

Agnieszka Szarkowska, Łukasz Dutka, Olga Pilipczuk, Krzysztof Krejtz

Respeaking crisis points. An exploratory study into critical moments in the respeaking process.

Abstract

In this paper we introduce respeaking crisis points (RCPs), understood as potentially problematic moments in the respeaking process, resulting from the difficulty of the source material and/or cognitive overload on the part of the respeakers. We present results of the respeaking study on Polish participants who respoke four videos intralingually (Polish to Polish) and one interlingually (from English to Polish). We measured the participants' cognitive load with EEG (Emotiv) using two measures: concentration and frustration. By analysing peaks in both EEG measures, we show where respeaking crisis points occurred. Features that triggered RCPs include very slow and very fast speech rate, visual complexity of the material, overlapping speech, numbers and proper names, speaker changes, word play, syntactic complexity, and implied meaning. The results of this study are directly applicable to respeaker training and provide valuable insights into the respeaking process from the cognitive perspective.

Keywords: respeaking; live subtitling; respeaking crisis points; cognitive load; audiovisual translation

1. Introduction

Respeaking, also known as voice-writing, is a method of producing live subtitles using speech recognition software (Lambourne 2006, Romero Fresco 2011). A respeaker listens to the original utterance in a live TV programme and repeats or rephrases the text if necessary, adding punctuation marks. Respeaker's spoken words are automatically turned into written text by speech-to-text software and displayed as subtitles with a delay of a few seconds. The main target audience of respoken subtitles are people who are deaf or hard of hearing, but in reality the group of recipients of subtitling is much more numerous and diverse, and includes for instance foreigners and language learners (Vanderplank 2013) as well as viewers in noisy environments like pubs or train stations.

While the vast majority of live subtitling produced with respeaking is intralingual, that is within the same language, respeaking can also be used to translate between different languages (den Boer 2001). Respeaking can thus be considered a type of either intra- or interlingual translation.

Regardless of the language combination, there are moments in the respeaking process – just like in the traditionally conceived translation – that are more difficult than others. Among the most problematic issues are fast speech rates of original speakers, high information density of the original text, overlapping speech in the case of multiple speakers, syntactic or pronunciation errors as well as poor quality of the transmission (Eugeni 2009, Romero Fresco 2011). Some genres are considered more difficult than others to respeak, for instance sports is thought to be easier than news or chat shows (Romero Fresco 2011). To the best of our knowledge, no previous study, however, has empirically verified these assumptions, which are the result of practical experience of researchers and respeakers rather than of experimental research into cognitive aspects of the respeaking process.

Although live subtitling through respeaking has been used in several countries since 2001, the actual empirical research conducted on respeaking so far has had a largely applied character and has mainly focused on developments in error analysis and accuracy rate of respoken subtitles (see Eugeni 2009, Lambourne 2013, Martínez & Lopez 2013, Romero Fresco 2011, Romero Fresco 2013), the quality of respoken text (Romero Fresco 2009, Romero Fresco & Martínez Pérez 2015, Martínez Pérez 2015), the reduction of text in live respoken subtitles (Luycks 2010, Sandrelli 2013), the reception of respoken subtitles (Eugeni 2008, Romero Fresco 2011) and the training of respeakers (Arumí Ribas & Romero Fresco 2008, Romero Fresco 2012, Eugeni 2008c).

In this paper we experimentally test difficult moments in the respeaking process, focussing on what we term ‘respeaking crisis points’ (RCPs). Our theoretical point of departure is the notion of ‘crisis points’, as defined by Lörscher (1991) and later discussed by a number of translation studies scholars (see Remael & Vercauteren 2010, Pedersen 2015). Translation crisis points can be thought of as stumbling blocks in the translation process. According to Lörscher (1991), ‘translation crisis points’ are “points in translation where the translator has to abandon his/her automated processed and resort to strategic behaviour”. In the same vein as Lörscher (1991), Mazur

(2013) and the ADlab project team (<http://www.adlabproject.eu>) talk about ‘audio description crisis points’, identifying critical moments in audio description and the strategies that could be used by audio describers. When examining crisis points, so far authors tended to focus mainly on the strategies used by translators and audio describers in tackling those difficult moments and trying to overcome the difficulties they pose.

Before we can talk about strategies to deal with difficult moments in respeaking, however, we need to first establish where those difficult moments occur. We therefore examine respeaking crisis points, which we understand as moments in respeaking when the respeaker is faced with a difficulty of some sort. RCPs can be thought of as stumbling blocks, or – as put by Gile (2009, p. 171) – “problem triggers”, forcing the respeaker to expend a larger cognitive effort to perform a task, which in turn may result in a longer pause in the respeaker’s output and/or a drop in the respeaker’s performance. We believe that RCPs may be caused by cognitive overload related to the difficulty of the task as well as to the respeaker’s competences. In what follows we examine two cognitive load indicators during the respeaking process with a view to identifying RCPs.

2. How to measure respeaking crisis points?

We believe that RCPs may result from an excessive cognitive load experienced by respeakers during a respeaking task. Cognitive load, or mental load, as it is sometimes referred to (see e.g. Xie and Salvendy 2000), is “the load imposed on working memory by the cognitive processes” (Antonenko et al. 2010, p. 426). Cognitive load is a complex notion, comprised of many elements of multifaceted nature. In their Subjective Workload Assessment Technique (SWAT), Reid and Nygren (1988) divide the load into three components: time load, mental-effort load and psychological-stress load. In the National Aeronautics and Space Administration Task Load Index (NASA-TLX), Hart and Staveland (1988) distinguish as many as six components of cognitive load: mental demand, physical demand, temporal demand, performance, frustration, and effort. In this

study, we focus on mental effort, operationalised as concentration, and on psychological-stress load, operationalised as frustration.

