The story of the tortoise and the hare – Speech rate in simultaneous interpretation and its influence on the quality of trainee-interpreters performance

This paper analyzes speech rate of the source text (ST) speaker and its implications on the rendition of the target text (TT) done by trainee-interpreters. Various studies (Gerver 1969, Pio 2003, Vik-Touvinen 2002) have shown that an increase in the presentation rate has an effect on the quality of the rendered text. The study was conducted on a corpus that comprised the recordings of simultaneous interpretation from English into Croatian (mother tongue). The paper will first look into the different ways of calculating speech rate, apply the optimal calculation on the ST, categorize the analyzed speeches, and finally, qualitatively verify the success of rendered texts. Most of the null hypotheses start from the point that the higher the speech rate the less successful the rendition of the target text will be. Nonetheless, the results of this study do not confirm these expectations.

Key words: speech rate; trainee interpreters; quality; categorization; simultaneous interpretation.

‘Festina lente’ Classical Latin adage

1. Introduction

Simultaneous interpretation (SI) is a complex cognitive activity, which presupposes several operations happening at or almost at the same time. Gile (2009) explains these operations by using the Effort Model, where simultaneous interpretation consists of Listening and Analysis Effort, Short-term memory Effort, Speech produc-
tion Effort and Coordination Effort. One, two or three efforts can be active at any time, but if we want SI to be successful, each of these efforts cannot exceed its processing capacity requirements and interpreter’s total available processing capacity (Gile 2009). Efforts function in a chain-reaction manner and overcapacity of requirements leads to saturation (Gile 2009: 170), which hinders successful rendition of the Target Text (TT). There are many problem triggers that can cause the overcapacity of requirements, such as thick or indiscernible accent of the source text speaker, enumerations of names and numbers, informational complexity of the speech, culture-bound terms, stress, fatigue, poor sound quality, to name but a few. This paper, however, observes fast speech rate of the Source Text (ST), as it is known to be a very common problem trigger in SI, especially for novices, which were subjects of this paper’s study.

In order to achieve the research objectives, this paper tries to determine the relation between speed of delivery of the ST speaker and its implications on the quality of the Target Text. The research addresses the following research questions:

1. What types of speech rate calculations are there? Which one should be applied to ST in SI?
2. Which ST should be categorized as slow, which as standard and which as fast?
3. Is there a correlation between the ST speech rate and the success of TT rendition?
4. To what extent does the high speed of delivery influence the quality of a rendered text?
5. Should the high speech rate of the ST speaker be a source of stress for novices?
6. Will the fast delivery rate cause problems in delivery?

Based on the research questions the null hypotheses of the research was that the faster the speech rate of the ST the poorer the quality of the rendered text.

2. Theoretical background

Speech rate has been investigated as the possible source of difficulty in interpretation. Gerver (1969) was the first to investigate high speed delivery of the ST and its impact on the TT quality and found them to be in inverse relation. Pöchhacker (1994: 113) understands that one of the most important aspects of performance in SI are temporale Redenmerkmale wie Sprechgeschwindigkeit und Verzögerungsphänomene ‘temporal features such as speech rate and time lag, also
known as ear-voice span’. Although it is important to stress that ‘some speakers may speak very rapidly but provide little information’ (Gile 1995: 173), speed of delivery still plays a very important role in SI, especially at the novices’ level. Various studies (Dahl 2006; Gile 2009; Pio 2003; Viaggio 1992; Vik-Touvinen 2002) have shown that an increase in the presentation rate has an effect on the performance of both professional interpreters and novices, the only difference being that the latter consider them a source of difficulty and carrier of additional stress. Renderation of such text is usually reflected in terms of various non-fluencies and tendency to omit phrases or different parts of the spoken segments (Gerver 1969). However, Shlesinger’s (2003) study provided converse results by showing that a higher presentation rate can have a positive effect on performance due to reduced strain on the interpreter’s working memory. Nonetheless, it must be noted that experts, who were subjects of her study, tend to have different strategies than novices in interpreting (Moser-Mercer 1997) and are able to ‘retain large amounts of information in accessible form while engaging in domain-relevant tasks’ (Shlesinger 2003: 41), unlike novices who are unable to successfully allocate attention while performing SI.

