



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
DIPARTIMENTO DI INTERPRETAZIONE E TRADUZIONE

**CORPUS-BASED INTERPRETING STUDIES:  
THE STATE OF THE ART  
FIRST FORLÌ INTERNATIONAL WORKSHOP  
FORLÌ, ITALY, 7-8 MAY 2015**

**BOOK OF ABSTRACTS**



# INDEX

## *PAPER PRESENTATIONS*

Acquiring the language of interpretese.....	7
GUY ASTON, UNIVERSITY OF BOLOGNA	
Corpus-based Interpreting Studies (CIS): past, present and future developments of a (wired) cottage industry .....	8
CLAUDIO BENDAZZOLI, UNIVERSITY OF TORINO, DEPARTMENT OF ECONOMIC AND SOCIAL STUDIES	
From quantity to quality: Using corpus-based methods to analyse interpreting performance.....	10
SABINE BRAUN, UNIVERSITY OF SURREY	
Lost in interpreting or...maybe not. Strings of nouns as a challenge for simultaneous interpreters working from Polish into Italian.....	11
RITA CAPPELLI, UNIVERSITY OF BOLOGNA	
Corpora worth creating: a pilot study on telephone interpreting .....	14
SARA CASTAGNOLI, UNIVERSITY OF BOLOGNA	
NATACHA NIEMANTS, UNIVERSITY OF BOLOGNA, UNIVERSITY OF MACERATA	
Looking up phrasal verbs in small corpora of interpreting: an attempt to draw out aspects of interpreted language .....	16
ANDREW J. CRESSWELL, UNIVERSITY OF BOLOGNA	
Politics interpreted on screen. A corpus-based investigation on the interpretation of televised political discourse.....	18
EUGENIA DAL FOVO, UNIVERSITY OF TRIESTE	
Distancing vs. simplification: where are the interpreter's priorities? .....	20
BART DEFRANCO, GHENT UNIVERSITY	
Discourse markers in police interpreting: illusions of invisibility .....	22
FABRIZIO GALLAI, UNIVERSITY OF BOLOGNA	
The Forlì final exams in interpreting: an invaluable corpus for interpreting research.....	24
CHRISTOPHER J. GARWOOD, UNIVERSITY OF BOLOGNA	
Corpus-based Interpreting Studies in Brazil: the design and retrieval of CEIS – Simultaneous Interpreting Learners' Corpora.....	26
LUCIANA GINEZI, UNIVERSITY OF SAO PAULO AND UNIVERSITY NOVE DE JULHO	

Interpreting vs. non-native language use .....	27
MARTA KAJZER-WIETRZNY, ADAM MICKIEWICZ UNIVERSITY IN POZNAŃ	
Cognitive Processing of Culture-Specific Items in Consecutive and Simultaneous Interpreting: A Corpus-Assisted Analysis.....	29
DEFENG LI, UNIVERSITY OF MACAU	
VICTORIA L.C. LEI, UNIVERSITY OF MACAU	
XIAODONG LIU, UNIVERSITY OF MACAU	
Panel: Benefits of working with corpora in community interpreting research: from qualitative analysis to quantitative verification – and back.....	30
CONVENORS: BERND MEYER, MAINZ UNIVERSITY & THOMAS SCHMIDT, MANNHEIM UNIVERSITY	
When and why do interpreters and translators at the EU Parliament use collocations? A target-oriented, bidirectional study of EPTIC. ....	31
MAJA MILIČEVIĆ, UNIVERSITY OF BELGRADE	
ADRIANO FERRARESI, UNIVERSITY OF BOLOGNA	
SILVIA BERNARDINI, UNIVERSITY OF BOLOGNA	
The development of technological resources based on exploitation of corpora for court interpreters in Spanish-Romanian, Arab, Chinese, French and English language pairs.....	33
MARIANA OROZCO-JUTORÁN, UNIVERSITAT AUTONÓMA DE BARCELONA	
<i>Over-uh-load</i> . The occurrence of <i>uh(m)</i> between elements of compounds during interpreting.....	34
KOEN PLEVOETS & BART DEFRANCO, GHENT UNIVERSITY	
Lexical patterns and gender: trends emerging from EPIC .....	36
MARIACHIARA RUSSO, UNIVERSITY OF BOLOGNA	
Mitigation and hedging in interpreter-mediated football press conferences: a case study.....	38
ANNALISA SANDRELLI, UNINT ROMA	
Studying figurative language in simultaneous interpreting: the IMITES ( <i>Interpretación de la Metáfora entre Italiano y Español</i> ) corpus .....	40
NICOLETTA SPINOLO, UNIVERSITY OF BOLOGNA	
Corpus-based Interpreting Studies in China.....	42
BINHUA WANG, THE HONG KONG POLYTECHNIC UNIVERSITY	
A Corpus-based Approach to Comparing the Use of Mandarin Discourse Particles in Spontaneous Speech and Simultaneous Interpreting.....	43

TA-WEI DAVID WANG, NEWCASTLE UNIVERSITY

POSTER PRESENTATIONS

Designing a multimodal corpus of English as a Lingua Franca (ELF) in Language for Specific Purpose (LSP) for the interpreting classroom: Challenges and Issues ..... 45

CHIARA BARBAGIANNI, SSML “A. MACAGNO” – CUNEO

VALENTINA BASELLI, IULM MILAN

CLARA PIGNATARO, IULM MILAN

Towards a purpose-specific interpreting corpus: *Anglintrad* and its possible exploitation for interpreters’ training..... 47

MICHELA BERTOZZI, UNIVERSITY OF BOLOGNA

A study of anaphoric relations in simultaneous interpreting..... 49

ANA CORREIA, UNIVERSITY OF MINHO

How can interpreting corpora extend our knowledge on intrusive “we” in SI? ..... 51

BARBARA DUMARA, UNIVERSITY OF WARSAW

Interpreting Universals: a study on explicitness in the intermodal EPTIC corpus..... 55

NICCOLÒ MORSELLI, UNIVERSITY OF BOLOGNA

Generating and compiling an interpreting corpus for a PhD dissertation: Challenges of corpus design, accessibility, data collection and processing ..... 56

DANUTA PRZEPIÓRKOWSKA, GESIS - LEIBNITZ INSTITUTE FOR THE SOCIAL SCIENCES / INSTITUTE OF APPLIED LINGUISTICS UNIVERSITY OF WARSAW

Interpreting Teaching in Libya: Challenges and Future Insights from a corpus-based perspective..... 58

MOHAMMED ZAGOOD, EL-MERGIB UNIVERSITY, KHOMS

What have we learned from CSI and what can others learn? ..... 59

JANA ZIDAR FORTE, UNIVERSITY OF LJUBLJANA

# PAPER PRESENTATIONS

# ACQUIRING THE LANGUAGE OF INTERPRETESE

GUY ASTON, UNIVERSITY OF BOLOGNA

[guy@sslmit.unibo.it](mailto:guy@sslmit.unibo.it)

There is a long tradition in the teaching of interpreting by such legendary figures as Ilg and Snelling of emphasis on useful formulae, phraseological items, multi-word expressions, lexical bundles or lexicalised sentence stems, to cite just some of the terms used in the linguistics literature to describe these features. It can be argued that these facilitate interpreters' fluency, reducing what Gile terms their "production effort" (Henriksen 2007; Aston 2015), thereby allowing greater attention to be dedicated to terminological issues (Gile, personal communication). This paper focuses on such items in a small (50k word) corpus of interpreting from the English booth at the European Parliament, and argues that while vastly larger corpora will be required in order to draw up even a tentatively satisfactory list of expressions to teach trainees in a syllabus, even small corpora can be used to draw learners' attention to the pervasiveness of formulaic items, enhancing their awareness of them, and providing them with a methodology for their acquisition.

Aston, G. 2015. "Corpus use and learning to translate - and to interpret?". Plenary lecture at AIETI7 conference, Malaga.

Henriksen, L. 2007. The song in the booth: Formulaic interpreting and oral textualisation. *Interpreting* 9(1): 1-20.

# CORPUS-BASED INTERPRETING STUDIES (CIS): PAST, PRESENT AND FUTURE DEVELOPMENTS OF A (WIRED) COTTAGE INDUSTRY

CLAUDIO BENDAZZOLI, UNIVERSITY OF TORINO, DEPARTMENT OF ECONOMIC AND SOCIAL STUDIES

[claudio.bendazzoli@unito.it](mailto:claudio.bendazzoli@unito.it)

Over the last ten years the corpus-based approach (Laviosa 1998) has been gaining momentum in both translation and, though to a lesser extent, interpreting studies, and has led to the creation of a number of language resources, i.e. corpora, across different interpreting modes and settings. In a comprehensive review of corpus-based research projects, Robin Setton (2011) highlighted the main features of CIS starting with Oléron and Nanpon's work in 1965, i.e. well before the seminal paper published by Miriam Shlesinger in 1998. At least seven more projects carried out before the new millennium were listed in the same review, but it is unlikely that such early corpora were machine-readable, and probably corpus linguistic tools were not part of those studies. On the other hand, it can be noted that later projects were based on fully machine-readable transcripts and, in many cases, took advantage of automatic extraction of occurrences. This distinguishing factor, due to the impact of information technology on CIS, is critical when having to define what is meant by 'corpus', a term that may trigger conflicting expectations even among experts. This is the first question I intend to raise in the present paper, so as to arrive at a clear understanding of CIS with respect to their past, present, and future developments. To this end, Setton's review of CIS, which he defined as "still a cottage industry" (*ibid.* 38), will be expanded to include subsequent works by scholars from both western and eastern research institutions, thus pointing out interesting trends in data collection and access (i.e. sharing a corpus once it has been created). Data collection is linked to access to and recording of interpreter-mediated communicative situations while data access and sharing can be linked to the process of transcription and annotation, which are (still) fundamental challenges in Interpreting Research. Considering the constant and growing prominence of information technology in CIS, such challenges will be contrasted with the fundamental features of Web 2.0, such as using the Internet as platform, content creation and sharing, collaborative practices and interaction, letting structures emerge without imposing patterns in advance and so on. The relevance of Web 2.0-induced effects has expanded from information technology to many other fields, as is the case with business and the so called Enterprise 2.0, which "is all about using technology to bring brains together effectively" (McAfee 2009: 16). Can the same be applied to Interpreting Studies? How can CIS benefit from the potential provided by Web 2.0? Possible answers to this question will shed light on effective future directions of CIS and their role in Interpreting Research and training.

## References

- Laviosa, Sara (ed.) (1998) *The Corpus-based Approach*. *Meta* 43/4. <http://www.erudit.org/revue/meta/1998/v/n4/> (accessed 7 January 2015)
- McAfee, Andrew (2009) *Enterprise 2.0. New Collaborative Tools for your Organization's Toughest Challenges*. Boston MA: Harvard Business Press.



Oléron, Pierre & Nanpon, Hubert (1965) Recherches sur la traduction simultanée. *Journal de Psychologie Normale et Pathologique* 62, 73-94.

Setton, Robin (2011) Corpus-based Interpreting Studies (CIS): Overview and prospects. In A. Kruger, K. Wallmach & J. Munday (eds.) (2011) *Corpus-based Translation Studies. Research and Applications*. London/New York: Continuum, 33-75.

# FROM QUANTITY TO QUALITY: USING CORPUS-BASED METHODS TO ANALYSE INTERPRETING PERFORMANCE

SABINE BRAUN, UNIVERSITY OF SURREY

[s.braun@surrey.ac.uk](mailto:s.braun@surrey.ac.uk)

This presentation will discuss the development and use of a corpus of interpreter-mediated police interviews to study differences between onsite and remote interpreting in legal settings. The study was prompted by the increasing use of remote interpreting in legal proceedings worldwide and the questions of viability and reliability that this method of interpreting raises.

The aim of the study required a comparative approach and led to the development of a corpus consisting of (simulated) bilingual police interviews and their renditions by multiple interpreters using different methods of interpreting (onsite and remote). The corpus includes video recordings of the interpreter-mediated events along with time-aligned transcripts and has been annotated with regard to interpreting problems. It has been compiled and annotated using Transana, a tool which has been developed primarily for the qualitative analysis of multimodal data, but which also provides for a basic quantitative analysis. The tool supports the integration of multiple video clips (e.g. multiple recordings of the same event with different cameras from different perspectives) and multiple transcripts.

The presentation will first discuss the development of the corpus design. Particular attention will be paid to existing corpus typologies and how they could be extended in order to include bilingual dialogue interpreting corpora, for which traditional labels such as 'parallel corpus' and traditional distinctions such as that between uni- and bidirectional corpora may need to be rethought.

