[This article is a post-print of 'Mahmood Yenkimaleki & Vincent J. van Heuven (2019). Prosody instruction for interpreter trainees: Does methodology make a difference? An experimental study. *Across Languages and Cultures* 20(2), 115–131.' DOI: 10.1556/084.2019.20.1. Please cite published version only.]

PROSODY INSTRUCTION FOR INTERPRETER TRAINEES: DOES METHODOLOGY MAKE A DIFFERENCE? AN EXPERIMENTAL STUDY

MAHMOOD YENKIMALEKI[1, 2], VINCENT J. VAN HEUVEN[1, 3]

[1]Leiden University, The Netherlands, [2]University of Nahavand

[3]University of Pannonia, Hungary

Abstract: This study investigates the effect of explicit vs. implicit prosody teaching on the quality of consecutive interpretation by Farsi-English interpreter trainees. Three groups of student interpreters were formed. All were native speakers of Farsi who studied English translation and interpreting at the BA level at the University of Applied Sciences, Tehran, Iran. Participants were assigned to groups at random, but with equal division between genders (6 female and 6 male students in each group). No significant differences in English language skills (TOEFL scores) could be established between the groups. Participants took a pretest of consecutive interpreting before starting the program. The control group listened to authentic audio tracks and did exercises in consecutive interpreting. The fi rst experimental group received explicit instruction of English prosody and did exercises based on the theoretical explanation which was provided by their Iranian instructor. The second experimental group received implicit instruction of English prosody through the use of recasts. The total instruction time was the same for all the groups, i.e. 10 hours. Students then took a posttest in consecutive interpretation. The results showed that explicit teaching of prosody had a significantly positive effect on the overall quality of interpreting from Farsi into English compared with that of implicit prosody instruction. These results have pedagogical implications for curriculum designers, interpreter training programs, material producers and all who are involved in language study and pedagogy.

Keywords: Consecutive interpreting, explicit teaching, implicit teaching, prosody instruction.

1. INTRODUCTION

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 (Kemper 2008). Implicit learning, on the other hand, refers to incidentally learning the structure of stimuli in the learner's environment, so that it is generally hard for the learner to

express what exactly this knowledge structure is (Cleeremans 1993; Berry 1997). Kemper (2008) states that the effectiveness of explicit and implicit instruction is determined by both the type of the learner and the rule that has to be learned.

Research shows that correct prosody (intonation, rhythm and stress) is important for successful EFL (English as a Foreign Language) pronunciation (Gut and Pillai 2014; Kang 2010; Xue and Lee 2014; Pickering 2004; Yoon 2014; Yenkimaleki and Van Heuven 2016a, 2016b, 2016c, 2016d). Gut and Pillai (2014) predict that second language learners will face problems in producing prosodic focus marking when their first language is different from the second language in the way it signals focus. In the present case, which studies the use of prosody by Iranian learners of English, the word and sentence prosody of the two languages involved, i.e. L1 = Farsi, L2 = English, are so different from one another that L2 input in English may be misinterpreted by Iranian listeners and that the intelligibility of the L2 English output may be compromised. Earlier research has in fact shown that especially the understanding of L2 English input is negatively affected by insufficient knowledge of the prosodic structure of English on the part of Iranian listeners (Yenkimaleki 2017). We have also shown that prosodic awareness training was beneficial for Farsi-English interpreter trainees.

In spite of the importance of explicit prosody teaching in second language acquisition, there are very few studies which have focused on the contribution of explicit prosody instruction in EFL contexts (Jang and Lee 2015). However, there are some studies which indicate the importance of segmentals over suprasegmentals (e.g. Flege, Murray and Ian 1995). Flege, Murray and Ian (1995) found that Italian EFL learners produced accurate vowels and consonants (segments) more than accurate prosody. Jilka (2007), writing on the difficulty associated with the teaching of prosody, points out that establishing comprehensive and universal rules and guidelines for speech production (including prosody) is difficult. It is easier to design rules that target the segmental pronunciation and prosodic problems that are specific to the specific combination of native and foreign language at issue. Jilka also says that nature of prosodic features is inherently complicated and because of the complexity of prosodic feature errors, no specific teaching methodology deals with them appropriately.

Moreover, most of the teaching methods focus on segmental aspects of the second language learner's pronunciation problems (reported in Yenkimaleki 2016a, 2016b). Ahrens (2004:10) states that, in order to solve some of the problems of instructors in prosodic feature awareness, technology should be called upon to solve the problems associated with this aspect. She claims that through computer-aided analysis of voice characteristics and prosody, we can get more information on the relationship of prosodic domains. She maintains that computer-aided analysis of voice characteristics is helpful in its present state but there should be more cooperation with experts in voice and signal processing. Hirschfeld and Trouvain (2007), discussing the current methodology in teaching prosody, suggest that suitable methods for prosody teaching be developed for second language learners. It demands the development of software that automatically recognizes segmental phonetic and prosodic deviations from the native norm, presents exercises in training programs and assesses the mastery of prosodic features for second language learners. Moreover, they state that systematic training awareness of prosodic features results in higher degree of intelligibility in the foreign language, which was illustrated by teaching practice in different academic settings.

