Linguistica Silesiana 27, 2006
ISSN 0208-4228

EWA GUMUL
University of Silesia, Katowice

CONJUNCTIVE COHESION AND THE LENGTH OF EAR-VOICE SPAN IN SIMULTANEOUS INTERPRETING. A CASE OF INTERPRETING STUDENTS

The present paper aims to examine the correlation between the presence of formal cohesive markers in the surface structure of the text and the length of the ear-voice span in simultaneous interpreting outputs. The types of cohesive devices analysed in the study are various types of conjunction: additive conjunction, adversative conjunction, causal conjunction, temporal conjunction, continuatives and emphatic conjunction. The model of classification of cohesive ties adopted in this study is that of Halliday and Hasan (1976), used previously by Shlesinger (1995) in her study of cohesion in simultaneous interpreting. The analysis of conjunctive cohesive ties rendition was performed using the criteria of Interpreting Constraints partly set in the framework of Information-processing Theory and relying heavily on Gile's Effort Models.

Introduction

The ear-voice span (EVS), also referred to as 'lag' or 'delay', can be defined as the time span elapsing between the reception of source input by the interpreter and its rendition in the output language (Setton 1999). This inherent aspect of simultaneous interpreting performance is subject to a number of factors including language combination, target-language speech features, both speaker's and interpreter's performance, and the interpreter's idiosyncratic preferences.

One of the questions addressed in Frauenfelder and Schriefers' (1997) account of psycholinguistic aspects of simultaneous interpreting is whether the availability or non-availability of certain cues in the source-language speech should lead to a mapping on a shorter or longer input string which is consequently reflected in the relative length of EVS (Frauenfelder and Schriefers 1997: 82). The potential cues mentioned by these authors include certain syntactic structures that are expected to influence the rendition of the incoming speech segment. Another type of text feature which might affect the length of EVS is its cohesion.

The crucial function of cohesion in interpreting has been emphasized by numerous studies and accounts (e.g. Hatim & Mason 1997, Shlesinger 1995, Mizuno 1999, Niska 1999, Łyda & Gumul 2002, Gumul 2004, Łyda 2004, Gumul – forthcoming). According to Hatim and Mason (1997), cohesive devices are one of the elements that build the texture of the target speech. Since, due to the presence of the linearity constraint in simultaneous interpreting, texture is the predominant domain of textuality in this mode, it is only via texture that an interpreter can gain access to the two remaining domains: structure and context. Thus, the relevance of the cohesive markers network in SI performance can by no means be underestimated. The question remains, however, whether this text feature should affect the length of EVS. Therefore, the aim of the present paper is to ascertain the correlation between the presence of formal cohesive markers in the surface structure and the length of the ear-voice span in simultaneous interpreting outputs.

EVS in simultaneous interpreting

The EVS phenomenon has been extensively examined in a number of studies conducted during the experimental psychology period in interpreting research history, resulting in the accounts of Treisman (1965), Oléron and Nanpon (1965), Gerver (1969), Kade and Cartellieri (1971), Barik (1973), Goldman-Eisler (1972), and Kirchhoff (1976). The sustained interest in this aspect of SI performance is also reflected in contemporary interpreting literature, particularly in the works of de Groot (1997), Yagi (2000), and two recent accounts of Lee (2002, 2003) constituting a major contribution to this area of interpreting research.

In his definition of EVS, Lee (2002) stresses another important feature of this phenomenon. He defines ear-voice span as the minimum time needed by an interpreter for information processing under heavy cognitive constraints. Under these constraints, interpreters are frequently forced to resort to the anticipation based on top-down processing, which enables them to begin a target-language sentence before the end of its source-language equivalent (Lee 2002: 598). The decision when to begin output production is part of the strategy of EVS regulation.

EVS regulation is probably the most frequently employed strategy in SI, as without keeping even the minimum time lag between reception of input and output production, simultaneous interpreting would be virtually impossible. It is worth noting, however, as indicated by Anderson (1994), that we should distinguish between a characteristic lag and a maximum lag. The former refers to the EVS inherent in the task of interpreting, while the latter, employed in cases of difficulty, is supposed to reflect the limits of human short-term memory.