The second part of the presentation will discuss different ways of exploiting the corpus to analyse and compare interpreting performance (interpreting quality) in onsite and remote interpreting. Although Zanettin (2012: 31) advocates qualitative approaches to the analysis of dialogue interpreting corpora, it will be argued that quantitative corpus-based methods can be fruitfully applied to dialogue interpreting corpora, especially when combined with qualitative approaches. By way of exemplification, it will be shown how a quantitative analysis based on the identification and annotation of interpreting problems was used to draw out important differences between onsite and remote interpreting and to establish the frequency of different types of problem. It will then be shown how a complementary qualitative analysis including a comparative analysis of different interpreted versions was used to corroborate, enrich and, equally important, challenge the quantitative findings; to identify strategies employed by the interpreters (with varying degrees of success) to overcome problems; and ultimately to paint a more comprehensive picture of the interpreters' performance than a quantitative analysis alone would have revealed.

The presentation will conclude with a reflection upon the main benefits and limitations of the approach reported, and a brief outline of considerations regarding corpus design and tools required to support further uses of dialogue interpreting corpora.

# LOST IN INTERPRETING OR...MAYBE NOT STRINGS OF NOUNS AS A CHALLENGE FOR SIMULTANEOUS INTERPRETERS WORKING FROM POLISH INTO ITALIAN

RITA CAPPELLI, UNIVERSITY OF BOLOGNA

[cappellirita@gmail.com](mailto:cappellirita@gmail.com)

Simultaneous interpreting from Polish into Italian has not been yet the object of many studies, even if this language combination is in demand since Poland joined the EU in 2004. This study aims at analysing the strategies used by experienced interpreters in order to overcome those difficulties which are peculiar for this language combination. In order to identify these elements both a contrastive analysis and a survey among interpreters working for the European Institutions were carried out. The contrastive analysis pointed out the main differences existing between the two language systems, namely: the inflection of the nouns, the free position of the elements in the sentence, the existence of the verbal aspect and of only three verbal tenses, the absence of the definite and indefinite articles and the long strings of nouns typical of the Polish syntax. The twelve Italian interpreters who took part in the survey mentioned the main difficulties they encountered interpreting from Polish into Italian, that is to say: the existence of seven cases (some of them with the same endings) the long string of nouns, the many tenses existing in Italian without equivalent in Polish, the presence of the verbal aspect in Polish and the complex Polish phonetics, where similar sounds have opposite meaning. In the light of the results of the contrastive analysis and of the survey, it seemed appropriate to focus our attention on the strings of nouns, since they represent a peculiar element of the Polish language and at the same time they include other features typical of this Slavic language. In order to carry out this study an ad-hoc corpus has been created. This bilingual corpus is composed of 313 speeches held in Polish during the plenary sessions of the European Parliament in the year 2011 and during the first semester of the year 2009 and their respective interpretations into Italian (around 560 minutes for each language were transcribed). The transcriptions were carried out following the guidelines developed and applied by the research group of the Sitlec Department of Forlì to compile the EPIC corpus (Monti, Bendazzoli, Sandrelli & Russo, 2005). The corpus can be investigated as a bilingual corpus in word or as two monolingual corpora using the software UAM Corpus Tool. Once the concept of string of nouns had been defined, all the string of nouns (312) that met the basic requirements (noun + two nouns in the genitive case + adjective or noun + three nouns in the genitive case) were extracted and listed in an excel document. Their respective renditions into Italian were then categorised according to Wadensjö's classification of renditions (1998) into close, substituted, reduced, expanded and zero rendition. Also the category divergent rendition introduced by Amato and Mack (2011) in their work has been considered. This represents the first study and corpus dedicated to the language combination PL>IT aiming at providing both students and trainers with practical guidelines in order to help them to overcome the hurdles posed by the dissimilarities of these different language systems.

## References

Amato, A. & G. Mack (2011) "Interpreting the Oscar Night on Italian TV: an interpreters' nightmare?" In *The Interpreters' Newsletter*, 16, 37-60.

- Bartłomiejczyk, M. (2006) "Strategies of simultaneous interpreting and directionality". In *Interpreting*, 8 (2), 149-74.
- Gile, D. (1995) "Interpretation research: a new impetus?". In *Hermes*, 14, 15-29.
- Gile, D. (1999) "Testing the effort models' tightrope hypothesis in simultaneous interpreting: a contribution" In *Hermes*, 23, 153-172.
- Gile, D. (2009) *Basic concepts and models for interpreter and translator training*. Amsterdam/Philadelphia: John Benjamins.
- Gile, D. (2011) "Errors, omissions and infelicities in broadcast interpreting. Preliminary findings from a case study". In Alvstad, C. & A. Hild & E. Tiselius (eds.), *Methods and strategies of process research*. Amsterdam/Philadelphia: John Benjamins, 201–218.
- Monti, C. & Bendazzoli, C. & Sandrelli, A. & M. Russo (2005) "Studying directionality in simultaneous interpreting through an electronic corpus: EPIC (European Parliament Interpreting Corpus)". In *Meta*, 50 (4).
- Pym, A. (2005) "Explaining explicitation". In A. Karoly & Á. Fóris (eds.) *New Trends in Translation Studies. In honour of Kinga Klaudy*. Budapest: Akadémia Kiadó, 29-34.
- Pym, A. (2008) "On omission in simultaneous interpreting. Risk analysis of a hidden effort". In Hansen, G. & A. Chesterman & H. Gerzymisch-Arbogast (eds) *Efforts and Models in Interpreting and Translation Research*. Amsterdam/Philadelphia: John Benjamins, 83-105.
- Riccardi, A. (1999) "Interpretazione simultanea: strategie generali e specifiche", in C. Falbo, M. Russo, F. Straniero Sergio (a cura di), *Interpretazione simultanea e consecutiva. Problemi teorici e metodologie didattiche*, Milano, Hoepli, 161-173.
- Setton, R. (1999) *Simultaneous interpretation: A cognitive-pragmatic analysis*. Amsterdam/Philadelphia: John Benjamins.
- Seeber, K.G. (2011) "Cognitive load in simultaneous interpreting: Existing theories - new models". In *Interpreting* 13 (2), 176-204.
- Straniero Sergio F. (1997) "La produzione della produzione prodotta ovvero tautologie e catene nominali nell'interpretazione simultanea dal russo in italiano. In L. Gran & A.
- Riccardi (a cura di) *Nuovi orientamenti negli studi sull'interpretazione*. Trieste: Università degli Studi, 159-177.
- Straniero Sergio, F. & Falbo, C. (eds.) (2012) *Breaking Ground in Corpus-based Interpreting Studies*. Bern: Peter Lang.
- Viaggio, S. (1991) "Teaching Beginners to Shut Up and Listen". In *The interpreters' Newsletter*, 4, 45-58.
- Wadensjö, C. (1998) *Interpreting as Interaction*. London/New York: Longman, Language in Social Life Series.



# CORPORA WORTH CREATING: A PILOT STUDY ON TELEPHONE INTERPRETING

SARA CASTAGNOLI, UNIVERSITY OF BOLOGNA  
[s.castagnoli@unibo.it](mailto:s.castagnoli@unibo.it)

NATACHA NIEMANTS, UNIVERSITY OF BOLOGNA, UNIVERSITY OF MACERATA  
[natacha.niemants@unibo.it](mailto:natacha.niemants@unibo.it)

Remote interpreting (RI) is a new form of interpreted-mediated communication where some or all the participants (here doctors, patients and interpreters) do not meet face-to-face and where interpreting is provided through a medium such as the telephone, videoconferencing, the internet, etc. RI is rapidly gaining momentum within public services, especially in healthcare and court settings, where RI technologies can help address a number of financial and logistical issues, e.g. locating qualified interpreters for minority languages and 24/7 availability (Ko 2006, Rosenberg 2007, Andres & Falk 2009). Despite the numerous advantages it offers vis-à-vis traditional public service interpreting (PSI) *in presentia*, RI is assumed to be harder for service users and providers, who arguably need to be trained to the different and additional challenges of interpreting *in absentia* (see e.g. Wadensjö 1999, Braun 2006, Kelly 2008, Hlvac 2013).

The availability of authentic data on RI is extremely limited compared with the (slowly) growing number of corpus-based studies of PSI (Vargas Urpi 2012), with which RI research shares a number of technical/practical problems and methodological concerns. To quote but a few, (a) the difficulty of accessing data and getting permission to use them for scientific purposes, which impacts on corpus design and representativeness, and ultimately on researchers' objectives; (b) the time-consuming nature of data collection and transcription, which limits corpus size and also influences analysis; (c) the problem of dealing with dialogic data including both monolingual and interpreted utterances, and characterised by overlaps and other conversational phenomena which can hardly be annotated or extracted automatically (Angermeyer et al. 2012, Straniero Sergio & Falbo 2012).

Against this background, the paucity of data available to the scientific community and the delay in the development of corpus-based PSI/RI studies are not surprising. However, we believe researchers can use available data to formulate and start testing hypotheses about the peculiarities of PSI/RI, and to justify the efforts needed to obtain larger and more full-fledged collections.

In this presentation we will introduce a pilot collection of telephone interpreting data, consisting of the transcriptions of 30 interpreted-mediated phone calls made at a local health authority of the Emilia Romagna region. As the collection was originally assembled for (non-linguistic) quality control purposes, we will first illustrate the difficulties we faced in accessing the relevant audio files and making do with the basic, inconsistent orthographic transcriptions provided. We will then analyse the preliminary insights that this small-scale *spoken* corpus can offer in relation to interactional aspects (such as meaning negotiation, turn-taking and topic management) that can be observed even through qualitative investigations of raw transcripts; and we will contrast them with what can additionally be observed, also at the linguistic level, when exploiting a larger *speech* corpus with aligned audio tracks.

By combining notions from Interpreting Studies and Conversation Analysis, we will identify similar and divergent courses of interaction in PSI *in absentia* and *in presentia*, and pave the way to further investigations.

## References

- Andres, D. & S. Falk (2009) "Remote and Telephone Interpreting". In D. Andres & S. Pöllabauer (eds), *Spürst Du wie der Bauch rauf runter?/Is everything all topsy turvy in your tummy? – Fachdolmetschen im Gesundheitsbereich/Health Care Interpreting*. München: Martin Meidenbauer. 9-27.
- Angermeyer, P.S., B. Meyer & T. Schmidt (2012) "Sharing community interpreting corpora. A pilot study". In T. Schmidt & K. Wörner (eds), *Multilingual Corpora and Multilingual Corpus Analysis*. Amsterdam/ Philadelphia: John Benjamins. 275-294.
- Braun, S. (2006) "Multimedia communication technologies and their impact on interpreting". In M. Carroll, H. Gerzymisch-Arbogast & S. Nauert (eds), *Audiovisual Translation Scenarios*. Proceedings of the Marie Curie Euroconference "MuTra: Audiovisual Translation Scenarios", Copenhagen, 1-5 May 2006. Online: [http://epubs.surrey.ac.uk/1291/1/Braun\\_2006\\_Mutra\\_final\\_draft.pdf](http://epubs.surrey.ac.uk/1291/1/Braun_2006_Mutra_final_draft.pdf) (last access: 09 Jan 2015).
- Hlvac, J. (2013) "Should Interpreters Be Trained and Tested in Telephone and Video- Link Interpreting? Responses from Practitioners and Examiners". *International Journal of Interpreter Education* 5(1): 34-50.
- Kelly, N. (2008) "A Medical Interpreter's Guide to Telephone Interpreting". International Medical Interpreters Association. Retrieved from: <http://www.imiaweb.org/uploads/pages/380.pdf> (last access: 09 Jan 2015).
- Ko, L. (2006) "The need for long-term empirical studies in remote interpreting research: A case study of telephone interpreting". *Linguistica Antverpiensia* 5: 325-338.
- Rosenberg, B.A. (2007) "A data driven analysis of telephone interpreting". In C. Wadensjö, B.E. Dimitrova & A. Nilsson (eds), *The critical link 4: Professionalisation of interpreting in the community*. Amsterdam/ Philadelphia: John Benjamins. 65-75.
- Straniero Sergio, F. & C. Falbo (2012) "Studying interpreting through corpora. An introduction". In F. Straniero Sergio & C. Falbo (eds), *Breaking ground in corpus-based interpreting studies*. Frankfurt: Peter Lang. 9-52.
- Vargas Urpi, M. (2012) "State of the art in community interpreting research. Mapping the main research topics". *Babel* 58(1): 50-72.
- Wadensjö, C. (1999) "Telephone Interpreting & the Synchronization of Talk in Social Interaction". *The Translator* 5(2): 247-264.

# LOOKING UP PHRASAL VERBS IN SMALL CORPORA OF INTERPRETING: AN ATTEMPT TO DRAW OUT ASPECTS OF INTERPRETED LANGUAGE

ANDREW J. CRESSWELL, UNIVERSITY OF BOLOGNA

[andrewjohn.cresswell@unibo.it](mailto:andrewjohn.cresswell@unibo.it)

Even at the time when Schlesinger (1998) called for the interpreting community to embrace corpus technologies, there was - alongside the assumption that vast amounts of data were needed - a growing awareness that interesting results and applications can be derived from corpora that are small (e.g. Aston 1997). This current has since become stronger in both practice (e.g. Ghadessy et al. 2001, Upton and Cohen 2009), and theory (e.g. Flowerdew 2004). Small corpus studies are specialised, as Flowerdew (2004) argues, because they are usually oriented towards uncovering and documenting information about particular language phenomena or particular genres. In the study of interpreted language, where large corpora are hard to assemble because of the time-consuming transcription process (see Bendazzoli 2010:66), concentrating research on small, specialised corpora should be particularly profitable.

