

Predictive processes during simultaneous interpreting from German into English

Ena Hodzik and John N. Williams

Notre Dame University–Louaize / University of Cambridge

Abstract

We report a study on prediction in shadowing and simultaneous interpreting (SI), both considered as forms of real-time, ‘online’ spoken language processing. The study comprised two experiments, focusing on: (i) shadowing of German head-final sentences by 20 advanced students of German, all native speakers of English; (ii) SI of the same sentences into English head-initial sentences by 22 advanced students of German, again native English speakers, and also by 11 trainee and practising interpreters. Latency times for input and production of the target verbs were measured. Drawing on studies of prediction in English-language reading production, we examined two cues to prediction in both experiments: contextual constraints (semantic cues in the context) and transitional probability (the statistical likelihood of words occurring together in the language concerned). While context affected prediction during both shadowing and SI, transitional probability appeared to favour prediction during shadowing but not during SI. This suggests that the two cues operate on different levels of language processing in SI.

Keywords: anticipation, prediction, latency, simultaneous interpreting, language processing

Introduction

The present study looks at prediction in two experimental conditions: during shadowing, or within-language repetition of speech with minimal delay; and during the task of simultaneous interpreting (SI), where speech is translated orally from a source language into a target language. Prediction, in the sense of the ability to create expectations during real-time, ‘online’ language processing, is characteristic of any communicative task. During SI interpreters are often observed to create expectations about what will follow in the source language, which may help manage the processing load imposed by the multiplicity and simultaneity of efforts (Gile 2009) in the interpreting task.

One of the fundamental debates in SI concerns the degree of deverbalization of the input. Proponents of the *Interpretive Theory* argue that this occurs, and that the intermediate stages in processing of the source input are thus associated at least in part with a language-independent message, irrespective of source and target language form (Lederer 1981; Seleskovitch 1984). By contrast, SI researchers following the *Information Processing* approach (Gerver 1976; Gile 2009; Moser 1978) attribute greater importance than interpretive theorists to language-specific form. This view mainly stems from psycholinguistic findings on language processing *per se* (i.e., reading and listening), indicating that language-specific syntactic structure and lexicalization patterns affect comprehension (Otten & Van Berkum 2008; Trueswell & Tannenhaus 1994; Van Berkum et al. 2005) – hence the argument that, if form affects comprehension in general, it should affect any language processing task, including SI. The resulting debate about the relevance of language specificity in the interpreting process will be addressed in the present study.

An important factor that has been thought to influence prediction is the difference in syntactic structure or word order between the source language and the target language (Gile 1992, 2009; Jörg 1995; Wilss, 1978). For example, when interpreting from German head-final sentence structures with the main verb at the end of the sentence into English head-initial sentence structures, where the main verb is always placed in second position in the sentence, interpreters often predict the sentence-final verb – and sometimes even produce it in the English output – before it becomes available in the German input. In the SI literature, this is an example of what is referred to more generally as anticipation (Gile 2009).

Two types of anticipation are identified in the SI literature: these are known as linguistic and extralinguistic anticipation (Lederer 1981). The first is assumed to rely more on transitional probability (TP), or the statistical likelihood of two or more words occurring together in a given language (Gile 2009). In an early study of this topic, Wilss (1978: 348) noted the example of the German word *Namens* ('On behalf of'), often introducing the standard expression of thanks *Namens ... darf ich ... danken* ('On behalf of ... I would like to thank ...'): hearing *Namens* thus enables the interpreter to anticipate *danken*, based on the statistical likelihood of the two words co-occurring in German. Setton (2005) characterizes such connecting devices as primary pragmatic factors that lead to anticipation during SI. Along with extralinguistic cues like background knowledge of the topic, these are used incrementally to draw inferences and anticipate what will follow in the unfolding speech (Setton 1999, 2005). This view suggests that SI is much more than transcoding of highly probable lexical items.

However, it is precisely this incremental or piecemeal nature of language processing that has been found to lead to TP effects on prediction during 'online' visual sentence processing (Frisson et al. 2005; McDonald & Shillcock 2003b) and during SI of syntactically symmetrical (head-initial) sentences from German into English (Hodzik 2014). By contrast, Hodzik (2014) found no effect of TP on prediction during SI between syntactically asymmetrical (from head-final to head-initial) sentence structures. Language-specific incremental or piecemeal processing of the input and output was posited to account for these findings, with the latency between the source and target languages considered as dependent on the relative probability of the words concerned co-occurring in at least one of the two languages. A TP effect was found only during SI of very short sentences, in which the context preceding the TP pair was minimal and neutral – e.g., *Sie bekam Unterstützung* ('She received support'), where TP was computed for the verb-noun pair comprised of *bekommen* ('to receive') and *Unterstützung* ('support').

However, it is arguable to what extent TP effects on prediction can be considered separate from the context as a whole. Studies on sentence processing (Frisson et al. 2005; McDonald & Shillcock 2003b; Otten & Van Berkum 2008; Trueswell & Tannenhaus 1994; Van Berkum et al. 2005) have found that prediction is an integral part of language comprehension, as a result of top-down language processing strategies: these occur as the already assimilated prior context and our background or world knowledge are progressively matched to incoming sensory cues that undergo bottom-up processing (Marslen-Wilson 1975, 1992; Tyler & Marslen-Wilson 1982). The types of strategies involved in integrating different levels of processing are thought to provide a basis for expectations about the upcoming linguistic input, and thus for possible anticipation of elements such as lexical items (Chernov 1994; Frisson et al. 2005; Kohn & Kalina 1996). In this way, a sufficiently constraining context can lead people to make predictive inferences as language unfolds: *The burglar had no trouble locating the family safe; of course it was situated behind a ... painting.*

McDonald and Shillcock (2003a, 2003b) found that upcoming words in discourse are predicted as a result of high TP during reading, this variable being computed on the basis of word frequencies from the British National Corpus. McDonald and Shillcock (2003a) discovered that high-TP collocations like *accept defeat* were read faster than non-collocates such as *accept losses*, which have a low TP. The authors not only looked at the statistical likelihood of the target word (*defeat/losses*) following *accept* – in other words, *forward* TP; they also calculated *backward* TP, or the likelihood that *accept* would precede *defeat/losses*. They found that both forward and backward TP had an effect on first fixation and gaze duration during reading.

