Effect of prosody awareness training on the quality of consecutive interpreting between English and Farsi

Yenkimaleki, M.

Chapter nine

Effects of directionality on consecutive interpreting between English and Farsi

Abstract

The present study investigates the effect of directionality on the quality of consecutive interpretation between English and Farsi by interpreter trainees. Two sets of experiments were run. In the first experiment, participants interpreted from English into Farsi., i.e., from foreign into native language, also called *recto* ('straight') interpreting. Two groups of student interpreters were formed. All were native speakers of Farsi who studied English translation and interpreting. At the beginning of the program all the participants took a pre-test of general English proficiency. No significant differences in English language skills (TOEFL scores) could be established between the groups. Both groups participated in the program for 18 sessions with each session 90 minutes. The control group received instruction and practice about the techniques of interpreting, different aspects of interpreting, and types of interpreting. The experimental group received not only the same type of instruction as provided to the control group, be it in less time, but also awareness training on prosodic features (stress at word and at sentence level) of English for 20 minutes in each session. Both experimental and control groups received exercises in interpreting by presenting authentic audio extracts. Three expert raters assessed quality measures of interpreting performance in the post-test. In the second experiment, 32 participants interpreted from Farsi into English, i.e., from native into foreign language, also called *verso* (or: inverse) interpreting. The same procedures were followed but an additional pre-test of interpreting performance was added to the experiment and one extra scale (accentuation) also was added to the rating instrument. In each of the two experiments the group that had received prosody training outperformed the control group, especially on prosody-related rating scales such as pace (fluency). Moreover, the results showed better scores overall when interpreting was done into the mother tongue of trainees. Finally, the gain in performance by the experimental groups was larger when interpreting into the target language (i.e., from Farsi into English) than when interpreting into the interpreters' native language. The pedagogical implications of the present study would pertain to interpreting programs in general. Policy makers in training programs of interpreting should take this perspective into account when designing the curriculum of interpreting. More effective and efficient training programs can be developed by highlighting prosodic differences between native and target languages, with emphasis on correct realization of prosody in the interpreter's non-native language.

Keywords: directionality, consecutive interpretation, foreign language, mother tongue

9.1 Introduction

Research shows that interpreter trainees make errors of presentation (e.g., wrong words, incorrect pronunciation, prosody, morphology and syntax) when interpreting from their native language into a foreign language (e.g., Hyang 2003, Gile 2005), which direction is often referred to as inverse (or *verso*) interpreting. This is in line with Lederer (1978) and Dejean Le Feal (2003), who claim that, in interpreting, production of the target speech requires more attention on the part of the interpreter than perception of the source speech, so that the interpreter will suffer more from linguistic interference from the source language both in retrieving the lexical units and in constructing syntactically acceptable sentences (Gile 2005). This is in contrast to what is often called recto ('straight') interpreting, i.e., from a foreign language into the interpreters' native language. Generally, interpreting into the first language of the interpreter is supposed to yield natural speech while interpretation into second is imbued with unnaturalness and untrustworthiness. In similar vein, Samuelsson-Brown (2010) claims that translators (rather than interpreters) will be able to translate correctly into the second language but it will still be evident that the produced text is not written by a native speaker (Zahedi 2013).

Duff (1989) is in line with Newmark about the directionality in translation and says that unnaturalness may result from translation into the second or foreign language. He claims that translation into a second language typically suffers from interference by the native-language source text. Crystal (1987) also mentions that translation into a second language will yield unnatural output text. Marmaridou (1996) takes the same perspective as Duff (1989) and Crystal (1987). Marmaridou states that the result of translation into mother tongue would be better than that of a foreign language. She argues that the asymmetry follows from general linguistic processing mechanisms – as explained in the opening of this section. Dollerup (2000) not only claims that translation into a foreign language will be non-authentic and unnatural to the audience but additionally states that this is due to the fact that the translator cannot adequately express the nuances of meaning in the second language. Therefore, he too insists that translation should be done into the translator's mother tongue. Grosman (2000) goes even further when he argues that all translation should be done by native speakers of the target language; translation into a second language is unacceptable. Moreover, Grosman states that *recto* translation will be axiomatically highly proficient (reported in Pavlovic 2013). Durban (2011) agrees that inverse translating generally yields work of inferior quality. Nevertheless, a survey of translation practice in different countries by Hunziker Heeb (2016) shows that *verso* translation is done in some countries, such as Croatia (Pavlović 2007), Slovenia (Pokorn 2008) and Poland (Whyatt & Kosciuczuk 2013).

The directionality issue has not been investigated systematically in interpreting studies, either. Bors-Brann (1976) points out that there are some interpreters who work into the second language but instructors and practitioners in the field of interpreting believe that true interpretation is only possible if it is done into the interpreter's mother tongue. As a case in point, Seleskovitch (1978) states that speech production is natural and idiomatic only in *recto* interpreting. So, in spoken interpretation the same view exists as in written translation studies, viz., that output of straight interpreting is superior to that of inverse interpreting (Gile 2005).

Gile (2005) estimates that in interpretation both production and comprehension demand 40% less processing capacity when the interpreting is done into the interpreter's mother tongue that when the interpreting is into the second language. Independently of the directionality, comprehension requires 30% processing capacity and production 70%. Based on these estimated parameter values, the attention required by the various interpreting tasks can be predicted by a straightforward computation, where PC unit stands for Processing Capacity Unit.