This research uses the text-retrieval program *Wordsmith Tools 6* to examine the occurrence of one particular area of English lexicogrammar, phrasal verbs, in several small annotated corpora of simultaneously interpreted language. The first of the corpora providing texts for this project is DIRSI (Bendazzoli 2010), compiled from conference interpretations from Italian by native-speaker and non-native speaker instructor-interpreters. The second contains interpretations from Spanish and Italian from EPIC, the corpus of European parliament interpretations (Monti et al. 2005, Sandrelli et al. 2010). The third is the English interpreted language component of "2249", a corpus of the English used in a single day at the European parliament, which I was given by Guy Aston, about which no research has yet to my knowledge been published. English phrasal verbs were chosen as the focus of study, first, because they are an exemplar of the formulaic language which marks out fluent speakers of a language (see, e.g., Wray 2000), and secondly, because they are notably hard for non-natives to acquire (Darwin & Gray 1999:65). For these two reasons it was thought that their relative range and frequency would reflect directionality, i.e. that in A direction texts in English there would be higher frequencies of both types and tokens than there would be in B direction texts in English.

Comparison of the frequencies of phrasal verb types in the interpreted language corpus with frequencies in a large reference corpus suggests the extent to which these frequencies correspond to expectations of the English language in general. Wordsmith's key word in context function is used to help identify the functions of phrasal verbs in the interpreted text, and whether these functions are associated (a) exophorically, either with the translated subject matter or with the conference talk genre, or (b) endophorically, either with the communication of discourse relations, or with management of the act of interpreting itself. In this way, an attempt is made to build up a picture of a small area of interpreted language and bring us to the point where it is possible to discuss its degree of resemblance to or divergence from "the norm", and whether there are suggestions in the interpreted language of A-direction interpreters that B-direction interpreters might usefully emulate.

## References



- Aston, G. 1997. Small and large corpora in language learnin'. Paper presented at the PALC conference, Lodz. <http://www.sslmit.unibo.it/~guy/wudj1.htm>
- Bendazzoli, C. 2010. *Corpora e Interpretazione Simultanea*. Alessandria. Asterisco Edizioni.
- Darwin, C. M., & Gray, L. S. (1999). Going after the phrasal verb: An alternative approach to classification. *TESOL Quarterly*, 33, 65–83.
- Flowerdew, Lynne 2004. The Argument for Using English Specialized Corpora to Understand Academic and Professional Language. In Connor, Ulla and Upton, Thomas (eds), *Discourse in the Professions*. Amsterdam: Benjamins, 11-33.
- Ghadessy, M., Henry, A., and Roseberry, R. (Eds). (2001). *Small Corpus Studies and ELT: Theory and Practice*. Amsterdam: John Benjamins.
- Monti, C., Bendazzoli, C., Sandrelli A. and M. Russo (2005). "Studying directionality in simultaneous interpreting through an electronic corpus: EPIC (European Parliament Interpreting Corpus)", *Meta*, 50 (4).
- Sandrelli, A., Bendazzoli, C. and M. Russo (2010). European Parliament Interpreting Corpus (EPIC): Methodological issues and preliminary results on lexical patterns in SI. *International Journal of Translation* 22 (1-2), 165-203.
- Shlesinger, M. (1998). "Corpus-based Interpreting Studies as an Offshoot of Corpus-based Translation Studies". *Meta*, 43 (4), 486-493.
- Upton, T. and Cohen, M.A. (2009). An approach to corpus-based discourse analysis: The move analysis as an example. *Discourse Studies*, 11/5:585-605
- Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics*, 21, 463–489.

# POLITICS INTERPRETED ON SCREEN. A CORPUS-BASED INVESTIGATION ON THE INTERPRETATION OF TELEVISED POLITICAL DISCOURSE

EUGENIA DAL FOVO, UNIVERSITY OF TRIESTE

[edalfovo@units.it](mailto:edalfovo@units.it)

The present paper illustrates a corpus-based analysis conducted on two multimedia parallel corpora, DEB.04 and OB.0811. Both corpora are collections of televised U.S. political speech events and their interpretations broadcast live on Italian television, and more specifically the three Bush/Kerry presidential debates, and the single Cheney/Edwards vice-presidential debate held during the 2004 U.S. presidential campaign in DEB.04 (ca. 17h of recordings), and a series of President Obama's speeches and remarks – both OTs and the relevant ITs, ranging from 5 to 45 min duration and addressing various topics of both domestic and foreign policy. Both corpora are sub-corpora of CorIT, the Italian Television Interpreting Corpus (Straniero Sergio & Falbo 2012) created by Straniero Sergio at the DIUSLIT of the University of Trieste. The aim of the study is identifying discourse elements in the OTs which operationalise the principle of infotainment, i.e., the merging of agendas between politics and television, and then verifying the presence –or absence– of such elements in the ITs broadcast on Italian TV.

The presentation starts with an overview of television interpreting (TI) as professional activity and research area: capitalizing on early contributions to the literature (*inter al.* Kurz 1985, 2003; Alexieva 1996, 2001; Pöchhacker 1997; Mack 2001; Bros-Brann 2002), the analyst reflects on issues such as constraints and setting-related factors that subsequently led scholars to identify specific norms and strategies, and eventually new quality standards and criteria to be applied to this particular field of reference. The presentation moves on to provide an illustration of the CorIT corpus (Falbo 2012) and the multiple and unique research opportunities it has provided in the past decade, both for researchers and MA students involved in the project (Dal Fovo 2011). In the second part of the presentation, data and analysis are presented, with particular attention devoted to the elaboration of the templates of analysis, based mainly on studies on conversation and discourse analysis (cf. Halliday & Hasan 1987; Heritage & Greatbatch 1991; Maley & Fahey 1991). The main rhetorical and structural aspects and features of political discourse are analysed and compared with aspects pertaining to television discourse and infotainment ideology. Outcomes are discussed in terms of significance of entertainment devices, and the extent, to which their presence or absence in the ITs affects what is said in the original and how it is said.

By means of conclusion, a broader view of the matter is taken into consideration, namely the implications of TV communication and relational constellations on interpreters' specific choices and behaviours in terms of discourse and translation attitudes and tendencies (cf. Katan & Straniero Sergio 2003). The preliminary outcomes suggest that, despite the considerable amount of research conducted in this field, TI still remains a very elusive subject, whose multi-faceted nature and diverse expressions have yet to be sufficiently identified and defined. As anticipated by Shlesinger (1998) in her seminal work fifteen years ago, and as has already been the case in numerous areas of interpreting studies, the corpus-based approach might prove a decisive tool in order to address and successfully answer some of these questions.

## References

- Alexieva, B. 1996. *Interpreting Mediated TV Events*. In Klaudy, K. and Kohn, J. (eds.) *Transferre Necessesse Est*. Budapest: Scholastica, 171-174.
- Alexieva, B. 2001. Interpreter-Mediated TV Live Interviews. In Gambier, Y. and Gottlieb, H. (eds.) *(Multi)media Translation. Concepts, Practices and Research*. Amsterdam/Philadelphia: John Benjamins, 113-124.
- Bros-Brann, E. 2002. Simultaneous interpretation and the media: interpreting live for television. <http://aiic.net/page/630/simultaneous-interpretation-and-the-media-interpreting-live-for-television/lang/1> (accessed on 05.01.2014).
- Dal Fovo, E. 2011. Through the CorIT looking glass - and what MA students found there. *The Interpreters' Newsletter* 16, Special Issue on Television Interpreting, 1-20.
- Falbo, C. 2012. CorIT (Television Interpreting Corpus): Classification Criteria. In Straniero Sergio, F. & Falbo, C. (eds.) *Breaking Ground in Corpus-based Interpreting Studies*. Bern: Peter Lang, 155-186.
- Halliday, M.A.K. & R. Hasan 1987. *Cohesion in English*. English Language Series; London/New York: Longman.
- Heritage, J. & D. Greatbatch 1991. On the institutional character of institutional talk: the case of news interviews. In D. Boden and Zimmerman, D.H. (eds.), *Talk and social structure*. Berkley: University of California Press, 93-137.
- Katan, D. & F. Straniero Sergio 2003. Submerged ideologies in Media Interpreting. In M. Calzada Perez (eds.), *Apropos of ideology*. Manchester: St. Jerome, 131-144.
- Kurz, I. 1985. Zur Rolle des Sprachmittlers im Fernsehen. In Bühler, H. (ed.) *Translators and their position in society*. Xth World Congress of FIT, Proceedings. Vienna: Braumüller, 213-215.
- Kurz, I. 2003. Live TV interpreting – A high-wire act?. In Collados Aís, À. and Sabio Pinilla, J. A. (ed.) *Avances en la investigación sobre interpretación*, Granada, Comares, 159-171.
- Mack, G. 2001. Conference Interpreters on the Air - Live Simultaneous Interpreting on Italian Television. In Gambier, Y. and Gottlieb, H. (eds.) *(Multi)Media Translation. Concepts, Practices and Research*. Amsterdam/Philadelphia: John Benjamins, 125-132.
- Maley, Y. & R. Fahey 1991. Presenting the evidence: Constructions of reality in court. *International Journal for the Semiotics of Law*, IV (10): 3-17.
- Pöchhacker, F. 1997. Clinton speaks German: a case study of live broadcast simultaneous interpreting. In Snell-Hornby, M., Jettmarová, Z. and Kaindl, K. (eds.) *Translation as intercultural communication. Selected papers from the EST Congress, Prague 1995*. Amsterdam/Philadelphia: John Benjamins, 207-216.
- Shlesinger, M. 1998. Corpus-based Interpreting Studies as an offshoot of Corpus-based Translation Studies. *Meta* 43 (4): 486-493.
- Straniero Sergio, F. & C. Falbo (eds.) 2012. *Breaking Ground in Corpus-based Interpreting Studies*. Bern: Peter Lang.

# DISTANCING VS. SIMPLIFICATION: WHERE ARE THE INTERPRETER'S PRIORITIES?

BART DEFRANCO, GHENT UNIVERSITY

[bart.defrancq@ugent.be](mailto:bart.defrancq@ugent.be)

The present paper aims at confronting two claims found in the literature on interpreter's linguistic strategies and to compare these with translator's strategies in a parallel intermodal corpus (PIC). The claims are as follows:

- (1) Simultaneous interpreters work under extreme cognitive load and, as a result, seek to simplify production by simplifying syntactic structures (Ilg 1978; Setton 1999; Seeber 2011)
- (2) Simultaneous interpreters have face concerns and, as a result, have a tendency to distance themselves from the speaker of the source text (Monacelli 2009)

I will investigate the two claims with data for which I believe they are contradictory: the use of causal connective items by Dutch interpreters. Indeed, Dutch possesses several subordinate causal connectives (*omdat, aangezien, doordat...*) and one coordinate causal connective (*want*). As word order in Dutch differs substantially between main (or coordinated) and subordinate clauses, claim (1) predicts that interpreters will transform subordinate clauses into coordinated clauses, so as to avoid the complex syntax of verb-final subordinate clauses in their production. As a consequence, the frequency of the coordinate connective (*want*) will be higher than that of its French counterpart in the source texts (*car*).

On the other hand, according to Degand & Maat (2003), Pit (2003) and Pit (2007), causal connectives can be distinguished by different degrees of subjectivity: Dutch *want* and *aangezien* score high on the subjectivity scale, whereas *omdat* and *doordat* are closer to the opposite end of the scale. In Pit's analysis, subjectivity is based on particular features of the causal participant (i.e. the participant affected by the causal relationship). From Monacelli's claim that interpreters have a tendency to distance themselves from the speaker, it follows that the interpreters' choice of connective items is likely to be less subjective than that of speakers. Frequencies of objective causal connectives (*omdat, doordat*) should therefore be higher in the interpreted texts than in the source texts.

The hypotheses are contradictory, as the coordinate connective *want* is expected at the same time to be more frequent in interpretations because it allows syntactic simplification, and to be less frequent because it is highly subjective.

In order to find out which tendency is given priority by interpreters, I analysed the use of a set of causal connective items in a so-called parallel intermodal corpus (PIC), i.e. a corpus of authentic source and target texts, both spoken and written (interpreted and translated) and delivered in comparable contexts (Bernardini 2014). The corpus I used is based on data retrieved from sittings of the European Parliament held between 1<sup>st</sup> September 2008 and 21<sup>st</sup> October 2008. The original French speeches and their Dutch interpretations are drawn from EPICG (*European Parliament Interpreting Corpus Ghent*), a larger corpus consisting of transcriptions of speeches held during plenary sessions of the European Parliament and of their interpretations. In its current shape, it comprises approximately 190,000 words. Written data were collected from the verbatim reports of the same French source texts and their translations in Dutch. The translations were used to verify whether interpreters' strategies diverged from translator's strategies.