Generally, the teaching of pronunciation in EFL settings is looked upon as a luxury and has received little attention (e.g. Goodwin, Brinton and Celce-Murcia 1994). In the pre-reform movement era in foreign-language teaching methodology, pronunciation had either a reputation as a subject language teachers tended to avoid (Fraser 2006; Macdonald 2002) or it was instructed implicitly depending on the learner's capability of imitating sounds and rhythms without any explicit instruction – i.e. an intuitive-imitative approach of teaching pronunciation (Celce-Murcia, Brinton and Goodwin 1996).

A number of scholars paid attention to this issue in EFL contexts by investigating the explicit teaching of phonological rules (e.g. Murakawa 1981; de Bot and Mailfert 1982; Leather 1990; Champagne-Muzar, Schneiderman and Bourdages 1993; Pennington 1998; Ahrens 2004; Derwing and Munro 2005; Venkatagiri and Levis 2007; Foote, Holtby and Derwing 2011; Derwing, Diepenbroek and Foote 2012; Robinson et al. 2012; Yenkimaleki and Van Heuven 2013, 2016a, 2016b, 2016c, 2016d, 2017; Suwartono 2014; Koike 2014; Yenkimaleki 2016, 2017). Some studies report a positive effect of implicit teaching of pronunciation

rules. For instance, Papachristou (2011) ran an experimental study investigating the effectiveness of pronunciation teaching of English to 16-year old Greek state school students, examining the production of English vowels. The implicit form of pronunciation instruction resulted in more native-like production of vowels.

On the basis of the findings reviewed above we have recommended that prosody teaching be included in the interpreter training curriculum (e.g. Yenkimaleki and Van Heuven 2016d; Yenkimaleki 2017). In the present article we aim to determine which instruction methodology (implicit vs. explicit) works better in training interpreters. The results will shed more light on the optimally effective choice of methodology for instructors and practitioners in teaching prosodic features for interpreter trainees.

Concretely we asked the following research question: does explicit or implicit method of prosody instruction enhance the quality of consecutive interpretation performance for student interpreter trainees?

To answer this question, we studied the development of interpreting performance in three groups of students. The first group, which served as the control group, was taught the routine curriculum. The second group spent less time on the routine exercises and received implicit prosody instruction instead. The third group was like the second but did fewer practical prosody exercises and received explicit explanation of prosodic differences between Farsi and English instead. The results of this should show that, first of all, the experimental groups, whether trained by implicit or explicit prosody instruction, should outperform the control group. It seems hazardous, however, to formulate a specific hypothesis as to the relative effectiveness of implicit versus explicit prosody teaching in general. The answer will depend on the working languages in different countries, the expectations the students have about the effective methodology of prosody instruction and the proficiency of instructors in implementing the methodological rules in teaching prosodic features. In the present case, given the quite large structural differences between English and Farsi, and given the high cognitive development of our (adult, educated and linguistically non-naïve) students, we expect explicit explanation of structural differences between the two languages to be more beneficial. Even these rather advanced language learners will not be able to quickly and effectively extract all the underlying regularities from the language materials presented to them; they will miss generalisations

and spend unnecessary time and effort trying to crack the code. These considerations lead us to the overarching hypothesis that the three groups will benefit from the teaching program in the order explicit > implicit > control.

2. METHOD

2.1. Participants

Thirty-six students of translation and interpreting between Farsi and English were chosen randomly from 100 sophomore students (i.e. in the second or higher year of the BA curriculum) at the University of Applied Sciences, Tehran, Iran. They were randomly divided into three groups of twelve students that each incorporated six male and six female students. The participants were native speakers of Farsi (New Persian) within an age range of 18 to 27 years. They participated in all sessions of the training program.

The students were admitted to the BA study in translation and interpretation only if they had a high school diploma and had passed the national entrance exam relevant to the university discipline of their choice. The national entrance exam specifically measures the applicants' knowledge of English (or other languages, e.g. German or French). Students at the BA level take both translation and interpretation courses. Graduates (after three years or 135 credits) may pursue a professional career as either interpreters or as translators depending on their interest and the amount of time they spent studying and practicing interpretation or translation. The language combination in a training program is always restricted to two working languages, where language A is the mother tongue (Farsi) and language B a non-native language (English, in most cases).

2.2. Procedure

The participants were divided into one control group and two experimental groups through the application of systematic random sampling. The control group received routine exercises, asking them to listen to authentic audio tracks in Farsi and then interpret these into

English. The first experimental group spent less time on these tasks and instead received explicit prosodic feature awareness instruction for 20 minutes during each session. The second experimental group also spent less time on interpretation tasks and instead received implicit prosodic instruction for 20 minutes during each session.

At the beginning of the program, all the participants took a test 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. The participants took part in the program for ten sessions (one hour per session) in five weeks, i.e. ten hours in all. Then, the control group and experimental groups took a pre-test on consecutive interpretation (for details, see below) so that their level of expertise in interpreting could be assessed before they received any type of training.