Regulating the EVS enables interpreters to control to a certain extent the processing-capacity requirements. However, this strategy has to be adopted with caution, as both reducing and increasing the time lag entail potential risk. Shortening the EVS is beneficial in terms of decreasing short-term memory requirements, but may produce an adverse effect, resulting in misunderstanding the propositional content or

embarking on a sentence which is difficult to complete. On the other hand, lagging too far behind the speaker does increase comprehension potential, but may impose an excessive strain on short-term memory resulting in processing capacity saturation (Gile 1995, 1997).

According to Kirchhoff (1976), 'the interpreter's optimum starting point would have to lie where a maximum amount of certainty and a minimum load on capacity are ensured (...) and would have to correspond to the respective limits of the smallest recoding unit' (Kirchhoff 1976: 115). De Groot (1997) stresses that 'the EVS should be as short as the prevailing circumstances permit' (de Groot 1997: 44). Kade and Cartellieri (1971) maintain that the optimal moment for an interpreter to start producing a source-language unit is immediately after all syntactic and semantic ambiguities in this unit have been resolved. Goldman-Eisler (1972) holds a similar view, claiming that 'the minimum EVS sequence is the NP + VP (...) and the VP is a crucial part of the information required' (Goldman-Eisler 1972: 72).

The emergence of the relevant information in the input unit is obviously not the only variable determining the EVS length. Apart from language combination and the interpreter's idiosyncratic preferences, it is essential to consider certain features of the target-language text, such as information density and syntactic complexity. The length of the ear-voice span is also claimed in some studies to be related to certain aspects of speaker performance, such as between-sentence pauses (e.g. Lee 2002) and the input rate measured in words per minute (e.g. Barik 1973, de Groot 1997).

The dependence of the EVS length on some of these variables will be explored in the study described in the present paper.

The study

The decision to analyse one particular aspect of textual cohesion, namely formal cohesive markers, stems from their inherently overt manifestation in the text structure. As emphasised by Blum-Kulka (1986:23), 'cohesion is an overt textual relationship objectively detectable', and therefore it 'lends itself to quantitative analysis'.

All the cohesive markers analysed in the present study fall into the category of conjunction, since the preliminary research revealed that this category is prone to EVS length variation to the largest extent in comparison with sentences which do not begin with a formal cohesive tie. The model of classification of cohesive ties adopted in this study is that of Halliday and Hasan (1976). However, it has to be emphasized that the category of conjunction in this classification is far wider than the term traditionally recognized by grammarians, since it encompasses virtually all kinds of linking words and expressions referring to the content of the preceding or following discourse. The types of conjunctions undergoing analysis belong to the subcategories of simple additive relations (*also*), complex additive relations (*in addition*), adversative relations (*but*), specific causal relations (*as a result*), temporal relations (*then, next*), sequential temporal relations (*first of all*), and conclusive temporal relations (*finally*).

In order to ascertain the correlation between the presence of cohesive markers in the surface structure of the text and EVS length, two types of sentences were analysed: those beginning with a conjunction, and those which do not begin with an overt cohesive marker. For the purpose of simplicity we shall refer to them as 'cohesive sentences' and 'empty sentences' respectively. Nine sentences of each type were selected for the analysis. They come from two recorded speeches, both written monologues intended for oral delivery.

The choice of source-language sentences to undergo analysis in the present study is subject to one major restriction. Only those preceded by the speaker's between-sentence pauses of over 4 seconds were selected in order to avoid the influence of the tail-to-tail span. This new variable in SI research was proposed by Lee (2003), and refers to the time span between the end of a source-language sentence and the end of a target-language sentence. Obviously, prolonged tail-to-tail span (TTS) is bound to influence adversely the EVS of the ensuing sentence, precluding an objective judgement of the effect of the presence or absence of the cohesive marker on the time delay. However, despite potentially ample speaker's between-sentence pauses, TTS overlap with the ensuing sentence was observed in some outputs. All such target-language sentences were excluded from the analysis. There were also some cases in which some source-language sentences were omitted altogether by the interpreting subjects. Therefore the total number of analysed speech chunks amounts to 370 sentences. EVS measurements were conducted using voice-editing software, which can measure up to one millisecond.