The results show that the frequency of the coordinate connective increases drastically in interpreters' performances, showing that the simplification of syntactic structure is given priority, in spite of increased subjectivity. This strategy sets interpreters apart from translators.

## References

- Degand, L & H. Pander Maat. 2003. 'A contrastive study of Dutch and French causal connectives on the Speaker Involvement Scale'. In A. Verhagen & J. van de Weijer (eds.) *Usage-Based Approaches to Dutch. Lexicon, grammar, discourse*. Utrecht : LOT.
- Ilg, G. 1978. L'apprentissage de l'interprétation simultanée. *Parallèles*, 1:69-99.
- Monacelli, C. (2009). *Self-Preservation in Simultaneous Interpreting*. Amsterdam: Benjamins.
- Pit, M. 2003. *How to express yourself with a causal connective*. Utrecht: USLC.
- Pit, M. 2007. 'Cross-linguistic analyses of backward causal connectives in Dutch, German and French'. *Languages in Contrast*, 7(1), pp. 53-82.
- Seeber, K. 2011. "Cognitive load in simultaneous interpreting: Existing theories – new models." *Interpreting* 13 (2): 176-204.
- Setton, R. 1999. *Simultaneous Interpretation: A cognitive-pragmatic analysis*. Amsterdam: Benjamins.

# DISCOURSE MARKERS IN POLICE INTERPRETING: ILLUSIONS OF INVISIBILITY

FABRIZIO GALLAI, UNIVERSITY OF BOLOGNA

[fabrizio.gallai@unibo.it](mailto:fabrizio.gallai@unibo.it)

Confusion appears to surround the definition of the police interpreter’s role by Police Studies scholars and authorities (e.g. Shepherd, 2007). A widespread assumption is literalism, whereby the formulaic questions in police interviews are expected to give rise to standard interpreter responses.

My analysis (Gallai, 2014; Blakemore & Gallai, 2014) of a large corpus of five interpreted police interviews (9 hours and 32 minutes) - involving four interpreters and two linguistic combinations (Portuguese-English and Italian-English) - shows that formal equivalence is untenable. In particular, I examine the overall effect of the police interpreter’s rendition of discourse markers such as *but*:

Police officer	Er: (.) it’s not normally too much of a problem really when we’ve got the- er an interpretation (.) taking place as well but (.) normally in this part I would ask you not to talk too fast
Interpreter	<p>°Okay° (.) er:: o Danny também está a tomar nota (.) do outro lado (.) e:: tá a to- a</p> <p><i>Danny is taking note as well (.) on the other side (.) and is ta-</i></p> <p>tomar notas da conversa (.) e normalmente ele ia-te pedir pra tu não falares muito</p> <p><i>taking notes of the conversation (.) and normally he would ask you not to talk too</i></p> <p>depressa (.)</p> <p><i>fast (.)</i></p>

In terms of methodology I follow the most recent trends in interpreting studies, as I adopt a qualitative approach, which rules out both an empiricist view relying mainly on factual evidence and a rationalist position which disregards data-based evidence in favour of theorising. Considering the short tradition of police interpreting studies, its analytical momentum has not been generated entirely from within the discipline’s existing paradigms. Rather, interpreting scholars have been looking to other disciplinary frameworks for relevant models and methods. My research aims at a dialectic synthesis of two theoretical frameworks: Sperber and Wilson’s (1986/1995) Relevance Theory, which offers a unified explanation across a large number of seemingly disparate uses of discourse “markers” (cf. Blakemore 2002); and Goffman’s (1981) analytical framework about the nature of social organisation, as first adopted in Wadensjö’s *Interpreting as Interaction* (1998).

On the basis of data transcribed, I show that interpreting practice is variable with respect to the approach to the inclusion of discourse markers: alongside cases in which the original speaker’s discourse marker is translated by the interpreter, there are examples in which:

the interpreter’s rendition does not include discourse markers found in the original;

the interpreter’s rendition includes discourse markers not found in the original but which are nevertheless understood as being attributed to the original speaker;

the interpreter's rendition includes a discourse marker not found in the original and which must be attributed to the interpreter rather than the original speaker.

My ultimate objective is that police interpreters and interviewers alike will draw out from this research the messages for practice development, implementation and operation.

## References

Blakemore, D. (2002) *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: CUP.

Blakemore, D. & Gallai, F. (2014). Discourse markers in free indirect style and interpreting. *Journal of Pragmatics*, 60, 106–120.

Gallai, F. (2014). *Discourse Markers in Interpreter-Mediated Police Interviews: an interdisciplinary approach*. University of Salford. (Ph.D. Thesis).

Goffman, E. (1981) *Forms of Talk*. Philadelphia: University of Pennsylvania

Shepherd, E. (2007) *Investigative Interviewing: The Conversation Management Approach*. Oxford: OUP.

Sperber, D. and Wilson, D. (1986/1995) *Relevance: Communication and cognition*. Oxford: Blackwell.

Wadensjö, C. (1998) *Interpreting as an interaction*. London: Longman.

# THE FORLÌ FINAL EXAMS IN INTERPRETING: AN INVALUABLE CORPUS FOR INTERPRETING RESEARCH

CHRISTOPHER J. GARWOOD, UNIVERSITY OF BOLOGNA

[christopher.garwood@unibo.it](mailto:christopher.garwood@unibo.it)

The use of corpora in interpreting studies has been growing steadily since Miriam Shlesinger highlighted the need for interpreting studies to follow the example of translation studies. Among the many problems in creating such corpora (Bendazzoli & Sandrelli 2005) is the limited access to authentic data (Pöchhacker, 2008), also because interpreters are not always willing to have their work assessed, making it “even harder to collect sufficiently large samples of representative and homogeneous interpreting data” (Bendazzoli & Sandrelli 2009). The first large-scale corpus of representative and homogeneous interpreting data, EPIC (European Parliament Interpreting Corpus), has resulted in many interesting studies in a variety of fields. This paper looks at another unique corpus in terms of homogeneity and the existence of parallel interpreted texts – the Final Exams of the University of Bologna’s Post-Graduate Course in Interpreting (Forlì). These exams are in the form of a one-day conference with authentic speakers (politicians, academics, experts etc) on a topic that is communicated to the students one month prior to the exams. Homogeneity is ensured by the fact that all the trainee interpreters have successfully completed their studies and by the format of the exams (the same length speeches for all candidates, the same Italian source speech interpreted simultaneously into various languages). The paper examines the post-graduate dissertations that have already been written using this mechanical corpus and examines other interesting areas of study, including interpreter training and assessment.

## References

- Bendazzoli, C. (2010). The European Parliament as a source of material for research into simultaneous interpreting: advantages and limitations. In Zybatow, L. N. a cura di. *Translationswissenschaft - Stand und Perspektiven*. Frankfurt: Peter Lang. 51-68.
- Bendazzoli, C. & Sandrelli, A. (2005-2007). “An approach to corpus-based interpreting studies: developing EPIC (European Parliament Interpreting Corpus)”, in Nauert S. (ed.) (2005-2007), *Proceedings of the Marie Curie Euroconferences MuTra: Challenges of Multidimensional Translation - Saarbrücken 2-6 May 2005*.
- Bendazzoli, Claudio / Sandrelli, Annalisa (2009). “Corpus-based Interpreting Studies: Early Work and Future Prospects”, *Revista Tradumàtica 07 L’aplicació dels corpus lingüístics a la traducció*. 2009. ISSN: 1578-7559.
- Monti, C., Bendazzoli, C., Sandrelli A. and M. Russo (2005). "Studying directionality in simultaneous interpreting through an electronic corpus: EPIC (European Parliament Interpreting Corpus)", *Meta*, 50 (4).



Russo, M. (2010). Reflecting on interpreting practice: graduation theses based on the European Parliament Interpreting Corpus (EPIC). In Zybatow, L. N. a cura di. *Translationswissenschaft - Stand und Perspektiven*. Frankfurt: Peter Lang. 35-50.

Shlesinger, M. (1998). "Corpus-based Interpreting Studies as an offshoot of Corpus-based Translation Studies", *Meta*, 43/4.

Setton, Robin. 2011. Corpus-based interpreting studies: Overview and prospects. In Alet Kruger, Kim Wallmach, Jeremy Munday (eds.) 2011. *Corpus-based Translation Studies. Research and Applications*. London & New York: Continuum.

Spinolo, N., 2014. Il linguaggio figurato e la cabina di simultanea: il progetto IMITES. PhD dissertation, University of Bologna.

# CORPUS-BASED INTERPRETING STUDIES IN BRAZIL: THE DESIGN AND RETRIEVAL OF CEIS – SIMULTANEOUS INTERPRETING LEARNERS' CORPORA

LUCIANA GINEZI, UNIVERSITY OF SAO PAULO AND UNIVERSITY NOVE DE JULHO

[luginezi@uol.com.br](mailto:luginezi@uol.com.br)

The aim of this presentation is to show the design of CEIS – Simultaneous Interpreting Learners' Corpora – and discuss if it answered the research question. We present part of a doctorate research, developed under University of Sao Paulo Translation Graduate Program, in which CEIS was retrieved to investigate interpreting training at Interpreting Undergraduate schools in the city of São Paulo (SP), Brazil. In SP there is a growing demand for interpreters and, as a consequence, a growing offer of Interpreting Undergraduate Schools. All Interpreting Training programs in SP were investigated, but only one out of seven was selected for the study, whose curriculum had at least 500 hours in Interpreting Training. The others offered only optional courses in Interpreting training, or less than 500 hours, so then they were excluded from this study, for quality reasons. This selected program has a curriculum focused in the different modes and situations of Interpreting, such as consecutive in Court Interpreting; consecutive in meetings, consecutive in patient and doctor's interaction, simultaneous in Conferences, Seminars etc. The courses take one semester each, and they follow a fixed order: sight translation as introductory mode, consecutive interpreting (question-answer, classical mode) and simultaneous. Although this is the most common order we may find around the world, as international literature in the area recommends Simultaneous Interpreting (SI) to be taught only after Consecutive Interpreting (CI), students spend most of their time taking sight-translation and consecutive lessons (around 70% of their whole program). The impact is that they don't have enough time to train simultaneous accordingly, as the simultaneous lessons take place only at the end of their program. Considering SI is the most required in SP, we decided to conduct an experimental study to investigate if students are able to learn SI before learning CI, or if Consecutive is prerequisite for Simultaneous. To answer this question, we designed parallel and comparative oral corpora, with interpreting students' rendition speeches, collected during their interpreting lessons, and additional classes for control group. The corpora are divided into two subcorpora, A and B, considering the same variables such as students' proficiency level, and no prior experience as interpreters. Group A (control group) is composed of second-semester students, who did not practice CI, but have already learned the modes in theory. Group B is formed by fourth-semester students, who practiced CI but not SI. Both groups perform the same task in SI, e.g. lectures taken from the web. The rendition of each group is recorded and transcribed, aligned with their original speeches and tagged, using different software to reach the results. Our study analysis shows the groups' rendition comparison, providing the interpreting strategies they have used, besides their lexical choices and the most common errors. From this analysis, we were able to make suggestions for curriculum changes and, besides, considered some further technological adjustments to CEIS for pedagogical use during Interpreting training lessons.

**Keywords:** Interpreting Training, Simultaneous Interpreting, Corpus Linguistics.

# INTERPRETESE VS. NON-NATIVE LANGUAGE USE

MARTA KAJZER-WIETRZNY, ADAM MICKIEWICZ UNIVERSITY IN POZNAŃ

[mkaizer@wa.amu.edu.pl](mailto:mkaizer@wa.amu.edu.pl)

KEYWORDS: interpretese, explicitness, corpus based interpreting studies, universals

Interpretese seems to be more spoken than translated (Shlesinger and Ordan 2012), and indeed it is hard to show in an undisputed manner and using the parameters traditionally applied to translation that the language of interpretation and the language of translation share the same characteristics. For example, Sandrelli and Bendazolli's (2005) analysis of lexical patterns showed that simplification was observed in one language pair, but not on another, and similar was the case with Kajzer-Wietrzny's (2012) investigations on yet different language pairs. But there is one parameter of explicitness (or explicitation, as originally suggested by Olohan and Baker 2000), which strongly distinguishes translations from non-translations and interpretations from non-interpretations – optional connective that.

Previous research shows that similarly to translations, simultaneous interpretations into English from four different languages also show a greater tendency to explicitness indicated by a more frequent use of optional connective that than in native English speeches (Kajzer-Wietrzny 2012). There are different plausible explanations of the increased frequency of optional that ranging from the subconscious process of explicitation, interference or the need for greater formality. But there could be more to it. As pointed out by Halverson (2003, 2007) certain characteristic features of translations might not be specific for translation only, but also for other cases of discourse production in multilingual contexts. The gravitational pull hypothesis associates features typical for translation to cognitive processes, which might be related to language use in particular constraints in general (Chesterman 2010 in reference to Halverson 2003, 2007). Following this line of thought, the present paper investigates explicitness in the corpus of speeches delivered at the European Parliament by MEPs speaking English as a foreign language and compares it to the corpus of simultaneous interpretations into English delivered also at the European Parliament and to the corpus of speeches made by the native English MEPs. It is hypothesized that the language used by non-native MEPs and by EP interpreters is more explicit than the language of native English MEPs.