The control group listened to 400 minutes of authentic audio tracks and did exercises in consecutive interpreting. Moreover, both the control group and the experimental groups listened during 200 minutes to the Iranian instructor who explained how to do exercises and also provided feedback on the students' consecutive interpreting performance. Both experimental groups altogether listened for 200 minutes to authentic audio tracks and did exercises in consecutive interpreting according the contents of the audio tracks. The first experimental group received 200 minutes of explicit instruction of English prosody and did the exercises based on the theoretical explanation which was provided by their Iranian instructor (for details, see Yenkimaleki 2016b). The second experimental group received 200 minutes of implicit instruction in English prosody through authentic audio tracks and did the exercises based on the tasks. This group received instruction of prosodic features implicitly through the use of "recasts", i.e. reformulating the learner's immediately preceding erroneous utterance while maintaining his or her intended meaning (for details, see Ammar and Spada 2006). The activities covered by the three participant groups and the time (in minutes) spent on them are summarized in *Table 1*.

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

Activity	Group		
	Control	Explicit	Implicit
Audio tracks/ exercises in interpretation	400	200	200
Listening to instructor for feedbacks	200	200	200
Explicit prosody instruction		200	
Implicit prosody instruction			200
Total time spent	600	600	600

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. Then, the control group and experimental groups took a posttest on consecutive interpretation so that the effect of treatment could be assessed. Both pretest and posttest were composed of three 30-seconds audio extracts that the participants were supposed to listen to and interpreted into Farsi after a one-minute interval. The pretest and posttest were the same in principle, except that different audio excerpts, with different speakers and contents, were used. The audio extracts were authentic English news bulletins spoken by professional native newscasters (for examples and details. see Yenkimaleki 2017; Yenkimaleki and Van Heuven 2017).

Three raters, who were native speakers of Farsi and lecturers in the Translation and Interpreting Department at the University of Applied Sciences, evaluated the participants' interpreting performance. The participants' performance was scored based on the criteria adapted from Sawyer (2004). They are:

Table 2. Eight evaluation criteria 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		

a. Accuracy: Interpreters should be faithful to the meaning of source language. An optimal and complete message should be transferred to the target language such that the content and intended in the source language be preserved without omission or distortion. Accuracy of interpretation should be a primary concern for interpreters and any change resulting in different meaning and intention in the source language when transferring it to the target language is unacceptable.
b. Omission: Jones (2014) pointed out that interpreters, in some situations, are not in a position to render exact and complete messages. Interpreters may omit part of the source text in order to deliver a coherent message to the audience. In fact, they may intentionally omit part of the source language because they want to transfer only the gist of the message so that the audience can perceive the message easily. Some omissions are considered errors but in other cases they are looked at as a technique which interpreters resort to in complicated situations, when they suffer from cognitive overload. In this study, if the interpretation preserved the content and intent of the source language, it was considered not deviating from the norms, otherwise it was looked upon as error.
c. Grammar: In this study the attempt was made to evaluate the speech production of the participants observing the standard structural rules of the target language. This criterion will not differentiate between interpreters who interpret into their native language; grammatical errors may be quite frequent when interpreting into a non-native language (so-called inverse interpreting).
d. Expression: Utterances should be appropriate regarding formality and informality with the target audience. Moreover, the utterances should be manifestation of appropriate use of target language.
e. Terminology: Interpreters should be familiar with technical terms of the subject matter that they are interpreting. In this study, the attempt was made to see to what extent the participants chose the correct technical terms when transferring the message.
f. Pace: It is widely recognized that a rate of delivery of speech between 100 and 120 words per minute (wpm) is optimal for English speech (Gerver 1969; Seleskovitch 1978; Lederer 1981; as cited in Chang

2005:12). In this study, the attempt was made to see how much the output of participants would be closer to this standard rate of delivery in speech production.

g. Additions: The interpreters should be faithful to the source language message and try to preserve the content of it; adding elements to the original message would be considered an error.

h. Accentuation is one important way to signal the information status of discourse constituents. In speech processing, listeners typically consider a sentence appropriate when new information is accented and old information remain unaccented (Birch and Garnsey 1995, reported in Li, Hagoort and Yang 2008). Appropriate accentuation speeds up sentence processing by listeners (e.g. Cutler 1976; Bock and Mazzella 1983; Terken and Nooteboom 1987; Van Donselaar and Lentz 1994, reported in Li, Hagoort and Yang 2008). In this study, the attempt was made to see how much accentuation is observed appropriately.

In the results section below, we will determine the reliability of the raters employed in the assessment of the students' interpreting performance. The interrater reliability, which is expressed in terms of Cronbach's alpha coefficient, in the present task turns out to be excellent. Given the high agreement in the scores of the three raters, the data analysis could then be based on the scores averaged over the three raters.

The experiment has a straightforward pretest-posttest design. The three groups should not differ from each other in the results of either the TOEFL placement test or in terms of their interpreting performance at the beginning of the training program. We will test the hypothesis that the three groups will differ in their scores on the posttest only such that the explicit group will have better scores than the implicit-learning group, which in turn will be better than the control group, i.e. explicit > implicit > control. The effect of the teaching method can be tested most cleanly by comparing the pretest and posttest scores within subjects. As will be seen later *(Figure 1)* the overall performance scores, though in the upper half of the range, are normally distributed and show no signs of ceiling effects, which allows us to use a repeated measures analysis of variance (RM-ANOVA) to test the statistical significance of the effects found.