3. Methodology

3.1. Corpus and database

This study was based on the simultaneous interpretation of 19 final-year graduate students who performed the interpretation of original speeches from English into Croatian (mother tongue) during three months of training. Subjects tested in the present study have never practiced SI before the training. The ST speeches were taken from the Speech Repository Portal, which is an e-Learning tool available to students of interpretation, teachers and professional conference interpreters working for European institutions. The corpus was recorded with the Cool Edit Pro 2 program, using 96000 sample rate, mono channel 16-bit resolution. The microphone was held at 30 cm from the speaker’s mouth and approximately 12 hours of recorded material was analyzed.

For this experimental study thirty-eight STs delivered at different speaking rate were analyzed, transcribed and compared to seventy-six TT speeches interpreted by novices in order to test the relation between TT quality and ST delivery rate. Discrepancy in number between STs and TTs is a result of more than one trainee-interpreter interpreting the same ST, but in a different group and at a different time. All the ST speeches were already evaluated by professional interpreters for the Speech Repository and divided into several categories: basic, beginner, intermedi-
ate, advanced, very advanced and test-type. For this study only beginner and intermediate speeches were used. The TT interpretations did not have the same duration, so an average of 5 minutes per student was achieved by having some students interpret more than one speech if the speech was short.

3.2. Evaluation criteria

The criteria chosen to evaluate novices’ performances was product-oriented (Pöchhacker 2001; Gile 2009), for it is our belief that only through product analysis can we understand and define seminal quality indicators. Quality in interpretation has been extensively observed and evaluated (Bühler 1986; Gile 1990; Hartley et al. 2004; Kalina 2002; Kopczynsky 1994; Kurz 1993; Viezzi 1996 among others) and it is not only the linguistic, but extralinguistic criteria that influence the quality parameter. Quality in interpretation is perceived as a balance between content (grammaticality, accuracy, fidelity, linguistic correctness, style etc.) and form/presentation (fluency, voice, diction, overall impression, delivery, comprehensibility etc.).

The evaluation criteria for TTs comprised of combination of Barik’s (1994), Riccardi’s (2002) and Vančura’s (2011) evaluation grids, where the TT and the ST were first evaluated on intratextual level that observed the TT as isolated product, basically the one that final users are receiving. It included observance of the semantic content of the message, illogical sentences, but primarily focused on delivery, recording non-fluencies like filled and unfilled pauses as well as false starts, repetitions, prosody, and speed of delivery. The second step was to include intertextual level, which is indispensable for quality observance, as it compares the TT with the ST. Here, the error categories to examine discontinuities in terms of meaning between TTs and STs were the following: omissions, additions, substitutions, and in greater detail observance of syntax, passive voice, wrong tense, wrong gender, non-concordance of subject-predicator in number, numbers and names, and culture-bound terms. Finally, overall assessment was done by assigning each student a mark ranging from 1, the worst to 5, the best.1

3.3. Calculation of speech rate

One of the methodology issues that unexpectedly came up in this research was calculation and categorization of speech rate. In order to explain why we calculated

---

1 A detailed reporting of these categories may be found in Vančura (2011).
speech rate in terms of words per minute (wpm) and syllables per second, a theoretical background is necessary.

In the past, different approaches were used for measuring the speech rate, beginning with the counting of words per time unit. Already in 1966, Caroll concluded that words vary in length and counting of the speech rate based on words per minute should be discounted because it can lead to misleading results. Déjean le Féal (1978) agreed, because if we observe words as letters between two empty spaces, then certain languages have generally longer words. Campbell (1988 qtd. in Siegler 1995: 10) adds the question of pauses, explaining that the word rate is unsatisfactory ‘because of unpredictability in the structure and length of a word, which may be monosyllabic or polysyllabic, and because of the indeterminacy of any pause durations between words’. Caroll (1966) proposed the syllable as the unit of speech output in measuring the speech rate and the same type of measurement was adopted by Déjean le Féal (1978). Roach (1998: 153) argued that the number of syllables is a much more reliable measure for the speech rate than the number of words, but that one ‘should bear in mind that different languages have very different syllable structures and that we should compare languages’ speaking rate by measuring the number of sounds produced per second, rather than the number of syllables’. Fenk-Oczlon and Fenk (2004) measured the mean number of syllables per clause in different languages and their results showed that Croatian and English have the same mean number of syllables, i.e. 5.77, which is important, especially for SI, because it is easier to interpret between languages which share approximately the same speed. Many speech rate researchers mentioned in Wang and Narayanan (2005), as well as the authors themselves, used syllables per second as a unit for measurement.