In addition, McDonald and Shillcock (2003a, 2003b) claimed that TP and contextual constraint affect prediction independently of each other, because they operate on different levels of language processing. While TP depends more on the processing of lower-level information in the sensory input, contextual constraints are associated above all with higher-level syntactic and semantic information (McDonald & Shillcock 2003a 2003b).

In order to test this claim, Frisson et al. (2005) used the same TP pairs as McDonald and Shillcock (2003b), but added contextual constraint as a factor to the experimental design. They did this by introducing semantic cues that would trigger prediction, similar to the above example of the burglar trying to find a *painting* located behind the safe. Contexts that did not contain such cues were deemed neutral (Frisson et al. 2005). The authors created four experimental conditions: (i) sentences with a constraining context and high TP; (ii) sentences with a constraining context and low TP; (iii) sentences with a neutral context and high TP; (iv) sentences with a neutral context and low TP. Gaze duration on the target word was found to be shorter in constraining contexts than in neutral contexts. There was also an effect of TP on prediction – i.e., reading times for the target word decreased in cases where TP was higher for its combination with the word just before it. Another finding was the lack of interaction between the effects of TP and contextual constraint, suggesting that the two are independent of each other in their possible effects on prediction. In an attempt to replicate these results with added semantic cues and an even more constraining context, Frisson et al. (2005) found no effect of TP. This led them to conclude that there is no clear-cut distinction between semantic and TP cues.

The present study will apply Frisson et al.'s (2005) experimental design, which attempts to separate TP from contextual constraints, to an investigation of prediction during SI: specifically, the aim is to see whether TP affects SI independently of contextual constraints. If so, a strong case can be made for considering that, as seems to be the case in reading, TP and contextual constraints operate on different levels of language processing in SI.

In the present study, a distinction is made between prediction and anticipation in SI. As explained above, prediction here means expectations regarding linguistic input that are created 'online' during language processing; consistent with widely accepted usage in the SI literature (Chernov 1994; Gile 1992, 2009; Jörg 1995; Kohn & Kalina 1996; Wilss 1978), anticipation will be used only in reference to actual utterance of target language equivalents before the interpreter is exposed to them in the source language input.

An effect of TP on the prediction of words during SI would suggest that analysis of the input is affected by individual lexical items and the language-specific likelihood of their co-occurrence. Prediction as a result of low-level processing of TP during SI would therefore be

more in line with language-dependent analysis of the input (Gerver 1976; Gile 2009; Moser 1978) than with a supposed deverbalization stage that would make the specificities of the source and target languages irrelevant (Lederer 1981; Seleskovitch 1984). However, a connection between specific cues to prediction and the information that the simultaneous interpreter actually predicts can be established only by investigating SI as it actually happens, or ‘online’.

The methodology of the present study, attempting as it does to separate TP from contextual effects as a whole, is adapted from research using an *English monolingual visual language processing task* (Frisson et al. 2005). Before applying it to the bilingual task of SI from German into English, it needed to be tested in a *German monolingual spoken language processing task*. The task selected for this purpose was shadowing, similar in some respects to SI but at the same time fundamentally different in many ways (Hyönä et al. 1995; Rinne et al. 2000; Hervais-Adelman et al. 2015). Shadowing involves repeating words as soon as they are recognized in continuous speech. Even though SI is a far more demanding and complex task than shadowing, they both involve simultaneous listening and production of spoken language. In addition, the continuity and dynamics of shadowing are like those of speech, meaning that prediction can occur only ‘online’. For these reasons, prediction was initially investigated by a shadowing task (Experiment I), followed by an SI task (Experiment II). This makes it possible to compare the effects of TP and context on prediction, taking into account important differences between the two tasks.

On the basis of the reading studies briefly discussed above (McDonald & Shillcock 2003a, 2003b; Frisson et al. 2005), it is expected that prediction will occur in shadowing and SI as an effect of both contextual constraints and TP. Taking into account McDonald and Shillcock (2003b) and Frisson et al. (2005), an independent effect of TP can be established only if there are main effects of contextual constraints and TP but no interaction between the two.

Experiment I: Shadowing in German

This experiment examined the effects of TP and contextual constraints on prediction during shadowing in German, before applying the same methodology to SI from German into English in Experiment II. To allow comparison between the results for shadowing and SI, native English speakers with advanced proficiency in German participated in both experiments. However, since the shadowing task involved only German, the sample for Experiment I also included native German speakers. The reason for their inclusion was solely methodological, in order to see whether the native English speakers would perform differently from a group doing the same shadowing task in their native language. Between-subject analyses of the data obtained in the native German and native English groups revealed no significant differences between the two groups regarding the effects of context and TP on shadowing latency. Accordingly, only the native English shadowing data were included in the comparison between the two tasks.

Method

Participants. Twenty native speakers of English who speak German at an advanced level participated in this experiment. All were students of German at the Faculty of Modern and Medieval Languages, at the University of Cambridge. Their level of proficiency was determined with a language background questionnaire, including years of study and

educational qualifications in German. Having studied the language for a minimum of seven years prior to the experiment, they had taken A-level German and, in some cases, subsequent university exams. They will be referred to below as ‘English-German bilinguals’, not as an indication that their L2 (German) is at the same level as their L1 (English), but as a simple confirmation of ability to speak their L2 at an advanced level.

Materials. Ninety-six noun-verb pairs with either high or low TP (48 and 48) were taken from the 100-million-word DWDS (*Digitales Wörterbuch der deutschen Sprache*) corpus. The mean number of occurrences of the noun-verb combinations in the corpus was 568.6 (range: 43-4300) for the high TP pairs, and 86.9 (range 5-1297) for the low TP pairs, $t(47) = 2.45$, $p < .01$.