In the data analysis we are interested in individual differences between participants. We will use the scores on the TOEFL placement test not

only to ascertain that the three groups are equal (and equally distributed) at the beginning of the training program but also to assess the level of each individual student's proficiency in English. This knowledge, in turn, will then be used to examine possible differences in the effects of the training programs on students with better or poorer overall command of the target language. The TOEFL scores, too, are evenly spread among the participants and show no traces of either bottom or ceiling effects, so that it is safe to use parametric correlation tests (Pearson's *r* coefficient) and tests of significance such as ANOVA. Since we find large differences in individual proficiency in English in each of the three groups, the cleanest test of the effect of the treatment (explicit teaching, implicit teaching, no prosody teaching) is obtained by examining the improvement (or 'gain') in the scores of the participants between pretest and posttest. The hypothesis here is that the gain is larger in the order explicit > implicit > control. Moreover, since the program targets the learning of prosody, we predict that the gain should be larger for rating scales that address prosodic subskills than for other subskills.

3. RESULTS

We will first present the results of the TOEFL English proficiency test, with separate modules testing the learner's (i) Listening comprehension, (ii) Reading comprehension and (iii) Structure and writing skills. *Table 3* shows the TOEFL scores (mean and standard deviation) for each of the three groups of participants.

Oneway analyses of variance were run for the three TOEFL component scores separately as well as for the overall (i.e. mean) TOEFL score with group (control, implicit, explicit) as a fixed factor. The very small differences in the scores were never statistically significant for any of the four dependent variables, $F(2, 33) < 1$ in all cases. We conclude that there were no differences between the three groups in terms of proficiency in English prior to the experiment.

Table 3. Raw component and overall (mean) scores on TOEFL proficiency test obtained by control and experimental (implicit instruction; explicit instruction) groups. For participants' individual scores see Yenkimaleki (2017:192) and *Figure 1* A-B-C.

TOEFL Component	Control group		Implicit instruction		Explicit instruction	
	Mean	SD	Mean	SD	Mean	SD
Listening comprehension	56.4	6.4	56.5	7.5	56.5	6.7
Structure & writing	56.7	5.2	57.1	6.1	56.2	5.8
Reading comprehension	56.0	6.9	54.8	7.5	56.0	6.5
Overall TOEFL	56.35	6.14	56.16	6.96	56.27	6.26

The three expert raters were in excellent agreement in their judgments of the interpreting performance of the 36 participants in the pretest. Cronbach's alpha computed on the overall scores given by the raters was as high as .969, while the coefficient never dropped below .935 when one rater was left out. On the strength of this finding all further analyses of the pretest scores were done on the ratings after averaging over the three experts.

Table 4 shows the scores obtained by the interpreter trainees on the pretest. These scores are the sum of the rating components as defined in *Table 2*. The scores range theoretically between 0 and 100. The individual trainees' score range between 65 and 93. The differences in scores on the pretest between three groups are very small, as they were in the TOEFL proficiency test. The TOEFL scores are very strongly correlated with the judged quality of the interpreting performance for each of the three groups of participants, with $r = .944$ for the control group, $r = .969$ for the implicit instruction group and $r = .997$ for the explicit instruction group ($N = 12$, $p < .001$ in all three cases). The correlation across all 36 participants was $r = .963$ ($p < .001$). Given this high correlation, we decided to evaluate the statistical significance of the differences in pretest scores among the three participant groups by a repeated measures one-way analysis of variance (RM-ANOVA) with participants matched on the basis of the TOEFL scores. Degrees of freedom were Huyhn-Feldt corrected (not shown here) whenever the assumption of sphericity was

violated. The RM-ANOVA revealed a significant effect of group, $F(2, 22) = 3.8$ ($p = .038$, $p\eta^2 = .257$). Post-hoc analysis of contrasts (with Bonferroni correction for multiple testing, $\alpha = .05$) indicated that only the difference between the control group (81.9) and the implicit-instruction group (80.7) was significant. This should not be a problem for the experiment, however. If the intervention (implicit or explicit prosody instruction) should be beneficial, we expect the experimental groups to outperform the control group in the posttest.

Table 4. Overall quality rating of interpreting performance in the pretest (on a scale between 0 and 100). Ratings are listed for each judge separately as well as averaged over judges, for participants in control group and two experimental groups. For participants' individual ratings see Yenkimaleki (2017:193) and *Figure 1*A.

Rater	Control group		Implicit instruction		Explicit instruction	
	Mean	SD	Mean	SD	Mean	SD
Rater 1	81.5	6.6	81.1	7.1	81.0	8.7
Rater 2	81.0	7.8	80.5	7.8	82.5	9.3
Rater 3	83.0	9.1	79.6	8.5	79.0	9.2
Mean rating	81.8	7.7	80.6	7.5	80.8	8.8

At the end of the training program, a posttest of interpreting was run to assess the effect of the treatment. We aimed to make the pretest and posttest equally difficult but with different fragments and items. The raters and procedures were the same as in the pretest.[1] The results of the posttest ratings for control group and experimental groups are presented in *Table 5*.

The overall scores obtained in the posttest were roughly the same as those obtained in the pretest for the control group as well as for the experimental group with implicit instruction. In fact, the score obtained by the control group had dropped .3 of a point, while the implicit-instruction group had gained 1.2 points. The second experimental group,

[1] In fact, the speech fragments produced in the pre-test and the post-test were rated in one large session. One rater (the first author) knew the students; the other two raters did not.

with explicit instruction of prosody, obtained a score of 86.1 points, which is a considerable (5.3 points) improvement vis-à-vis the pretest. The effect of group on the posttest scores was statistically significant by the same type of RM-ANOVA as was used in the pretest, $F(2, 22) = 47.8$ ($p << .001$, $p\eta^2 = .813$). Post-hoc analyses revealed that the difference between the control group and the implicit-instruction group was not significant; the explicit instruction group, however, differed from both other groups.