The notion of cognitive load can be approached from two perspectives: task-oriented, where the load stems from the characteristics of the task, and human-oriented, where cognitive load is conceived of in terms of the effects that performing a task induces on a participant. In this study, we combine both approaches: by analysing participants' concentration and frustration (human orientation), we aim to find which characteristics of the respeaking task were most challenging (task orientation).

Cognitive load can be measured using subjective techniques like rating scales with, and/or objective physiological techniques like electroencephalography (EEG). Rating scales usually rely on the participants assessing the overall cognitive load they experienced when performing a task. However, as argued by Xie and Salvendy (2000), overall load measured through rating scales used in many studies is inadequate as it does not make it possible to study fluctuations in cognitive load over time, i.e. at particular points in time when carrying out the task. Instead of rating scales, according to Antonenko et al. (2010: 425), EEG can be used “as a physiological index that can serve as an online, continuous measure of cognitive load detecting subtle fluctuations in instantaneous load”. In this study, we used EEG to examine peaks and drops in cognitive load over time during different respeaking tasks. Although the analysis below is based on numerical data derived from EEG, our approach is largely qualitative as we are mainly interested in identifying and discussing RCPs from a linguistic and translation perspective.

3. Method

Departing from the hypothesis that respeaking crisis points may be caused by cognitive overload related to the characteristics of the task, we examined two EEG indicators of cognitive load: concentration and frustration. Concentration, also sometimes referred to as ‘engagement’, is defined as “the conscious direction of attention towards task-relevant stimuli” (Emotiv 2014, p.

31). It is detected by increased presence of beta waves in the brain alongside with attenuated alpha waves. “The greater the attention, focus and cognitive workload, the greater the output source reported by the detection” (Emotiv 2014, p. 31). Frustration is defined as a negative feeling of annoyance. All the parameters are measured on a scale from 0 to 1; the higher the score, the more concentrated/frustrated a participant is.

We hypothesised that respeaking crisis points may occur during peaks in concentration and frustration. A peak is here understood as a moment in which the indicator measured (concentration and/or frustration) was larger by at least two standard deviations (SD) from the mean for that person in that clip.

Participants’ brain activity was measured with the frequency of 128 Hz (every 8 ms). We divided the results into 10-second time intervals and looked whether there were any peaks in each interval. If at least one peak was found for at least 10 participants in that interval, we manually examined that moment in the clip in terms of its textual, audio and visual content with a view to finding the cause for the peak, and by extension, for the respeaking crisis point.

3.1. Participants

A total of 57 people (50 women, 7 men) participated in the study. Their mean age was 27.48 (SD=5.71), ranging from 21 to 51.

Because at the time this study was conducted, respeaking in Poland was still in infancy, there were no professional respeakers we could recruit. We therefore trained the participants during two-day workshops carried out by professional respeaker trainers from Switzerland, UK, and Italy. Participants were divided into three groups: professional interpreters and interpreting trainees (N=22), translators and translation trainees (N=23) and people with no experience in either translation or interpreting (N=12), but fluent in English.

Due to incomplete or missing EEG data in some clips and participants (no variation recorded in at least one of the measured indices), the final sample analysed in this paper was: 11 subjects in the interpreter group, 11 subjects in the translator group and 7 participants in the control group.

3.2. Procedure

Participants were tested individually in a research lab (see Fig. 1). They were seated at a distance of about 60 cm from the monitor. Their brain activity was monitored using Emotiv EPOC EEG headset with the frequency of 128 Hz. Another laptop with a custom-built software (Prompter) was used to record the two synchronised audio channels (the original speech and the respeaker's output).

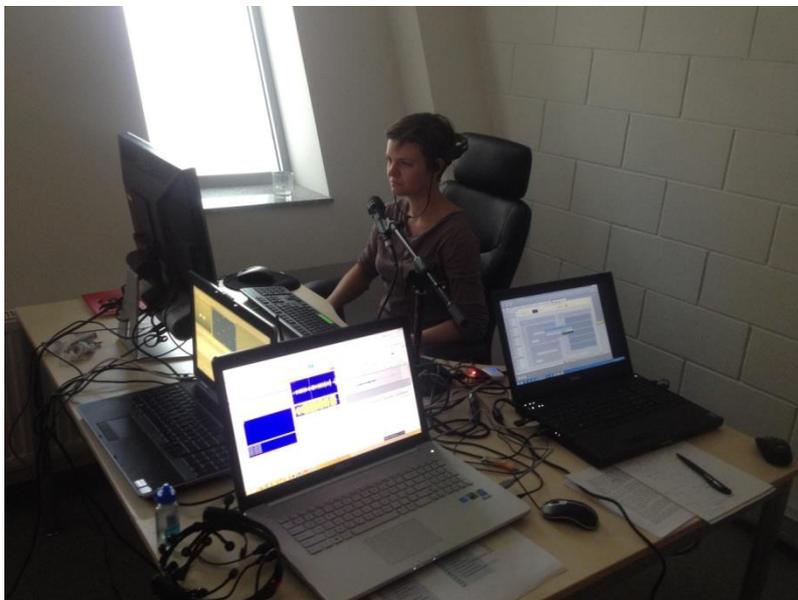


Fig. 1. Experimental set-up in the respeaking test.

The speech recognition software used in this study was Newton Dictate for the Polish language (version 4.0) manufactured by Newton Technologies. Participants created their own voice profiles during the respeaking workshops, and these profiles were later used in the respeaking test.

Prior to commencing the study, all participants signed informed consent forms. The respeaking test began with the procedure being explained to the participants. Then, EEG calibration was performed. After the calibration, participants went through a mock respeaking task to

familiarise themselves with the procedure (the data for this part were not recorded). In the main part of the experiment, participants were asked to respeak intralingually four Polish videos, whose order was randomised, and finally, to respeak one video interlingually from English to Polish. After the respeaking tasks, participants were asked about how they assess particular tasks in a semi-structured interview.

