Title: Searching for evidence of Gile's Effort Models in retrospective protocols of trainee simultaneous interpreters

Author: Ewa Gumul

Citation style: Gumul Ewa. (2018). Searching for evidence of Gile's Effort Models in retrospective protocols of trainee simultaneous interpreters. "Między Oryginałem a Przekładem" T. 42 (2018), s. 17-39. doi: 10.12797/MOaP.24.2018.42.02

UNIWERSYTET ŚLĄSKI W KATOWICACH

Biblioteka Uniwersytetu Śląskiego

Ministerstwo Nauki i Szkolnictwa Wyższego

Ewa Gumul
University of Silesia, Katowice, Poland
ewa.gumul@us.edu.pl

Searching for Evidence of Gile's Effort Models in Retrospective Protocols of Trainee Simultaneous Interpreters[1]

The aim of this study is to find further empirical evidence for Gile's Effort Models through retrospection. The main objective is to investigate simultaneous interpreting trainees' perception of the cognitive processes underlying the SI task, their decision-making process when faced with increased cognitive load and processing capacity saturation, as well as their coping mechanisms under the constraints of constant requirement of attentional resources allocation. The study triangulates process analysis (retrospective protocols) with product analysis (manual comparison of source and target texts). The article begins with the presentation of the model in question and an overview of the related studies.

1. Gile's Effort Models

The Effort Models formulated by Gile [1985, 1995, 2009, 2017] are among the best-known accounts of cognitive load in simultaneous interpreting, which is evidenced both by their salient presence in numerous

---

[1]   I am very grateful to Professor Daniel Gile for his insightful comments on an earlier version of this paper.

interpreting studies literature overviews [see e.g. Pöchhacker 2004, 2015, 2016, Tryuk 2007, Setton 2015, etc.] and a wealth of empirical research adopting this conceptual framework [e.g. Petite 2005; Vik-Tuovinen 2005, quoted in Gile 2008; Gumul 2006, 2017; Chang & Schallert 2007; Pym 2008; Korpal 2012, Koshkin et al. 2018, etc.]. Gile's initial intention in developing this model was only to create a pedagogical tool to help students and practitioners understand the problems they encountered in the process that could not be attributed to linguistic incompetence, deficient extralinguistic knowledge or poor quality of source-text delivery. However, although Gile stresses that, unlike other cognitive processing models, his "rough holistic models" [Gile 2003: 20] are in fact not meant to describe the interpreting process and function as a research tool, the Effort Models evolved beyond their initial training purpose and are among the most popular conceptual frameworks adopted in interpreting research to explain cognitive processes involved in SI.

The foundations of Gile's Effort Models rest on the assumption that there is an intrinsic difficulty in interpreting, which lies in the cognitive tasks involved [Gile 1997b: 197]. Thus, the approach in question focuses on explaining this difficulty inherent in the interpreting process in "a way that should facilitate the selection and development of strategies and tactics toward better interpreting performance" [Gile 1995: 156]. Although the model is firmly rooted in cognitive psychology and information-processing theory, the EMs [Gile 1983] emerged as intuitive, observation- and introspection-based conceptual structuring of SI, which only later proved to be a valid construct in terms of cognitive psychological thinking [see Gile 1995, 1997].

1.1 The Effort Model for simultaneous interpreting

Gile models simultaneous interpreting as a process consisting of three concurrent operations requiring processing capacity:

the Listening and Analysis Effort – (L)
consisting of all comprehension-oriented operations, from the analysis of the sound waves carrying the source language speech that reach the interpreter's ears, through the identification of words, to the final decisions about the meaning of the utterance;

the Production Effort – (P)
defined as the set of operations extending from the initial mental representation of the message to be delivered, through speech planning, and up to the implementation of the speech plan;
the Memory Effort – (M)
described as all the mental operations related to storage in memory of heard segments of discourse or the high demand on short-term memory during simultaneous interpreting [Gile 1997: 198]

The first version of The Effort Model [Gile 1983, 1985] listed only the three above operations and the assumption was made that the sum of the three efforts should not exceed the interpreter's total available processing capacity. In later refinements of the model Gile [1995, 1997, 1999] complements the equation with the Coordination Effort (C) stressing that the coordination function (or „executive" function) also consumes attentional resources and that coordinating the efforts is an important prerequisite for satisfactory rendering. However, incorporation of the Coordination Effort is not the only requirement for successful performance. Two more conditions must be satisfied. In the first place, L + P + M + C must be less than the Total Available Processing Capacity (TAPC). Secondly, the processing capacity management condition must be satisfied, which means that "the capacity available for each effort (LA, MA, PA, and CA) must be equal to or larger than its requirements for the task at hand" [Gile 1997: 199]. There are three operational assumptions underlying this model: the competition hypothesis, the tightrope hypothesis, and a largely non-automatic nature of the three efforts.

1.2 The Competition Hypothesis

The competition hypothesis postulates competitive allocation of effort between the three non-automatic processes. The model assumes that even though the three efforts may be to some extent cooperative, their coexistence will invariably increase processing capacity requirements [Gile 1999: 156].

In an ideal interpreting situation, each effort is working on a separate subsequent speech segment [Gile 1995: 171]. Thus, when dealing with a succession of segments A, B, C, etc., production may be involved in A, memory may be responsible for B, while listening and analysis may be working on C. However, this linearity assumption is not always valid.

Two or even three of the efforts might have to be active simultaneously and there is no single rule that accounts for the allocation of individual efforts to specific speech segments.