1. Straight interpreting (i.e., from foreign language into mother tongue):
   Comprehension + production requirements: $30 \times 1.0 + 70 \times 0.6 = 72$ PCUs
2. Inverse interpreting (i.e., from mother tongue into foreign language):
   Comprehension + production requirements: $30 \times 0.6 + 70 \times 1.0 = 88$ PCUs

Accordingly, Gile states that it is better for the interpreters to work from the foreign language into their mother tongue rather than the other way around in order to reduce cognitive load. At the same time, if follows from Gile's model that the effect of directionality will be relatively small: inverse interpreting would require 88/72 = 22 percent more processing capacity than straight interpreting

In interpreting studies, it is necessary to have valid theories based on systematic studies about directionality so that curriculum developers may design a reliable pedagogical model in training the next generation of interpreters. In this regard, Pavlović (2007) states that, in order to construct valid theories and pedagogical models of directionality, experimental studies should be performed to investigate this aspect systematically. The present study responds to this challenge and investigates the issue of directionality experimentally.

To conclude this introduction we explicitly state the research question under investigation:

Do interpreter trainees perform better when they interpret into their mother tongue in consecutive interpreting?

Our expectation is that interpreter trainees' output will have better quality when the interpretation is into their mother tongue (*recto*) than when the interpretation is into the foreign language (*verso*). The difference, however, will be modest – but significant.

9.2 Methodological considerations

The experiments we are about to compare were part of a larger study on the effect of prosody awareness training on the quality of consecutive interpreting into English by native speakers of Farsi. In the first experiment (Yenkimaleki & Van Heuven 2013, 2017, this dissertation Chapter 3), we tested an experimental group that had received specific explanation of, and training in, differences between Farsi and English in the area of prosody, with emphasis on differences in word and sentence stress and the proper use of these phenomena in English. The results were compared with those obtained for a control group of similar participants who had been taught consecutive

interpreting by the traditional curriculum, which can basically be described as 'learning by doing'. The experimental and control groups largely underwent the same training; the difference was restricted to only 20 minutes of explicit prosody awareness training in each of 18 sessions of 90 minutes of teaching. The participants were randomly selected from a larger group of interpreter trainees, and distributed over two groups (one experimental, one control) such that they were matched for gender and general proficiency in English (based on TOEFL scores) at the beginning of the experiment. The students' interpreting performance was tested at the end of the training program by having them interpret (the same) passages of spoken English newscasts into Farsi. The trainees' interpreting performance was then rated by three expert judges, i.e., native speakers of Farsi who were employed as English-Farsi interpreting instructors. The students' performance was rated along eleven scales, which aimed to cover all relevant aspects of interpreting quality. A mean rating was finally computed from differentially weighted component rating scores as an index of overall interpreting quality. The results indicated that the experimental group received better ratings than the control group overall and that they specifically outperformed the control group in scales related to prosody, in particular accentuation and pace of delivery. Separate counts of hesitations and pauses, as well as acoustic measurements revealed that the experimental group produced fewer filled and unfilled pauses, a larger speech-pause ratio and faster speech rate. To some extent this result was surprising, since the participants in both groups produced output speech in their native language, i.e., Farsi. We assumed that part of the advantage of the experimental group was caused by the interpreter's heightened awareness of the importance of prosody. A second explanation was that the experimental group was better prepared to meet the challenges of unexpected (from a Farsi point of view) locations of word and sentence stressed in the English source fragments. This, in turn, would yield better word recognition and better comprehension of the source text. Later experiments (with new groups of students) showed that these effects indeed obtain as a result of the prosody awareness training (Yenkimaleki & Van Heuven 2016a, b, this dissertation Chapters 4 and 6, respectively).

We reasoned that the effect of the prosody awareness training program should be more readily noticeable when the students interpreted from Farsi into English – so that the learned prosodic skills could be used to directly improve the quality of the output speech (in English) rather than serve to improve the recognition and interpretation of the input speech (in Farsi). So a new series of experiments was carried out, with a fresh group of interpreter trainees (Yenkimaleki & Van Heuven 2016c, this dissertation Chapter 8). Again an experimental and a control group were formed by matching students in terms of gender and TOEFL scores. This time we did not only administer a post-test (similar to the first series of experiments) but also a pre-test – using the same method of testing as before but with fragments excerpted from news bulletins in Farsi rather than in English as input speech. The students' interpreting performance in the English output in both the pre-test and the post-test was judged by three raters, using the same rating instrument as before. The raters were native speakers of Farsi and experienced Farsi-English interpreters and instructors. One rater (the author of the present dissertation) had participated earlier as a judge in the first series of experiments. The results of this experiment reveal a strong effect of the prosody awareness training. The experimental group improved more in terms of voice and, especially, in correctness of accentuation. On the latter scale an impressive difference in gain was obtained by the

experimental group, in the amount of 3.8 points on the 10-point rating scale. At the same time, however, we noticed that the overall score the interpreter trainees obtained in the second series of tests, was considerably lower than in the first.