Because both forward and backward TP have been found to affect predictability (McDonald & Shillcock 2003b), the two parameters were computed for each TP pair. This confirmed consistently high or low mean forward and backward TP values for each noun-verb pair, depending on which group it was in. The rationale for ensuring consistency between the forward and backward TP values of any given pair was that both values contribute to TP as a cue to prediction. Forward TP, the statistical likelihood of the verb following the noun in German, was calculated by the following equation: $p[\text{verb}|\text{noun}] = \text{frequency}[\text{noun, verb}] / \text{frequency}[\text{noun}]$, where ‘p’ stands for the probability with which the verb follows the noun and ‘frequency’ is the number of times the verb and/or the noun occur in the corpus (McDonald & Shillcock 2003b; Perruchet & Peereyman 2004). In descriptive terms, the probability that the verb will follow the noun is equal to the frequency of co-occurrence of the noun and verb in the corpus, divided by the frequency of occurrence of the noun in the corpus.

Backward TP, the statistical likelihood of the noun preceding the verb in German, was calculated by the following equation: $p[\text{noun}|\text{verb}] = \text{frequency}[\text{noun, verb}] / \text{frequency}[\text{verb}]$. In descriptive terms, the probability that the noun will precede the verb is equal to the frequency of co-occurrence of the noun and the verb in the corpus, divided by the frequency of occurrence of the verb in the corpus. The mean forward TP values were .06978 and .00823 for high-TP pairs and low-TP pairs respectively. Backward TP values were .02401 for high TP pairs and .01379 for low TP pairs.

Constraining and neutral contexts were created, so that each high and low TP noun-verb pair could be incorporated into a constraining and a neutral context following Frisson et al.’s (2005) experimental design. This resulted in 48 experimental items, all including a subitem for each of four experimental conditions. Table 1 shows one such item, comprising the four experimental conditions and a subitem for each.

Table 1. Sample experimental item in four conditions

Experimental condition	Example in German (with English gloss and translation)
C-H: Constraining context, High transitional probability	<p>Obwohl er kein Geld hatte, war er entschlossen, in Oxford zu studieren. Although he no money had was he determined in Oxford to study Da das Studium zu teuer für ihn war, hatte er von der Universität Because the studies too expensive for him were had he from the university finanzielle Unterstützung bekommen. financial support received</p> <p>(‘Although he did not have any money he was determined to study in Oxford. Because the studies were too expensive for him, he had received financial support from the university.’)</p>

<p>C-L: Constraining context, Low transitional probability</p>	<p>Obwohl er kein Geld hatte, war er entschlossen, in Oxford zu studieren. Although he no money had was he determined in Oxford to study Da das Studium zu teuer für ihn war, musste er von der Because the studies too expensive for him were had to he from the Universität finanzielle Unterstützung verlangen. university financial support request</p> <p>('Although he did not have any money he was determined to study in Oxford. Because the studies were too expensive for him, he had to request financial support from the university.'))</p>
<p>N-H: Neutral context, High transitional probability</p>	<p>Obwohl er kein Auto hatte, war er entschlossen, uns zu besuchen. Although he no car had was he determined us to visit Gestern hat er uns einen Brief geschickt, in dem er uns erzählt hat, dass er Yesterday has he us a letter sent in which he us told has that he Unterstützung bekommen hatte. support received had</p> <p>('Although he did not have a car, he was determined to visit us. Yesterday he sent us a letter in which he told us that he had received financial support.')</p>
<p>N-L: Neutral context, Low transitional probability</p>	<p>Obwohl er kein Auto hatte, war er entschlossen, uns zu besuchen. Although he no car had was he determined us to visit Gestern hat er uns einen Brief geschickt, in dem er uns erzählt hat, dass er Yesterday has he us a letter sent in which he us told has that he Unterstützung verlangen musste. support request had to</p> <p>('Although he did not have a car, he was determined to visit us. Yesterday he sent us a letter in which he told us that he had to request financial support.')</p>

The target word in each experimental condition was the sentence-final main verb in German (the second German word in bold, for each of the four conditions in Table 1). The experimental conditions were mixed, so that every participant would be exposed to 12 sets of four items (one condition per item). As an example, each set contained one condition from the experimental item in Table 1. Within each set, conditions were presented in a pre-established random order. The experimental materials were examined beforehand and, if necessary, corrected by three native speakers of German. A pilot study involving a reading task was also conducted, in order to ensure that the sentences concerned would prove appropriate. To this end, thirteen native German speakers were instructed to read the experimental items and guess the target word (i.e., the sentence-final verb). This was meant to test the effects of the two featured cues (contextual constraints and TP) on ability to predict the target words. ANOVAs were carried out by item and by subject, both treated as random factors. A main effect of TP on first correct guesses was found, $F_1(1, 12) = 102.855, p < .001, F_2(1, 47) = 27.001, p < .001$. Similarly, a main effect of contextual constraint was also observed, $F_1(1, 12) = 15.567, p < .01, F_2(1, 47) = 22.810, p < .01$.

For the shadowing task, the item sets were recorded by a native speaker of German. Each story sequence in an item set (representing one of the four experimental conditions) was followed by a comprehension question concerning the context preceding the noun-verb pair. The purpose of this was to ensure that participants would pay attention to the content of the whole sequence, rather than limit themselves to simply repeating the phonetic input, which is what a shadowing task entails. An example of a comprehension question (for the C-H

condition, in Table 1) is the following: *Was brauchte er um in Oxford zu studieren?* (What did he need in order to study in Oxford?).

Procedure. The experimental material was presented on a computer with Superlab 4.0. Participants were instructed to repeat each sentence they heard, in German, staying as close as possible to the original. Their shadowing output was recorded on a digital TASCAM HD-P2 portable stereo audio recorder, connected to the computer through two channels (one to record the original input, the other for the shadowing).