In certain circumstances the pattern is bound to be broken. For instance, employing the coping strategy of anticipation leads to producing a target-language segment before the corresponding source-language segment has been heard [Gile 1997: 199]. There are also other possible deviations from the linearity assumption. For example, when faced with ambiguity at the beginning of the sentence, the interpreter might decide to wait until more information is available to decode the segment and is therefore forced to store in memory more than one translation unit [Gile 2009: 168]. Gile observes that "the simultaneity of two efforts is associated with heavier load than operation of one (such as the Listening Effort only), and the simultaneity of three Efforts leads to higher cognitive pressure than the simultaneous operation of two" [Gile 2009: 169].

A recent study by Koshkin et al. [2018] provides electrophysiological evidence of the competition for cognitive resources between the Memory Effort and the Listening and Analysis effort. Their analysis performed with the aid of the event-related potential (ERP) technique confirmed that a higher load on working memory may lead to a deficit of attention to the auditory stream.

1.3 The Tightrope Hypothesis

Through metaphorical comparison of the process of interpreting to the difficulty of performing acrobatic feats on a tightly stretched rope, the third operational assumption of the Effort Models – the tightrope hypothesis – presupposes that most of the time interpreters work close to cognitive saturation. Gile emphasizes that:

> Without the Tightrope Hypothesis, the natural assumption would be that available processing capacity is sufficient to cover all the needs and that interpreting failures are due to insufficient linguistic or extralinguistic knowledge or mistakes rather than to chronic cognitive tension between processing capacity supply and demand. [Gile 2009: 182]

Most of the critique levied against Gile's Effort Models regards the tightrope hypothesis. It is mostly criticized for lack of empirical evidence to support this claim [see e.g. Seeber 2011, Seeber & Kerzel 2011, Setton

2003]. Some researchers even claim that it is infeasible to obtain tangible evidence [see e.g. Pöchhacker 2016], given that the model is not precise enough to reflect real-time processing. In their study aiming at pupillo-metric analysis of cognitive load, Seeber and Kerzel [2012] found larger pupil dilation attesting to increased cognitive effort only towards the end of the sentences with verb-final structures in German-English language pair, whereas most of the time the pupil dilation was lower than maximum. They interpreted the results as refuting Gile's tightrope hypothesis, according to which simultaneous interpreters are expected to experience cognitive load bordering on the limit of their cognitive capacity during most of the process. However, these findings do not seem to refute Gile's hypothesis entirely. Given that pupillometric measurements are only relative, not absolute, the proportions cannot indicate with precision how close interpreters are to cognitive saturation.

In recent works Gile [2008, 2009, 2017] responds to the critique of his model pointing out that it is unrealistic to expect tangible quantitative analysis of cognitive architecture in the case of a construct which was formulated as holistic and intuitive and which, in principle, was meant not to postulate any specific process architecture. The nature of his hypothesis is thus misunderstood [Gile 2017: 10].

Gile also indicates that the claim about no empirical support for his tightrope hypothesis is largely unfounded. Leaving aside extensive anecdotal evidence from practitioners and interpreting trainers, there are some studies which lend support to Gile's claim [e.g. Barranco Droege 2015]. Among them are also Gile's own [1999, 2011]. One of his experimental studies [Gile 1999] was conducted to provide evidence supporting his initially intuitive claim. In this experiment 10 subjects were asked to interpret the same source text twice, the second rendition following shortly the first one. Multiple renditions of the same text revealed that not only were there substantial inter-subject disparities in terms of errors and omissions, but also the two outputs of the same subjects differed in this respect. The second translation rendered errors in different segments than the first one and there were some infelicities in those segments which were interpreted correctly at the first attempt.[2] Gile [2008] acknowledges

---

[2]   Pym [2008] reanalyzes part of Gile's data (only cases of omission) from a pragmatic and socio-cultural perspective and arrives at a conclusion that most of the omissions performed during the second rendition are low-risk omissions, which suggests another kind of causal patterning apart from the intrinsically random

that these results do not entirely corroborate the claim about the cognitive saturation at global level, but rather point to local deficits at micro-level. He stresses, however, that a possibility of a more global saturation cannot be discarded. Gile [2008, 2011] also suggests that a tendency to cognitive saturation may reflect to some extent language-pair specific idiosyncrasies and thus the tightrope hypothesis might be applicable to larger degree to language pairs marked with syntactic asymmetry. This is one of the findings of his case study on authentic broadcast interpreting [Gile 2011], which also lend some support to his tightrope hypothesis.

1.4 Problem triggers

Unlike most of the existing cognitive models of simultaneous interpreting, which do not take into account the external factors that further hinder the process of the cognitively demanding interpreting task, Gile identifies several sources of SI processing failure [Gile 2009: 192ff] referring to them as problem triggers. He defines them as "factors and conditions which increase processing capacity requirements or make the interpreter more vulnerable to attention lapses and attention management errors" [Gile 2015: 136]. Problem triggers are grouped according to processing capacity requirements imposed by individual efforts.

First of all, processing failure might be due to high density of the speech associated either with fast delivery rate or high density of the information content, the latter encountered even in relatively slowly delivered texts. Additionally, densely informative speech segments may be rich in enumerations or/and devoid of hesitations and natural pauses (e.g., read out written texts which are much denser than spontaneously formulated speech). Such limitations affect both the Listening and Analysis Effort and the Production Effort [Gile 2009: 192-193]. Failure in SI can also be attributed to such external factors as deterioration of sound quality or some outside interferences taxing the Listening and Analysis Effort. The processing capacity requirements of this particular effort are also increased when the interpreter is faced with technical terms which cannot be immediately recognised, strong accents, or ungrammatical linguistic structures and incorrect lexical usage. Two additional factors that

---

distribution of processing capacity mismanagement. However, as emphasized by Pym, these findings do not deny Gile's Effort Models their explanatory power, but merely offer a wider perspective.

are bound to affect the Listening and Analysis Effort include unusual linguistic style or unconventional reasoning style [Gile 2009: 193].