Table 5. Posttest scores. For more information see *Table 4*. For participants' individual ratings see Yenkimaleki (2017:194) and *Figure 1*B.

Rater	Control group		Implicit instruction		Explicit instruction	
	Mean	SD	Mean	SD	Mean	SD
Rater 1	81.7	6.3	82.0	7.0	86.8	7.3
Rater 2	80.8	8.2	81.5	8.2	86.4	7.6
Rater 3	82.2	8.5	82.2	8.4	85.1	7.5
Mean rating	81.5	7.6	81.9	7.6	86.1	7.4

We assume that, in spite of our precautions to make the pretest and the posttest equally difficult, the posttest has turned out to be somewhat more difficult, i.e. yielded lower ratings. Possibly, the speakers of the posttest materials talked faster, articulated less clearly or used less common words and phrases than their counterparts in the pretest. Yet, it would be hard to imagine that ten hours of practice and feedback with interpreting tasks (see *Table 1*) would not yield any positive results for the control group. It is therefore probably better to depart from the assumption of equal pretest and posttest, and evaluate the effect of the intervention (implicit or explicit instruction) by adopting the gain, i.e. the difference between the posttest and the pretest score obtained by the same individual, as the optimal dependent variable. The gain values are listed in *Table 6.*

Table 6. Gain (difference between posttest and pretest scores) for three groups of participants. For participants' individual gain scores see Yenkimaleki (2017:195) and *Figure 1*C.

Control group		Implicit instruction		Explicit instruction	
Mean	SD	Mean	SD	Mean	SD
–0.275	1.176	1.250	1.254	5.300	1.719

An RM-ANOVA on the gain-scores reveals a highly significant effect of participant group, $F(2, 22) = 55.9$ ($p << .001$, $p\eta^2 = .836$). Moreover, post-hoc analyses show that the differences in gain between all three groups of participants are significant. The explicit-instruction group (+5.3) outperformed the implicit instruction group (+1.3), which in turn gained significantly more by the treatment than the control group (−0.3).

*Figure 1*A-B-C shows how the TOEFL scores, which were used as the matching criterion, and the group membership together determine the participant's interpreting performance.

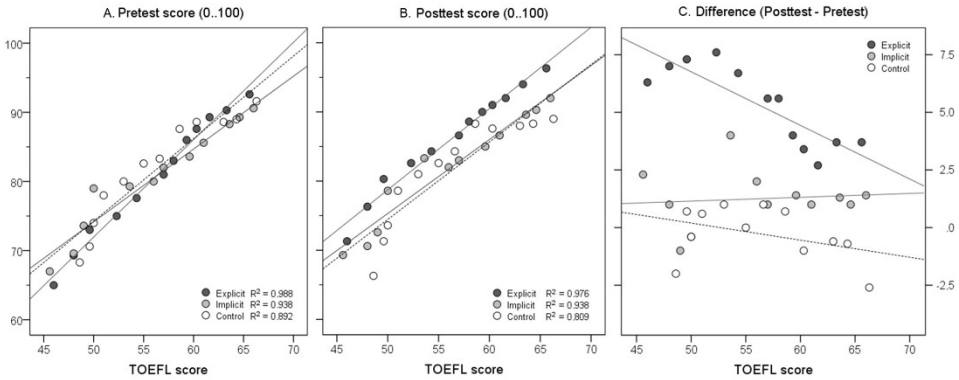

Figure 1. Pretest score (A), posttest score (B) and Gain (difference between posttest and pretest score, C) plotted as a function of the individual interpreter trainee's TOEFL score prior to the experiment, broken down by three groups of participants.

*Figure 1*A clearly illustrates the very strong dependence of the pretest scores on the individual trainee's TOEFL score as established just before the experiment started. It also shows that there is no difference in scatter between the participant groups. In panel B the strong correlation between TOEFL score and posttest performance has hardly changed but at the

same time is obvious that the group that received explicit instruction on prosody has better scores overall. Panel C clearly illustrates that all three participant groups differ from one another when the trainees' performance is expressed in terms of gain between pretest and posttest. The improvement in interpreting quality is largest for the experimental group with explicit instruction, intermediate for the implicitly instructed group, and no gain is seen for the control group. Moreover, we point out that explicit instruction in the use of prosody affects the trainees differentially. There is a significant but inverse correlation ($r = -.848$, $N = 12$, $p < .001$) between an individual's TOEFL score and the size of the benefit gained by the treatment: the poorer (in terms of the TOEFL proficiency score) the students at the start of the experiment, the more they benefit by the explicit instruction. No such inverse relationship is observed for the other two groups.

As a final exercise, we computed the gain obtained between pretest and posttest for each of the eight rating scales separately. Since the training of the experimental groups was focused on prosody, we would expect those rating scales evaluating prosodic aspects of the interpreters' performance to improve more than other aspects – relative to the control group. *Figure 2* plots the gain for each of the eight scales for the three participant groups separately. *Table 7* is a summary of the RM-ANOVAs which were run to test the effect of participant group on each of the eight rating scales separately. Braces in *Figure 2* include participant groups that do not differ significantly from each other by post-hoc analyses (with Bonferroni correction for multiple comparisons, $\alpha = .05$).