The present chapter analyses the differences between the scores obtained in *recto* (straight) interpreting and those found in *verso* (inverse) interpreting in depth and relates the findings to the directionality issue. We will begin by summarizing the two experiments in § 9.3 and refer for details to Chapters 3 (for *recto* interpreting) and 8 (for *verso* interpreting). We will not reiterate the presentation of the results as they were obtained from the two experiments but again refer the reader to the said chapters. Instead we will proceed immediately to a statistical comparison of the results obtained in the two chapters.

9.3 Summary of methods

9.3.1 Participants

Sixty-two students of translation and interpreting studies at the BA level at Arak University, Iran, were chosen randomly. In the first experiment, which we will refer to in this chapter as Experiment R (for *recto*), 30 senior students were chosen. They were divided into two groups of 15 each incorporating 7 male and 8 female participants. The participants, aged 20-22 years, were all native speakers of Farsi. In the second experiment, which will be referred to as Experiment V (for *verso*) 32 students of translation and interpreting studies at the BA level from the same university were chosen randomly. These were 16 male and 16 female students. They were divided into two groups each incorporating 8 male and 8 female participants. Their age range was between 18-26 years old and they took part in all sessions of the program.

9.3.2 Materials

A TOEFL test was administered as the pre-test to measure the general English proficiency of the students and was used as an index of homogeneity of the groups as well. The reason for using the TOEFL was its pre-established indices of reliability and validity as a standard test. The authentic recorded extracts used in the instructional sessions were news, political discussions and social interviews. A pre-test and post-test of interpreting performance were developed around the points practiced during sessions of instruction and exercise in interpreting English and Farsi utterances. They were administered to both control and experimental groups in similar fashion.

9.3.3 Procedures

In Experiment R (*recto*), the 30 interpreter trainees were divided into two groups each including 7 male and 8 female participants through systematic random sampling. One of these groups was considered the control group, which received routine instruction in interpreting (i.e., the routine curriculum or the syllabus which is used by instructors in

academic settings in teaching different courses). The other was considered the experimental one, which received awareness training on prosodic features of English to practice in interpreting. Before any instruction all participants took a pre-test of general English proficiency. For the control group, the techniques of interpreting, different aspects of interpreting, and types of interpreting were normally instructed and practiced. For the experimental group students received not only the same instruction as provided to the control group but also information on prosodic features (stress at word and at sentence level) of English and their effect on their performance. Altogether each group took part in 18 sessions (two hours per session and one session every week) for a total of 36 hours of instruction. In both classes authentic extracts from spoken English were presented to the students and they interpreted the verbal utterances consecutively. Typically, the control group practiced consecutive interpreting in each session in role plays during 20 minutes while the experimental group received explanation of prosodic concepts and did prosodic exercises. In addition to the formative quizzes administered from time to time during sessions of program, a post-test test was administered at the last session to assess the performance of both groups for further analysis.

In Experiment V (*verso*), 32 two interpreter trainees participated. The same procedure was employed as in the first experiment but with two modifications. We decided to add a pre-test of interpretation at the beginning of the program to assess the basic level of the participants' consecutive interpreting skills before the start of the program. Moreover, in order to disambiguate potentially confusing rating scale called 'accent', we split this criterion into two separate scales, i.e., 'accentedness of pronunciation' and 'appropriateness of accentuation by word and sentence stresses'.

The participants' performance, both in Experiment R and Experiment V, was scored based on the criteria adapted from Sawyer (2004). Criteria 1-10 and a description of the rating procedures were elaborated in Chapters 3 (§ 3.5.2) and 8 (§8.3.2). These are:

Table 9.1 Eleven evaluation criteria subdivided into three domains used in the quality judgment of interpreting performance. Weights add up to 110. After Sawyer (2004). Note that in experiment 1 'Foreign accent' and 'accentuation' were collapsed into a single scale (7 points) called 'accent'.

Meaning		**Language use**		**Presentation**	
Accuracy	20	Grammar	7	Pace	10
Omissions	15	Expression	7	Voice	10
Overall coherence	10	Word choice	7	Accentuation	10
		Terminology	7		
		Foreign accent	7		

The intersubjective quality ratings were supplemented by more objective correlates of interpreting quality. These were of two kinds: (i) counts of errors that were observed in written protocols of the students' interpretations and (ii) acoustic measures of fluency

in the oral delivery of the interpretation. These counts and acoustic parameters were explained in detail in Chapter 7 (§§ 7.3.3-4); here they are just summarized in Table 9.2.

Table 9.2. Summary of objective correlates of interpreting quality. For details and explanations of parameters see Chapter 7 (§§ 7.3.3-4).

Counts determined from written protocols		
1.	N key concepts omitted	normalized per 100 words
2.	N key concepts incorrectly represented	normalized per 100 words
3.	N grammatical errors	normalized per 100 words
4.	N filled pauses (+ length of filled pause)	normalized per 100 words
5.	N false starts (+ length of false stretch)	normalized per 100 words
6.	N repetitions (+ length of repetition)	normalized per 100 words
Parameters measured from acoustic signal		
7.	N IPUs 'fluent runs')	Uninterrupted speech bounded by pauses
8.	N silent pauses	Silence > 200 ms
9.	N filled pauses	*eh*, *ehm*, *mm*, *mmm*
10.	Articulation time	Duration of all IPUs added together
11.	Pause time	Duration silent + filled pauses added together
12.	Filled pause time	Duration of all filled pauses added together
13.	%pause	Pause time / (articulation time + pause time)
14.	SD IPU duration	SD IPU duration
15.	SD pause duration	SD pause duration
16.	W-Speaking rate	(articulation time + pause time) / N words
17.	S-Speaking rate	(articulation time + pause time) / N syllables
18.	W-Articulation rate	articulation time / N words
19.	S-Articulation rate	articulation time / N syllables