In turn, the capacity requirements of the Memory Effort are substantially raised when coping with unknown names, such as, for instance, composite proper nouns [see also Gile 1984]. This effort might also be taxed heavily when the language pair involved is marked by syntactic differences requiring reordering or when the interpreter has to wait until the meaning of a given segment is disambiguated [Gile 1995: 174].

Problems resulting from temporary lack of capacity for one of the tasks are also associated with signal vulnerability. This means that the handled segments do not necessarily demand a great deal of processing capacity. These segments are vulnerable to lapses of attention due to their low redundancy and short duration, as in the case of non-contextual information, such as short names, acronyms, or numbers, whose information content is difficult to recover when the interpreter experiences a momentary lapse of attention in the Listening Effort [Gile 2015: 136]. The research conducted by Gile [1984] confirms high rate of failure in rendering proper names with low morphological redundancy.

It is also worth noting that the very presence of problem triggers does not necessarily engender problems with processing capacity. Problem triggers can only be treated as potential sources of errors or omissions, but whether they occur or not depends on the context. For instance, an informationally dense segment may come at the end of the sentence and additionally be followed by a pause. Then, the Listening and Analysis Effort is no longer active, and the whole processing capacity can be directed to the Memory and Production components [Gile 1995: 174].

1.5 Failure sequences

Because of the time pressure inherent in SI, the way one segment is processed affects the availability of processing capacity for handling further incoming segments. This makes the interpreter working in this mode prone to temporary overload or saturation, which might result in erroneous performance. Thus, to account for different errors and omissions, Gile [1995] introduces the idea of failure sequences, which assumes that a given problem trigger might potentially give rise to EOIs (errors, omissions, or infelicities – see Gile 2015) as a result of processing capacity mismanagement rather than only because of its inherent difficulty.

One possible scenario of a potential failure sequence is when the interpreter is faced with an incoming speech segment requiring additional capacity resources for production (e.g., a speech segment of high density in terms of either fast delivery rate or dense information content). As a result he or she may be forced to delay producing target-language version until more processing capacity is available for the Production Effort. Obtaining extra processing capacity is possible after the interpreter has been freed from the Listening Effort, that is, working on the incoming speech segment. This, however, may impose excessive strain on the Short-Term Memory Effort because of the backlog of incoming input segments that has accumulated in the meantime. If the interpreter tries to deal with the problem by directing more processing capacity to the Short-Term Memory Effort, this may lead to losses in the capacity aimed for the Listening and Analysis Effort, putting at risk comprehension of another incoming segment [Gile 1997b: 200]. Many other types of failure sequences are possible [see Gile 2009: 173].

The above simulation clearly shows that failure sequences do not necessarily affect the problematic segment that triggered them, but may occur at a distance, influencing the rendition of those segments that pose no particular difficulty [Gile 1995: 175]. Therefore, identifying the exact source of failure is not always possible by analysing the corresponding input segment. It may be more productive to attempt to trace it by looking globally at larger portions of discourse.

<p style="text-align:center">***</p>

Gile emphasizes that his main aim in developing the model was not to provide a description of the interpreting process, but only "to highlight the theoretical and practical consequences of the limited availability of processing capacity (attentional resources) on the process" [Gile 2015: 135]. As could be seen in the course of this section, consistently with this aim, Gile does not postulate any specific process architecture and does not propose any objective quantitative measures to validate it. Nevertheless, the model proves to be a useful tool in explaining both inter-subject and intra-subject variation in interpreting performance in terms of interpreters' management of their mental energy, to which the present study is hoped to contribute.

2. Research design and aims of the study

The data for the analysis presented in this paper have been gathered in a quasi-experimental study triangulating process analysis in the form of retrospective protocols with product analysis. In fact, the corpus of the study was recorded for the purpose of another more extensive project – a study into explicitation in simultaneous interpreting [Gumul 2017], but the research design of the study makes it possible to use the obtained data to investigate other aspects of the simultaneous interpreting process. In the retrospective session the participants were simply asked to comment on any decisions taken consciously during the task of interpreting, which already allows for a much wider scope of the analysis. Additionally, one of the inherent weaknesses of this method, which is low degree of selectivity of reports, has proved to be an advantage when analysing other aspects, as a large proportion of the obtained verbalizations simply report on problems encountered during the process of interpreting. Also the majority of the participants when reporting on strategic processing mention explicitly the problems which led them to adopt specific coping or preventive tactics. Thus, the corpus of retrospective protocols constitutes legitimate material to investigate processing capacity management and find empirical evidence for Gile's Effort Models.

The method of retrospection, as any other form of introspection, has its inherent weaknesses. First of all, it is not an objective method of obtaining evidence of cognitive load like methods of physiological response (e.g., pupillometry or eye-tracking). Moreover, unconscious processes which are also part of the cognitive effort cannot be observed by means of retrospection, as this method, by definition, can reveal only the conscious part of the observed task. This means that the obtained picture is incomplete. Retrospection is undoubtedly a highly biased method which attests more to how the cognitive load is perceived by the interpreters rather than how they actually experience it and what are the real-time processing events. Therefore, the aim of the present study is not directly to validate or test Gile's effort models, but rather to investigate trainees perception of the cognitive processes underlying the SI task and their coping mechanisms. Despite its drawbacks, retrospection gives another perspective on the construct of cognitive load in SI and definitely much more insight into the decision-making process of interpreters than objective quantifiable methods.