Analysis. Latency was used to measure *prediction* of the sentence-final verb, or the ‘facilitation’ of its production as an effect of contextual constraints and/or TP. If the latency between the sentence-final verb in the input and output was negative (i.e., if the verb was uttered in the output before it became available in the input), this was considered as *anticipation*. The analysis of the data was carried out on two audio recording and editing programs: Audacity 2.0.2 and Praat. Latency was measured between the onset of the sentence-final verb in the input and the onset of its repetition in the shadowing. This value will be referred to below as shadowing latency. Experimental recordings containing an incorrectly shadowed verb, or none, in the shadowed output were excluded from the analysis of shadowing latency.

Results

There were no instances of anticipation (i.e., negative shadowing latency).

Figure 1 shows the overall mean shadowing latency for each experimental condition. A within-subject ANOVA was carried out, with contextual constraint and TP as within-subject factors, to determine their effects on shadowing latency. The ANOVA was conducted in relation to the random factors subject (F_1) and item (F_2).

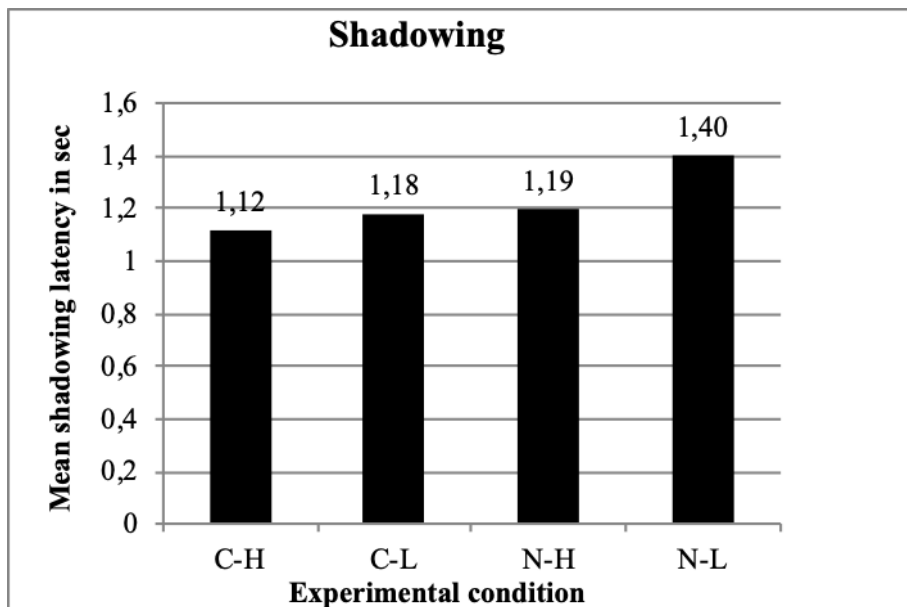


Figure 1. Mean shadowing latency in English-German bilinguals, in four experimental conditions

(C-H = Constraining context, High TP; C-L = Constraining context, Low TP; N-H = Neutral context, High TP; N-L = Neutral context, Low TP)

Both subject and item ANOVAs revealed a main effect of contextual constraint on shadowing latency, $F_1(1, 19) = 13.060, p < .01, F_2(1, 46) = 8.973, p < .01$. A main effect of TP was also found in both analyses, $F_1(1, 19) = 10.133, p < .01, F_2(1, 46) = 8.595, p < .01$. The interaction between contextual constraint and TP was not significant, $F_1(1, 19) = 1.993, p = .174, F_2(1, 46) = 2.847, p = .098$.

Discussion

The absence of negative shadowing latencies suggests that participants waited for lower-level phonetic information on the identity of the sentence-final verb in the sensory input before producing the verb in the output. This is not surprising, given that the shadowing task requires participants to repeat what they hear.

However, the significant difference between shadowing latency in constraining and neutral contexts suggests that higher-level contextual information in the input leads to faster output, or shorter shadowing latency between the input and output. The next step was to see what types of cues trigger prediction (and possibly anticipation) during SI.

Experiment II: Simultaneous interpreting from German into English

The methodology for the shadowing task was applied to SI in Experiment II, which involved some English-German bilinguals with previous experience in SI and others with none. Results for the subjects with no previous experience in SI will be compared with those obtained during shadowing in Experiment I. As mentioned above, an effect of contextual constraint during a bilingual task like SI would not be surprising; what was of interest here was to examine whether SI latency would also show a TP effect, which arguably relies on lower-level processing of information (McDonald & Shillcock 2003a, 2003b) and might thus be thought more likely in shadowing.

In addition, it was important to see whether SI results would differ between participants with previous experience in SI and those with none. Data will therefore not only be analysed and discussed for each of the two groups of participants separately, but also compared between groups. Results obtained in participants with previous experience in SI could provide a basis for further investigation of prediction, as an integral part of language comprehension, by practising interpreters.

Method

Participants. There were two groups of participants: a) English-German bilinguals with no previous experience in SI ('bilinguals'); and b) English-German bilinguals with previous experience in SI ('interpreters').

The first group ('bilinguals') included twenty-two native English speakers who spoke German at an advanced level. All were students at the University of Cambridge, the majority of them studying German at the Faculty of Modern and Medieval Languages. As in the sample for the shadowing task, their level of German was determined with a language background questionnaire, including years of study and educational qualifications in German. Having

studied the language for a minimum of seven years prior to the experiment, they had taken A-level German and, in some cases, subsequent university exams.

The second group of participants ('interpreters') comprised eleven native speakers of English with experience in interpreting from German into English and vice versa. Of these, seven were student interpreters enrolled on Master's courses in Conference Interpreting at the Universities of Leeds (3) and Bath (4). The remaining four were professional interpreters. One was a lecturer in Interpreting at the University of Leeds, also practising as a freelance interpreter. The other three worked freelance, mostly interpreting for the Council of Europe, and were contacted and recruited through AIIC (*Association Internationale des Interprètes de Conférence*).