9.4 The effect of directionality in consecutive interpreting

We will now compare the overall quality ratings obtained in Experiment R ('straight' interpreting from English into the interpreter's native language, Farsi) and in Experiment V ('inverse' interpreting from native Farsi into English). As was explained in our introduction, inverse translation and interpreting is expected to be more difficult and cognitively more demanding than straight translation/interpreting. The participants in the two experiments were different individuals, which should normally be treated as independent samples. In the present study, however, the participants can be matched with respect to their command of English at the beginning of the experiment, i.e., on the basis of their TOEFL scores – which were obtained by administering the same standardized test to both groups of participants at the same (relative) point in time. Moreover, we will not only test the effect of directionality (straight versus inverse) on the quality of interpreting at the end of the training period (i.e., on the basis of the post-test scores) but we will also examine the potential interaction with the prosody awareness training. Here we test the hypothesis that better awareness of prosody will have a

greater yield in inverse interpreting. As a result of this, inverse interpreting is expected to be relatively more difficult for the control group than for the experimental group.

Since the number of participants differed between experiment R (*recto*, 2 × 15) and Experiment V (*verso*, 2 × 16), the matching procedure requires that one participant in each of the control and experimental group be dropped from Experiment V. The optimal matching in terms of TOEFL scores between the straight and the inverse interpreting groups was obtained by excluding the individual with the highest TOEFL score in the control group (AlB in Tables 3.4-5-6), and the person with the lowest TOEFL score in the experimental group (ZaS in Tables 7.4-5-6). By doing so we minimized the mean difference in TOEFL scores between the participants in the two experiments: 518 versus 526 for the control groups in experiments R and V, respectively, and an even 536 for each of the experimental groups.

An additional problem compromising a direct comparison between the two experiments is the difference in rating instrument. In experiment V a scale was added raising the maximum score from 100 to 110. In order to make the scores comparable across experiments, the overall ratings obtained in experiment V were therefore multiplied by 100/110 = .91. The results are shown in Figure 9.2.

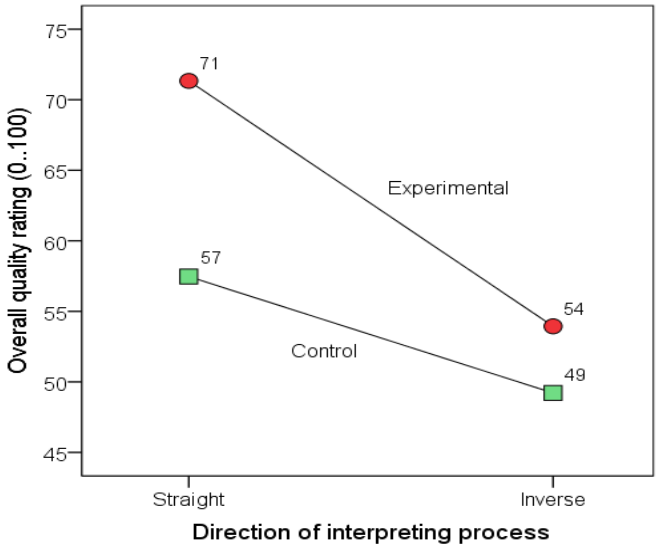

Figure 9.2. Overall quality rating of interpreting for participants with (Experimental) and without (Control) prosody awareness training broken down by direction of the interpreting process (straight: from foreign English into native Farsi; inverse: from native into foreign language).

A two-way Repeated Measures Analysis of Variance with Direction (Straight, Inverse; mean ratings 64 versus 52, respectively) and Intervention (Control, Experimental; mean ratings of 53 versus 63, respectively) then shows that both main effects are highly

significant, $F(1, 14) = 110.8$ ($p < .001$, $p\eta^2 = .888$) for Direction and $F(1, 14) = 92.0$ ($p < .001$, $p\eta^2 = .868$) for Intervention. Also, the interaction between Direction and Intervention is highly significant, $F(1, 14) = 30.2$ ($p < .001$, $p\eta^2 = .683$). Post-hoc analysis of contrasts (with Bonferroni correction for multiple comparisons) reveals that all pairs of conditions in Figure 9.2 differ from one another ($p < .05$) with the exception of the pair {54, 57}.

We provisionally conclude from these results that inverse interpreting yields lower ratings than straight interpreting overall ($\Delta = -12$ points), as well as for students with ($\Delta = -17$ points),and without prosody awareness training ($\Delta = -8$ points) separately. The effects of prosody awareness training, of course, are hardly influenced by the matching of the two experiments. The awareness training improves the overall quality of the interpreting performance by 10 points. Breaking the effect down by direction of the interpreting process, we observe that (counter to our prediction) the students benefit more from the prosody training in the straight direction (i.e., into the mother tongue, $\Delta = 14$ points) than in the inverse direction (into foreign English, $\Delta = 5$ points).