The participants in this study were 120 advanced interpreting students from three Polish universities: the University of Silesia in Katowice, the University of Gdańsk, and the University College of Social Sciences in Częstochowa. Prior to the experiment, the students received between 120 and 150 hours of conference interpreting training, All the subjects were native speakers of Polish with English as language B in their language combination.

The corpus of the study consists of approximately 75 hours of recordings of simultaneous interpreting outputs. The source texts belong to three different genres in order to minimise the influence of genre specificity. They comprise typical text types, routinely interpreted using simultaneous mode: conference presentations (delivered during the same seminar on medical ethics), commencement addresses (delivered at artistic schools: Berklee College of Music and The Academy of Fine Arts in Wrocław), and political speeches (both concerning the Iraqi conflict).[3] Each text was interpreted by 40 interpreters working in both directions of interpreting, which amounts to a total of 240 target texts. The source texts used in the experiment are authentic speeches. However, the texts were recorded in laboratory conditions by native lectors to control the rate of delivery and keep it at the same level in each text. For English source texts, it is on the average of 110–120 words per minute,[4] whereas in the case of source texts delivered in Polish, the rate is 80–90 words per minute.[5]

The experiment was recorded in standard laboratories used for teaching simultaneous interpreting. Prior to the interpretation, the subjects received a thorough briefing concerning the pragmatic setting of each speech, that is, the details concerning the identity of the speaker, the profile of the target audience, time, venue, and the subject matter of each text. Verbalisations were elicited with the recording without any prompting of the researcher other than the prior instructions. Immediately following the interpretation, the subjects were asked to listen to the dual-track recording

---

[3]   For a more detailed description of the corpus see Gumul [2017].

[4]   The average of 120 words per minute is the rate of delivery of the source text in English which is generally believed to be the most comfortable and optimal for simultaneous interpreters to ensure the best quality of interpreting [e.g., Gerver 1975, Pöchhacker 2004].

[5]   The difference in word count stems from the systemic differences between Polish and English, since Polish words tend to be longer. These values are believed to be roughly equivalent on the basis on the syllable count per minute [Gumul 2017].

of their own outputs and the source text and to report all decisions taken consciously during the interpreting task.[6]

3. Results and discussion

The retrospective protocols have been grouped according to four aspects of The Effort Models: problem triggers, failure sequences and two operational principles, competition hypothesis and the tightrope hypothesis. Obviously some remarks refer to more than one aspect and thus have been counted as belonging to more than one category. The table below presents the number of segments from retrospective protocols referring to a particular feature and provides examples illustrating each category. The first figure provided for each aspect (under NoP) refers to the total number of participants reporting a given cognitive load-related problem, while the second figure (under NoRC) is the total number of retrospective comments identified in the corpus referring to these problems. 531 retrospective remarks out of the total number of 5,005 comments obtained in the study are related to Gile's Effort Models, operationalised in this study as problem triggers, failure sequences, competition hypothesis, and the tightrope hypothesis:

Table 1. **Cognitive load-related processing difficulties reported in the study**

	NoP	NoRC	Example
Problem trigger	92	157	RC21/P25/T2/A-B:[1] In this fragment when he enumerates all these actors, musicians, painters, and sculptors, because the pace is quite fast, I only managed to mention three of them. I didn't want to lose the incoming text. The same happened a bit further. I simplified the text because the speaker accelerates and I had to cope with it somehow.

Failure sequence	39	52	RC20/P25/T2/A-B: In the fragment about scientists I missed the very beginning and I omitted it. I had no opportunity to go back to that, so I only mentioned the scientists. Simplifying it was my way of saving the fragment. Also the next sentence wasn't translated because I didn't manage to reenter the text after this terrible failure at the beginning. This omission wasn't a deliberate, conscious decision.
Competition hypothesis	103	231	RC5/P24/T2/A-B: When listening to this sentence I wasn't sure if there was a negation. At that moment I was busy translating the previous sentence and I didn't manage to hear that. Because of that, just to leave myself more options open and the possibility of inserting a negation later on, I decided to make the sentence affirmative – "You may think that..." – and only later added "but that's not true". That was a sort of "just-in case" strategy.
Tightrope hypothesis	73	91	RC13/P25/T2/A-B: In this fragment I simplified the text a bit because I got tangled up again and unfortunately that affected the quality of my translation.

As can be inferred from the above table, the analysed retrospective protocols provide ample evidence for all four aspects, as there is a significant number of comments referring to each of the analysed aspects. The results reveal that the participants mostly experience, or are most aware of, competing efforts: as many as 103 out of 120 trainee interpreters reported problems related to the coordination of the three concurrent efforts, some of them repeatedly, as there are 231 comments referring to this aspect. The second most frequently verbalised aspect are problem triggers, which were commented upon by 92 interpreters in 157 remarks.