Materials. The audio files, including experimental items and comprehension questions, were those used in Experiment I: co-occurrence and TP values for the German noun-verb pairs were thus the same as indicated above (see Experiment I Materials). The experimental material had been devised in such a way that TP was consistent between each German noun-verb pair and its English equivalent (i.e., either high TP in both, or low TP in both). According to the British National Corpus (BNC), the mean number of occurrences was 166.3 (range: 7-1318) for the translations of the high-TP pairs and 60.9 (range: 1-1310) for the translations of the low-TP pairs, $t(47) = 2.419, p < .05$. Forward TP values were .01032 for translations of high-TP pairs, and .00872 for translations of low-TP pairs. Backward TP values were .02997 for translations of high-TP pairs, and .00608 for translations of low-TP pairs.

Procedure. The procedure was similar to that in Experiment I. Participants were instructed to simultaneously interpret each sentence from German into English, doing so as soon as possible after hearing the original. Recordings were made on two channels: one for the original input in German, the other for the participants' interpreting output in English on another channel.

Analysis. Latency was measured between the onset of the target word in the source language and the onset of its translation in the target language. This will be referred to as SI latency. The programs used to analyse the data were, again, Audacity 2.0.2 and Praat. In addition to within-subject ANOVAs, between-subject ANOVAs were also carried out in order to identify any differences in the effects of predictive cues between shadowing and SI, as well as between the two groups doing the SI task. Experimental recordings containing an incorrectly interpreted verb, or none, in the interpreting output were excluded from the analysis of SI latency.

Results

English-German bilinguals with no previous experience in SI ('bilinguals')

Figure 2 shows the mean SI latency in English-German bilinguals with no experience in SI, for each experimental condition.

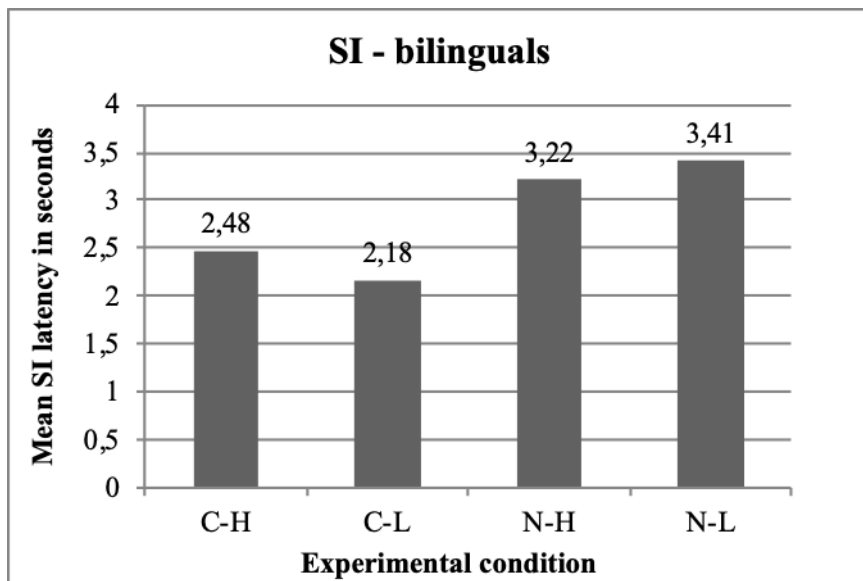


Figure 2. Mean overall SI latency in 22 bilinguals, in four experimental conditions

(C-H = Constraining context, High TP; C-L = Constraining context, Low TP; N-H = Neutral context, High TP; N-L = Neutral context, Low TP)

ANOVA by subject (F_1) and by item (F_2), with contextual constraints and TP as within-subject factors, showed a main effect of contextual constraint on SI latency, $F_1(21) = 31.721$, $p = .00$, $F_2(46) = 33.217$, $p = .00$. Contrary to the shadowing task, the effect of TP on SI latency was not found to be significant during SI, $F_1(21) = .107$, $p = .75$, $F_2(46) = .663$, $p = .42$. The by-subject analysis showed statistically significant interaction between contextual constraint and TP, $F_1(21) = 5.071$, $p < .05$; the by-item analysis did not, $F_2(46) = 2.236$, $p = .14$.

Within-subject analysis for the bilinguals with no experience in SI showed a difference between results for shadowing and SI. A between-subject ANOVA was therefore carried out on the data for the two tasks, in order to determine whether the statistical differences observed were related to the between-subject factor 'task'.

Between-subject ANOVAs showed a main effect of contextual constraint, $F_1(40) = 36.049$, $p = .00$, $F_2(92) = 41.004$, $p = .00$; no effect of TP was identified, $F_1(40) = .614$, $p = .44$, $F_2(92) = .007$, $p = .93$. Interaction between contextual constraint and TP was significant in the subject analysis, $F_1(40) = 6.553$, $p < .05$; it was not far below significance in the item analysis, $F_2(92) = 3.776$, $p = .05$. There was interaction only between contextual constraint and task, $F_1(40) = 20.815$, $p = .00$, $F_2(92) = 20.894$, $p = .00$. *No interaction was observed between TP and task*, $F_1(40) = 1.871$, $p = .18$, $F_2(92) = 2.686$, $p = .10$. Finally, there was no interaction between contextual predictability, TP and task, $F_1(40) = 2.108$, $p = .15$, $F_2(92) = .812$, $p = .37$.

As expected, there were some instances of actual anticipation of the sentence-final verb, with negative SI latency between the sentence-final verb in the German input and its equivalent in the English output. This occurred in only 2.4% of all interpreted sequences (25 out of 1056). Of these, 52% were in the C-H condition, 32% in the C-L condition, and the remaining 16% in the N-H condition. There were no anticipations of the sentence-final verb in the N-L condition. Despite the low overall proportion of anticipations, a within-subject ANOVA was

carried out on the SI data in order to see whether contextual constraint or TP had an effect on anticipation. A main effect of context on anticipation was identified, $F(21) = 9.900, p < .01$. The effect of TP on anticipation was not significant, $F(21) = 2.072, p = .16$. In addition, there was no interaction between contextual constraint and TP, $F(21) = .023, p = .88$.