Following the presented information, everything indicated that we should calculate speech rate in terms of syllables, not words. So why did we opt for calculating both syllables per second and words per minute? The answer is quite simple. It is because most of the previous research done on the subject of speech rate, especially speech rate in SI, and which served as reference for this study, was measured in words per minute.

Therefore, the speech rate in this study was measured in two ways:

a) the mean rate of syllables per second, which was obtained by counting the number of all uttered syllables for every ten seconds in the first minute of the ST. The total number of syllables was divided by six in order to calculate the average speech rate for the first 60 seconds and an average was calculated. After the average number for each of the 10 seconds was obtained, it was divided by 10 to get the
average number of syllables per second in the first 60 seconds of the ST speech.

b) Words per minute (wpm) rate was calculated on the basis of transcription of the first minute of the ST. To simplify the transcription, there were no punctuation marks other than full stops, commas and apostrophes. Taking into account the unpredictability of a word’s length, each word was standardized to five characters or keystrokes, since this is the average length of an English word in characters including spaces, numbers and other printable characters (Arif and Stuerzlinger 2009). Consequently, the number of characters with spaces was divided by five to obtain the average number of words per minute.

4. Results and discussion

The arithmetic mean of all the ST speeches is 149.44 wpm with the total span between 91 wpm (the slowest) and 211 wpm for the fastest speech. Beginner speeches range from 91 to 181 wpm with the average 142.79 wpm. Intermediate speeches range from 126.88 to 211 wpm and have an average of 149.44 wpm. As the results show, beginner and intermediate speeches that were analyzed have a very large span of speech rate, which indicates that the categorization of speeches in the Speech Repository Portal was not based on the speed parameter. We can assume that information density, topic and complexity of the speeches were crucial for their categorization.
Diagram 1. ST speeches in words per minute.

4.1. Categorization of speech rate

Once the ST were calculated it was necessary to categorize them, which posed another challenge. Most of the speech rate categorizations are done for discourse in normal, spontaneous speech. Thus, Goldman-Eisler (1961, qtd. in Ahrens 2005: 5) quotes an average of 5–8 syllables per second as the normal speech rate. Lenneber (1967, qtd. in Daró 1994) says that the mean rate of adults is 7 syllables per second and Škarić (1991) lists 4–7 syllables per second as the normal speech rate of discourse, with included pauses.

In the analyzed speech rate of the ST, the arithmetic mean of all 27 speeches was 3.542 syllables per second which is, when compared to the previous results, a little bit slower than the usual speech rate. It is important to know that the speech rate of syllables per second does not correspond to words per minute, because there are some very fast speeches in wpm and slow in terms syllables per second. This can be explained with the English not being a syllable-time but stressed-time language (Richards 1983). What this means is that the time between stressed syllables in the utterance, no matter how many there are, must accommodate to the rhythm established by the stressed syllables.

For word per minute calculation, Rivers (1981) cites the following classification: fast above 220 wpm, moderately fast: 190–220 wpm, average speech rate: 160–220 wpm, moderately slow: 130–160 wpm and slow: everything below 130 wpm. However, many authors confirm that the categorization for normal speech cannot be applied to SI and the speech rate must be slower due to ‘processing capacity requirements, because more information must be processed per unit of time’ (Gile 1995: 172). Optimal speech rates for SI drop significantly when compared to normal speech. A rate of 100 to 120 wpm had been suggested as comfortable for SI (Pöchhacker 2004). Moser-Mercer (1994) mentions in her aptitude testing that the candidate should be able to comprehend material at normal speaking speed of 90-110 wpm and Horga (1995: 389) confirms that srednji tempo izgovora koji iznosi 120 riječi/minuta najpogodniji za tumača ‘medium speech rate of 120 wpm is optimal for the interpreter’. Gerver (2002: 64) found out that ‘at input rates of 120 wpm and over, the interpreters were lagging further and further behind and making more and more errors. They spoke at a steady rate but only after long pauses’.