3.3. Materials

5 clips of different genres from Polish TV channels were used in the study (see Table 1): news (*Fakty* evening news broadcast on private television channel TVN, reporting on miners' protests, fast-paced with a number of speakers, but no overlapping speech), political chat show (an excerpt from *Kropka nad i* on TVN, fast-paced with numerous cases of overlapping speech), entertainment chat show (*Fakty po Faktach* on TVN24 with an actress and a film critic discussing the movie *Ida* after it was awarded an Oscar for best foreign language film, medium-paced) and two speeches, a New Year's address by Prime Minister Ewa Kopacz televised by public television TVP1 and an excerpt from President Obama's speech at the Castle Square in Warsaw celebrating 25th anniversary of free elections in Poland, a broadcast by public service broadcaster TVP1. Both speeches were slow-paced with one speaker and no overlapping speech. Each video to be respoken lasted about 5 minutes. Four clips were in Polish and President Obama's speech was in English.

Table 1. Characteristics of the clips used in the study

	Duration	Number of words	Speech rate	Number of sentences
Intralingual				
Speech	4:47	520	108 wpm	40
News	6:10	935	152 wpm	99
Entertainment show	5:20	707	133 wpm	74
Political chat show	5:35	916	165 wpm	107
Interlingual				
Speech	4:55	458	94 wpm	35

4. Results and discussion

4.1. Overall results

Table 2 presents the number of peaks in concentration and frustration in each clip. The highest number of peaks was found in the fast intralingual clips (news and political chat show), while the clip with the lowest number of peaks was the interlingual one, which was the slowest of all the clips in terms of speech rate (see Table 1). This may suggest an importance of the speech rate in respeaking.

Table 2. Number of peaks per clip

	Frustration	Concentration
1-speaker slow (speech)	118	123
Many speakers slow (entertainment chat show)	148	144
1-speaker fast (news)	140	205
Many speakers fast (political chat show)	206	207

Generally, peaks in frustration and concentration frequently occurred side by side. The only exception being the news, which required more concentration without inducing frustration in many participants. Highest frustration was found in the fast clip with overlapping speech.

4.2. RCPs in intralingual respeaking

Below we report on respeaking crisis points as measured by peaks in concentration and frustration in four different genres in intralingual Polish-to-Polish respeaking.

In all clips, frequent peaks were found at the beginning of the videos. According to Emotiv (2014, p. 31), events which “result in a peak in the detection are difficult tasks requiring concentration, discovering something new, and entering a new area”. This is in line with our findings, where participants experienced higher concentration when embarking on a new task.

4.2.1. Speech

This video was a speech, delivered by former Polish Prime Minister Ewa Kopacz at the end of 2014. As is usually done on such occasions, the Prime Minister was reading out a pre-scripted text from a prompter. There were no camera movements; the figure of the PM was in the centre of the screen in medium and medium close-up shots. This type of speech is characterised by high lexical density, low redundancy with few oral features, and highly structured content, which makes it difficult to respeak and to edit information out without losing important meaning.

Proper names and numbers. Concentration peaks were found in fragments containing proper names and numbers, such as dates (see Example 1 below). Increased occurrence of peaks may be an indication of a higher demand on working memory, and in consequence, of higher cognitive load experienced by respeakers.

- 1) W dwa tysiące czternastym roku podczas uroczystych obchodów siedemdziesiątej piątej rocznicy wybuchu drugiej wojny światowej i siedemdziesiątej rocznicy wybuchu powstania warszawskiego oraz bitwy o Monte Casino, oddaliśmy hołd odwadze i bohaterstwu Polaków.

In two thousand fourteen, during ceremonial celebrations of the seventy fifth anniversary of the outbreak of World War Two and the seventieth anniversary of the outbreak of the Warsaw Uprising and the Battle of Monte Casino, we paid tribute to the courage and heroism of Poles.

Numbers and proper names are sometimes referred to as ‘non-contextual information’, i.e. information that cannot be inferred or predicted from the context (Chmiel 2015, p. 235). They are usually elements of low redundancy (Gile 1984), low predictability (Braun and Clarici 1996) and high informative content (Alessandrini 1990). As stated by Jones (1998, p. 130), “numbers have an objective meaning and are in no way open to linguistic interpretation”. As suggested in interpreting literature (Mazza 2001, p. 91, Braun and Clarici 1996), one possible way of releasing the load on working memory is to note the numbers down. In our study, we did not allow the participants to make notes, as it is not typically done in respeaking. Another method the participants could use to tackle increased cognitive load, as suggested by Mazza (2001), was to reduce *décalage*, or ear-voice span (EVS). In other words, they could unload their working memory by saying the numbers as soon as possible after the speaker (Jones 2002). While this method enables shorter

retention of a unit of information in working memory, it also requires faster processing and production. Another strategy, frequently used in interpreting, is to generalise the number and focus on the order of magnitude rather than details (Jones 2002; Shlesinger 2000). Allocating too many attentional resources to processing numbers, as noted by Shlesinger (2000), may result in the spill-over effect, whereby the interpreter, or the respeaker, will be forced to lose some information in the chunk of meaning that follows. This was indeed sometimes the case in our study, where some participants got lost in the dates and lost track of the following parts of the sentence.

Implied meaning. Concentration peaks were also found when the original speaker used some implied meaning, as shown in Example 2:

- 2) Znaczenie Polski na arenie międzynarodowej zostało także potwierdzone nominacją naszego rodaka na jedną z najważniejszych funkcji w Unii Europejskiej.

The importance of Poland on the international scene was also confirmed by the nomination of our compatriot on one of the most important posts in the European Union.

