.net/1887/44206

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*(1), 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*, 82–91. http://hdl.handle.net/1887/44132

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*, 11–22. https://www.academia.edu/29805751/

Yenkimaleki, M., & van Heuven, V. J. (2017). The effect of memory training on consecutive interpreting performance by interpreter trainees: An experimental study. *FORUM: International Journal of Interpretation and Translation*, 157–172. https://doi.org/10.1075/forum.15.1.09yen

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

Yenkimaleki, M., & van Heuven, V. J. (2019a). 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., & van Heuven, V. J. (2019b). 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., & van Heuven, V. J. (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

Appendices

Appendix 1. Individual scores and ratings.

Group	Partic. nr.	Gender	TOEFL	Accu.	Omis.	Add.	Gram.	Expr.	Term.	Pace	Acce.	Total
Control	1	M	643.0	16.2	13.0	13.3	8.5	8.5	8.5	8.4	8.0	84.4
Control	2	F	630.0	16.6	13.0	13.0	9.0	9.0	8.0	9.0	9.0	86.6
Control	3	F	590.0	16.0	12.4	14.0	8.2	8.6	8.5	8.1	8.0	83.8
Control	4	M	566.6	15.0	13.0	13.0	8.5	8.5	8.4	8.0	8.0	82.4
Control	5	M	563.3	16.0	13.0	13.0	8.7	8.0	8.5	8.3	8.0	83.5
Control	6	F	563.3	14.0	13.0	12.4	8.0	8.0	8.4	7.8	7.9	79.5
Control	7	M	553.3	13.0	13.0	12.0	8.0	9.0	8.6	7.0	7.0	77.6
Control	8	F	543.3	13.0	11.0	11.0	8.0	9.0	8.0	6.9	7.0	73.9
Control	9	M	536.6	13.0	11.0	10.0	7.6	8.0	8.0	6.5	6.7	70.8
Control	10	F	536.0	11.0	10.0	11.0	7.2	8.0	8.0	6.0	5.7	66.9
Control	11	M	516.6	12.0	11.0	11.0	7.4	7.4	8.0	6.3	6.2	69.3
Control	12	F	510.0	11.0	9.4	10.0	7.3	7.4	7.0	6.0	5.5	63.6
Control	13	F	500.0	10.0	9.5	10.0	7.0	7.2	6.8	5.2	5.3	61.0
Control	14	M	486.0	10.0	9.3	8.9	6.0	7.0	6.0	5.0	5.0	57.2

(*Continued*)

Continued.

Group	Partic. nr.	Gender	TOEFL	Accu.	Omis.	Add.	Gram.	Expr.	Term.	Pace	Acce.	Total
Control	15	M	486.0	10.0	9.4	9.0	7.0	7.4	6.5	5.0	5.1	59.4
Mean			548.3	13.1	11.4	11.4	7.8	8.1	7.8	6.9	6.8	73.3
Exp. Segments	1	F	633.0	18.2	14.1	14.0	8.9	9.0	8.8	8.9	8.2	90.1
Exp. Segments	2	M	616.0	17.9	12.2	13.7	9.0	8.8	9.0	8.5	8.1	87.2
Exp. Segments	3	M	603.0	17.0	13.1	13.0	8.8	8.4	7.9	7.5	7.8	82.7
Exp. Segments	4	M	570.0	17.8	13.4	13.1	9.0	8.7	9.0	7.8	7.9	86.7
Exp. Segments	5	F	566.6	16.3	13.7	12.7	8.9	7.9	7.9	7.3	7.2	81.9
Exp. Segments	6	F	563.3	15.2	12.3	12.0	8.1	7.3	7.7	7.0	7.0	76.6
Exp. Segments	7	M	553.3	16.0	12.5	12.4	8.0	7.8	8.0	7.2	7.1	79.0
Exp. Segments	8	F	543.3	14.0	12.0	12.0	8.0	7.2	7.6	7.0	6.8	74.6
Exp. Segments	9	M	536.6	13.0	12.0	12.0	7.9	7.0	7.0	6.8	6.6	72.3
Exp. Segments	10	M	516.6	11.0	11.0	11.0	8.0	6.8	6.6	6.3	6.0	66.7
Exp. Segments	11	M	510.0	12.0	11.0	12.0	7.6	6.9	6.8	6.7	6.3	69.3
Exp. Segments	12	F	506.6	11.0	11.0	11.0	7.8	6.8	6.3	6.0	5.9	65.8
Exp. Segments	13	F	496.6	10.0	10.6	10.0	7.3	6.0	6.0	6.0	5.3	61.2
Exp. Segments	14	F	490.0	11.0	10.9	10.6	7.7	6.6	6.0	6.0	5.5	64.3
Exp. Segments	15	M	487.3	10.2	10.3	10.0	7.0	5.6	5.6	5.3	5.0	59.0
Mean			546.1	14.0	12.0	12.0	8.1	7.4	7.3	7.0	6.7	74.5
Exp. prosody	1	F	646.0	17.0	13.0	14.6	8.0	9.2	8.0	9.1	9.2	88.1
Exp. prosody	2	F	636.0	17.0	14.5	13.0	8.6	8.9	8.0	9.2	9.3	88.5
Exp. prosody	3	M	610.0	17.0	13.0	13.5	9.0	9.4	8.0	8.9	9.2	88.0
Exp. prosody	4	M	593.3	17.9	13.0	13.0	8.0	8.0	8.0	8.7	8.9	85.5
Exp. prosody	5	M	566.6	17.9	13.0	13.3	8.0	8.0	9.0	8.8	9.1	87.1
Exp. prosody	6	F	563.3	16.6	13.5	13.7	8.0	8.0	8.0	8.7	8.6	85.1
Exp. prosody	7	F	543.3	16.9	13.4	12.4	9.0	8.0	8.0	8.8	8.9	85.4
Exp. prosody	8	M	536.0	17.0	12.6	11.6	8.0	8.1	8.7	8.8	8.8	83.6
Exp. prosody	9	F	516.6	14.0	11.0	11.3	7.0	8.5	7.5	8.6	8.7	76.6
Exp. prosody	10	M	506.6	16.0	11.3	12.4	8.0	7.7	8.3	8.6	8.7	81.0
Exp. prosody	11	F	500.0	13.0	11.0	11.2	8.0	7.0	7.0	8.5	8.0	73.7
Exp. prosody	12	M	496.6	12.0	10.0	11.0	7.0	6.1	6.0	8.4	8.0	68.5
Exp. prosody	13	M	493.3	11.0	10.9	10.7	6.6	6.3	5.6	8.0	7.9	67.0
Exp. prosody	14	F	487.3	12.0	10.0	10.0	6.9	6.8	6.9	8.0	7.9	68.5
Exp. prosody	15	M	486.0	10.0	10.8	9.4	6.0	6.0	6.1	7.6	7.7	63.6
Mean			545.4	15.0	12.1	12.1	7.7	7.7	7.5	8.6	8.6	79.3