English-German bilinguals with previous experience in SI ('interpreters')

Figure 3 shows mean SI latency in English-German bilinguals with previous experience in SI, for each experimental condition.

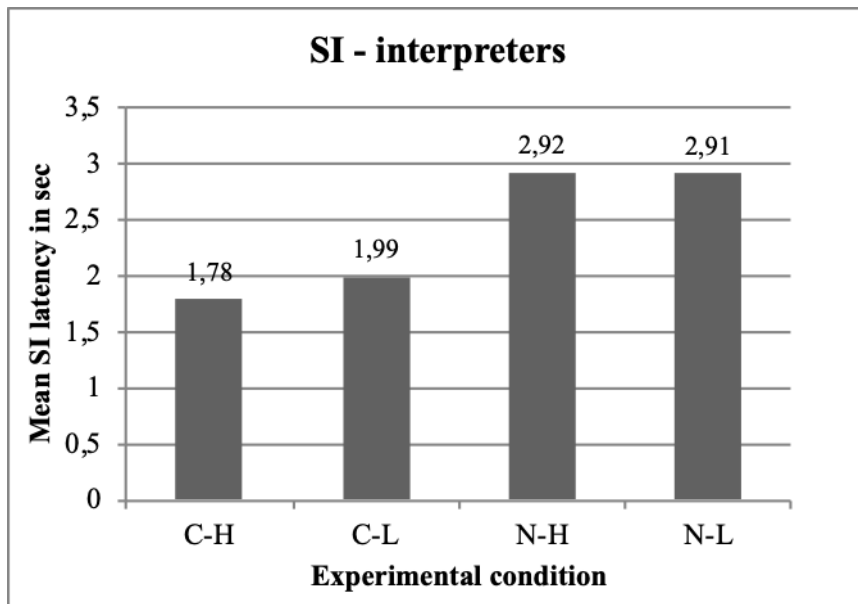


Figure 3. Mean overall SI latency in 11 interpreters, in four experimental conditions

(C-H = Constraining context, High TP; C-L = Constraining context, Low TP; N-H = Neutral context, High TP; N-L = Neutral context, Low TP)

English-German bilinguals with previous experience in SI ('interpreters') showed a main effect of contextual constraint, $F_1(10) = 74.056, p = .00, F_2(35) = 23.493, p = .00$. Like the bilinguals without experience of interpreting, they showed no effect of TP on interpreting latency, $F_1(10) = .819, p = .39, F_2(35) = .500, p = .48$. There was no interaction between contextual constraint and TP, $F_1(10) = .345, p = .57, F_2(35) = .104, p = .75$.

In order to identify any statistically significant differences between the results for the bilinguals and interpreters, between-subject analyses were carried out: contextual constraint and TP were within-subject factors, while 'group' (i.e. bilinguals vs interpreters) was a between-subject factor. Consistent with the trend noted in the previous paragraph for the within-subject analyses, the between-subject results again showed a similar pattern for bilinguals and interpreters. The between-subject ANOVA revealed a main effect of contextual constraint on SI latency, $F_1(31) = 58.215, p = .00, F_2(81) = 55.872, p = .00$. No interaction was observed between contextual constraint and group, $F_1(31) = .016, p = .90, F_2(81) = .036, p = .85$. No effect of TP was found, $F_1(31) = .156, p = .70, F_2(81) = .002, p = .96$. In addition, significant interaction was identified in neither of the following analyses: between TP and group, $F_1(31) = .643, p = .43, F_2(81) = 1.153, p = .29$; and between contextual

constraint and TP, $F_1(31) = .451, p = .51, F_2(81) = .290, p = .59$. Finally, the interaction between contextual constraints, TP and group was not significant, although approaching significance in the ANOVA by subject, $F_1(31) = 3.076, p = .09, F_2(81) = 1.236, p = .27$.

SI by the interpreters showed anticipation of the sentence-final verb in 4% of all the interpreted sequences in the data set (21 out of 528). Of these, 57% were in the C-H condition, 38% in the C-L condition, and 5% in the N-H condition. No anticipation was observed in the N-L condition. To identify any significant differences that participants might show from one experimental condition to another, a within-subject ANOVA was carried out on the interpreters' anticipation data. A main effect of contextual constraints on anticipation was found, $F(10) = 10.198, p = .01$. No effect of TP was observed, $F(10) = 1.000, p = .34$. Finally, no interaction was observed between contextual constraints and TP, $F(10) = .405, p = .54$.

Despite the similar pattern in the bilinguals' and interpreters' results, the latter did show proportionally higher anticipation. Data for the two groups on this parameter were subjected to a Z-test for differences between proportions (Woods et al. 1986): the difference in the overall occurrences of anticipation between bilinguals and interpreters was *not significant*, $z = 1.437, p = .151$. There was also no significant difference between the groups when data were analysed by experimental condition: C-H, $z = 1.607, p = .108$; C-L, $z = 1.444, p = .149$; N-H, $z = 1.263, p = .206$. Neither the bilinguals nor the interpreters anticipated any verbs in the N-L condition.

Discussion

The results obtained in Experiment II, for English-German bilinguals with and without previous experience in SI, showed a main effect of contextual constraints on SI latency. The effect of TP was not significant during SI for either group, despite the main effect of TP during shadowing in Experiment I. Between-subject analysis of the shadowing data and the SI data in English-German bilinguals with no previous experience in SI nevertheless showed no interaction between the effect of TP and the task: this is surprising, considering that the within-subject analyses showed an effect of TP only in the shadowing task, but not in SI.

When comparing mean latencies between the two tasks, the lowest mean latency during shadowing is observed in the C-H condition (high contextual constraints and high TP: see Figure 1); during SI by English-German bilinguals with no previous experience in SI, the lowest mean latency is seen in the C-L condition (high contextual constraints and low TP: see Figure 2). This could explain the observed interaction between contextual constraints and TP in the subject analyses for SI by bilinguals with no previous experience in SI. However, this interaction was not confirmed by the item analyses. By contrast, the bilinguals with previous experience in SI showed the expected pattern – i.e., shorter mean latency in the condition with both TP and contextual constraints than in the condition with only the latter offering a cue to prediction (see Figure 3); however, this difference was not statistically significant.