Thus the classification for this study was primarily focused on lower numbers suggested by previously mentioned authors. However, when speech rate for the
speeches was calculated, it implied that most of the speeches should be categorized as fast or extremely fast if we follow findings by other SI authors. Taking into account distribution of the results, our categorization was as follows: everything over 160 wpm was considered fast, average/standard was between 120–160 wpm, and slow was below 120 wpm.

Table 1. Beginner speeches presentation rate.

<table>
<thead>
<tr>
<th>Speaker – title</th>
<th>Speech rate (syllables/s)</th>
<th>Speech rate (wpm)</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Biofuels</td>
<td>2.00</td>
<td>91.00</td>
<td>Slow</td>
</tr>
<tr>
<td>- India untouchables</td>
<td>2.60</td>
<td>114.00</td>
<td>Slow</td>
</tr>
<tr>
<td>Martin Wooding - Restaurant bill</td>
<td>3.10</td>
<td>130.00</td>
<td>Standard</td>
</tr>
<tr>
<td>- Maoris and Aborigines</td>
<td>3.20</td>
<td>139.40</td>
<td>Standard</td>
</tr>
<tr>
<td>Helen Campbell - United Airlines rewards fittest people</td>
<td>3.50</td>
<td>140.00</td>
<td>Standard</td>
</tr>
<tr>
<td>Helen Campbell - Multitasking</td>
<td>3.20</td>
<td>142.00</td>
<td>Standard</td>
</tr>
<tr>
<td>Helen Campbell - Disenchantment at work</td>
<td>3.20</td>
<td>142.00</td>
<td>Standard</td>
</tr>
<tr>
<td>Zsuzsa Lang - Note taking in consecutive interpreting</td>
<td>3.20</td>
<td>143.00</td>
<td>Standard</td>
</tr>
<tr>
<td>- Red Cross and Red Crescent</td>
<td>3.70</td>
<td>144.00</td>
<td>Standard</td>
</tr>
<tr>
<td>Helen Campbell - The beneficial effects of exercising</td>
<td>3.20</td>
<td>147.00</td>
<td>Standard</td>
</tr>
<tr>
<td>Alison Graves - Education</td>
<td>3.50</td>
<td>151.00</td>
<td>Standard</td>
</tr>
<tr>
<td>Caterine Pearson - President Bush and Katrina</td>
<td>3.60</td>
<td>152.60</td>
<td>Standard</td>
</tr>
<tr>
<td>Helen Campbell - The story of the cat and the python</td>
<td>3.20</td>
<td>154.00</td>
<td>Standard</td>
</tr>
<tr>
<td>Helen Campbell - Unemployment in the age of mobility</td>
<td>4.00</td>
<td>154.80</td>
<td>Standard</td>
</tr>
<tr>
<td>Alison Graves - Children: asset or liability?</td>
<td>3.50</td>
<td>158.80</td>
<td>Standard</td>
</tr>
<tr>
<td>Ann Ford - The Hippo Story</td>
<td>4.10</td>
<td>181.00</td>
<td>Fast</td>
</tr>
</tbody>
</table>

As is clearly seen from the table most of the beginner speeches are categorized as standard (81%) followed by two slow (13%) and one fast speech, which makes 6% of the total number of categorized beginner speeches.
Table 2. Intermediate speeches.