The potential problem triggers identified in the source texts constituting the corpus of the present study are: high density of speech in terms of density of information, proper names, enumerations (including enumerations of proper names), metaphorical expressions in A-B interpreting, and ambiguous or unclear sentence beginnings. The most frequently reported problem trigger are proper names. Two of the source texts (T3 and T4) abound in composite names of medical associations, medical schools and journals, such as the sequence in the following segment:

(1)
ST (T3): and this is something that has been written in what is called a physician's charter which has been developed by the American Board of Internal Medicine Foundation / and the American College of Physicians Foundation and European Federation of Internal Medicine / and published in two thousand and two

A total of 14 out of 40 subjects interpreting this source text verbalized having experienced increased cognitive load due to this problem trigger and in further 5 cases the product data suggests similar problems. When dealing with this text segment, one of the participants reports having resorted to transcoding and admits that despite adopting this facilitating technique, the operation was cognitively demanding. The product data also reveals cognitive effort, which is evidenced not only by hesitation marks, false starts and a mistranslation or mispronunciation ("Ward") in the segment quoted below, but also further omissions and errors in the subsequent target-text segments where there are no obvious problem triggers:

TT (P65): i yy istnieje coś takiego jak mm kodeks mm kodeks lekarski który został w y American charter który został wynaleziony / a właściwie skonstruowany przez American Ward of internati Internal Medicine Foundation / mmm American College of Physicians eee Foundation i European Federation of Internal Medicine w roku dwa tysiące drugim

RC6/P65/T3/B-A: It was a very difficult fragment for me because I had to remember all these names of institutions. I decided to resort to foreignization because these names aren't really that important for the listener. Besides, I could have translated it incorrectly. I would have had to process mentally each of these complex names and formulate them nicely. Each of them is composed of many items and they were delivered rapidly. Additionally, there is a date in this fragment. Well, it is just a year, a very simple one – two thousand and two, just two zeros and two twos, but I think that I wouldn't have managed to give Polish equivalents for these complex names. That's why I decided to just repeat them and it cost me a lot of effort anyway.

The same problem trigger has been reported by another subject. In this case it resulted in omissions (incomplete proper names – "American Board of Internal Medicine Foundation" rendered as "American Association") and mistranslations ("European Federation of Internal Medicine" as

"European Association of Doctors"). This participant considers transcoding of these names as a more adequate solution, which she however discards because she is unable to follow the rate of delivery of the speaker:

(2)
TT (P66): / to zostało zapisane w tak zwanym / w tak zwanej karcie lekarza / która była rozwijana która była zapisana przez yyy amerykańskie stowarzyszenie oraz amerykańską szkołę wyższą lekarzy oraz stowarzyszenie europejskie dot yy lekarzy i zostało opublikowane w dwa tysiące drugim roku
RC67/P66/T3/B-A: I didn't use here the original names of the associations. I could've repeated them, but I wasn't able to catch up with the speaker. That's why I tried to translate these names in the most general way using the equivalent "associations" not to risk any error and still try to render at least part of their meaning in Polish.

Another coping tactic adopted when dealing with this particular problem trigger is generalisation. One of the participants, who also refers explicitly to a cognitive load imposed by the sequence of composite proper names, reports having used this technique:

(3)
RC72/P66/T3/B-A: There was no chance to translate these names. I wasn't able to find correct equivalents. I was in doubt whether they should be translated at all. But even if I had wanted just to repeat them, I wouldn't have managed. These are three long names concentrated in one sentence. That's why I used generalization and translated it as "two foundations and a federation." It was a very tiring sentence.

Another relatively frequently reported problem trigger was the rate of delivery, although it was kept at about 110-120 wpm, which is supposed to be the optimal speed for simultaneous interpreters. So it should not constitute a problem trigger for advanced students of interpreting. However, 32 out of 120 participants mentioned this aspect as a problem trigger at least once. Looking at the product data, it becomes clear that in some cases it is the high informational density of the speech rather than its actual pace of delivery, which creates this subjective impression of an excessively fast delivery rate. This impression might also be due to inefficient processing capacity management between individual efforts and imported cognitive load from preceding text segment.

The very density of information has also been reported as a problem trigger. One of the participants reports having simplified a densely informative text segment and omitted some items in order to save the most relevant information:

(4)
RC18/P69/T3/B-A: This fragment has been simplified because it contained a lot of information and if I wanted to translate them all precisely, some would be negatively affected. So thanks to that I could give the essential information although it wasn't as precise as in the original.

Some of the problem triggers led to failure sequences. One of the participants reports imported cognitive load from a previous segment due to apparently poor coordination of the Production Effort and the Listening and Analysis Effort:

(5)
RC4/P78/T3/B-A: Here appears this phrase "American Board of Internal Medicine". Because a few moments earlier I had a problem with the word "charter" I didn't hear this name and I knew I wouldn't be able to translate it and didn't even try to do it.

In the above example, both items are problematic, but in fact as suggested by Gile [1995, 2009], failure sequences may affect segments with no particular difficulty, ones which are not demanding in terms of processing capacity management. Such a case is reported by one of the participants who directed all the available resources to numbers processing and as a result was unable to render correctly the subsequent segment:

(6)
RC5/P8/T1/B-A: In this fragment I concentrated mainly on the date, as I usually have problems with numbers. I focused on it so much that I didn't understand completely the part that followed. I mispronounced the name as well.

A failure sequence caused by the conflict between these two efforts (The Production and The Listening and Analysis) is also verbalised by another subject:

(7)
TT (P74/T4): gdyby / na przykład zostali y poproszeni przez pacjenta / o przepisanie tabletki y poronnej stosowanej do 72 godzin po / czy / y czy zabiliby

pacjenta tak by mu pomóc / pytanie brzmiało // czy czy czy udzieliliby takiej pomocy czy zgodziliby się na to / i byli zobligowani wypowiedzieć się na ten temat

RC35: I wasn't sure what is the equivalent of „the morning-after pill" and because of that I decided to translate it descriptively.

RC36: I didn't hear well the second part of the sentence and that's why I had to omit it and resort to a slight repetition of what was said in the first part.