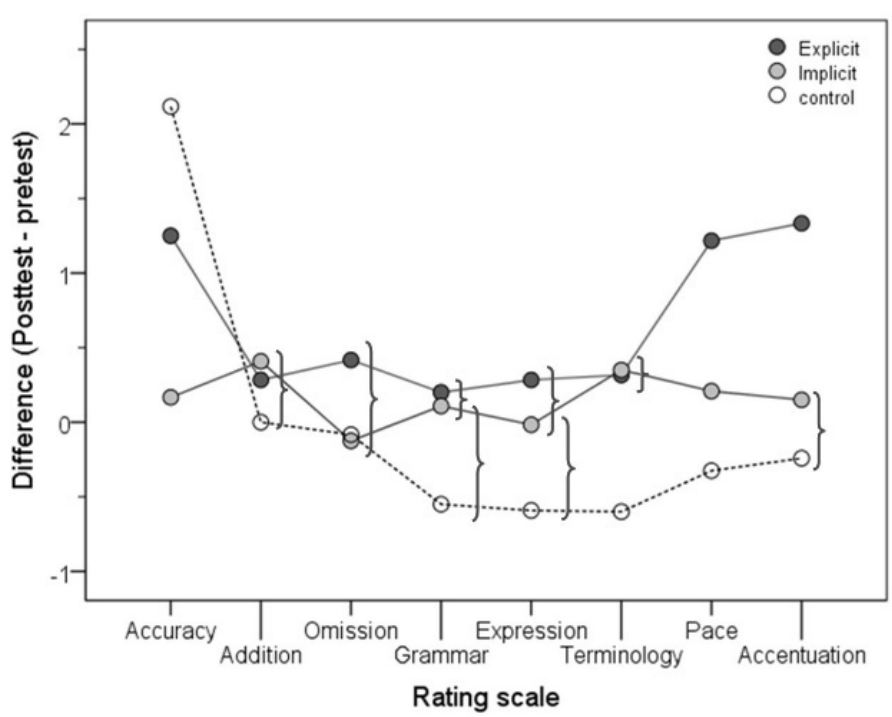

Figure 2. Gain (difference between posttest and pretest) for each of eight rating scales, broken down by participant group. Maximum difference is 10 points, except for Accuracy (20 points), Omissions (15 points) and Additions (15 points). Braces include group levels which do not differ significantly for the scale at issue by a Bonferroni test.

Table 7 shows that the effects of participant group are fairly small, or absent, for most rating scales, with the exception of three. The two scales that pertain to prosody prove highly sensitive to the group effect, with large percentages of the variance in the judgments accounted for ($p\eta^2 >$ .750). The largest effect is obtained for judged optimality of pace (i.e. fluency of interpreting). All three groups differ from each other such that the explicit-instruction group E gains more than the implicit instruction group I, which in turn gains more than the control group C. The second-largest effect of group is observed for the accentuation scale, where E outperforms I and C, which do not differ significantly from each other. Interestingly, the effect of participant group is reversed for the Accuracy scale.

Here we find that C > E > I. This finding suggests that focusing on prosody (whether by implicit or by explicit instruction) diverts the interpreter's attention away from accuracy. When the training program does not specifically draw the students' attention to prosodic aspects of the interpreting task, the traditional method would appear to impress on the students that accuracy in interpreting is the most important aspect to attend to – which is in fact made explicit by the fact the this scale is weighted more heavily than any other rating scale in the judgment procedure.

Table 7. Summary of RM-ANOVA on each of eight rating scales with participant group as the factor. P-values are based on Huyhn-Feldt corrected degrees of freedom (not indicated in the table).

Rating scale	*F*	*p*	pη^2
Accuracy	28.5	<< .001	.722
Omissions	1.6	.224	.127
Additions	1.5	.239	.122
Grammar	5.5	.011	.334
Expression	6.1	.008	.356
Terminology	13.3	<< .001	.547
Pace	35.0	<< .001	.761
Accentuation	34.1	<< .001	.756

4. CONCLUSION

This study investigated the effect of explicit vs. implicit prosody teaching on the quality of consecutive interpreting by Farsi-English interpreter trainees. The results showed that the teaching of prosody had a significantly positive effect on the overall quality of interpreting even when the time spent on prosody training could not be devoted to the traditional interpreting practice. The results also revealed that explicit instruction in the use of prosody leads to a greater improvement of interpreting quality than implicit instruction and that the gain yielded by explicit instruction was especially beneficial as the trainee was less proficient in English at the start of the training program. Moreover, the

results showed that the effect of explicit prosody teaching was especially strong as far as the interpreter's use of accentuation is concerned, i.e. on the scale that should be most sensitive to the intervention.

The results of this study converge with Yenkimaleki and Van Heuven (2016a, b, c, d), who argued that the explicit teaching of prosodic features should improve interpreter trainees' speech perception and production, which in turn should result in better performance in interpreting tasks. The results of this study are also in line with Fullana (2006), who stated that second-language learners cannot achieve native-like pronunciation without the help of explicit instruction. However, there are some studies which hold that implicit teaching of strategy instruction would help students reinforce their awareness of the language rules and would impact more strongly on students' developing pronunciation skills than explicit instruction (e.g. Griffiths 2003; Papachristou 2011). In light of such conflicting experimental results, this issue demands more investigations with other learners and different combinations of source and target languages.