A more detailed view of the results is seen in Figure 9.3, where we plot the overall quality of interpreting as rated by the experts as a function of the individual TOEFL score that each participant obtained at the start of the experiment, with separate markers for participants in the R-experiment (*Recto*/Straight = green circles) and those in the V-experiment (*Verso*/Inverse = red squares). Within each experiment, no further breakdown is made into participants in experimental and control groups.

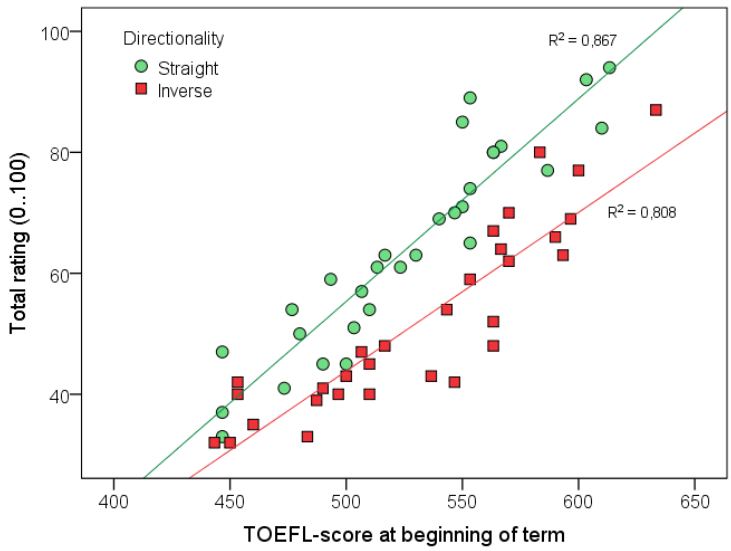

Figure 9.3. Total quality rating of interpreting (on a scale from 0 to 100) plotted as a function of the individual participant's TOEFL score obtained at the start of the training, broken down by the direction-ality of the interpreting training and tests (recto interpreting: green circes, verso interpreting: red squares). Linear regression lines are drawn separately for the recto and verso groups.

Figure 9.3 provides a clear visual confirmation that the English proficiency of the participants in the two experiments is virtually identical in terms of mean and scatter. The figure also illustrates that the participants in Experiment R obtain better quality ratings for their interpreting after the training program than their counterparts in experiment V. Finally, it can be seen that the TOEFL proficiency scores are very strong predictors of the student's interpreting performance, even though the predictive power of the TOEFL test is somewhat better in Experiment R (R^2 = .87) than in Experiment V (R^2 = .81).

So far, we have only analysed the overall ratings of the student interpreters' performance. The overall rating is the sum of the component scores the students obtained for each of ten separate rating scales. The ratings for the component scales were presented in Tables 3.4 and 8.4 for the *recto* and *verso* experiments, respectively. Figure 9.4 shows the mean ratings for each of the ten scales for the two directions side-by-side. For each experiment the means are based on the combined experimental and control groups, with N = 30 for each experiment, after elimination of two students with extreme TOEFL scores from the *verso* experiment (see above). T-tests for correlated samples, with pairs of students matched on their TOEFL scores, revealed that the differences between the *recto* and *verso* ratings were significant, with $p < .01$ (one-tailed), for each of the ten scales.

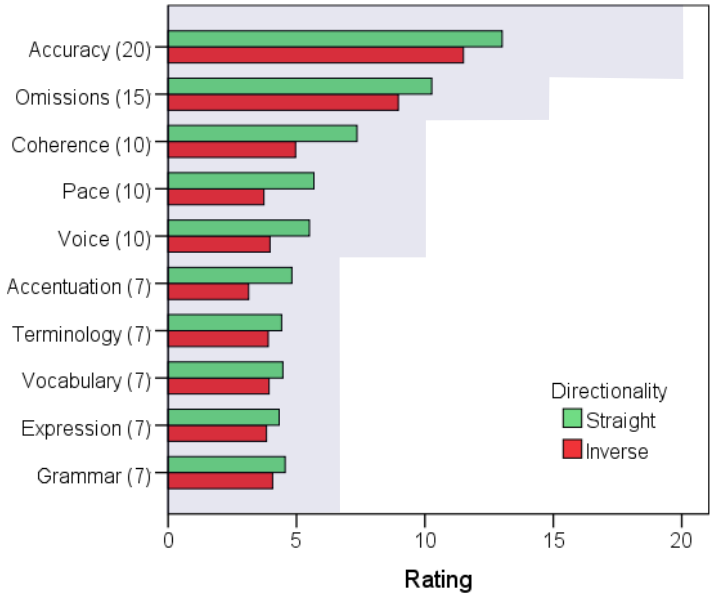

Figure 9.4. Ratings (averaged over three expert raters) for ten scales used to assess quality of interpreting in the post-test, broken down by direction of the interpreting task (straight/recto: from foreign into native language; inverse/verso: from native into foreign language). Each mean is based on 30 students. The contour of the grey polygon delineates the maximum number of points available for the rating scale. The effect of direction is significant for all ten scales (p < .01).