The subjects in the study were English Philology students (Translation and Interpreting Programme) of the 3rd year, referred to as Novices, and 5th year, whom we shall call Graduates. The two groups consisted of twenty four subjects each. The first group received a nine-month training in simultaneous interpreting, the other a twenty-seven month training of the same kind. All subjects were Polish native speakers having English as language B in their language combination. Research conducted on interpreting students is often criticised as lacking ecological validity. However, first of all, it would be impossible to obtain multiple renditions of the same speech in real life conditions. Secondly, the experimental design of the study, employing students as subjects, offers the additional benefit of insight into the training process and thus carries potential didactic implications. Therefore, one of the research questions is to determine whether the length of EVS in rendering sentences with cohesive markers is experience-related.

Discussion of results

A quantitative analysis of output sentences revealed a statistically significant difference between the sentences that begin with a cohesive marker and those which do not. T-tests carried out between these two groups of sentences indicate that the average EVS length in the case of 'cohesive sentences' is 2.16 seconds, whereas the result for the so-called 'empty sentences' is 1.5 seconds ($p<0.05$). The results obtained for

the group of novice interpreters reveal an even larger discrepancy. The average ear-voice span for the sentences opening with an explicit cohesive tie is 2.35 seconds, while the score for the other type of sentences is 1.62 second ($p<0.05$). The EVS length seems to be experience-related for both types of sentences, as the t-test results for the graduate interpreters are 1.96 seconds and 1.38 seconds ($p<0.05$) for 'cohesive' and 'empty' sentences respectively.

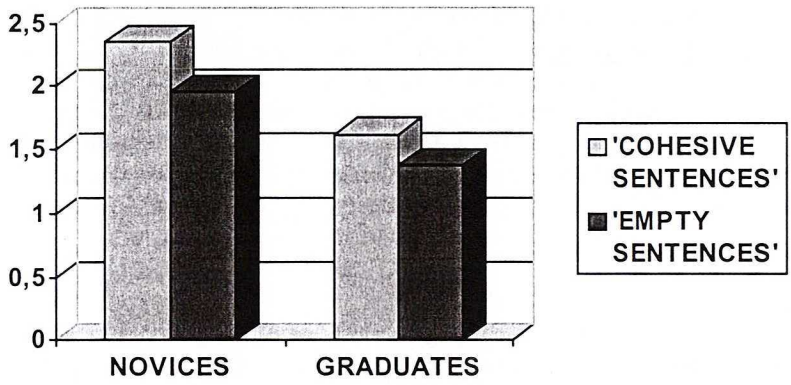

Fig. 1. EVS values of the renditions of 'cohesive sentences' and 'empty sentences'

The average EVS calculated for both types of sentences, which amounts to 1.84 seconds, seems to be excessively short in comparison with the average length of ear-voice span reported in other studies. Notwithstanding the fact that these studies were conducted for different language combinations, which appears to be a significant factor determining the length of the temporal delay, the value of 1.84 seconds is well below the results obtained by Barik (1973) – 2 to 3 seconds, Lederer (1978) – 3 to 6 seconds, and Lee (2002) – 3 seconds. However, considering the nature of the analysed material, it has to be emphasized that the average of 1.84 seconds cannot be regarded as the mean EVS for the English-Polish combination. There are two factors that contribute significantly to shortening the ear-voice span in the analysed output samples. As indicated before, all the selected sentences are preceded with a speaker's between-sentence pause ranging from 4 to 5 seconds. Previous research (Lee 1999, 2002) indicates that longer between-sentence pauses tend to shorten EVS, since the load of sharing attentional capacity is eliminated and consequently the interpreter is free from the production effort while processing the current input unit.

Another factor contributing to ear-voice span reduction is the input rate. Some studies (e.g. Barik 1973, de Groot 1997) report that increasing the wpm ratio lengthens the EVS.[1] According to de Groot (1997), high input rate means the time span

[1] There exists some conflicting evidence concerning the influence of the input rate on the length of the EVS. Barik (1973) and de Groot (1997) claim that the higher the input rate, the

over which the individual words are presented in the input is relatively short, which impedes their effective processing (de Groot 1997: 45). The average input rate of the texts analysed in this study is approximately 120 words per minute, the ratio which is generally recognized in interpreting literature to be optimal for successful processing (e.g. de Groot 1997, Seleskovitch 1978). Therefore, this factor might be perceived as attesting to the substantially shortened EVS in the case of the discourse samples analysed in the present study.