## References

- Chesterman Andrew. 2010. "The space between the data and the concepts", *MikaEL* 4. (URL: <http://www.sktl.fi/toiminta/seminaarit/mikael-verkkajulkaisu/vol-4-2010/>) ( accessed 15<sup>th</sup> December 2014),
- Halverson, Sandra. 2003. "The cognitive basis of translation universals", *Target* 15:2, 197-241.
- Halverson, Sandra. 2007. "Investigating gravitational pull in translation: the case of the English progressive construction", In Riitta Jääskeläinen, Tiina Puurtinen and Hilka Stotesbury (eds.), *Texts*,

*process and corpora: research inspired by Sonja Tirkkonen-Condit*. Joensuu: Publications of the Savonlinna School of Translation Studies 5, 175-195.

Kajzer-Wietrzny, Marta. 2012. Interpreting universals and interpreting style (unpublished PhD dissertation) (URL: <https://repozytorium.amu.edu.pl/jspui/handle/10593/2425?mode=full>) (accessed 15<sup>th</sup> December 2014)

Olohan, Maeve – Mona Baker. 2000. "Reporting *that* in Translated English: Evidence for subconscious processes of explicitation?", *Across Languages and Cultures* 1(2): 141-158,

Sandrelli, Analisa – Claudio Bendazzoli. 2005. "Lexical patterns in simultaneous interpreting: a preliminary investigation of EPIC (European Parliament Interpreting Corpus)", Proceedings from the Corpus Linguistics Conference Series, 1(1), ([www.corpus.bham.ac.uk/PCLC](http://www.corpus.bham.ac.uk/PCLC)) (accessed 15 October 2013),

Shlesinger Miriam – Noam Ordan. 2012. "More spoken or more translated?: Exploring a known unknown of simultaneous interpreting", *Target* 24 (1), 43-60.

# COGNITIVE PROCESSING OF CULTURE-SPECIFIC ITEMS IN CONSECUTIVE AND SIMULTANEOUS INTERPRETING: A CORPUS-ASSISTED ANALYSIS

DEFENG LI, UNIVERSITY OF MACAU  
[defengli@umac.mo](mailto:defengli@umac.mo)

VICTORIA L.C. LEI, UNIVERSITY OF MACAU  
[viclcl@umac.mo](mailto:viclcl@umac.mo)

XIAODONG LIU, UNIVERSITY OF MACAU

Corpus-assisted Studies of Interpreting has attracted considerable research interest in recent years. This presentation attempts to make a corpus-assisted analysis of cognitive processing in consecutive interpreting (CI) and simultaneous interpreting (SI) between English and Chinese. More specifically, culture-specific items in English/Chinese translation are examined in these two modes of interpreting. The methodology consists of corpus analysis of natural on-site interpreting data and lab experiments involving professional vs. novice interpreters followed by retrospective interviews. The research findings indicate that, in terms of cognitive processes, SI diverges significantly from CI and there are marked differences between professional interpreters and novice interpreters. Pedagogical implications are then explored. The presentation concludes with suggestions for further studies, including triangulating corpus analysis with “wet methods”, e.g. the use of fNIRS in the investigation of the cognitive processes of interpreters.

Keywords: Corpus-assisted method, Cognitive Processes, Consecutive interpreting, Simultaneous interpreting, Culture-specific items, Translation pedagogy

## PANEL: BENEFITS OF WORKING WITH CORPORA IN COMMUNITY INTERPRETING RESEARCH: FROM QUALITATIVE ANALYSIS TO QUANTITATIVE VERIFICATION – AND BACK

CONVENORS: BERND MEYER, MAINZ UNIVERSITY & THOMAS SCHMIDT, MANNHEIM UNIVERSITY

[meyerb@uni-mainz.de](mailto:meyerb@uni-mainz.de), [thomas.schmidt@ids-mannheim.de](mailto:thomas.schmidt@ids-mannheim.de)

Participants of this panel all work with empirical data from interpreter-mediated interactions in community settings. Analyses in this field usually follow the premises of qualitative approaches to interaction, based on pragmatic, conversation analytic and ethnomethodological frameworks. Sections of discourse are investigated in detail in order to pinpoint systematic challenges and pitfalls of community interpreting, such as language alternation, dyadic sequences, explanations of technical terms, role conflicts, etc. However, rapidly developing corpus technologies seem to open paths to combine such analyses with quantitative approaches. Panelists will briefly present their data, outline what they expect from corpus technologies, and discuss how quantitative analyses may enhance their research. A separate section of the panel will be devoted to issues of data sharing and data sustainability.

Philipp Angermeyer (Toronto), Bernd Meyer (Mainz) and Thomas Schmidt (Mannheim) will present the Community Interpreting Database (<http://www.yorku.ca/comindat/comindat.htm>) and outline possible uses of the database, as well as new directions in research that emerge from it. (15 min.)

Sabine Braun (Surrey) will use a small corpus of (simulated) interpreter-mediated police interviews to demonstrate how quantitative and qualitative methods can be combined to investigate and systematise a number of issues arising in legal interpreting. (15 min.)

Claudio Baraldi, Laura Gavioli and Natacha Niemants (Modena) will present the AIM corpus of healthcare interpreter-mediated interaction. They will outline some issues in which automatic extraction of lexical items might be useful, e.g. instances of reported speech (“she said”, “he told me”); directives (“let’s tell her”, “please explain her”, etc.), personal markers (“doctor”, “interpreter/mediator”, “lady”). (15 min.)

Claudia V. Angelelli (Edinburgh) will discuss the California Hope-corpus of healthcare-mediated communication. She will present examples of discourse bundles and discuss the challenges and opportunities of using tools (e.g. EnViVo to analyse data). (15 min.)

Bernd Meyer (Mainz) and Thomas Schmidt (Mannheim): Roundtable discussion on data sharing and data sustainability. (45 min.)

### References

Angermeyer, P.; Meyer, B. & Schmidt, T. 2012. Sharing community interpreting corpora – a pilot study. In: Schmidt, T. & Wörner, K. (ed.): *Multilingual Corpora and Multilingual Corpus Analysis*, Hamburg Studies in Multilingualism 14, 275-294. Amsterdam: John Benjamins.

# WHEN AND WHY DO INTERPRETERS AND TRANSLATORS AT THE EU PARLIAMENT USE COLLOCATIONS? A TARGET-ORIENTED, BIDIRECTIONAL STUDY OF EPTIC.

MAJA MILIČEVIĆ, UNIVERSITY OF BELGRADE

[m.milicevic@fil.bg.ac.rs](mailto:m.milicevic@fil.bg.ac.rs)

ADRIANO FERRARESI, UNIVERSITY OF BOLOGNA

[adriano.ferraresi@unibo.it](mailto:adriano.ferraresi@unibo.it)

SILVIA BERNARDINI, UNIVERSITY OF BOLOGNA

[silvia.bernardini@unibo.it](mailto:silvia.bernardini@unibo.it)

Spearheaded by the work of the late Miriam Schlesinger and colleagues (e.g. Schlesinger and Ordan 2012), intermodal comparisons are emerging as a new methodological approach in corpus-based translation/interpreting studies. By comparing translations and interpretations, intermodal analyses may shed light both on the features that distinguish one form of language mediation from the other, thus helping us to better understand them, and on the common traits that may be taken to be characteristic of language mediation at large, and thus potential candidates for universality.

One of the few intermodal corpora available is EPTIC, the *European Parliament Translation and Interpreting Corpus* (Bernardini et al provisionally accepted). EPTIC builds on the well-known EPIC (*European Parliament Interpreting Corpus*; Sandrelli et al 2010), pairing EPIC's transcripts of interpreted speeches and their sources with the corresponding translated versions and respective source texts. The corpus is part-of-speech tagged, lemmatised and indexed with the Corpus WorkBench (<http://cwb.sourceforge.net/>). Each text is aligned (at sentence level) with its source/target and with the corresponding text in the other mode (oral/written). At the time of writing EPTIC consists of Italian and English texts only, for a total of 284 texts, and about 120,000 words. The English>Italian portion contains four versions of 81 texts, while the Italian>English portion has four versions of 61 texts. A trilingual sub-corpus including French is currently under construction thanks to collaboration with colleagues at Institut Libre Marie Haps, Brussels, and further expansion is foreseen for the near future.

Relying on bidirectional data from EPTIC and on a mix of quantitative and qualitative techniques, in this paper we present a comparative analysis of the choices made by translators and interpreters that result in the presence of collocations in the target texts. Collocation candidates (bigrams) based on pre-selected part of speech patterns are extracted from the four target language corpora (interpreted English, translated English, interpreted Italian, translated Italian). To evaluate their collocation status, frequency data about these bigrams are obtained from two large reference corpora of English and Italian (itWaC and ukWaC, Baroni et al 2009) and used to calculate lexical association scores (Mutual Information and T-Score). A total of 150 collocations per corpus are randomly extracted for manual analysis from among those with association scores above the EPTIC median for the relevant language. Parallel concordances are then examined to see if the presence of a collocation in the target text corresponds to the presence of a collocation in the source text, or not.

As well as offering quantitative data about the effect of directionality and mediation mode on the use of collocations, the manual perusal of parallel concordances allows us to start sketching a categorisation of collocation-yielding shifts, relying on classifications provided by Berman (1985 [2000]) and Marco (2012, in turn referring back to Delabastita 1996), and adapted to the issue at hand. The paper concludes by briefly considering the descriptive/theoretical implications of these results and the methodological potential of intermodal comparisons for corpus-based translation/interpreting studies.

## References

- Baroni, M., Bernardini, S., Ferraresi, A. and Zanchetta, E. 2009. "The WaCky Wide Web: A Collection of Very Large Linguistically Processed Web-Crawled Corpora". *Language Resources and Evaluation* 43, 3: 209-226.
- Berman, A. 1985 (2000). "Translation and the Trials of the Foreign". In Venuti, L. (ed.) *The Translation Studies Reader* (1<sup>st</sup> ed.). London: Routledge. 284- 297. [Tr. of "La traduction comme épreuve de l'étranger", *Texte* 77-81. (Translated by L. Venuti)].
- Bernardini, S., Ferraresi, A. and Miličević, M. (provisionally accepted). "From EPIC to EPTIC – Exploring simplification in interpreting and translation from an intermodal perspective". *Target*.
- Delabastita, D. 1996. "Introduction". *The Translator* 2, 2: 127-139.
- Marco, J. 2009. "Normalisation and the Translation of Phraseology in the COVALT Corpus". *Meta* 54, 4: 842-856.
- Sandrelli, A., Bendazzoli, C. and Russo, M. 2010. "European Parliament Interpreting Corpus (EPIC): Methodological Issues and Preliminary Results on Lexical Patterns in Simultaneous Interpreting." *International Journal of Translation* 22 (1-2): 165–203.
- Shlesinger, M., and Ordan, N. 2012. "More Spoken or More Translated? Exploring a Known Unknown of Simultaneous Interpreting." *Target* 24 (1): 43–60.



# THE DEVELOPMENT OF TECHNOLOGICAL RESOURCES BASED ON EXPLOITATION OF CORPORA FOR COURT INTERPRETERS IN SPANISH-ROMANIAN, ARAB, CHINESE, FRENCH AND ENGLISH LANGUAGE PAIRS

MARIANA OROZCO-JUTORÁN, UNIVERSITAT AUTONÓMA DE BARCELONA

[mariana.orozco@uab.cat](mailto:mariana.orozco@uab.cat)

In Spain, in contrast with its development in other countries, court interpreting has been an under-researched area until recently. In the last two decades, however, research in court interpreting has emerged as a major topic in Europe. In fact, within the Horizon 2020 Programme, the Directorate-General for Justice of the European Commission, through the Justice Programme 2014-2020, is offering grants to undertake research in this area. Nevertheless, the European projects which have been carried out so far have focussed on the analysis of language barriers in the access to Justice and on the need to set similar standards regarding the training, accreditation and recruitment of interpreters.

In Spain, a new bill has just been published in the Official Gazette of the Spanish Parliament (BOCG, 5 September 2014) which will modify Spain's Code of Criminal Procedure. As stated therein, this new legislative text "significantly reinforces procedural guarantees in criminal proceedings, as it provides detailed regulation on the right to translation and interpreting in criminal proceedings as well as on the right to information of an accused person in relation to the subject of the criminal proceedings so that they can exercise efficiently their right to self-defence." This turns translation and interpreting into an essential element in the rights to effective legal protection in the exercise of lawful rights and interests before the courts in order to avoid any state of defencelessness, namely the right to be informed of the accusation against them, the right to a public process with all procedural guarantees, and the right to self-defence, as enshrined in Article 24 of the Spanish Constitution.