A potential problem with the interpretation of the results is that the raters in both Experiment R and V were native speakers of Farsi, who are university instructors in English as a foreign language but who are not native speakers of English. It is not clear if these non-native raters apply the same standards when judging the students' interpreting performance in the foreign language as native English listeners would. Native English listeners might be more lenient than the Iranian evaluators. But then again, the target listeners might not be native listeners of English but members of an international audience who have to rely on English as a lingua franca.

Given that the ratings have to be viewed with some caution, let us turn to the objective measures of interpreting quality based on counts of errors and disfluencies in the written protocols of the students' interpretations, and on acoustic measurements of the recordings. Table 9.3 summarizes the results.

Table 9.3. Mean values computed for ten objective correlates of interpreting quality broken down by direction of the interpreting task (recto, verso). The difference Δ between the means (recto minus verso), the t-statistic and the p-value are given for each parameter. Significant effects of direction are in bold face in the rightmost column.

Variables	Recto	Verso	Δ	t(29)	p
Count (per 100 words of transcribed text)					
Keywords wrong	9.89	6.82	−3.07	−4.8	**< 0.001**
Keywords omitted	12.10	9.15	−2.95	−1.4	0.193
Syntax errors	0.00	1.73	1.73	1.2	0.256
False starts or repetitions	4.80	24.31	19.5	9.9	**< 0.001**
Filled pauses	1.98	5.36	3.38	2.1	**0.047**
Acoustic variables					
Speech rate (syll/s)	5.00	3.22	1.78	6.7	**< 0.001**
Articulation rate (syll/s)	6.42	4.26	2.16	13.1	**< 0.001**
Percent speech	70.13	75.56	−5.43	−2.4	**0.023**
Percent filled pause	1.85	3.25	1.40	1.2	0.256
Percent silent pause	28.02	21.19	−6.83	3.4	**0.002**

The first two count parameters relate to accuracy of interpreting. The results show that accuracy is better when the interpreters work from the foreign into their native language. The difference between *recto* and *verso* is highly significant by a t-test for correlated samples (individuals in Experiments R and V matched on TOEFL score) when it comes to the number of incorrectly interpreted words. Note that we computations were not performed on the absolute number of errors but on the prevalence of errors per unit text produced by the interpreter, i.e., per 100 words. This normalization was necessary in order to be able to compare across individuals, whose output texts differed considerably in length. The same tendency is seen in the number of keywords (or key concepts) that were omitted in the interpretation but this time the effect failed to reach significance.

The number of syntax errors differed substantially between the two directions. No syntax errors were observed when the students interpreted into their native language, while the evaluator (the present author) noted an average of 1.73 syntax errors when the participants interpreted from native Farsi into foreign English. The difference did not reach significance, however.[1]

More than twice as many filled paused were produced in the *verso*-experiment than in the *recto* direction, the difference being just significant. The largest effect by far is seen in the number of false starts and repetitions, which for the purpose of the present analysis were simply added up.[2] Five times more false starts and repetitions occurred (per 100 words of text produced) in the *verso* direction (24.3) than in *recto* interpreting (4.8). The difference is very significant.

Speech rate (i.e., including silent pauses) and articulation rate (excluding silent pauses) were significantly faster when interpreters worked into their native language. This effect will be mainly due to the greater difficulty the speakers experience when having to produce speech in the non-native language, but a secondary reason for the difference may be in the more complex syllable structure of English, which would slow the native English speaker down as well. Generally, languages with only a simple CV syllable structure have a faster articulation rate than languages that allow complex syllables with many consonants in the onset and coda (especially if lexical tones are analysed as adding to syllable complexity, e.g., Pellegrino, Coupé & Marsico 2011).

The speech-pause ratio parameters yield somewhat unexpected results. It is seen that interpreting into the native language is characterised by a relatively large percentage of silent pause time (28%) together with a relatively small percentage of speech time (70%), while the opposite is found for *verso* interpreting – with 21% silent pause time and 76% speech time. These differences in the speech-pause ratio parameters are (highly) significant between the two directions of interpreting. No significant difference is for filled pause duration (although the prevalence of filled pauses is larger for *verso* interpreting, see above). We suspect that the longer total silent pause duration in the native language should be interpreted as a sign of competence on the part of the interpreter. The silent pauses should then coincide with deeper prosodic boundaries. When interpreting into the native language, the interpreter will typically articulate relatively fast, but will also insert relatively long pauses at deep prosodic boundaries, allowing the addressee to process the incoming chunk of information. This would be an excellent communicative strategy. When interpreting into the foreign language, articulation rate will be relatively slow, which the interpreter may then try to compensate for by economising on silent pauses. The high incidence of false starts and repeats in *verso* inter-

[1] Syntax errors were only counted if they were deemed to comprise intelligibility of the interpretation. Small deviations (such as the wrong choice of determiner) and morphological infelicities (such as inflections) were ignored.

[2] False starts typically occur at the beginning of a prosodic utterance. The speaker produces a number of words, then stops and resumes speech production from the beginning of the prosodic unit. In the present analysis the second production is not counted as a repetition. It does add, however, to the total speaking time. Repetitions were scored only when one or more words were repeated verbatim without a repair (i.e., without substitution of any element).

preting will also engender short pauses followed by longer uninterrupted second attempts.