<table>
<thead>
<tr>
<th>Speaker – title</th>
<th>Speech rate (syllables/s)</th>
<th>Speech rate (wpm)</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrew Upton - Common fisheries policy reform</td>
<td>2.6</td>
<td>126.8</td>
<td>Standard</td>
</tr>
<tr>
<td>Andrew Upton - James Dyson’s inventions</td>
<td>2.8</td>
<td>126.8</td>
<td>Standard</td>
</tr>
<tr>
<td>Helen Campbell - Early days: Big Brother</td>
<td>2.7</td>
<td>136</td>
<td>Standard</td>
</tr>
<tr>
<td>Paul Brennan - Communion silver</td>
<td>3.4</td>
<td>139</td>
<td>Standard</td>
</tr>
<tr>
<td>Peter Midgley - Early days: Advertising</td>
<td>3.3</td>
<td>145.5</td>
<td>Standard</td>
</tr>
<tr>
<td>Robin Hazelhurst - Humans versus planet Earth</td>
<td>3.4</td>
<td>148</td>
<td>Standard</td>
</tr>
<tr>
<td>Postal Services</td>
<td>3.3</td>
<td>154.8</td>
<td>Standard</td>
</tr>
<tr>
<td>Andrzej Kopczynsky - Errors in interpreting</td>
<td>3.5</td>
<td>155</td>
<td>Standard</td>
</tr>
<tr>
<td>Caterine Pearson - Weapons of mass destruction</td>
<td>3.7</td>
<td>156</td>
<td>Standard</td>
</tr>
<tr>
<td>Martin Wooding - Speech repository</td>
<td>3.5</td>
<td>156</td>
<td>Standard</td>
</tr>
<tr>
<td>David Baker - Early days: Foxhunting</td>
<td>3.6</td>
<td>161</td>
<td>Fast</td>
</tr>
<tr>
<td>Paul Brennan - Early days: weights and measures</td>
<td>3.8</td>
<td>162.4</td>
<td>Fast</td>
</tr>
<tr>
<td>Robin Hazelhurst - Wildlife in regions of war</td>
<td>3.6</td>
<td>163</td>
<td>Fast</td>
</tr>
<tr>
<td>Martin Wooding - Of nerds and interpreters</td>
<td>3.6</td>
<td>166</td>
<td>Fast</td>
</tr>
<tr>
<td>Alison Graves - The sun: friend or foe?</td>
<td>3.9</td>
<td>168</td>
<td>Fast</td>
</tr>
<tr>
<td>Robin Hazelhurst - The most remote hotel in Australia</td>
<td>3.7</td>
<td>172</td>
<td>Fast</td>
</tr>
<tr>
<td>Robin Hazelhurst - 300 club (Sauna …)</td>
<td>3.8</td>
<td>175</td>
<td>Fast</td>
</tr>
<tr>
<td>Anne Ford - European Governance</td>
<td>4.1</td>
<td>179</td>
<td>Fast</td>
</tr>
<tr>
<td>Brian Fox - Globalisation</td>
<td>4.1</td>
<td>183.8</td>
<td>Fast</td>
</tr>
<tr>
<td>- Energy</td>
<td>4.6</td>
<td>185</td>
<td>Fast</td>
</tr>
<tr>
<td>Brian Fox - Multilingualism</td>
<td>4.5</td>
<td>199</td>
<td>Fast</td>
</tr>
<tr>
<td>Brian Fox - Knowledge in interpreting</td>
<td>4.5</td>
<td>211</td>
<td>Fast</td>
</tr>
</tbody>
</table>

Intermediate speeches are categorized only as standard (55%) and fast (45%). Still, as the following results will show, their speed did not significantly influence
the quality of rendered material.

4.2. **Comparison of the ST speech rate and trainee-interpreters’ overall assessment success**

In a cross reference between whether the fast ST speech rate negatively influences the overall evaluation mark of the student’s rendition of the TT, there was no positive correlation. On a scale from 1 to 5, average grade for standard speeches was 3.79, followed by 3.47 for fast speeches and 3.16 for slow speeches.