Accordingly, the research group MIRAS (<http://grupsderecerca.uab.cat/miras/es>) has recently launched a research project aiming to create a computer application which can include in only one interface all the necessary resources to facilitate court interpreters' performance: a code of good practice, a protocol for conduct and behaviour in the most frequent situations for a court interpreter, a set of guidelines for Justice personnel on interpreters' role and on how to interact with interpreters, and lastly a database containing the terms which are most frequently used in criminal proceedings with comments and two-way translation options in the most frequently translated languages, i.e. Arab, Chinese, English, French and Romanian. All these resources will be based on the data obtained from the exploitation of a corpus compiled, for the first time, in real criminal proceedings (not simulations) in five language-pairs.

The paper will explain the design and the methodology developed for the corpus compilation and exploitation as well as the results expected from the outcomes of this project, namely a new computer application, called ReTIPp, which will help court interpreters to perform their tasks more accurately and efficiently.

## OVER-UH-LOAD. THE OCCURRENCE OF $UH(M)$ BETWEEN ELEMENTS OF COMPOUNDS DURING INTERPRETING.

KOEN PLEVOETS & BART DEFRANCO, GHENT UNIVERSITY

[koen.plevoets@ugent.be](mailto:koen.plevoets@ugent.be), [bart.defrancq@ugent.be](mailto:bart.defrancq@ugent.be)

Cognitive load is a major source of processing difficulties and disfluencies, such as *uh* or *uhm*, in both spontaneous speech (Levelt 1983; Clark & Fox Tree 2002; Bortfeld et al. 2001; Watanabe et al. 2008) and interpreting (Goldman-Eisler 1967; Mead 2000, Tissi 2000; Cecot 2001). Gile (1995) lists the interpreter's lack over the content and his reduced background knowledge as potential obstacles during interpreting, and for simultaneous interpreting he adds the lack of control over the original speech rate as well as the interference of speaking and listening. In Plevoets & Defrancq (2014) and Plevoets & Defrancq (submitted), the occurrence rate of  $uh(m)$  was analysed with respect to the delivery rate, the lexical density and the proportion of numbers in both interpretations and in non-interpreted speeches. An interesting qualitative finding was that some  $uh(m)$ 's occurred between the elements of lexical compounds, e.g. *onderzoeks-uh-gelden* ('research-uh-funds'), *voedsel-uh-middelen* ('food-uh-resources'), *uit-uh-breiden* ('ex-uh-pand') or *Cohn-uh-Bendit*. As such constructions can shed more light on cognitive load, this paper will investigate them further.

This paper will make a comparison between interpreted language and spontaneous speech in two corpora. The corpus of interpreted language was compiled at Ghent University between 2010 and 2013. It consists of French, Spanish and Dutch interpreted speeches in the European Parliament from 2006 until 2008, which were transcribed according to the VALIBEL guidelines (Bachy et al. 2007). For the purposes of this analysis, a sub-corpus of French source speeches and their Dutch interpretations was selected (amounting to a total of 140 000 words), which has been annotated for lemmas, parts-of-speech and chunks (Van de Kauter et al. 2013) as well as sentence-aligned with WinAlign (WinAlign 2014). The reference corpus for spontaneous speech is the sub-corpus of political debates of the Spoken Dutch Corpus (Oostdijk 2000), compiled between 1998 and 2003. This sub-corpus contains 220.000 words of Netherlandic Dutch and 140.000 words of Belgian Dutch and is annotated for lemmas and parts-of speech.

In both corpora, the number of  $uh(m)$ 's between elements of compounds were counted for each sentence. The difference proved to be statistically significant, in that interpreters produce these constructions more often than spontaneous speakers. A subsequent analysis pointed to two interesting determinants. First, the compounds by the spontaneous speakers often involved a self-repair of a mispronunciation, e.g. *fundu-uh-mentalisme* ('fundu-uh-mentalism'), *geïmplende-uh-menteerd* ('implende-uh-plemented'), *oorsprangsk-uh-land* ('\*origin-uh-country'), *samenwerkingsfron-uh-verbanden* ('collaboration fron-uh-relations'), *tariefafspr-uh-aanpassing* ('tariff agree-uh-adaptation'), *topcro-uh-crimineel* ('top cro-uh-criminal'), or *voortgaand-uh-voorgaand* ('persistent-uh-previous'). Second, the compounds by the spontaneous speakers frequently consisted of a combinations of a word with non-word such as an acronym, a proper noun or a number, e.g. *CBS-uh-cijfers* ('CBS-uh-figures'), *E-uh-313* ('E-uh-313'), *SG-uh-niveau* ('SG-uh-level'), *vrijstellings-uh-AMVB* ('license-uh-AMVB'), *Millinx-uh-buurt* ('Millinx-uh-neighbourhood'), *Polaris-uh-systeem* ('Polaris-uh-system'), *post-94-uh-traject* ('post-94-uh-trajectory'), and *15-procent-uh-eis* ('15 percent-uh-demand'). These results suggest that the occurrence of  $uh(m)$  with spontaneous speakers is more related to issues of phonetic realisation of the message than to the memory limitations which are at stake with interpreters.

## References

- Bachy, S., Dister, A., Francard, M., Geron, G., Giroul, V., Hambye, P., Simon, A.C. and Wilmet, R. (2007). *Conventions de transcription régissant les corpus de la banque de données VALIBEL*. University of Louvain-la-Neuve. Available online at [http://www.uclouvain.be/cps/ucl/doc/valibel/documents/conventions\\_valibel\\_2004.PDF](http://www.uclouvain.be/cps/ucl/doc/valibel/documents/conventions_valibel_2004.PDF).
- Bortfeld, H., S.D. Leon, J.E. Bloom, M.F. Schober & S.E. Brennan (2001). "Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender". *Language and Speech* 44: 123-147.
- Cecot, M. (2001). "Pauses in simultaneous interpretation: a contrastive analysis of professional interpreters' performances". *The interpreters' newsletter* 11: 63-85.
- Clark, H.H. and Fox Tree, J.E. (2002). "Using uh and um in spontaneous speaking". *Cognition* 84: 73-111.
- Gile, D. (1995). *Basic concepts and models for interpreter and translator training*. Amsterdam: John Benjamins.
- Goldman-Eisler, F. (1967). "Sequential temporal patterns and cognitive processes in speech". *Language and Speech* 10 (3): 122-132.
- Levelt, W. (1983). "Monitoring and self-repair in speech". *Cognition* 14: 41-104.
- Mead, P. (2000). "Control of pauses by trainee interpreters in their A and B languages". *The Interpreters' Newsletter* 10: 89-102.
- Oostdijk, N. (2000). "The Spoken Dutch Corpus: Overview and first evaluation". *Proceedings of the Second International Conference on Language Resources and Evaluation*: 887-894.
- Plevoets, K. & B. Defrancq (2014). "Informational load as a trigger for disfluencies in interpreting". Paper presented at the *UCCTS4* conference, Lancaster, UK.
- Plevoets, K. & B. Defrancq (submitted). "The effect of informational load on disfluencies in interpreting". *Translation and Interpreting Studies*.
- Tissi, B. (2000). "Silent pauses and disfluencies in simultaneous interpretation: A descriptive analysis". *The Interpreters' Newsletter* 10: 103-127.
- Van de Kauter, M., Coorman, G., Lefever, E., Desmet, B., Macken, L. and Hoste, V. (2013). "LeTs Preprocess: The multilingual LT3 linguistic preprocessing toolkit". *Computational Linguistics in the Netherlands Journal* 3: 103-120.
- Watanabe, M., Hirose, K. Den, Y. and Minematsu, N. (2008). "Filled pauses as cues to the complexity of up-coming phrases for native and non-native listeners". *Speech Communications* 50: 81-94.
- WinAlign (2014). *SDL Trados WinAlign Tutorial*. Available at: <http://www.translationzone.com/resources/downloads/winalign-tutorial.html>.

# LEXICAL PATTERNS AND GENDER: TRENDS EMERGING FROM EPIC

MARIACHIARA RUSSO, UNIVERSITY OF BOLOGNA

[mariachiara.russo@unibo.it](mailto:mariachiara.russo@unibo.it)

Corpus-based Translation and Interpreting Studies have so far explored a large variety of phenomena pertaining to different dimensions of language production and reception in different modalities and settings. In particular, some studies focused on the use of lexicon by investigating lexical variety (LV) and lexical density (LD) in English original and translated texts and showed a higher degree of both features in the former vs the latter (Laviosa 1998). Inspired by these results, another study set out to investigate whether the same patterns emerged when comparing English, Italian and Spanish source speeches delivered in the European Parliament and their target speeches in these three languages produced by simultaneous interpreters; this study obtained less clear-cut, language-dependent results (Russo et al. 2006).

These preliminary results on lexical patterns could shed light on possibly different ways to access mental lexicon, that is our mental store of words and basic information about them (Bock and Levelt 1994). The prevalence of content words (lexical density) and a wider vocabulary (lexical variety) in either original speakers or interpreters could be not only language- but also gender-dependent, a study perspective which is one of the major preoccupations of contemporary social sciences and humanities (Shlesinger et al. 2009). In particular, speech styles have been analysed from cultural difference and power relation approaches (Tannen 1994) or problematising gender polarization (Bergvall et al. 1996).

Against this backdrop, a first quantitative study was carried out on our trilingual European Parliament Interpreting Corpus which consists of approximately 190.000 tokens (Sandrelli et al. 2010). The study showed that women generally display greater lexical variety and lexical density than men (Russo forthcoming).

The objective of the present study is to take the research a step forward and explore the qualitative dimension of these preliminary findings in an attempt to highlight meaningful differences in female and male use of lexicon in EPIC speakers and interpreters.

## References

- Bergvall V., Bing J., and Freed A. (eds.) (1996) *Rethinking Language and Gender Research*, London/New York: Longman.
- Bock, K., and Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*, London: Academic Press, 945-984. (pdf)
- Laviosa S. (1998). Core patterns of lexical use in a comparable corpus of English narrative prose. *Meta*, 43/4, 557-570.

Russo, M., Bendazzoli, C. and Sandrelli, A. (2006). Looking for lexical patterns in a trilingual corpus of source and interpreted speeches: extended analysis of EPIC (European Parliament Interpreting Corpus), *FORUM*, 4/1, 221-254.

Russo, M. (forthcoming). Oralità e genere in interpretazione simultanea: uno studio basato su corpora. Paper presented at the international conference "Oraliter", Forlì 3-4 December 2014, University of Bologna at Forlì.

Sandrelli, A., Bendazzoli, C. e Russo M. (2010) "European Parliament Interpreting Corpus (EPIC): Methodological issues and preliminary results on lexical patterns in SI", *International Journal of Translation* 22 (1-2), 165-203.

Shlesinger, M., Koppel, M., Ordan, N. and Malkiel, B. (2009) Markers of translator gender: do they really matter?, *Copenhagen Studies in Language* 38, 138-198.  
[http://u.cs.biu.ac.il/~koppel/papers/Shlesinger%20et%20al\\_183-198.pdf](http://u.cs.biu.ac.il/~koppel/papers/Shlesinger%20et%20al_183-198.pdf)

Tannen, Deborah (1994) *Gender and Discourse*, New York: Oxford University Press.

# MITIGATION AND HEDGING IN INTERPRETER-MEDIATED FOOTBALL PRESS CONFERENCES: A CASE STUDY

ANNALISA SANDRELLI, UNINT ROMA  
[annalisa.sandrelli@unint.eu](mailto:annalisa.sandrelli@unint.eu)

Football is currently the sport with the greatest global appeal. In recent years, media exposure of international competitions such as the FIFA World Cup and the UEFA European championships has grown exponentially, with football games being broadcast live worldwide. Therefore, the press conferences organised during competitions have become full-blown international media events, requiring interpreting services for the international press in attendance (*primary audience*) and for football fans everywhere (*secondary audience*).

The FOOTIE (Football In Europe) corpus is made up of the 16 press conferences involving Italy and its opponents during the EURO 2008 football championships i.e. Holland, Romania, France and Spain. For each game there was one pre-match and one post-match press conference for both teams: the official languages used always included English and the languages of the two teams. Simultaneous interpreting was carried out in both translation directions (A to B and B to A) and English was used as a *pivot* language whenever necessary. The resulting parallel corpus includes Italian, English, Spanish and French source texts and their matching target texts (Sandrelli 2012a, 2012b). The proposed paper only focuses on the Italian-English component of the corpus (Italian source texts and English target texts): the overall duration of the Italy press conferences was roughly 150 minutes, and the two subcorpora run to about 22,000 words each.

Football press conferences are examples of dialogic *institutional communication* (Orletti 2000, Sandrelli 2012a): they are highly ritualized, with pre-established roles for participants and a limited range of acceptable topics. The overall function of these communicative events is to allow accredited journalists to collect the information they need and to obtain quotable answers from coaches and players who, on their part, are obliged to take part by the competition regulations (Sandrelli 2012a). When dealing with coercive and/or probing questions, interviewees may try to find ways to resist or sidestep them, whilst at the same time seemingly collaborating, since their institutional role is to answer questions (as has been shown to happen in court, political interviews, news conferences and media briefings. See Berk-Seligson 2009; Galatolo 2002; Partington 2001; Clayman & Heritage 2002).