In this utterance, a reference is made to Donald Tusk, President of the European Council. The speaker used the words ‘our compatriot’ instead of Donald Tusk, and ‘one of the most important functions of the EU’ instead of stating explicitly what the function is. While respeaking, some participants employed the explicitation strategy (see Klaudy 2009) by replacing the implied terms with the name ‘Donald Tusk’ and ‘the President of the European Council’ or ‘the EU President’, respectively. The processing necessary to perform this operation probably required expending higher mental effort (concentration), as shown by the peaks in EEG, and allows us to classify such situations as RCPs.

Long and complex sentences. As opposed to impromptu utterances, a typical feature of pre-scripted speech are long sentences with embedded clauses, some of which raised problems in respeaking, as shown by peaks in concentration in the example below:

- 3) Proces integracji europejskiej i obecność Polski w NATO pomogły w wykorzystaniu wielkiego potencjału naszego kraju, który od lat kroczy drogą szybkiej modernizacji i ciągłego rozwoju, dając przy tym obywatelom poczucie bezpieczeństwa.

The process of European integration and the presence of Poland in NATO helped to make use of the great potential of our country, which has been following the path of swift modernization and continuous growth for years, granting its citizens with a sense of safety.

The high propositional density of such sentences, together with its complex syntactic structure and short duration of pauses in the delivery may result in a respeaking crisis point. A possible strategy to deal with a RCP of this type is, as suggested by Romero Fresco (2011, p. 110), to use the salami technique (see also Jones 2002): respeakers can ‘chop’ and ‘slice up’ original sentences, making them shorter and simpler. According to Jones (2002, p. 104), “the salami technique can help with such Russian doll-like sentences with subordinate clauses”. It is also a good solution from the point of view of live subtitling – the end product of respeaking – given that long sentences are difficult to read, particularly considering the lack of perfect subtitle-to-image synchrony in live subtitling.

4.2.2. News

News programmes are characterised by high speech rates, high lexical density, highly structured content, and low redundancy. As such, they are quite demanding for respeakers. Unlike in the case of the speech discussed above, there are more speakers in the news programme than just the main news presenter (though not speaking at the same time): live reporter, interviewees, etc. The lines spoken by journalists are usually accompanied by supporting images and, sometimes very loosely connected, footage (see Romero Fresco 2011, p. 126), which makes news programmes more complex and cognitively demanding to respeak than speech.

Visual complexity. Among a number of issues that triggered an increased occurrence of concentration peaks, which we interpret are RCPs, were moments when complex visual material was presented in the programme, such as information graphics or split screen (see Fig. 2).



Fig. 2. Example of split screen in the news programme

This was the case with a report on negotiations related to a crisis in the Polish mining industry. At one point, apart from the national government officials talking with miners, another meeting was held between local government officials from regions where the protests were taking place. The fact that there were two parallel meetings and the issue who spoke with whom, combined with the split screen showing footage from the two meetings, turned out to require increased concentration in many participants and can be classified as a respeaking crisis point:

- 4) Co prawda nie było trzaskania drzwiami i żadna ze stron nie zerwała rozmów, ale trudno mówić o jakimkolwiek postępie. Bo premier wspierana przez swoich ministrów przysłała z gotową propozycją i zaczęła negocjacje od tego, że jej nie zmieni. Jak widać, w pewnym momencie były już nawet dwa równoległe spotkania, bo oprócz spotkania górników z rządem rozmawiali samorządowcy, prezydenci i burmistrzowie z rejonów, gdzie dziś protestują górnicy pracujący jeszcze w szkole w kopalniach z rządowej listy kopalń nierentownych.

Although there was no door slamming and none of the parties decided to break the talks, it's hard to talk of any progress because the prime minister accompanied by her ministers came to the meeting with a ready-made proposal and started the negotiations by saying that she wouldn't change it. As you can see, at some point there were even two simultaneous meetings, as, apart from the meeting between the miners and the government, talks were also held between the local government politicians and mayors of cities and towns from the areas where the miners who still work in companies which are on government's list of unprofitable mines are protesting today.

This points to a possible increase in cognitive load, as the incoming input was becoming more complex, respeakers had to concentrate more to make sense of the multiple sources of information they were exposed to.

Speaker changes and change in delivery. A number of concentration peaks were found in the moment of speaker changes, especially with speakers other than the news anchor, as illustrated by

Ex. 5 and 6. RCPs may have also been also related to a poorer quality of sound and different accents of people being interviewed by the reporter.

- 5) U nas w biurze padł pomysł. Choinka, spróbujemy tę kopalnię wziąć sami. W pięć osób założyliśmy firmę pod nazwą Przedsiębiorstwo Górnicze Silesia. Nie mieliśmy żadnego doświadczenia.

At our office, we had this idea. It's Christmas, let's try to run this mine ourselves. There were five of us and we set up a company called Przedsiębiorstwo Górnicze Silesia [Silesia mining enterprise]. We didn't have any experience.

- 6) Na początku nie miał kto dla nas pracować. Ściągaliśmy ludzi z firm górniczych. Ściągaliśmy ludzi z ulicy. I ci ludzie, którzy tu zostali, namawiali na przyjście do pracy, na obdarzenie zaufaniem. Teraz sytuacja jest taka, że to ludzie przychodzą do nas do pracy, a my już mamy komplet.

At first, we didn't have enough people who could work for us. We were hiring people from other mining companies. We were hiring people from the street. And the people who stayed with us persuaded others to come to work for us, to place their trust in us. Now the situation is different. There're are people who come and ask about jobs but we already have a full set of staff.

Such impromptu speech constituted a change in information density, redundancy, speech rate, and sound quality from the main news presenter. Talking about the difficulties of a news report given by a correspondent live, Romero Fresco (2011, p. 128), explains that “it has not been prepared in advance [...], so features of higher grammatical intricacy (with more features of orality), lower lexical density (the content being less condensed), a slower speech rate and often poor quality of transmission. The challenge here is to understand the source text and turn its oral nature into written subtitles”.