The pedagogical implications of this study would pertain to interpreting programs all over the world. The policy makers, curriculum developers, practitioners and administrators need to make a number of changes in their overall approach in methodology choice in teaching prosody at interpreter training programs. Producers of teaching materials for interpreter training programs should be in contact with researchers in the field of phonetics, take publications of phonetics into consideration and include methodological issues of prosody teaching in the textbooks for interpreting programs.

REFERENCES

Ahrens, B. 2004. *Prosodie beim Simultandolmetschen*. Frankfurt am Main: Peter Lang.

Ammar A. & Spada, N. 2006. One Size Fits All? Recasts, Prompts and L2 learning. *Studies in Second Language Acquisition* Vol. 28. 543−574.

Berry, D. C. 1997. *How Implicit is Implicit Learning?* Oxford, UK: Oxford University Press.

Birch, S. L. & Garnsey, S. M. 1995. The Effect of Focus on Memory for Words in Sentences. *Journal of Memory and Language* Vol. 34. 232−267.

Bock, J. K. & Mazzella, J. R. 1983. Intonational Marking of Given and New Information: Some Consequences for Comprehension. *Memory & Cognition* Vol. 11. 64−76.

Bot, K. de & Mailfert, K. 1982. The Teaching of Intonation: Fundamental Research and Classroom Applications. *TESOL Quarterly* Vol. 16. 71−77.

Celce-Murcia, M., Brinton, D. M. & Goodwin J. M. 1996. *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages.* New York: 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* Vol. 31. 143–160.

Chang, C. C. 2005. Directionality in Chinese/English simultaneous interpreting: impact on performance and strategy use. PhD thesis. Graduate School of the University of Texas at Austin.

Cleeremans, A. 1993. *Mechanisms of Implicit Learning. Connectionist Models of Sequence Processing.* Palatino, MA: MIT Press.

Cutler, A. 1976. Phoneme-monitoring Reaction Time as a Function of Preceding Intonation Contour. *Perception & Psychophysics* Vol. 20. 55−60.

Derwing, T. M., Diepenbroek, L. G. & Foote, J. A. 2012. How Well Do General-skills ESL Textbooks Address Pronunciation? *TESL Canada Journal* Vol. 30. 22−44.

Derwing, T. M. & Munro, M. J. 2005. Second Language Accent and Pronunciation Teaching: A Research-based Approach. *TESOL Quarterly* Vol. 39. 379−397.

Donselaar, W. van & Lentz, J. 1994. The Function of Sentence Accents and Given/New Information in Speech Processing: Different Strategies for Normal-hearing and Hearing-impaired Listeners? *Language and Speech* Vol. 37. 375−391.

Flege, J. E., Murray, J. M. & Ian, R. A. M. 1995. Factors Affecting Strength of Perceived Foreign Accent in a Second Language. *Journal of the Acoustic Society of America* Vol. 97. 3125−3134.

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* Vol. 29. 1−22.

Fraser, H. 2006. Helping Teachers Help Students with Pronunciation: A Cognitive Approach. *Prospect* Vol. 21. 80−86.

Fullana, N. 2006. The Development of English (FL) Perception and Production Skills: Starting Age and Exposure Effects. In: Muñoz. C. (ed.) *Age and the Rate of Foreign Language Learning*. Clevedon: Multilingual Matters. 41−64.

Gerver, D. 1969. The Effects of Source Language Presentation Rate on the Performance of Simultaneous Conference Interpreters. In: Foulke, E. (ed.) *Proceedings of the Second Louisville Conference on Rate and/or Frequency-Controlled Speech.* Louisville, KY: Center for Rate Controlled Recordings, University of Louisville. 162−184.

Griffiths, C. 2003. Patterns of language learning strategy use. *System* Vol. 31. 367−383.

Goodwin J. D., Brinton, M. & Celce-Murcia, M. 1994. Pronunciation Assessment in the ESL/EFL Curriculum. In: Morley, J. (ed.) *Pronunciation Pedagogy and Theory: New Views, New Directions*. Alexandria, VA: Teachers of English to Speakers of Other Languages.

Gut, U. & Pillai, S. 2014. Prosodic Marking of Information Structure by Malaysian Speakers of English. *Studies in Second Language Acquisition* Vol. 36. 283−302.

Hirschfeld, U. & Trouvain, J. 2007. Teaching Prosody in German as Foreign Language. In: Trouvain, J. & Gut, U. (eds) *Non-native Prosody*. Berlin: Mouton de Gruyter. 171−189.

Jang, J. & Lee, J. 2015. Comparing Two Types of Explicit Pronunciation Instructions on Second Language Accentedness. *Linguistic Research* Vol. 32. 15−32.

Jilka, M. 2007. Different Manifestations and Perceptions of Foreign Accent in Intonation. In: Trouvain, J. & Gut, U. (eds) *Non-native Prosody,* Berlin: Mouton de Gruyter. 77−97.

Jones, R. 2014. *Conference Interpreting Explained*. Routledge: New York.