In Experiment II, instances of negative latency (anticipation) between the sentence-final verb in the input and its equivalent in the output were found in SI by both groups (with and without experience in SI). Results for this parameter were consistent with the pattern found in the rest of the SI latency data: anticipation showed a main effect of contextual constraint, but no effect of TP. However, no interaction was found between contextual constraint and TP in the anticipation data of either group.

A between-subject analysis of the results in SI for bilinguals, according to whether they had previous experience in SI, revealed no statistically significant differences between the two groups. The interpreters nevertheless showed non-significant lower mean latencies across conditions (see Figures 2 and 3), and proportionally more anticipations (again non-significant), than the bilinguals.

General discussion

It was established that a highly constraining context facilitates the processing of the sentence-final verb during shadowing in German and SI from German into English: in both tasks, this is reflected in significantly lower latency between the sentence-final verb in the input and the output. This suggests that, while unfolding speech is being processed during the shadowing and SI tasks, semantic information is used to build expectations about the sentence-final verb in German. This higher-level information is matched to lower-level phonetic information emerging more or less in real time from the sensory input, which results in facilitated recognition and, subsequently, production. In some cases, like actual anticipations of the sentence-final verb observed during SI, production may be based solely on the higher-level semantic information. Even though few such cases were observed in this study, they suggest that there is at least some reliance on top-down information during SI.

The anticipation and prediction data obtained during SI by the bilinguals and interpreters showed the same pattern – i.e., an effect of context, but not of TP. This could potentially seem to justify the argument that both prediction and anticipation should be considered as manifestations of predictive processes during SI. However, anticipations accounted for a very small proportion of interpretations in both groups. More instances of anticipation would be needed to draw any valid conclusions.

The interaction of the between-subject factor ‘group’ (bilinguals vs interpreters) with contextual constraints and TP did not show statistically significant differences in SI latency between the two groups: both of them used contextual constraint, and not TP, as a cue to prediction.

While contextual constraint had an effect on both shadowing and SI latency, an effect of TP was evident only during shadowing. There was no interaction between TP and contextual constraint in either of the tasks. Contrary to our initial study rationale, the combination of the two cues (in the C-H condition, with high contextual constraints and high TP) did not determine an additive effect or result in a significant decrease in SI latency. If anything, the effect of TP seems to have become much weaker during SI than during shadowing.

Based on the present findings, it seems that no clear-cut distinction can be made between contextual and TP cues during shadowing or SI: as integral parts of the context, both can to a certain extent be considered as contextual cues to prediction. This is in contradiction with McDonald and Shillcock’s (2003a, 2003b) view that contextual constraint and TP require different levels of sentence processing, more dependent on higher- and lower-order statistical information respectively.

While empirical studies on SI from German head-final into English head-initial structures have shown an effect of collocations – as highly probabilistic expressions – on anticipation

(Jörg 1995; Van Besien 1999; Wilss 1978), our methodological focus on separating the effects of TP and contextual constraints in the present study did not enable us to identify any effect of TP on SI latency. This suggests that the presumed effects of TP (or collocations) on prediction and anticipation in SI can arguably be seen to depend in part on the practical difficulty of keeping TP distinct and separate from context.

What do these findings suggest about language processing during SI from head-final German sentence structures into head-initial English sentence structures?

It has been found that segmentation of the input during SI, also referred to as *chunking* (Gile 2009), results in longer processing segments or chunks than during shadowing (Goldman-Eisler 1972). This is to be expected, considering that the input has to be processed for translation during SI and only for repetition during shadowing. Possibly the different segmentation or chunking of the input during shadowing and SI leads to correspondingly different effects of TP. In other words, perhaps low-level information (McDonald and Shillcock 2003a, 2003b) like TP (in the present study, between the target word and the word just before it) does not affect processing during SI, based as it is on segmentation into larger meaningful units (comprised of clauses or sentences). This suggests that some degree of deverbalization of the input is required during SI (Setton 1999, 2005), at least to the point where statistical probabilities based on language-specific lexicalization patterns do not affect the interpreter's cognitive processing.

Even in SI of short speech segments, such as those in the present study, semantic cues in the context have an effect on prediction and, in a few instances, on anticipation, which seems to override any low-level effect of TP. In the SI task, while the verb was in sentence-final position in the German input, English head-initial word order was mandatory in the output. The need to change from the noun-verb source language structure to the verb-noun target language structure may have entailed some delay in the SI, which would have overridden any decrease in SI latency caused by the high TP between the noun and verb.

Future research on SI between syntactically symmetrical source and target language structures (e.g., SI from SVO German structures into SVO equivalents in English) would offer further clues as to whether language-specific sentence structure affects the predictive role of TP during SI. This in turn would afford better insight into the effect of language-specific factors, such as syntactic (a)symmetry between the source and target languages, on the analysis of the input during SI. An effect of TP which is dependent on symmetrical sentence structure between the source and target languages has been established in the absence of contextual cues (Hodzik 2014). It would be relevant to determine whether such an effect would also be obtained independently, in the presence of contextual cues, during SI between syntactically symmetrical source and target languages.

In conclusion, the findings presented above show that language-specific probabilistic information may not lead to prediction during 'online' language processing in a SI task, but semantic information in the context as a whole does. Despite efforts to methodologically separate the effects of TP and of context as a whole during the SI task, TP effects seem to be an integral part of semantic effects in the context, at least in the present case of SI between syntactically asymmetrical sentence structures. It could therefore be argued that the specific (asymmetric) structures of the source and target languages used for the SI in the present study were the reason for the lack of TP effect on prediction. Further investigations of prediction

during SI are needed, involving both symmetrical and asymmetrical sentence structures in the source and target languages.