Numbers. Similarly to speech, peaks were also found in the case of numbers in the news programme, as shown in examples below.

- 7) Kopalnia pracuje dwadzieścia cztery godziny na dobę, siedem dni w tygodniu.

The mine operates twenty four hours a day, seven days a week.

- 8) Nie mamy co prawda czternastki, ona jest uzależniona od wyniku finansowego.

Even though we don't get a fourteenth salary, which depends on the financial results.

Info graphics with numbers. Numbers were not only used in the original spoken text, but sometimes they were also present in information graphics on the screen, as was the case with Ex.

- 9) Ewa Kopacz musi się także liczyć z tym, że zdecydowanie większe poparcie w tym sporze mają górnicy niż jej rząd. W sondażu dla Faktów i TVN24 przytłaczająca większość, 68 procent badanych, stoi po stronie protestujących. Tylko 15 procent popiera tu Ewę Kopacz.

Ewa Kopacz must take into account that the miners have a far greater social support than her government. According to a poll conducted for Fakty and TVN24, the overwhelming majority, 68 per cent of the respondents, backs up the protesters. Only 15 per cent support Ewa Kopacz.

Another factor, which was mentioned by some of the participants later in the post-test semi-structured interview, was the fact that the info graphic differed slightly (68.5%) from the text that was read by the news presenter (68%) (see Fig. 3).



Fig. 3. Info graphics with numbers from the news programme

Here, the two channels of information: visual and auditory, instead of being complementary and redundant, may have been the source of slight confusion. Such situations can be classified as respeaking crisis points, as they triggered peaks in both concentration and in frustration. One possible strategy to deal with numbers is to use approximation, i.e. rounding the number up or down, respecting the order of magnitude (Mazza 2001). Indeed, this is what we observed here, as most participants decided to respeak 68 per cent, and not 68.5 per cent.

Syntactic complexity. We found both concentration and frustration peaks in the following fragment featuring journalist off-screen narration illustrated with footage of coal train (not much related).

- 10) I właśnie wtedy [adverbial of time], zupełnie jak w filmie [adverbial of manner], garstka ludzi niemająca niczego, a już na pewno niczego do stracenia [subject complement], postanowiła wziąć sprawę w swoje ręce.

And just then [adverbial of time], as if in a movie [adverbial of manner], a bunch of people who had nothing, and certainly nothing to lose [subject complement], decided to take the matter into their own hands.

A closer syntactic analysis of this sentence shows its atypical grammatical structure. The sentence contains as many as three grammatical inserts, i.e. adverbials of time and manner as well as subject complement. As noted by Gile (2009, p. 296), “embedded structures, in particular, seem to impose increased pressure on the comprehender”. They may disrupt the flow of the sentence, adding extra information. At the same time, owing to space and time limitations, these elements are typically omitted in subtitling. Aware of those limitations, respeakers may have noticed them and strove to remove them.

4.2.3. Entertainment programme

This programme was a fragment of *Fakty po faktach*, an interview between an actress Grażyna Szapołowska and a film critic Tomasz Raczek, following the 2015 Oscars ceremony where a Polish film *Ida* won in the best foreign language film. The show featured live slow-rate spontaneous dialogue, which was unscripted, and contained typical features of spoken language like hesitations, false starts, repetitions, and a higher level of redundancy compared to speeches and news programmes.

Proper names. Similarly to previous tasks, we found concentration peaks in the case of proper names (Ex. 11-12).

- 11) Fakty po Faktach. Justyna Pochanke, witam państwa gorąco. Grażyna Szapołowska, Tomasz Raczek, czyli goście warci Oscara.

Fakty po Faktach [Facts after Facts, a name of a current affairs chat show] Justyna Pochanke, a warm welcome to all of you. Grażyna Szapołowska, Tomasz Raczek, guests worthy of an Oscar.

- 12) Musiał przyjechać facet z Oksfordu, żebyśmy dostali tego Oscara", tak powiedziała Agnieszka Holland o Pawle Pawlikowskim.

"A guy from Oxford had to come so that we could get this Oscar," that's what Agnieszka Holland said about Paweł Pawlikowski.

Spoken language. Another area, typical of unscripted speech, where we found peaks were elements of spoken language, particularly false starts and reformulations, as shown in Ex. 13-14.

- 13) Tak, tak, ale pewnie ten jego taki... ta jego światowość wypływa stąd, że spędził większość życia na Zachodzie.

Right, right, but probably his... this worldliness of his comes from the fact that he spent most of his life in the West.

- 14) Zastanawiam się, wspomniął pan, panie Tomaszu, ten deszcz nagród... Historia "Idy" to była historia biegu z przeszkodami tak naprawdę.

I wonder, Tomasz, you mentioned this rain of awards... The story of "Ida" is really a story of an obstacle race.

Here, the respeaking problem consisted in a dilemma which strategy to choose: either follow the speaker closely and fall into the short décalage trap, by repeating the unfinished sentences, or wait longer for a speaker to complete a sentence, thus increasing the burden on working memory and the ear-voice span. Given that the end product of respeaking is subtitling, where unnecessary elements of spoken language with no propositional value tend to be omitted (see Díaz Cintas & Remael 2007), it seems that a better strategy would be to increase the EVS and wait for the speaker to complete their utterance.

Word play. Another fragment which induced peaks in frustration and concentration, and may therefore be classified as a potential RCP, was an instance of word play (Ex. 15). When discussing the success of *Ida* directed by Paweł Pawlikowski, the speakers emphasised that the director is not considered to be Polish, as he spent many years in the UK, where he had already established his position as an independent film-maker.

- 15) - Łatwiej jest nakręcić oscarowy film, niż z nim się przebić w Stanach. Jak się przebili?