Having doubts about the Polish equivalent of „a morning-after pill", the interpreter uses a more extensive descriptive equivalent. Because of this time-consuming and processing-costly description, the interpreter was unable to hear the subsequent segment, which in turn resulted in omission of the propositional content. The problems with text processing and production are reported in the retrospective protocol, but in fact, they are also visible in the very product of translation. The substantial pause (marked in the transcript with a double slash) just before the reiteration used instead of the omitted fragment, the false start, and the resulting redundant repetition indicate that the interpreter was working close to processing capacity saturation, which appears to be a convincing example illustrating the tightrope hypothesis proposed by Gile [1995, 2009].

In the corpus of the study there are some more remarks referring to this operational principle of Gile's Effort Models that presupposes that most of the time interpreters work close to cognitive saturation. Testimonials of excessively high cognitive load are quite frequent, but tend to be more frequent towards the end of the texts, as evidenced by the following retrospective comments made near the final part of 19-minute and 22-minute long texts (examples 8 and 9):

(8)
RC15/P65/T3/B-A: As far as I remember, in this fragment and the subsequent ones I tried to infer what was the sense of the text on the basis of just some words because I was at the limits of my endurance.

In the above case it is apparently the Listening and Analysis Effort which is most affected due to a prolonged cognitive effort, which, as the product data shows, has led to a total processing capacity saturation. Similar problems are reported by another participant:

(9)

RC30/P6/T1/B-A: I interpreted the last sentences of the speech very generally because my concentration was already so low that I was only able to transmit the most general sense.

The competition hypothesis is another operational principle of Gile's Effort Models for which there is ample evidence in the retrospective protocols verbalised by the participants of the present study. This finding is consistent with the results obtained by Koshkin et al. [2018], who, as mentioned earlier in this paper [see section 1.2] found electrophysiological evidence of the brain activity attesting to the correlation between the load on working memory and the deficit of attention while listening to the source text. The subjects in the present study also often report the conflict between the Production Effort and the Listening and Analysis Effort which should be already working on the subsequent text segment, as in the example quoted below in which the interpreter admits the Listening Effort was not active because too much processing capacity was expended on the production:

(10)

RC8/P24/T2/A-B: I didn't hear the beginning of this sentence. At that moment I was busy translating the previous part of the text and I didn't know how the sentence began.

The subjects also report that the very coordination of the efforts is often coupled with additional problem triggers. It is the case with an unknown expression in the following example, in which the interpreter did not manage to engage the Listening and Analysis Effort in time to hear the incoming new proposition. As reported by the subject this was due to the difficulties with lexical search, which, according to Gile's Gravitational Model [Gile 1995, 2009] is more demanding in SI since the cognitive load of the interpreting task reduces the availability of the lexis and impedes accessing linguistic knowledge in long-term memory.

(11)

RC14/P74/T3/B-A: I didn't hear the first part of the sentence beginning with "professional societies" because I was still interpreting the previous sentence. Additionally I had a problem with the phrase "to come down with the fact" and while trying to figure out how to say it in Polish I lost the next fragment.

In the retrospective protocols reporting the competition hypothesis there are frequent references to coping or preventive tactics, which the interpreters adopt in order to coordinate the efforts and manage successfully their processing capacity. In the following example, the interpreter reports having omitted a speech segment intentionally to be able to engage the Listening and Analysis Effort in the subsequent segment:

(12)
RC1/P64/A-B: I omitted the phrase because the speaker was already beginning a new sentence. I knew that if I hadn't focused on the new sentence, I wouldn't have been able to catch its sense.

The comment refers to sentences opening the speech, so giving priority to the Listening and Analysis Effort appears to be a reasonable decision given that due to the linearity constraint, in simultaneous interpreting the interpreter needs more effort to disambiguate the meaning at the beginning of the text. Thus it appears that whereas at the end of the text, the coordination of the efforts is a more demanding task due to accumulated cognitive load and fatigue, it also relies heavily on processing capacity at the beginning because the interpreter has fewer clues as to the meaning and still has not formed any mental models facilitating recall and comprehension [see Setton 1999].

4. Conclusions

Although retrospection does not provide an objective method of measuring the cognitive load and does not reflect real-time processing, it does provide first-hand evidence of the conflicting efforts and problems with processing capacity management described by Gile. The total of 531 comments refer to the aspects related to Gile's Effort Models and 108 out of 120 interpreters participating in this experiment made at least one remark reporting them, which clearly confirms Gile's observations about the nature of the simultaneous interpreting process. The number of comments is surprisingly high given that the subjects received only general instructions during the retrospective sessions and this number might be expected to be higher if the subjects were asked specifically about the examined features.[7] We also have to take into account the delay factor which cer-

---

7   Some researchers believe (as I do) that in order to ensure ecological validity of the experimental study, at no stage should the subjects be informed about the actual

tainly made interpreters forget certain decisions and difficulties as well as the unwillingness of some subjects to reveal problems they experienced. Thus the actual cognitive load might be expected to be higher than the retrospective evidence suggests. We also have to take into account that the Polish-English pair is not marked by syntactic asymmetries (in neither direction of interpreting) to the same extent as German-English and Japanese-English, which provides further support for the intrinsic difficulty of cognitive processing in simultaneous interpreting in general, irrespective of the language pair involved. Substantial inter-subject differences in terms of distribution of processing capacity management problems also suggest that there are reasons at play which are beyond the features of source-text.