Kang, O. 2010. Relative Salience of Suprasegmental Features on Judgments of L2 Comprehensibility and Accentedness. *System* Vol. 38. 301−315.

Kemper, M. 2008. *Implicit and explicit instruction of spelling rules.* PhD dissertation. Radboud Universiteit Nijmegen.

Koike, Y. 2014. Explicit Pronunciation Instruction: Teaching Suprasegmentals to Japanese Learners of English. In: Sonda, N. & Krause, A. (eds) *JALT2013 Conference Proceedings*. Tokyo.

Leather, J. 1990. Perceptual and Productive Learning of Chinese Lexical Tone by Dutch and English Speakers. In: Leather, J. & James, A. (eds) *New Sounds*. *Proceedings of the 1990 Amsterdam Symposium on the Acquisition of Second Language Speech*. University of Amsterdam. 72−97.

Lederer, M. 1981. *La traduction simultanée: Expérience et théorie.* Paris: Minard.

Li, X., Hagoort, P. & Yang, Y. 2008. Event-related Potential Evidence on the Influence of Accentuation in Spoken Discourse Comprehension in Chinese. *Journal of Cognitive Neuroscience* Vol. 20. 906−915.

Macdonald, S. 2002. Pronunciation – Views and Practices of Reluctant Teachers. *Prospect* Vol. 3. 3−15.

Murakawa, H. 1981. Teaching English pronunciation to Japanese adults. PhD dissertation, University of Texas at Austin.

Papachristou, V. 2011. Explicit vs. Implicit Pronunciation Teaching to Greek Children: the Case of the Acquisition of English Vowels. *Selected Papers from the 19th ISTAL.* 371−381.

Pennington, M. C. 1998. The Teachability of Phonology in Adulthood: A Re-examination. *International Review of Applied Linguistics* Vol. 36. 323−341.

Pickering, L. 2004. The Structure and Function of Intonational Paragraphs in Nonnative Speaker Instructional Discourse. *English for Specific Purposes* Vol. 23. 19−43.

Robinson, P., Mackey, A., Gass, S. M. & Schmidt, R. 2012. Attention and Awareness in Second Language Acquisition. In: Gass, S. & Mackey, A. (eds) *The Routledge Handbook of Second Language Acquisition*. London: Routledge. 247−267.

Sawyer, B. 2004. *Fundamental Aspects of Interpreter Education. Curriculum and Assessment*. Amsterdam: John Benjamins.

Seleskovitch, D. 1978. *Interpreting for International Conferences: Problems of Language and Communication*. Washington, DC: Pen and Booth.

Suwartono, S. 2014. Enhancing the Pronunciation of English Suprasegmental Features through Reflective Learning Method. *TEFLIN Journal* Vol. 25. 80−93.

Terken, J. M. B. & Nooteboom, S. G. 1987. Opposite Effects of Accentuation and De-accentuation on Verification Latencies for "Given" and "New" Information. *Language and Cognitive Processes* Vol. 2. 145−163.

Venkatagiri, H. S. & Levis, J. M. 2007. Phonological Awareness and Speech Comprehensibility: An Exploratory Study. *Language Awareness* Vol. 16. 263−277.

Xue, X. & Lee, J. K. 2014. The Effects of Non-native Listeners' L1 Background and L2 Proficiency on the Perception of Foreign Accent for Pitch-manipulated Native and Chinese Accented English Speech. *Linguistic Research* Vol. 31. 275−303.

Yenkimaleki, M. 2016. Why Prosody Awareness Training is Necessary for Training Future Interpreters. *Journal of Education and Human Development* Vol. 5. 256−261.

Yenkimaleki, M. 2017. *Effect of Prosody Awareness Training on the Quality of Consecutive Interpreting between English and Farsi*. LOT International Dissertation Series, 459. Utrecht: LOT.

Yenkimaleki, M. & Heuven, V. J. van 2013. Prosodic Feature Awareness Training in Interpreting: An Experimental Study. In: Gómez Chova, L., López Martínez, A. & Candel Torres I. (eds) *Proceedings of the 6th International Conference on Education, Research and Innovation, November 18–20, 2013, Seville, Spain*. 4179−4188.

Yenkimaleki, M. & Heuven, V. J. van 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* Vol. 4. 32−41.

Yenkimaleki, M. & Heuven, V. J. van 2016b. The Effect of Prosody Teaching on Developing Word Recognition Skills for Interpreter Trainees: An Experimental Study. *Journal of Advances in Linguistics* Vol. 7. No. 1. 1101−1107.

Yenkimaleki, M. & Heuven, V. J. van 2016c. Prosody Teaching Matters in Developing Speaking Skills for Farsi–English Interpreter Trainees: An Experimental Study. *International Journal of English Language and Linguistics Research* Vol. 4. No. 5. 82−91.

Yenkimaleki, M. & Heuven, V. J. van 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.* Vol. 3. No. 3. 235−251.
DOI: 10.1080/23306343.2016.1233930

Yenkimaleki, M. & Heuven, V. J. van 2017. 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* Vol. 26. No. 1. 84−99.
DOI: 10.1080/0907676X.2017.1315824

Yoon, T. J. 2014. Speaker Variation in English Prosodic Boundary. *Linguistic Research* Vol. 31. 1−23.