- Ja myślę, że on nie jest traktowany tak do końca jako reżyser polski. Paweł Pawlikowski jest już przebity. On się przebił jako twórca brytyjski.

- It is easier to make a movie worthy of an Oscar than it is to actually get it noticed in the [United] States. How did they get noticed?

- I think he is not considered to be a Polish director. Paweł Pawlikowski has already been noticed. He was noticed as a British filmmaker.

The speakers used various grammatical forms of the word 'przebijać się' ('to make it', 'to be noticed'), one of which was a neologism past participle ('przebity'). It is possible that the respeakers had to expend extra processing effort to be able to repeat all the instances of this word, as substituting it with a synonym or paraphrasing it would probably not work best in this case.

4.2.4. Political chat show

The video was a fragment of a political chat show *Kropka nad i* by a famous journalist Monika Olejnik, where she interviewed with two politicians: left-wing Kazimiera Szczuka of *Your Movement* and right-wing Przemysław Wipler from *The Congress of the New Right*. As declared by the participants in the post-test interview, this video was the most difficult one to respeak owing to overlapping speech and fast speech rates. It also had the highest number of peaks among all of the videos in the test.

Proper names. Again, proper names turned out to be problematic, as they also triggered peaks in frustration (Ex. 16-17):

- 16) Jeżeli pan będzie wyciągał te stare rzeczy, to mogę powiedzieć, że dzisiaj na listach Twojego Ruchu jest Anna Grodzka¹, Robert Biedroń².

If you keep bringing up those old things, then I can say that today Twój Ruch [a political party] has Anna Grodzka and Robert Biedroń on its lists.

- 17) To tyle Korwin-Mikke³. I o Hitlerze, i o gwałcie. Kazimiera Szczuka.

That's what Korwin-Mikke said. About Hitler and about rape. Kazimiera Szczuka.

As noted by some participants in the post-test interview, some people had problems with the names, as they were not interested in politics, and were not familiar with the names or events mentioned in the programme. In the professional context – not a research study –, respeakers would have to prepare beforehand, so this problem could at least partially be eliminated.

Reactions to the content. It is also possible that some peaks in frustration might have been triggered by the sheer content of the discussion, which related to one politician's highly controversial view on rape:

- 18) - Czy pan uważa, że kobiety chcą być gwałcone? Dobrze zrozumiałam?

- Co znaczy gwałcone? Co znaczy gwałcone? Kobiety zawsze udają, że pewien opór stawiają. To jest normalne chyba, nie?

- Co to jest według pana gwałt?

- Pokonywanie pewnego... bo trzeba wiedzieć kiedy można, kiedy nie można.

¹ Anna Grodzka, born as Krzysztof Bęgowski, was the first openly transgender woman elected Member of the Polish Parliament in 2011.

² Robert Biedroń is a LGBT activist and the first openly homosexual Member of the Polish Parliament, elected in 2011. He is now the mayor Słupsk, a town in northern Poland.

³ Korwin Mikke, a Polish politician and Member of the European Parliament, famous for his controversial views on women, the disabled, the Holocaust, and the EU, among others.

- *Do you think that women want to be raped? Have I understood correctly?*
- *What does it mean «raped»? What does it mean «raped»? Women always pretend some resistance. That's normal, right?*
- *What is a rape according to you?*
- *It's overcoming some... because you need to know when you can and when you can't.*

In the interview, some participants declared their lack of appreciation for this political figure, which – together with the controversial content of the video – may have affected their EEG frustration patterns.

4.3. RCPs in interlingual respeaking

In order to find RCPs, we also examined one clip with interlingual English-to-Polish respeaking. The clip was a fragment of US President Barack Obama speaking at the 25th Anniversary of the Freedom Day in Warsaw in June 2014. In his speech⁴, he made frequent references to Poland, its culture, history, historic figures as well as to the European situation back in 1989 after the fall of the Berlin Wall and the collapse of the communist regimes across Eastern Europe.

Similarly to intralingual respeaking, we found frequent peaks at the very beginning of the clip, particularly in concentration. Increased concentration was also found in the case of Polish words pronounced by Obama with a strong American accent, as shown in Ex. 19-21:

- 19) The faithful come together at churches like Saint Stanislaus Kostka.
- 20) And every summer, we celebrate the Taste of Polonia, with our kielbasa and pierogies, and we're all a little bit Polish for that day.
- 21) The Righteous Among the Nations, among them Jan Karski, who risked all to save the innocent from the Holocaust.

Interestingly, frustration peaks were triggered when the speaker made long pauses in long sentences (Ex. 22-23), resulting in very low speech rates (roughly 90-100 wpm).

- 22) [00:33] Distinguished guests, [00:36] people of Poland, [00:40] thank you for your extraordinary welcome ... and for the privilege of joining you here today... [00:46] I bring with me the greetings and

⁴ The entire text of the speech can be accessed here: <https://www.whitehouse.gov/the-press-office/2014/06/04/remarks-president-obama-25th-anniversary-freedom-day>

friendship of the American people... [00:53] and of my hometown of Chicago... [00:55] home to so many proud Polish Americans... [01:01]

- 23) [02:01] The Communist regime thought an election would validate their rule... [02:06] or weaken the opposition... [02:10] Instead, Poles turned out in the millions... [02:14] And when the votes were counted... it was a landslide victory for freedom.[02:22]

This goes against the commonly held assumption that RCPs can be expected in material with a high speech rate. As noted by Alpbach with reference to interpreting, “slow input can disrupt processing as much as fast input” (1968, p. 2, cited after Pöchhacker 2004, p. 129). When talking about slow delivery of speeches, Gile (2009) also argues that “if it is too slow, information elements have to be kept longer in short-term memory before they can be integrated into target speech sentences, which may cause cognitive saturation” (p. 200). In this case, the very slow speech rate of this clip was probably causing a lot of strain to the participants’ working memory, which in turn could increase the cognitive load at this point in respeaking. This also shows another area of convergence between interpreting and respeaking.