Above all, the evidence obtained through retrospection clearly confirms the operational principle of non-automaticity of the Efforts. The very fact of verbalizing the problems with processing capacity management means that these processes are in many cases[8] fully conscious. Additionally the process data provides some insight into the various stages of the decision-making process when the interpreters experience increased cognitive load and the reasons behind adopting certain coping tactics in favour of the others.

The evidence obtained from retrospective protocols obviously does not corroborate Gile's claim about the cognitive saturation at the global level underlying the tightrope hypothesis. It would clearly be unrealistic to expect that kind of confirmation from an intrinsically selective method like retrospection. However, although the vast majority of the obtained protocols refer to local problems at the micro level of the text, there are some protocols reporting more recurrent problems with processing capacity management. Both types appear to provide evidence for what Gile refers to as occasional reasons for cognitive saturation and failure rather than chronic reasons [see Gile 2009: 192].

---

research object. This is supposed to help avoiding unnatural behaviour on the part of the participants and provoking the comments made to please the researcher rather than reflecting the actual processes underlying the interpreting task [see Gumul 2017: 103].

[8]   We cannot rule out that some comments were *a posteriori* observations provoked by listening to one's own output, which might in certain cases install false memories [see Gumul 2017: 103].

It is also worth stressing the advantage of testing the cognitive load with introspective tools. Unlike objective methods, like pupillometry, which only measure cognitive load in attended tasks, retrospective protocols may provide evidence of the tasks which the interpreter decided not to attend to because of their predictably excessive cognitive cost. Such cases were found in the analysed corpus (as exemplified in this paper) and they would not have been captured in pupillometry tests.

Finally, it is important to emphasise that the results of the present study concern trainee interpreters, who, although in this case were at the advanced stage of their interpreting training, cannot be compared with experienced professionals. Thus the result cannot be extrapolated to simultaneous interpreting in general. Higher level of expertise would most probably mean fewer problems with coordination of the efforts and more automatised strategies when dealing with problem triggers. It would be therefore interesting to conduct a similar study using professional interpreters as subjects to see how their testimonials of experiencing a cognitive load and processing capacity management differ from those of trainees.

References

Barranco Droege, R. (2015), Minimización de „problem triggers" y optimización de la calidad en interpretación simultánea, Doctoral dissertation, unpublished, University of Granada.

Chang, C., and Schallert D. (2007), "The Impact of Directionality on Chinese/English Simultaneous Interpreting", *Interpreting*, 9(2), pp. 137-176, [on-line:] https://doi.org/10.1075/intp.9.2.02cha.

Gile, D. (1983), "Des difficultés de langue en interprétation simultanée", *Traduire*, 117, pp. 2-8.

Gile, D. (1984), "Les noms propres en interprétation simultanée", *Mutlilingua*, 3(2), pp. 79-86, [on-line:] https://doi.org/10.1515/mult.1984.3.2.79.

Gile, D. (1985), "Le modèle d'efforts et l'équilibre d'interprétation en interprétation simultanée", *Meta*, 30(1), pp. 44-48, [on-line:] https://doi.org/10.7202/002893ar.

Gile, D. (1995), *Basic Concepts and Models for Interpreter and Translator Training*, John Benjamins, Amsterdam.

Gile, D. (1997), "Conference Interpreting as a Cognitive Management Problem", in: J.E. Danks, G.M. Shreve, S.B. Fountain and M.K. McBeath (eds.), *Cognitive Processes in Translation and Interpreting*, Sage, London, pp. 196-214.

Gile, D. (1999), "Testing the Effort Models' Tightrope Hypothesis in Simultaneous Interpreting – A Contribution", *Hermes*, 23, pp. 153-172, [on-line:] https://doi.org/10.7146/hjlcb.v12i23.25553.

Gile, D. (2003), "Cognitive Investigation into Conference Interpreting: Features and Trends", in: Á. Collados Aís and J.A. Sabio Pinilla (eds.), *Avances en la investigación sobre interpretación*, Comares, Granada, pp. 1-27.

Gile, D. (2008), "Local Cognitive Load in Simultaneous Interpreting and Its Implications for Empirical Research", *Forum* 6(2), pp. 59-77, [on-line] https://doi.org/10.1075/forum.6.2.04gil.

Gile, D. (2009), *Basic Concepts and Models for Interpreter and Translator Training. Revised edition*, John Benjamins, Amsterdam.

Gile, D. (2011), "Errors, Omissions and Infelicities in Broadcast Interpreting. Preliminary Findings from a Case Study", in: C. Alvstad, A. Hild and E. Tiselius (eds.), *Methods and Strategies of Process Research. Integrative Approaches in Translation Studies*, John Benjamins, Amsterdam, pp. 201-218.

Gile, D. (2015), "Effort Models", in: F. Pöchhacker (ed.), *Routledge Encyclopedia of Interpreting Studies*. Routledge, London, pp. 135-136.

Gile, D. (2017), "The Effort Models and Gravitational Model. Clarifications and Update. 2017.02.16 Version with Updates on the Social Situatedness of the EMs and Risk Assessment", [on-line:] http://www.cirinandgile.com/powerpoint/The-Effort-Models-and-Gravitational-Model-Clarifications-and-update.pdf – 10.12.2017.

Gumul, E. (2006), "Conjunctive Cohesion and the Length of Ear-Voice Span in Simultaneous Interpreting. A Case of Interpreting Students", *Linguistica Silesiana*, 27, pp. 93-103.

Gumul, E. (2017), *Explicitation in Simultaneous Interpreting. A Study into Explicitating Behaviour of Trainee Interpreters*, Wydawnictwo Uniwersytetu Śląskiego, Katowice.