Appendix 2. Sample of training program for the prosody group

Time	Opening	Monitor/Feedback
5 mins.	Greetings with the students. Asking some question based on previous discussions to get feedback to see how things were going with the program.	
30 mins.	Activities Codes of conduct and good practice guides were explained to the students. Guides to good practice for the range of working contexts were elaborated to the students. Students were divided into five groups (each group 3 students). They were asked to listen to audio extracts from IRIB (Islamic Republic of Iran broadcasting). Each extract lasted 30 s and the groups were given one minute intervals to interpret it into English. Each time one member of the group could speak and sometimes students could consult with each other before interpreting.	I. moved around the class and helped some students when needed.[4]
20 mins.	Suprasegmental awareness training Prosodic theory: I. explained to the students that change of stress in English would result in different interpretations. Prosodic practice: Marking syllables: I. played a list of words/sentences and had learners count syllables and mark which syllables were stressed. Examples: Words: *deport, demarcation, campsite, cardiologist, carnival, catastrophe, cavalry, champion, charger, cheery, chowder.* Sentences: *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*	I. asked students to mark the syllables on work sheet and hand in to him 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 in practice they were able to produce the correct stress patterns of words and sentences.

	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.	
5 mins.	Homework Students were asked to listen to three minutes of VOA news, transcribe it, and deliver it the next week.	I. instructed the students to do this observation outside the classroom and present it to the class the next session.

Appendix 3. Sample of training program for the segmental group

Time	Opening	Monitor/Feedback
5 mins.	Greetings with the students. Asking some question based on previous discussions to get feedback to see how things were going with the program.	
30 mins.	Activities Codes of conduct and good practice guides were explained to the students. Guides to good practice for the range of working contexts were elaborated to the students. Students were divided into five groups (each group 3 students). They were asked to listen to audio extracts from IRIB (Islamic Republic of Iran broadcasting). Each extract lasted 30 s and the groups were given one minute intervals to interpret it into English. Each time one member of the group could speak and sometimes students could consult with each other before interpreting.	I. moved around the class and helped some students when needed.
20 mins.	Segmental training Segmental theory: I. explained to the students the segmental features of English speech (vowels, consonants, syllable structure). Segmental practice: minimal pairs: Examples: *sheep-ship*; *green-grin*; *least-list*. Sentences: *Don't sit in that seat; Did you at least get the list?; Don't slip on the floor- Don't sleep on the floor.* The students were asked to contextualize the minimal pairs; The students were asked to have a practice on vowel shifts and stress shifts by affixation.	I. asked students to mark the syllables on work sheet and hand in to him 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 in practice they were able to produce the correct stress patterns of words and sentences.
5 mins.	Homework Students were asked to listen to three minutes of VOA news, transcribe it, and deliver it the next week.	I. instructed the students to do this observation outside the classroom and present it to the class the next session.

Appendix 4. Sample of training program for the control group.

Time	Opening	Monitor/Feedback
5 mins.	Greetings with the students. Asking some question based on previous discussions to get feedback to see how things were going with the program.	
50 mins.	Activities Codes of conduct and good practice guides were explained to the students. Guides to good practice for the range of working contexts were elaborated for the students. Students were divided into five groups (each group 3 students). They were asked to listen to audio extracts from IRIB (Islamic Republic of Iran broadcasting). Each extract lasted 30 s and the groups were given one minute intervals to interpret it into English. Each time one member of the group could speak and sometimes studentscould consult with each other before interpreting.	I. moved around the class and helped some students when needed.
5 mins.	Homework Students were asked to listen to three minutes of VOA news, transcribe it, and deliver it the next week.	I. instructed the students to do this task outside of the classroom and present it to the class the next session.