Frustration peaks were also found in the case of historical and geographical references (Ex.

24). Some of them were not named directly, but only implied:

- 24) The victory of 1989 was not inevitable. It was the culmination of centuries of Polish struggle, at times in this very square. The generations of Poles who rose up and finally won independence.
- 25) The images of that year are seared in our memory.

First, as stated in the post-test semi-structured interview, not all the participants managed to understand the sentence in Ex. 24 and 25. Others had problems with the idiomatic expression ‘seared in our memory’ and struggled to find an appropriate phrasing in Polish, which increased the EVS and the load on the working memory, possibly resulting in cognitive load. The actual year that is being referred to in this sentence (1989) was not mentioned by Obama before in his speech. In the previous paragraph, he said: “twenty-five years ago”, which was the only indication of the time period, apart from the participants’ general knowledge. This may also have constituted another stumbling block for the participants. Those who did not manage to grasp the idea behind this

sentence struggled with the next few sentences, as President Obama went on to give a few examples of the images from that year:

- 26) Citizens filling the streets of Budapest and Bucharest. Hungarians and Austrians cutting the barbed wire border. Protesters joining hands across the Baltics. Czechs and Slovaks in their Velvet Revolution. East Berliners climbing atop that wall.

These images can also be considered RCPs, as we registered peaks in frustration in EEG activity.

Some of the peaks may have stemmed from difficulties in translation. Since the participants did not know what material they were going to respeak, they could not prepare in advance for this task, as interpreters/translators would normally do. In the semi-structured interview after the respeaking test, a number of participants mentioned problems with finding the right equivalents to some English terms, mostly proper names, such as *the Righteous Among the Nations*, *the Velvet Revolution*, *the Baltics*.

Another problematic issue to translate was a delayed referent in Ex. 27:

- 27) Mister President, Mister Prime Minister, Madam Mayor, heads of state and government, past and present, including the man who jumped that shipyard wall to lead a strike that became a movement, the prisoner turned president who transformed this nation, thank you, Lech Walesa, for your outstanding leadership.

President Obama enumerated a few facts related to Lech Wałęsa, like jumping the shipyard wall, leading a strike, being prisoner and later becoming president, without actually stating his name until the end of the sentence. This long and syntactically complex sentence – combined with Obama’s pronunciation of the name *Lech Wałęsa* – turned out quite difficult for the participants to respeak, which was also demonstrated by an increased the EVS between the original utterance and the respeaker’s output.

5. Conclusion

In this paper we proposed the notion of respeaking crisis points and presented a number of examples where they can be found, based on the results of qualitative analysis of cognitive load based on EEG data.

We hope to have shown that EEG can be used to find peaks in concentration and frustration during the respeaking process, and as such, can be a useful tool to detect RCPs. Apart from examining peaks in EEG as cognitive load indicators, other measures that could potentially be used as indicators of RCPs are length of pauses in respeaker's output, fixation duration, and performance indicators.

Knowing where RCPs occur is important to understand the process of respeaking from a cognitive perspective. The results of this study can also inform respeaker training. Once we know what the crisis points are, we can examine the strategies which can be adopted to deal with them and have research-based evidence to inform respeaker training in terms of strategies to deal with RCPs.

In this study, we found a number of areas where RCPs can occur. One crucial factor conducive to RCP is the pace of the original dialogue: both very slow and very fast speech rate. Another factor is the number of speakers and overlapping speech – the more speakers talking simultaneously, the more difficult respeaking becomes, and the more frustration may be experienced by a respeaker. Other problematic issues include numbers and proper names, complex syntactical structures with embedded clauses, visual complexity of the material, speaker changes, word play and implied meaning.

It needs to be stressed that different people may experience different cognitive load in the same task. Training and experience can lead to developing expertise, which helps in reducing the load and improving the performance.

It is important to note that respeaking crisis points do not necessarily have to lead to actual problems or a drop in respeaking performance. Increased cognitive effort, for instance in concentration, may result in a respeaker continuing his/her performance on the same level, adopting a strategy s/he finds helpful at this point. One of the next steps in the respeaking process research could be to match RCPs with online performance indicators.

Acknowledgements

This work was supported by the National Science Centre Poland under Grant no. 2013/11/B/HS2/02762. Many thanks to Ewa Nowik-Dziewicka for her linguistic insights, to Agnieszka Chmiel and to Katarzyna Stachowiak for their help with interpreting sources.