The best results are for speeches that were categorized as standard. Surprisingly, speeches categorized as fast were generally evaluated better than slow speeches. Still, in order to show how individual the coping of ST with fast presentation rate is, the following example will be presented. Two students interpreted the fastest speech. One received the highest grade, signifying optimum quality of the rendered text, both for the intratextual and intertextual evaluation, while the other one quit after 40 seconds by saying ‘I can’t, he’s too fast’ and thus received the lowest grade. The possible reason why one of the students from the example successfully rendered the TT could be the student’s very fast speech rate in ‘natural circumstances’. She was advised by the instructor to slow down on many occasions. Nevertheless, other students quite successfully interpreted fast speeches as well, which is supported by high average evaluation mark. Unexpected success in rendition of the fast ST material can be matched with Shlesinger’s (2003) findings that a higher presentation rate can have a positive effect on performance because of the reduced strain on the interpreter’s working memory. Also, intermediate speeches, which contain majority of speeches categorized as fast, were introduced later during the training, when students learned the essential strategies for a successful rendition of the TT. Unexpectedly, slow speeches were interpreted with the lowest quality for which the trainees received bad marks in evaluation. This is very surprising, knowing that the more pauses a ST speaker makes, the more time the interpreter has to render his/her own output as well as to listen to the input without interference. Although it was only natural to assume that the student who interpreted the speech with the slowest speech rate would get the best grade, that did not happen. She received a grade 3 in the overall assessment. Tentative explanation can be found in Shlesinger’s (2003) findings that a lower presentation rate implies a risk of trace decay. Although Shlesinger’s 2003 study is in contrast with other studies, it is clear that the success of rendition of the TT is not dependent on the speech rate, as it is on various factors and the student’s personal skill to deal with a faster ST rendition.
5. Limitations and future research

Since the research was conducted as a part of another teacher’s course, there was no possibility to choose the speeches that the students were interpreting. This was due to time and space constraints and accessibility of the interpreting equipment. Another study which would have all the students interpret slow, standard and fast speeches should follow and would provide us with information that would lead to more objective conclusions. With a sufficiently large number of examples one could test the validity of the experimental results reported above. Another study that would test speech rate of students in natural discourse and possible correlation of their success in rendition of fast STs in SI is planned for the future.

6. Conclusion

The study shows that fast speech rate in ST did not affect TT quality, although trainees disliked speeches that they intuitively perceived as fast. Success of rendition of the fast ST is individual. Some trainee-interpreters presented instances of filled pauses, audible comments like Što je brza (...), pa prebrza mi je ‘She’s fast (...) she’s too fast for me’, omissions, illogical sentences and change of pace according to the presentation rate, when their task was to interpret fast ST. Trainees in this study quite often tried to keep up with the original speaker, which is one of novices’ characteristics since the speaking speed of an interpreter mainly depends on the pace of the incoming message (Daró 1994).

To conclude, success in interpreting is susceptible to many factors, speech rate being only one of them. Although the results from this study didn’t confirm the null hypothesis, it is still fairly reasonable to expect that trainee students, especially at the beginning of training will observe speed as additional stress carrier. Therefore, it is essential to highlight that speed is not something they should be afraid of. Coping with the fast ST is very individual and training should aim at teaching students to listen to the sense of the message, not word-for-word repetition. Information density i.e. complexity of the speech, preparation and familiarity with the topic are more important than the speed of delivery, since high speaking speed is typical of SI. One of the instructor’s main tasks should be sensitizing trainees to fast ST. Hopefully, it will result in one problem-trigger less to worry about.

References


Bühler, Hildegund (1986). Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters’ Multilingua 5.4: 231–235.


Alma Vančura:
The story of the tortoise and the hare – Speech rate in simultaneous interpretation and its influence on the quality of trainee-interpreters performance

Pittsburgh: Carnegie Mellon University


Author’s address:

Department of English
Faculty of Humanities and Social Sciences
Josip Juraj Strossmayer University
Lorenza Jägera 9
31000 Osijek, Croatia
avancura@ffos.hr
PRIČA O KORNJAČI I ZECU:
BRZINA GOVORA U SIMULTANOM PREVOĐENJU
I NJEN UTJECAJ NA KVALITETU IZVEDBE STUDENATA PREVOĐENJA


**Ključne riječi:** brzina govora; studenti prevodenja; kvaliteta; kategorizacija; simultano prevodenje.