According to Lee (2002), short EVS indicates that processing was smooth and speedy. However, although this might be true for professional interpreters, who were the subjects in Lee's research, it is clearly not always the case with interpreting students. In this study, substantial interpreter's between pauses following the initial word or phrase, entirely independent of speaker's between pauses, indicate that the onset of a source-language chunk does not always mean that this particular unit has been fully processed:

(1)
Source text:
[...] first of all / we applied ourselves to identifying the root causes of our national ailments / examining contemporary evidence / and refusing to be slaves to outmoded doctrinaire beliefs [...][2]

Target text:
[...] **(EVS = 1.31s)** po pierwsze **(IBP[3] = 3.86s)** / chcieliśmy zidentyk / zidentyfikować / przyczyny / świadomości narodowej / odmówiliśmy być niewolnikami / dok / przestarzałych / pog / poglądów [...]

Considerable interpreter's between pauses, up to 4 seconds, were observed in both types of analysed sentences. However, they were 40% more frequent in the case of novice interpreters, attesting to the lack of experience and poor mastery of the strategy of EVS regulation in this group.

Moreover, excessively short EVS, below 0.8 seconds, almost invariably led to transcoding, especially in the case of novice interpreters, who frequently ventured to produce a target-language chunk which had not been processed completely:

---

longer the EVS. On the other hand, the results obtained by Gerver (1969) and Lee (2002) indicate that the opposite is true. Lee (2002) discovered that 'high speech rate and SP of the speaker reduced EVS while EVS was lengthened when the speech rate and SP decreased' (Lee 2002: 601).

[2] The transcriptions have been marked for pauses within the respective utterances. A single slash (/) denotes a short pause, and a double slash (//) denotes a long pause. Apart from the capitalised proper names, only lower case letters have been used in the transcription. There is no punctuation.

[3] IBP stands for Interpreter's Between Pause(s).

(2)
Source text:
[...] these new projects / have enabled us / to lay a firm foundation for better things / it is at this stage that we may confidently begin to examine the route which we wish to follow in the future

Target text:
[...] **(EVS = 0.78)** te wszystkie nowe programy / umożliwiły nam położyć solidny fundament pod nowe lepsze życie / już teraz na tym poziomie możemy zacząć rozpoznawać / możemy zacząć rozpoznawać tę drogę / zająć się tą drogą którą chcemy podążać w przyszłości

On the other hand, longer EVS or longer interpreter's between pauses in this group resulted in a number of omissions or failures to render the propositional content of the source text:

(3)
Source text:
[...] first of all / we applied ourselves to identifying the root causes of our national ailments / examining contemporary evidence / and refusing to be slaves to outmoded doctrinaire beliefs / secondly / we embarked on a reasoned policy to ensure steady economic growth / the modernisation of industry / and a proper balance between public and private expenditure [...]

Target text:
[...] **(EVS = 3.1s)** po pierwsze **(IBP = 6.83s)** / postanowiliśmy obrać drogę // naszego działania i / i odrzucić doktryny które nam narzucono / po drugie / post / postanowiliśmy prowadzić sensowną politykę / powolnego wzrostu ekonomicznego / oraz równowagi pomiędzy publicznymi i prywatnymi wydatkami [...]

The performance of the graduate interpreters differed substantially from that of the novice group. Their EVSs were significantly shorter and there were fewer cases of omissions and transcoding. Instead of long between pauses, they tended to employ the strategies of padding (4) or chunking when faced with processing difficulties. Moreover, the time lost on prolonged EVS or long between pauses immediately after the cohesive marker was in many cases compensated for by increasing the output rate and shortening between sentence pauses in the course of producing that particular segment:

(4)
Source text:
[...] as a result of those immediate measures // and aided by the tremendous effort / which they evoked from the British people [...]

Target text:
[...] **(EVS = 1.47s)** w wyniku / tych natychmiastowych środków przez nas podjętych // wspomaganych ogromnym wysiłkiem ze strony brytyjskiego narodu [...]