9.6 Conclusion and discussion

The present study investigated the effect of directionality on the quality of consecutive interpreting between English and Farsi by interpreter trainees. Our results showed better results when speech fragments are interpreted into the interpreter's mother tongue. This result is line with Gouadec (2007), who observed that translations into the mother tongue contain clear, effective and natural language, indistinguishable from non-translated texts that are originally produced in that language. The results of our study also agree with Pokorn's (2008) assertion that translators should only work into their mother tongue since translation into the mother tongue guarantees good quality. Moreover, our results converge with Gile's (2005) claim that interpreters working from a foreign language into their mother tongue will have the advantage of a lower cognitive load than in the case of inverse translation. The smaller cognitive load, in turn, should lead to better quality of the interpretation. In § 9.1 we reasoned that, based on Gile's processing capacity model, *verso* interpreting would require 22 percent more processing units than *recto* interpreting, all else being equal. The comparison of the results obtained in Experiments R and V bear out that the interpreting quality, as assessed by expert raters for the two directions, is 64 (for R) against 52 (for V) on a quality scale between 0 and 100, which is a difference of 23 percent in favour of *recto* interpreting. It would appear, then, that our experimental results match the theoretical prediction from Gile's processing model quite closely.

The overall conclusion, then, would appear to be that, irrespective of the modality, whether translating written text or interpreting spoken fragments, the quality of the product is better when working from a foreign language into one's native language than when working into a non-native language (also called inverse translation or inverse interpreting).

The second hypothesis was that our students would benefit more from the prosody awareness training when performing inverse interpreting (i.e., from native Farsi into foreign English) than when engaged in straight interpreting (from English into native Farsi). Our results indeed show a statistically significant interaction between direction and prosody awareness training, but the effect runs counter to our prediction: the gain due to prosody awareness is larger for straight interpreting.

These conclusions, however, should be viewed with considerable caution. At least two aspects of the results are problematic. The first is that the quality ratings in the post-test with inverse interpreting are poorer than those of the corresponding pre-test. It seems highly improbable that the student interpreters did not improve their interpreting skill at all after half a year of training, including 18 sessions lasting 90 minutes each filled with intensive practice and theoretical explanation. This seems to suggest that the Iranian news bulletins used in the post-test were just more difficult to translate or interpret than the pre-test fragments. This issue could potentially be addressed by asking a number of qualified experts to rate the linguistic and conceptual difficulty of the frag-

ments used in the tests. In hindsight, of course, it would have been better to block the fragments across pre-test and post-test so that differences in difficulty would average out.

A more important concern is that the judges who rated the interpreter trainees' performance in both series of experiments (i.e., into the native language and inversed) were native speakers of Farsi, who had learnt English as foreign language after the age of puberty. It is unclear at this time whether these judges used the same criteria with the same severity when listening to their students in the shared native language (i.e., in Farsi) as when they had to evaluate the students' performance in English – which is a foreign language to both the students and the evaluators. In future research this problem might be tackled by offering the interpretations into English to native English judges who would then be asked to rate the intelligibility of the English text in comparison to a rendition of the same news bulletin produced by native speakers of English. Similarly, the interpretations into Farsi should be compared with ratings of original Farsi news bulletins. In order to do such an experiment with perfectly balanced materials, the news bulletins in Farsi and in English should be translation equivalents produced by the best translators in the field. An alternative approach would be to test the adequacy of the interpretation in functional tests, e.g., by asking native listeners of the recipient language to perform comprehension tasks on the interpretation. As a last resort we may count the number and severity of disfluencies, omissions and grammatical errors and compute speech rate and speech-pause ratios as indexes of interpreting quality. We would test the hypothesis that inverse interpreting (i.e., into the foreign language) yields more (severe) disfluencies and errors, a lower speaking rate and a less favorable speech-pause ratio than straight interpreting (i.e., into the native language).

This would be an agenda for the future. As matters stand currently, we will have to resign to the assumption that the Farsi raters' intuitions of what constitutes a good interpretation of a Farsi text into English is a valid reflection of a native English listener's judgment.

The practical implication of the present research would be to accept as a general policy that interpreting should be done in principle only from the foreign language into the native language, that is, the direction that we call straight interpreting. It does not mean, of course, that no attention should be paid at all to teaching the skill of inverse interpreting, i.e., from native into foreign language. There may always be situations in which inverse interpreting is the only option, simply because no straight interpreter is available – but the quality of the product will be noticeably poorer and the interpreter will get tired sooner because of the heavier cognitive demands of inverse interpreting. To alleviate the problems, especially in inverse interpreting, interpreting curricula should incorporate substantial explanation of and training in the prosodic peculiarities of the target language.

References

Beeby L. A. (2009). Direction of translation (directionality). In M. Baker & G. Saldanha (Eds.), *Routledge Encyclopedia of Translation Studies*. London, New York: Routledge, 63–67.

Birch, S. & Garnsey, S. (1995). The effect of focus on memory for words in sentences. *Journal of Memory and Language*, 34, 232–267.

Bock, K. & Mazzella, J. (1983). Intonational marking of given and new information: Some consequences for comprehension. *Memory & Cognition*, 11, 64–76.

Bros-Brann, E. (1976). Critical comments on H.C. Barik's article Interpreters Talk a Lot, Among Other Things. *AIIC Bulletin*, 4, 16–18.

Chang, C. (2005). Directionality in Chinese/English simultaneous interpreting: Impact on performance and strategy use. Ph. D. thesis, Graduate School of the University of Texas at Austin.