As a result, mitigation and hedging phenomena are noticeable in both questions and answers, and are important aspects of the message which must be preserved in translation. At the same time, hedging additions have also been shown to be used by interpreters in press conferences and in media interpreting as “stalling” devices, emergency strategies or face-saving strategies (Straniero Sergio 2003, 2007). The proposed paper focuses on mitigation and hedging in the source texts and investigates how these features are reflected in the interpreters’ renditions. Moreover, cases of mitigation and hedging found in the target texts which do not have a corresponding source text item are also discussed. Given the small size of the sample under analysis, the proposed paper is only meant as a corpus-based case study, to be further developed on a larger scale.

## Selected references

Berk-Seligson, Susan (2009) *Coerced Confessions. The Discourse of Bilingual Police Interrogations*, Berlin, Mouton De Gruyter.

Clayman, Steven E. & Heritage, John (2002) *The News Interview. Journalists and Public Figures on the Air*. Cambridge: Cambridge University Press.

Galatolo, Renata (2002) "La comunicazione in tribunale", in Bazzanella Carla (ed.) *Sul dialogo. Contesti e forme di interazione verbale*, Milano, Edizioni Angelo Guerini e Associati SpA: 137-152.

Orletti, Franca (2000) *La conversazione diseguale*, Roma, Carocci.

Partington, Alan (2001) "Spin and Counter-Spin at the White House: the Reformulation of Questions and Answers in a Corpus of Spoken Political English", In Guy Aston & Lou Burnard (eds) *Corpora in the Description and Teaching of English*. Papers from the 5th ESSE conference. Bologna: CLUEB: 103-117.

Sandrelli, Annalisa (2012a) "Introducing FOOTIE (Football in Europe): Simultaneous Interpreting at Football Press Conferences", in Straniero Sergio, Francesco & Falbo, Caterina (eds) *Breaking Ground in Corpus-based Interpreting Studies*. Bern: Peter Lang: 119-153.

Sandrelli, Annalisa (2012b) "Interpreting football press conferences: the FOOTIE corpus", in Kellett Bidoli, Cynthia J. (ed.) *Interpreting across Genres: Multiple Research Perspectives*, Trieste, EUT: 78-101.

Straniero Sergio, F. (2003). Norms and quality in media interpreting: the case of formula one press conferences. *The Interpreters' Newsletter*, 12: 135-165.

Straniero Sergio, F. (2007). *Talkshow Interpreting. La mediazione linguistica nella conversazione spettacolo*. Trieste: Edizioni Universitarie Trieste.

# STUDYING FIGURATIVE LANGUAGE IN SIMULTANEOUS INTERPRETING: THE IMITES (*INTERPRETACIÓN DE LA METÁFORA ENTRE ITALIANO Y ESPAÑOL*) CORPUS

NICOLETTA SPINOLO, UNIVERSITY OF BOLOGNA

[nicoletta.spinolo@unibo.it](mailto:nicoletta.spinolo@unibo.it)

The translation of figurative language has always been a major issue and, often, a problem in the translation process (see, among others, Álvarez 1993, Azar 1989, Dagut 1976 and 1987, Dobrzyńska 1995, Menacere 1992, Newmark 1981, 1988, 1993, 1998, Pliego Sánchez 1993, Prandi 2010, Van Besien & Pelsmaekers 1988, Van den Broeck 1981). But what happens when a choice on how to translate a figurative item needs to be made in fractions of a second? The natural difficulties lying in the translation process of figurative expressions (among others, interpretation of meaning, interlingual knowledge, cultural implications) become even more complicated to overcome if applied to the simultaneous interpreting process, where interpreters only have fractions of a second to decide which strategy to apply when dealing with a certain figurative expression (Turrini 2004, Spinolo & Garwood 2010).

Samaniego Fernández (2011:263) claims that “the proper task of translation theory would not be to specify how metaphor should be translated, but to describe and account for actual renderings of metaphors”; believing that the same could be held true for Interpreting Studies, the IMITES (*Interpretación de la Metáfora entre Italiano y Español*) project was devised in order to study how professional interpreters actually deal with metaphors. The IMITES corpus includes a total of 23 original speeches (287 minutes) delivered in Italian and Spanish at conferences held at the European Commission, and their interpreted versions in Spanish and Italian respectively (287 minutes).

The purpose of this presentation is to describe the methodology used to build the IMITES corpus and extract a set of 1135 figurative expressions which were then aligned and compared with the interpreters' versions, with the twofold goal of a) understanding which ones cause most problems to interpreters and b) analysing the interpreting strategies applied by professionals such as those of the Directorate General for Interpretation (SCIC) of the European Commission in interpreting metaphors.

## References

- Álvarez A., (1993). “On Translating Metaphor”, *META*, vol.38, n.3, 479-490.
- Azar M., (1989) “La Métaphore traduisible”, *META*, vol.34, n. 4, 794-796.
- Dagut M., (1987). “More about the Translatability of Metaphor”, *Babel*, n. 2, Vol. 33, 78-83.
- Dagut, M., (1976). “Can Metaphor Be Translated?”. *Babel: International Journal of Translation*, XXII (1), 21-33.
- Dobrzyńska T. (1995). “Translating Metaphor: Problems of Meaning”, *Journal of Pragmatics*, no. 24, 595-604.



- Menacere M. (1992). "Arabic Metaphor and Idiom in Translation", *META*, vol.37, n. 3, 567-572, 1992.
- Newmark P. (1981 [1984]). *Approaches to Translation*, Oxford: Pergamon Press.
- Newmark P. (1988). *A Textbook of Translation*, New York: Prentice Hall.
- Newmark P. (1998). *More Paragraphs on Translation*, Clevedon: Multilingual Matters.
- Newmark P., (1993). *Paragraphs on Translation*, Clevedon: Multilingual Matters.
- Pliego Sánchez I., (1993) "La Traducción de la Metáfora", *Essays on Translation*, n. 1, 97-103.
- Prandi M., (2010). "Typology of Metaphors: Implications for Translation". *Mutatis Mutandis*. Vol. 3, No. 2.2010, 304-332.
- Samaniego Fernández, E. (2011). "Translation Studies and the Cognitive Theory of Metaphor". *Review of Cognitive Linguistics*, Vol. 9, n. 1, pp. 262-279.
- Spinolo N. e Garwood C.J. (2010). "To Kill or Not To Kill: Metaphors in Simultaneous Interpreting", *Forum* 8:1, 181-211.
- Turrini, C. (2004), "Metafora e dintorni: l'interpretazione simultanea del linguaggio non letterale al Parlamento europeo", in Bersani Berselli, Mack, Zorzi (2004).
- Van Besien F. e K. Pelsmaekers, (1988). "The Translation of Metaphor", in Nekeman (a cura di), 140-146.
- Van Den Broeck R. (1981). "The Limits of Translatability Exemplified By Metaphor Translation", *Poetics Today*, vol.2, No. 4, 73-87 (1981).

# CORPUS-BASED INTERPRETING STUDIES IN CHINA

BINHUA WANG, THE HONG KONG POLYTECHNIC UNIVERSITY

[wangbinhua@hotmail.com](mailto:wangbinhua@hotmail.com)

For a long time in its history, interpreting studies has focused on the cognitive process in interpreting, which is typical of “decontextualized” experimental methods and small-scale of research data. The emergence of Corpus-based Interpreting Studies (CIS) in recent years has helped to strengthen the ecological validity and significance of interpreting studies. The inception of CIS in China occurred in 2009 (Hu, 2009; Zhang, 2009; Wang, 2009b) as a natural result of the fast growth of corpus linguistics and corpus-based translation studies in the region. In spite of various difficulties in data collection, transcription and annotation, CIS, ‘as an alternative and complement to intuition, surveys or laboratory experiments’ (Setton, 2012), has shown great potential in describing the features of the interpreting product and in revealing complexities of the interpreting process with large-scale authentic data.

This presentation reviews the development of CIS in China based on a survey of interpreting corpora developed in the region and journal articles and published books (or book chapters) published in the region. It summarises the techniques used in the construction of several representative corpora such as the CECIC (Hu, 2009; 2010; 2012; 2013) and CEIPPC (Wang, 2009a; 2012; 2013) and presents an overview of the research topics explored and methods utilised by the researchers. It also discusses the issues to be addressed further in CIS, including: 1) How can the analytic tools of corpus linguistics and corpus-based translation studies be better adapted to the peculiarities and complexities of interpreting studies? 2) How can CIS go beyond the description of data and provide explanation about phenomena in interpreting? 3) How can CIS contribute to the conceptualisation of theories?

## References

- Hu, Kaibao. 2009. Explicitation in the Chinese-English Conference Interpreting and Its Motivation – A Study Based on Parallel Corpus. *Journal of PLA University of Foreign Languages*, 32 (4): 67-73.
- Setton, Robin. 2011. Corpus-based interpreting studies: Overview and prospects. In Alet Kruger, Kim Wallmach, Jeremy Munday (eds.) 2011. *Corpus-based Translation Studies. Research and Applications*. London & New York: Continuum.
- Wang, Binhua. 2009b. Constructing a Corpus for Interpreting Training: Theory and Practice. *Foreign Language World*, 131: 23-32.
- Wang, Binhua. 2009a. *Description of Norms in Interpreting and Its Application – A Study Based on the Corpus of Consecutive Interpreting in Chinese Premier Press Conferences*. PhD dissertation, Guangdong University of Foreign Studies.
- Wang, Binhua. 2012. A Descriptive Study of Norms in Interpreting: Based on the Chinese-English Consecutive Interpreting Corpus of Chinese Premier Press Conferences. *Meta: Translators’ Journal*, 57 (1): 198-212.
- Zhang, Wei. 2009. Interpreting corpus: Some theoretical and practical issues. *Chinese Translators Journal*. 30 (3): 54-59.

# A CORPUS-BASED APPROACH TO COMPARING THE USE OF MANDARIN DISCOURSE PARTICLES IN SPONTANEOUS SPEECH AND SIMULTANEOUS INTERPRETING

TA-WEI DAVID WANG, NEWCASTLE UNIVERSITY

[t.w.wang@newcastle.ac.uk](mailto:t.w.wang@newcastle.ac.uk)

Simultaneous interpreting (SI) is a fundamentally different form of language use from spontaneous speech. Discourse particles (words like *well, so, like* etc.) are widely known to be sensitive to discourse structure, speech planning, and interactional format. Among them, those that help clarify speakers' intentions with the use of words are defined as discourse markers (Chaume, 2004). We would thus expect their use in SI to be different from their use in spontaneous speech. My study aims at making a contribution towards finding out any such differences. A pilot study revealed that people tended to use more conjunction particles in Mandarin spontaneous speech, whereas in English-to-Mandarin SI, determiner particles were used more frequently. Given the issues unaddressed in the pilot study such as what caused the differences and how the differences could affect interpreting output in Mandarin, a corpus-based approach with the combination of listener survey was employed to form the methodological basis in the current study. To create the corpora for comparison, 7 trainee interpreters were recruited to take part in interviews aimed at creating the corpora for Mandarin spontaneous speech and the simulated interpreting conference aimed at creating the corpora for English-to-Mandarin SI. The data created from both corpora was compared and analyzed in terms of frequency and meaning via a semantic parser. A frequency count aimed at locating the most frequently utilized particles in Mandarin spontaneous speech for analysis in terms of meaning. Semantic parsing aimed at investigating what semantic roles these most frequently utilized particles assumed in a given Mandarin context to investigate any changes in their meanings (Xu, 2008), which serves as a crucial indicator in identifying potential Mandarin discourse markers from regular discourse particles. Lastly, the listener survey was employed not only to validate the results by the parser but also to assess the effects these particles may have on the interpreting output in terms of flow to see if they can provide contextual cues (Hansen, 1987) the way discourse markers do to listeners so that the rendered texts are easier to follow. The results of the current study will be presented in the conference.

**Keywords:** spontaneous speech, discourse marker, semantic, parser

# POSTER PRESENTATIONS

# DESIGNING A MULTIMODAL CORPUS OF ENGLISH AS A LINGUA FRANCA (ELF) IN LANGUAGE FOR SPECIFIC PURPOSE (LSP) FOR THE INTERPRETING CLASSROOM: CHALLENGES AND ISSUES

CHIARA BARBAGIANNI, SSML “A. MACAGNO” – CUNEO

[chiarabargianni@gmail.com](mailto:chiarabargianni@gmail.com)

VALENTINA BASELLI, IULM MILAN

[valentina.baselli@fastwebnet.it](mailto:valentina.baselli@fastwebnet.it)

CLARA PIGNATARO, IULM MILAN

[clara.pignataro@iulm.it](mailto:clara.pignataro@iulm.it)

The spread of English as a Lingua Franca (ELF) counts as one of the major causes for the dramatic English language prominence in specialized communication scenarios (Albl-Mikasa 2013:4; Nickerson 2013:445). In the literature, there is some evidence that interpreting from ELF speakers is generally more demanding than interpreting from native speakers of English in terms of cognitive load during the in-conference performance (Albl-Mikasa 2010; 2013). This additional processing cost in source-text comprehension, analysis, short-term memory management and target text production when interpreting from non-native speakers deteriorates the interpreter’s performance (Gile 1999; 2005; Albl-Mikasa 2010). This is particularly true in Language for Specific Purposes (LSP) conferences where a high degree of accuracy and precision is required (cf. Nobs et al. 2011).