Closer examination of individual renditions of both types of sentences reveals that interpreter's between pauses tend to be more frequent in those beginning with a cohesive tie. The pause, which we shall refer to as 'post cohesive marker lag' (PCML) usually occurs immediately after the cohesive marker has been rendered:

(5)
Source text:
[...] next / we instituted the largest programme of educational expansion / that the country has ever seen [...]

Target text:
[...] **(EVS = 1.02s)** następnie **(PCML = 2.74s)** / staraliśmy się wdrożyć największy program edukacyjny który kiedykolwiek / widziano w tym kraju [...]

The average PCML is 3.15 seconds for the novice group, and 2.32 seconds for graduate interpreters. It has to be noted, however, that PCML was calculated only for those sentences in which a between pause longer than 1 second occurred following a cohesive marker. This was the case in 42% of novice outputs and 23% of graduate renditions. In the remaining cases, the target-language version indicated smooth processing or the rendition of the whole sentence was preceded by a substantial EVS.

Another conclusion reached when analysing sentences opening with a cohesive marker is that in some cases, longer between pauses and errors stemmed from ignoring the function of cohesive markers. The cohesive element was translated automatically, but the following segment was found to be cohesive only in formal terms; the content was not rendered in line with the cues provided by the cohesive tie.

The two types of cohesive markers which seem to be prone to the longest EVS are specific causal relations, with an average EVS of 2.57 seconds, and adversative conjunctions, with a mean value of 2.39 seconds ear-voice span. The shortest average EVS was observed in the case of temporal relations (1.67 seconds) and conclusive temporal relations (1.63 seconds). The considerably shorter EVS while rendering sentences beginning with a temporal link is probably due to their transparent function in discourse. Those temporal cohesive markers undergoing analysis in the present study serve a simple function of enumeration. Prior processing of the following input is not essential for correct rendition of such sentences.

Table 1. The EVS value for different types of conjunctions

CONJUNCTIONS	EVS length
SIMPLE ADDITIVE RELATIONS	2.49
COMPLEX ADDITIVE RELATIONS	2.38
ADVERSATIVE RELATIONS	2.39
SPECIFIC CAUSAL RELATIONS	2.57
TEMPORAL RELATIONS	1.67
SEQUENTIAL TEMPORAL RELATIONS	1.89
CONCLUSIVE TEMPORAL RELATIONS	1.63

Conclusion

The findings of the study indicate that the presence of overt cohesive marker in the surface structure of the text does influence the length of the EVS. Interpreters appear to need more time to disambiguate sentences beginning with a cohesive device, irrespective of their level of training. However, the analysis of the outputs of the interpreting students at the final stage of their training reveals that the quality and accuracy of such speech segments does not seem to be adversely affected. The time lost on longer EVS is compensated for by shortening interpreter's between pauses and increasing the output rate. The delivery of the chunk is smooth. Therefore, it might be assumed that texture-creating cohesive devices facilitate source-language speech processing.

The differences in performance observed between the two groups of subjects attest to the need to put greater emphasis on the function of cohesive devices during interpreters' training.

Finally, it must be emphasized that in order to generalize the results of this paper, analysis of the output samples of professional interpreters not only in the English--Polish combination, but also in other language combinations would be necessary.

References

Anderson, L. (1994). Simultaneous Interpretation: Contextual and Translation Aspects. In: S. Lambert and B. Moser-Mercer (eds.), *Bridging the Gap. Empirical Research in Simultaneous Interpretation.* Amsterdam/Philadelphia: John Benjamins Publishing Company. 101–120.

Blum-Kulka, S. (1986). Shifts in Cohesion and Coherence in Translation. In: J. House and S. Blum-Kulka (eds.), *Interlingual and Intercultural Communication: Discourse and Cognition in Translation and Second Language Acquisition Studies.* Tübingen: Gunter Narr. 17–35.

Davy, D. (1966). *Advanced English Course.* London: Linguaphone.

Frauenfelder, U.H., H. Schriefers (1997). A psycholinguistic perspective on Simultaneous Interpretation, *Interpreting* 2(1/2), 55–89.

Gerver, D. (1969). The Effects of Source Language Presentation Rate on the Performance of Simultaneous Conference Interpreters. In: F. Pöchhacker and M. Shlesinger (eds.) (2002), *The Interpreting Studies Reader.* London and New York: Routledge. 52–66.