References

- Chernov, V. M. (1994). Message redundancy and message anticipation in simultaneous interpreting. In S. Lambert & B. Moser-Mercer (Eds.), *Bridging the gap: Empirical research in simultaneous interpretation*. Amsterdam: John Benjamins.
- Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31 (5), 862-877.
- Gerver, D. (1976). Empirical studies of simultaneous interpretation: A review and a model. In R. W. Brislin (Ed.), *Translation: Applications and research*. New York: Gardner Press, 165-207.
- Gile, D. (1992). Predictable sentence endings in Japanese and conference interpretation. *The Interpreters' Newsletter*, Special Issue 1, 12-23.
- Gile, D. (2009). *Basic concepts and models for interpreter and translator training*. Amsterdam: John Benjamins.
- Goldman-Eisler, F. (1972). Segmentation of input in simultaneous translation. *Journal of Psycholinguistic Research* 1 (2), 127-140.
- Hervais-Adelman, A., Moser-Mercer, B., Michel, C. M. & Golestani, N. (2015). fMRI of simultaneous interpretation reveals the neural basis of extreme language control. *Cerebral Cortex* 25 (12), 4727-4739.
- Hodzik, E. (2014). The effect of word order on predictive processes during simultaneous interpreting from German into English. In W. Baur, B. Eichner, S. Kalina, N. Kessler, F. Mayer & J. Oersted (Eds.), *Man versus machine. Vol. II: The future of translators, interpreters and terminologists*. Proceedings of the XXth FIT World Congress, Berlin.
- Hyönä, J., Tommola, J. & Alaja, A. M. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *Quarterly Journal of Experimental Psychology* 48A, 598-612.
- Jörg, U. (1995). Bridging the gap: Verb anticipation in German-English simultaneous interpreting. In M. Snell-Hornby, Z. Jettmarová & K. Kaindl (Eds.), *Translation as intercultural communication*. Amsterdam: John Benjamins, 217-228.
- Kohn, K. & Kalina, S. (1996). The strategic dimension of interpreting. *Meta* 41 (1), 118-138.
- Lederer, M. (1981). *La traduction simultanée – expérience et théorie*. Paris: Minard Lettres Modernes.
- Marslen-Wilson, W. (1975). Sentence perception as an interactive parallel process. *Science*, 189(4198).
- Marslen-Wilson, W. (1992). The mental lexicon in language processing. In W. Bright (Ed.), *International encyclopedia of linguistics*. New York: Oxford University Press, 273-276.
- McDonald, S. A. & Shillcock, R. C. (2003a). Eye movements reveal the online computation of lexical probabilities during reading. *Psychological Science* 14 (6), 648-652.
- McDonald, S. A. & Shillcock, R. C. (2003b). Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research* 43, 1735-1751.

- Moser, B. (1978). Simultaneous interpretation: A hypothetical model and its practical application. In D. Gerver & H. W. Sinaiko (Eds.), *Language interpretation and communication*. New York/London: Plenum Press, 353-368.
- Otten, M. & Van Berkum, J. J. A. (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes* 45 (6), 464-496.
- Perruchet, P. & Peereman, R. (2004). The exploitation of distributional information in syllable processing. *Journal of Neurolinguistics* 17, 97-119.
- Rinne, J. O., Tommola, J., Laine, M., Krause, B. J., Schmidt, D., Kaasinen, V., et al. (2000). The translating brain: cerebral activation patterns during simultaneous interpreting. *Neuroscience Letters* 294, 85-88.
- Seleskovitch, D. (1984). Les anticipations dans la compréhension. In D. Seleskovitch & M. Lederer (Eds.), *Interpréter pour traduire*. Paris, Didier, 273-283.
- Setton, R. (1999). *Simultaneous interpretation: A cognitive-pragmatic analysis*. Amsterdam: John Benjamins.
- Setton, R. (2005). So what is so interesting about simultaneous interpreting? *SKASE Journal of Translation and Interpretation* 1 (1), 70-84.
- Trueswell, J. C. & Tanenhaus, M. K. (1994). Toward a lexicalist framework for constraint-based syntactic ambiguity resolution. In C. Clifton, L. Frazier & K. Rayner (Eds.), *Perspectives in sentence processing*. Hillsdale, NJ: Lawrence Erlbaum Associates, 155-179.
- Tyler, L. K. & Marslen-Wilson, W. (1982). Speech comprehension processes. In J. Mehler, E.C.T. Walker & M. Garrett (Eds), *Perspectives on mental representations*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 169-184.
- Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V. & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31 (3), 443-467.
- Van Besien, F. (1999). Anticipation in simultaneous interpretation. *Meta* 44 (2), 250-259.
- Wilss, W. (1978). Syntactic anticipation in German-English simultaneous interpreting. In D. Gerver & H. W. Sinaiko (Eds.), *Language interpretation and communication*. New York/London: Plenum Press, 335-343.
- Woods, A., Fletcher, P. & Hughes, A. (1986). *Statistics in language studies*. Cambridge: Cambridge University Press.

Authors' address

Ena Hodzik
 Department of English and Translation
 Faculty of Humanities
 Notre Dame University–Louaize
 Zouk Mosbeh
 Lebanon

ena_hodzik@yahoo.com, ehodzik@ndu.edu.lb

About the authors

Ena Hodzik is an Assistant Professor at Notre-Dame University–Louaize in Lebanon. She did her PhD in English and Applied Linguistics at the University of Cambridge, under the supervision of Dr John Williams (the co-author of this article). The topic of her thesis was ‘Predictive processes during simultaneous interpreting from German into English’, her experimental methodology being based on the use of speech latency measurements. Her research interests also include bilingual language processing and production, bilingual lexical processing, morphological representation, and German sentence processing.

Dr John N. Williams is a Reader in Applied Psycholinguistics, specializing in the cognitive mechanisms of second language learning and second language lexical and syntactic processing. His work draws on theoretical concepts and experimental methodologies from cognitive psychology, which he applies to laboratory-based study of second language processing and learning. He has published numerous articles in journals such as *Language Learning*, *Studies in Second Language Acquisition*, and *Applied Psycholinguistics*.
jnw12@cam.ac.uk