Korpal, P. (2012), "Omission in Simultaneous Interpreting as a Deliberate Act", in: A. Pym and D. Orrego Carmona (eds.), *Translation Research Projects 4*, Intercultural Studies Group, Tarragona, pp. 103-111.

Koshkin, R., Shtyrov, Y., Myachykov, A., Ossadtchi, A. (2018), "Testing the Efforts Model of Simultaneous Interpreting: An ERP Study", *PLoS ONE*, 13(10), e0206129, [on-line:] https://doi.org/10.1371/journal.pone.0206129.

Petite, Ch. (2005), "Evidence of Repair Mechanisms in Simultaneous Interpreting: A Corpus-Based Analysis", *Interpreting*, 7(1), pp. 27-49, [on-line:] https://doi.org/10.1075/intp.7.1.03pet.

Pöchhacker, F. (2004), *Introducing Interpreting Studies,* Routledge, London.

Pöchhacker, F. (ed.) (2015), *Routledge Encyclopedia of Interpreting Studies*, Routledge, London.

Pöchhacker, F. (2016), *Introducing Interpreting Studies. Second edition*, Routledge, London.

Pym, A. (2008), "On Omission in Simultaneous Interpreting. Risk Analysis of a Hidden Effort", in: G. Hansen, A. Chesterman and H. Gerzymisch-Arbogast (eds.), *Efforts and Models in Interpreting and Translation Research*, John Benjamins, Amsterdam, pp. 83-105.

Seeber, K.G. (2011), "Cognitive Load in Simultaneous Interpreting: Existing Theories – New Models", *Interpreting*, 13(2), pp. 176-204, [on-line:] https://doi.org/10.1075/intp.13.2.02see.

Seeber, K.G., Kerzel, D. (2011), "Cognitive Load in Simultaneous Interpreting: Model Meets Data", *International Journal of Bilingualism*, 16(2), pp. 228-242, [on-line:] https://doi.org/10.1177/1367006911402982.

Setton, R. (1999), *Simultaneous Interpretation: A Cognitive-Pragmatic Analysis*, John Benjamins, Amsterdam.

Setton, R. (2003), "Models of the Interpreting Process", in: Á. Collados Aís and J.A. Sabio Pinilla (eds.), *Avances en la investigación sobre interpretación*, Comares, Granada, pp. 29-89.

Setton, R. (2015), "Models", in: F. Pöchhacker (ed.), *Routledge Encyclopedia of Interpreting Studies*. Routledge, London, pp. 263-267.

Tryuk, M. (2007), *Przekład ustny konferencyjny,* Wydawnictwo PWN, Warszawa.

ABSTRACT

The aim of this study is to find empirical evidence for Gile's Effort Models through retrospection. The main objective is to investigate simultaneous interpreting trainees' perception of the cognitive processes underlying the SI task, their decision-making process when faced with increased cognitive load and processing capacity saturation, as well as their coping mechanisms under the constraints of constant requirement of attentional resources allocation. The study triangulates process analysis (retrospective protocols) with product analysis (manual comparison of source and target texts). The corpus of the study consists of about 75 hours of recordings of 240 interpreting outputs in both directions of interpreting (Polish-English and English-Polish) and the recordings of retrospective protocols (5,005 remarks). The analysed retrospective comments provide evidence of the conflicting efforts and problems with processing capacity management described by Gile. The total of 531

verbalisations refer to the aspects related to Gile's Effort Models and 108 out of 120 interpreters participating in this experiment made at least one remark reporting them, which clearly confirms Gile's observations about the nature of the simultaneous interpreting process.

Key words: Effort Models, cognitive load, retrospective protocols, simultaneous interpreting, process research

Streszczenie

W poszukiwaniu dowodów na istnienie Modeli Wysiłkowych Gile'a w protokołach retrospektywnych tłumaczy symultanicznych

Badania przedstawione w niniejszym artykule mają na celu poszukiwanie dowodów empirycznych na istnienie Modeli Wysiłkowych Gile'a za pomocą narzędzia retrospekcji. Głównym celem jest zbadanie, w jaki sposób studenci tłumaczenia symultanicznego postrzegają procesy kognitywne składające się na ten rodzaj przekładu. Analiza obejmuje także percepcję własnych procesów decyzyjnych w obliczu zwiększonego wysiłku kognitywnego u badanych osób, a także stosowanych przez nie strategii, aby przeciwdziałać ograniczeniom, jakie nakłada konieczność ciągłej koordynacji czynności w przekładzie symultanicznym. Badany jest zarówno produkt przekładu (przez porównanie tekstów wyjściowych z docelowymi), jak i sam proces tłumaczenia symultanicznego (przez analizę protokołów retrospektywnych). Korpus pracy stanowią nagrania i transkrypcje 240 tekstów docelowych wykonanych przez 120 tłumaczy, co liczy łącznie około 75 godzin nagrań. Analizowane protokoły retrospektywne potwierdzają problemy tłumaczy z zarządzaniem procesami przetwarzania informacji oraz konieczność koordynacji procesów opisanych w Modelach Wysiłkowych Gile's. Aż 531 komentarzy retrospektywnych odnosi się do aspektów związanych z Modelami Wysiłkowymi. Tego rodzaju uwagi zostały zwerbalizowane przez 108 ze 120 tłumaczy biorących udział w badaniu. Potwierdza to założenia Gile'a co do specyfiki procesu tłumaczenia symultanicznego.

Słowa kluczowe: Modele Wysiłkowe Gile'a, obciążenie kognitywne, protokoły retrospektywne, tłumaczenie symultaniczne, badania nad procesem przekładu