Gile, D. (1995). *Basic Concepts and Models for Interpreter and Translator Training.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

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

Goldman-Eisler, F. (1972). Segmentation of Input in Simultaneous Translation. In: F. Pöchhacker and M. Shlesinger (eds.) (2002), *The Interpreting Studies Reader.* London and New York: Routledge. 69–76.

de Groot, A.M.B. (1997). The Cognitive Study of Translation and Interpretation. In: J. Danks, G.M. Shreve, S.B. Fountain and M.K. McBeath (eds.), *Cognitive Processes in Translation and Interpreting.* London: Sage Publications. 196–214.

Gumul, E. (2004). *Cohesion in Interpreting.* Unpublished PhD dissertation, University of Silesia.

Gumul, E. (2005). EVS: czynnik czasu w tłumaczeniu symultanicznym. In: J. Arabski, E. Borkowska and A. Łyda (eds.), *Czas w języku i kulturze.* Katowice: Para. 176–183.

Gumul, E. (forthcoming). Explicitation in Conference Interpreting. In: M. Thelen and B. Lewandowska-Tomaszczyk (eds.), *Translation and Meaning. Part 7.* "Maastricht: Department of Translation and Interpreting, Maastricht School of International Communication, Zuyd University."

Gumul, E. (forthcoming). Conjunctive Cohesive Markers: Translational Shifts in English-Polish SI and CI. In: M. Thelen and B. Lewandowska-Tomaszczyk (eds.), *Translation and Meaning. Part 8.* "Maastricht: Department of Translation and Interpreting, Maastricht School of International Communication, Zuyd University."

Halliday, M.A.K., R. Hasan (1976). *Cohesion in English.* London and New York: Longman.

Hatim, B., I. Mason (1997). *Translator as a Communicator.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Kade, O., C. Cartellieri (1971). Some methodological aspects of simultaneous interpreting, *Babel* 17, 12–16.

Kirchhoff, H. (1976). Simultaneous Interpreting: Interdependence of Variables in the Interpreting Process, Interpreting Models and Interpreting Strategies. In: F. Pöchhacker and M. Shlesinger (eds.) (2002), *The Interpreting Studies Reader.* London and New York: Routledge. 110–119.

Lederer, M. (1978). Simultaneous Interpretation – Units of Meaning and other Features. In: F. Pöchhacker and M. Shlesinger (eds.) (2002), *The Interpreting Studies Reader.* London and New York: Routledge. 130–140.

Lee, T.H. (1999). Simultaneous Listening and Speaking in English into Korean Simultaneous Interpretation, *Meta* 44(4), 560–572.

Lee, T.H. (2002). Ear Voice Span in English into Korean Simultaneous Interpretation, *Meta* 47(4), 596–606.

Lee, T.H. (2003). Tail-to-Tail Span: A New Variable in Conference Interpreting Research, *Forum* 1:1, 41–62.

Łyda, A., E. Gumul (2002). Cohesion in Interpreting. In: D. Stanulewicz (ed.), *PASE Papers in Language Studies: Proceedings of the Ninth Annual Conference of the Polish Association for The Study of English*, Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego. 349–356.

Łyda, A. (2004). On Causal Connectives in English-Polish Simultaneous and Consecutive Interpreting, *Linguistica Silesiana* 25, 103–116.

Mizuno, A. (1999). Shifts of Cohesion and Coherence in Simultaneous Interpretation from English into Japanese, *Interpreting Research* 8:2 (no 16), 31–41.

Niska, H. (1999). *Text Linguistics Models for the Study of Simultaneous Interpreting*. Stockholm University (http://www.geocities.com/~tolk/lic/licproj_w2wtoc.htm).

Oléron, P., H. Nanpon (1965). Research into Simultaneous Translation. In: F. Pöchhacker and M. Shlesinger (eds.) (2002), *The Interpreting Studies Reader*. London and New York: Routledge. 42–50.

Seleskovitch, D. (1978). *Interpreting for International Conferences. Problems of Language and Communication*. Washington: Library of Congress.

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

Shlesinger, M. (1995). Shifts in Cohesion in Simultaneous Interpreting, *Translator* 1(2), 193–212.

Yagi, S. M. (2000). Studying Style in Simultaneous Interpretation, *Meta* 45(3), 520–547.