References

- Alessandrini M.S. (1990). Translating numbers in consecutive interpretation: an experimental study. *The Interpreters' Newsletter* 3, 77-80.
- Alpbach (1968). Report on Session "Interpreters and Interpreting". European Forum Alpbach, 19-24 August 1968, unpublished typescript.
- Antonenko, P., F. Paas, R. Grabner and T. van Gog (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review* 22, 425-438.
- Arumí Ribas, M. & P. Romero Fresco (2008). A Practical Proposal for the Training of Respeakers. *The Journal of Specialised Translation* 10, 106-127.
- Braun S. and Clarici A. (1996). Inaccuracy for numerals in simultaneous interpretation: neurolinguistic and neuropsychological perspectives. *The Interpreters' Newsletter* 7, 85-102.
- Chmiel, A. Przetwarzanie w tłumaczeniu symultanicznym ['Processing in simultaneous interpreting']. Chapter 11. In A. Chmiel and P. Janikowski (eds.) *Dydaktyka tłumaczenia ustnego* ['Interpreting training'] (pp. 227- 247). Katowice: Stowarzyszenie Inicjatyw Wydawniczych.
- den Boer, C. (2001). Live interlingual subtitling. In Y. Gambier & H. Gottlieb (Eds.) (*Multi*)*media translation: concepts, practices and research* (pp. 167-172). Amsterdam/Philadelphia, Benjamins.
- Díaz Cintas, J. & A. Remael (2007). *Audiovisual translation: subtitling*. Manchester: St Jerome Publishing.
- Emotiv (2014). Emotiv EPOC User Manual.
- Eugeni, C. (2007). Il rispeakeraggio televisivo per sordi. Per una sottotitolazione mirata del TG. *Intralinea* 9.
http://www.intralinea.org/archive/article/Il_rispeakeraggio_televisivo_per_sordi
- Eugeni, C. (2008). Respeaking the TV for the Deaf: For a Real Special Needs-Oriented Subtitling. *Studies in English Language and Literature* 21, 37-47.
- Eugeni, C. (2009). Respeaking the BBC News. A strategic analysis of Respeaking on the BBC. *The Sign Language Translator and Interpreter* 3(1), 29-68.
- Gile, D. (1984). Des difficultés de la transmission informationnelle en interprétation simultanée. *Babel* 28,18-25.
- Gile, D. (2009). *Basic Concepts and Models for Interpreter and Translator Training*. Amsterdam/New York: John Benjamins.
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.), *Human Mental Workload* (pp. 139-183). Amsterdam: Elsevier Science.
- Jones, R. (2002). *Conference interpreting explained*. Manchester: St. Jerome Publishing.
- Klaudy, K. (2009). Explicitation. In M. Baker and G. Saldanha (Eds.) *Routledge Encyclopedia of Translation Studies*. 2nd edition (pp. 104-108). London & New York: Routledge.

- Lambourne, A. (2006). Subtitle respeaking. A new skill for a new age. *Intralinea* 8. http://www.intralinea.it/specials_old/respeaking/eng_more.php?id=447_0_41_0_M
- Lambourne, A. (2013). How to identify and fix the causes of transcription errors. Paper presented at the *4th International Symposium on Live Subtitling* in Barcelona, March 2013.
- Lörscher, W. (1991). *Translation Performance, Translation Process, and Translation Strategies. A Psycholinguistic Investigation*. Tübingen: Narr.
- Luyckx, B., Delbeke, T., Van Waes, L., Leijten, M., Remael, A. (2010). Live subtitling with speech recognition. Causes and consequences of text reduction. *Across Languages and Cultures* 14 (1), 15-46. DOI: <http://dx.doi.org/10.1556/Acr.14.2013.1.2>
- Martínez Pérez, J. & O. Lopes (2013). Tool for Assessing Accuracy Rate of Live Subtitles. Paper presented at the *4th International Symposium on Live Subtitling* in Barcelona, March 2013.
- Martínez Pérez, J. (2015). New Approaches to Improve the Quality of Live Subtitling on TV. Paper delivered at the Respeaking, Live Subtitling and Accessibility Conference, Rome, 12 June 2015.
- Mazur (2013). Audio Description Crisis Points: Insights from the ADLAB project. Paper presented at the *Media For All* conference in Dubrownik, 25-27 September 2013.
- Mazur I. (2014). Projekt ADLAB i funkcjonalizm w przekładzie – w stronę strategii audiodeskrypcyjnych. *Przekładaniec* 28: 11-22.
- Mazza, C. (2001). Numbers in simultaneous interpretation. *The Interpreters' Newsletter* 11, 87-104. Trieste, EUT Edizioni Università di Trieste <http://www.openstarts.units.it/dspace/bitstream/10077/2450/1/05.pdf>
- Pedersen, J. (2015). On the subtitling of visualized metaphors. *Journal of Specialized Translation* 23, 162-180.
- Pöschhacker, F. (2004). *Introducing Interpreting Studies*. London & New York: Routledge.
- Reid, G. B., & Nygren, T. E. (1988). Subjective workload assessment technique: A scaling procedure for measuring mental workload. In P. A. Hancock and N. Meshkati (Eds.), *Human Mental Workload* (pp. 185–217). Amsterdam: Elsevier.
- Remael, A. & Vercauteren, G. (2010). The translation of recorded audiodescription from English into Dutch. *Perspectives. Studies in Translatology* 18(3), 155-171.
- Romero Fresco, P. (2009). Quality in *Respeaking: The Reception of Respoken Subtitles*. Presentation delivered at the *Media for All 3* conference, Antwerp, 22-24 October 2009.
- Romero-Fresco, P. (2012). Respeaking in Translator Training Curricula. Present and Future Prospects. *The Interpreter and Translator Trainer* 6(1), 91-112.
- Romero-Fresco, P. (2011). *Subtitling through Speech Recognition: Respeaking*. Manchester: St Jerome.
- Romero-Fresco, P. (2013). Quality and NER in the UK. Paper presented at the *4th International Symposium on Live Subtitling* in Barcelona, March 2013.
- Romero-Fresco, P. Martínez Pérez, J. (2015). Accuracy rate in live subtitling: the NER Model. In Díaz Cintas, J., Baños Pinero, R. (eds) *Audiovisual translation in a global context. Mapping an Ever-changing Landscape* (pp. 28-50). Palgrave Macmillan, London.
- Sandrelli, A. (2013). Reduction strategies and accuracy rate in live subtitling of weather forecasts: a case study. Paper presented at the *4th International Symposium on Live Subtitling* in Barcelona, March 2013.
- Shlesinger, M. (2000). *Strategic allocation of working memory and other attentional resources in simultaneous interpreting*. Unpublished doctoral thesis.
- Vanderplank, R. (2016). Effects of and effect with captions: How exactly does watching a TV programme with same-language subtitles make a difference to language learners. *Language Teaching* 49 (02), 235-250. doi: 10.1017/S0261444813000207.
- Xie, B., & Salvendy, G. (2000). Review and reappraisal of modeling and predicting mental workload in single- and multi-task environments. *Work and Stress*, 14, 74–99.