Crystal, D. (1987). *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press.

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception & Psychophysics*, 20, 55–60.

Déjean Le Féal, K. (2003). Impact of the international status of the interpreting students' mother tongues on training. *Forum*, 1, 63–76.

Dollerup, C. (2000). English: Axes for a target language. In M. Grosman, M. Kadric, I. Kovačič & M. Snell-Hornby (Eds.), *Translation into non-mother tongues in professional practise and training*. Tübingen: Stauffenburg Verlag, 61–70.

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*, 37, 375–391.

Duff, A. (1989). *Translation*. Oxford: Oxford 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?* Ghent: Communication and Cognition, 9–26.

Gerver, D. (1969). The effects of source language presentation rate on the performance of simultaneous conference interpreters. In E. Foulke (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.

Gouadec, D. (2007). *Translator as a profession*. Amsterdam, Philadelphia: John Benjamins.

Grosman, M. (2000). Non-mother tongue translation – An open challenge. In M. Grosman, M. Kadric, I. Kovačič & M. Snell-Hornby (Eds.), *Translation into non-mother tongues in professional practice and training*. Tübingen: Stauffenburg Verlag, 21–34.

Hunziker Heeb, A. (2016). Professional translators' self-concepts and directionality: indications from translation process research. *The Journal of Specialised Translation*, 25, 74–88.

Jones, R. (2014). *Conference interpreting explained*. Routledge: New York.

Kearns, J. (2007). Translate and be damned! Inverse translation and professional recognition. In R. Sokoloski, H. Duda, K. Klimkowski & J. Klimek (Eds.), *Warsztatytranslatorskie [Workshop on translation*.] Vol. 4. Lublin/Ottawa: Towarzystwo Naukowe Katolickiego Uniwersytetu Lubelskiego Jana Pawła II/Slavic Research Group at the University of Ottawa, 171–191.

Kelly, D., Martin, A., Nobs, M.L., Sánchez, D. & Way, C. (2003). La traducción A-B en el mercado professional. In D. Kelly, A. Martin, M.-L. Nobs, D. Sánchez & C. Way (Eds.), *La direccionalidad en traducción e interpretación: Perspectivas teóricas, profesionales y didácticas*. Granada: Atrio, 43–64.

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*, 20, 906–915.

Liu, M. (2001). Expertise in simultaneous interpreting: a working memory analysis. Ph. D. dissertation, University of Texas at Austin.

Marmaridou, S. (1996). Directionality in translation processes and practice. *Target*, 8, 49–73.

Newmark, P. (1981). *Approaches to translation*. Oxford, New York, Toronto, Sydney, Frankfurt: Pergamon Press.

Newmark, P. (1988). *A textbook of translation*. New York: Prentice Hall.

Padilla, B. (1995). Procesos de memoria y atención en la interpretación de lenguas. Doctoral dissertation, University of Granada.

Pavlović, N. (2007). Directionality in translation and interpreting practice. Report on a questionnaire survey in Croatia. *Forum*, 5, 79–99.

Pavlović, T. (2013). Exploring directionality in translation studies. *Explorations in English Language and Linguistics*. 2, 149–165.

Pellegrino, F, Coupé, C. & Marsico, E. (2011). A cross-language perspective on speech information rate. *Language*, 87, 539-558

Pokorn, N. K. (2000). Translation into a non-mother tongue in translation theory: Deconstruction of the traditional. In A. Chesterman, N. Gallardo San Salvador & Y. Gambier (Eds.), *Translation in context*. Amsterdam: John Benjamin, 61–72.

Pokorn, N. K. (2008). Translation and TS research in a culture using a language of limited diffusion: The case of Slovenia. *The Journal of Specialised Translation*, 10, 2–9.

Samuelsson-Brown, G. (2010). *A practical guide for translators*. Clevedon, OH: *Multilingual Matters.*

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

Stewart, D. (1999). Translators into the foreign language: Charlatans or professionals? *International Journal of Translation*, 4, 41–67.

Seleskovitch, D. (1978). *Interpreting for international conferences: Problems of language and communication.* Washington, DC: Pen and Booth.

Terken, J. & Nooteboom, S. (1987). Opposite effects of accentuation and accentuation on verification latencies for given and new information. *Language and Cognitive Processes,* 2, 145–163.

Whyatt, B. & Tomasz, K. (2013). Translation into a nonnative language: The double life of the native-speakership axiom. *mTm. A Translation Journal*, 5, 60–79.

Yenkimaleki, M. & Heuven, V. J. van (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*, 4179–4188.

Yenkimaleki, M. & Heuven, V. J. van (2016a). The effect of prosody teaching on developing word recognition skills for interpreter trainees: An experimental study. *Journal of Advances in Linguistics*, 7, 1101–1107.

Yenkimaleki, M. & Heuven, V. J. van (2016b). Effect of explicit teaching of prosodic features on the development of listening comprehension by Farsi-English interpreter trainees: An experimental study. *International Journal of English Language Teaching*, 4, 32–41.

Yenkimaleki, M. & Heuven, V. J. van (2016c). 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, 1–17.

Zahedi, S. (2013). Translation at the periphery: A meta-analysis of current views on translation directionality. *TranscUlturAl*, 5, 43–60.