As EFL and LSP appear to be intrinsically intertwined phenomena in the interpreting practice, they should not be overlooked in the interpreter training. The use of authentic material, such as audiovisual material recorded LSP conferences in ELF scenarios can contribute to expose students to real life situations and raise their awareness about the relevant quality benchmarks. The choice of building a multimodal corpus (Straniero Sergio & Falbo 2012:31) is key to provide a comprehensive understanding of real life interpreting settings and discourse practices (Kellett Bidoli 2006). In addition to audio and video layers, captions should be manually transcribed and aligned to allow comprehensive terminology and stylistic analyses.

The project is still at its design stage and a variety of corpus building issues are being considered (cf. Knight 2011:324). The first is discourse representativity, which in turn is dependent on the ELF speakers' willingness to collaborate, as an explicit permission from conference participants is required to record, store and process any content or data in which they are involved as speakers. Second, the generalizability of subsequent quantitative and qualitative analyses relies on the creation of significant samples to fully represent the relevant conference setting and discourse. In addition, some key pragmatic aspects are still to be further refined, such as 1) practical corpus deployment with interpreting students and 2) actual corpus impact on the students’ learning curve.

Albl-Mikasa, M. (2010). “Global English and English as a Lingua Franca (ELF): implications for the interpreting profession”, *Trans-kom*. Retrieved from <http://www.trans-kom.eu> (09/01/2015).

Albl-Mikasa, M. (2013). "Teaching Globish: the need for an ELF pedagogy in Interpreter Training", *International Journal of Interpreter Education* 5(1), 4-16.

Gile, D. (1999). "Testing the Effort Model's Tightrope Hypothesis in simultaneous interpreting: a contribution", *Hermes* 23, 153-172.

Gile, D. (2005). "Directionality in conference interpreting: a cognitive view", in Godijns, R. & Hindedaël, M. (eds) *Directionality in Interpreting. The retour or the Native?* Ghent: Communication and Cognition, 9-26.

Straniero Sergio, S., Falbo, C. (2012). "Studying interpreting through corpora", in Straniero Sergio, F. & Falbo C. (eds) *Breaking Ground in Corpus-Based Interpreting Studies*. Berlin: Peter Lang, 9-52.

Kellett Bidoli, J. (2006). "Glossary compilation of LSP including a signed language: a corpus-based approach", in the *Workshop Proceedings of the 5<sup>th</sup> International Conference on Language Resources and Evaluation, LREC 2006, 2<sup>nd</sup> Workshop on the Representation and Processing of Sign Language: Lexicographic Matters and Didactic Scenarios*, Pisa: ILC-CNR, 37-42.

Knight, D. (2011). "The future of multimodal corpora", *RBLA* 11 (2), 391-415.

# TOWARDS A PURPOSE-SPECIFIC INTERPRETING CORPUS: *ANGLINTRAD* AND ITS POSSIBLE EXPLOITATION FOR INTERPRETERS' TRAINING

MICHELA BERTOZZI, UNIVERSITY OF BOLOGNA

[michela.bertozzi6@unibo.it](mailto:michela.bertozzi6@unibo.it)

Keywords: interpreting, anglicism, corpus, Italian, Spanish

Corpus-based interpreting studies have proven to be an effective tool to analyze important phenomena that had been neglected by research (Shlesinger 1998), but little attention has been attached to the possible exploitation of these tools and the benefits of corpus-derived insights for interpreters' training (Bendazzoli 2010). The aim of this (ongoing) experimental project is to provide some guidelines for the construction of a purpose-specific multimodal Italian-Spanish corpus and to suggest a possible exploitation for interpreters' training and practice.

The idea of *Anglintrad* multimodal bilingual corpus came from the practical need to shed light on a particularly challenging phenomena in Italian-Spanish simultaneous interpreting: the strong presence of unmodified English loanwords in Italian political speeches (Marzocchi 2007) and the different Spanish loanwords integration mechanisms (Tonin 2010). For this specific purpose, a bilingual intermodal corpus is currently being compiled, whose main aims are observing this peculiar phenomenon in Italian original speeches, the way it is managed by simultaneous interpreters (through cancellation, adaptation, modification, etc.) and the related potential problems (silent/filled pauses, false starts, carry-over effects, etc.).

In the light of their accessibility and above all their comparability and homogeneity, the choice fell on the plenary sessions of the European Parliament, whose original Italian and interpreted Spanish versions are being transcribed (and some of them are being taken from the *EPIC* Corpus) and then compared to the related translated Spanish official reports available online. The possibility of comparing the same phenomena (unmodified English loanwords) from two different perspectives (interpreting and translation) represents an unprecedented opportunity entailing unparalleled possible didactic exploitations for interpreters' training and practice.

## References

Bendazzoli, C., 2010a. Il corpus DIRSI: creazione e sviluppo di un corpus elettronico per lo studio della direzionalità in interpretazione simultanea. PhD dissertation, University of Bologna.

Bendazzoli, C., 2010b. Corpora e interpretazione simultanea. Bologna: Asterisco, In: Alma-DL. [http://amsacta.unibo.it/view/series/Alma-DL=2E\\_Saggi.html](http://amsacta.unibo.it/view/series/Alma-DL=2E_Saggi.html)

Görlach, M. (ed.) 2005. A dictionary of European Anglicisms : a usage dictionary of Anglicisms in 16th European languages. Oxford : Oxford University Press.

Gusmani, R., 1981. *Saggi sull'interferenza linguistica*. Firenze: Le Lettere.

Iglesias Fernández, E., 2007. *La didáctica de la interpretación de conferencias. Teoría y práctica*. Granada: Comares.

Marzocchi, C., 2007. Translation — Transcript — Interpretation. Notes on the European Parliament verbatim report of proceedings. *Across Languages and Cultures* (8:2), pp. 249-254.

Petite, C., 2005. Evidence of repair mechanisms in simultaneous interpreting: A corpus-based analysis, *Interpreting* (7:1), pp. 27-49.

Russo, M., Bendazzoli, C., *et al.* 2012. "The European Parliament Interpreting Corpus (EPIC): Implementation and developments", in F. Straniero Sergio and C. Falbo (eds.) *Breaking Ground in Corpus-Based Interpreting Studies*, Bern: Peter Lang, 35-90.

Russo, M., Sandrelli, A., Bendazzoli, C., 2010. European Parliament Interpreting Corpus (EPIC): Methodological Issues and Preliminary Results on Lexical Patterns in Simultaneous Interpreting, *International Journal of Translation* (22:1-2).

Shlesinger, M., 1998. Corpus-Based Interpreting Studies as an Offshoot of Corpus-Based Translation Studies, *Meta* (43: 4), pp. 486-493.

Spinolo, N., 2014. La resa del linguaggio figurato in interpretazione simultanea: Una sperimentazione didattica. PhD dissertation, University of Bologna.

Straniero Sergio, F. and C. Falbo (eds) (2012). *Breaking Ground in Corpus-based Interpreting Studies*. Bern: Peter Lang.

Timarova, S., 2005. Corpus linguistics methods in interpreting research: a case study. *The Interpreter's Newsletter* (13), pp. 65-70.

Tonin, R., 2010. *El vaivén de las palabras. Los anglicismos en español y en la traducción al italiano*. Rome: Aracne.



# A STUDY OF ANAPHORIC RELATIONS IN SIMULTANEOUS INTERPRETING

ANA CORREIA, UNIVERSITY OF MINHO

[anav.moutinho@gmail.com](mailto:anav.moutinho@gmail.com)

The European Parliament Interpreting Corpus opened up a wide range of possibilities for research in the field of corpus-based interpreting studies as it drew attention to the massive sources of material that are the European Parliament plenary sessions. In the wake of EPIC, we now wish to put up for discussion a project for the creation and analysis of a bidirectional corpus of simultaneous interpreting, made up of speeches delivered by Portuguese and English MEPs at plenary sessions of the European Parliament and the respective interpretations. This initiative is part of a doctoral research, which in turn stemmed from a broader project for the compilation of (written) corpora known as *Per-Fide*.

This corpus will afford a glimpse into the anaphoric links that interpreters establish in their renditions. This type of cohesive device is a basic condition for coherence (Halliday & Hasan, 1976). If not properly established, anaphoric links may hinder a text's communicative intelligibility, which is regarded as the main goal of any translation-related activity, particularly interpreting. Therefore the issue of cohesion and coherence has held the interest of researchers in translation and interpreting studies alike (Blum-Kulka, 1986; Gerzymisch-Arbogast, 2000; Shlesinger, 1995).

For this study, we will rely on corpus linguistics methods. It is important to clarify that the process of transcribing the source and target speeches has already provided some useful insights into the problems raised by anaphoric relations, which are mainly attributable to the use of pronouns and determiners. Once our corpus has been annotated with part-of-speech tags, we will be able to isolate occurrences of these markers in the source language of each subcorpus (Portuguese source texts-English target texts; English source texts-Portuguese target texts). As the corpus will feature parallel alignment, this will allow us to compare original and interpreted versions and determine whether or not the anaphoric links are preserved in the target speeches (Araújo & Correia, in press). If so, what strategies are implemented by interpreters to successfully reformulate the anaphoric chain? If not, to what extent does the rupture of the anaphoric chain impair the conveyance of the speakers' message? And which alternative strategies might have been employed?

While the corpus is not yet available in electronic format, preliminary results have shown that anaphoric chain rupture often leads to loss of coherence of the interpretations, but in some cases interpreters manage to solve anaphoric inconsistencies detected in original speeches. We have observed that problems occur when a) the antecedent is misidentified, b) the anaphoric term is misidentified, b) the anaphoric term is not identified at all, and c) when the identification of the anaphoric term raises ambiguity. These conclusions, tentative at this point, will be complemented by quantitative data to be yielded from the corpus.

## References

Araújo, S. and Correia, A. (in press). "Corpus y formación de intérpretes". In M. T. Sánchez, S. Álvarez, V. Arnáiz, T. Ortego, L. Santamaría and R. Fernández (eds) *Metodologías y aplicaciones en la investigación en traducción e interpretación con corpus*. Valladolid: Universidad de Valladolid, Secretariado de Publicaciones e Intercambio Editorial, 165-181.

Blum-Kulka, S. (1986). "Shifts of cohesion and coherence in Translation". In J. House & S. Blum-Kulka (eds) *Interlingual and Intercultural communication: Discourse and cognition in Translation and Second Language Acquisition Studies*, Tübingen: Narr, 17-35

Gerzymisch-Arbogast, H. (2000). "'Text-bound interpretation': Zum Aufschlußwert der Textdimension für die Dolmetschforschung". In P. A. Schmitt (ed.) *Paradigmenwechsel in der Translation. Festschrift für Albrecht Neubert zum 70. Geburtstag*, Tübingen: Stauffenberg, 101-119.

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

Shlesinger, M. (1995). "Shifts in Cohesion in Simultaneous Interpreting". *The Translator*, Vol. 1, No. 2, 193-214.

# HOW CAN INTERPRETING CORPORA EXTEND OUR KNOWLEDGE ON INTRUSIVE “WE” IN SI?

BARBARA DUMARA, UNIVERSITY OF WARSAW

[barbaradumara@gmail.com](mailto:barbaradumara@gmail.com)

Intermodal corpora are widely believed to be the future of translation/interpreting studies. In 2010–2012, the present author compiled a set of 105 English and 107 Polish speeches in four versions (original speech, official transcription, translation and interpreters’ output) (with videos) given in the European Parliament to analyse the distribution of *we* markers in original texts, translated texts and interpreted texts. As a result, strong evidence was gathered to show in what circumstances interpreters’ use *we* structures even though such structures do not exist in the original speech and are not required linguistically, a phenomenon that has been already observed by Morven Beaton-Thome who called them  $\emptyset > we$  shifts. Since numerous questions still need to be answered so that the nature of intrusive *we* may be better understood, this paper is aimed to propose certain paths that may be followed with automated corpora. By describing a process behind the research, which was still rather manual in its nature but based on a representative intermodal sample and designed in accordance with the rationale of corpus linguistics, the author intends to discuss opportunities brought by automated corpora as well as her first attempts and experiments to use the EPIC corpus to find answers for further questions related to shifts to the first person plural pronoun, including challenges encountered along the way. The author equally wishes to draw the attention of scholars studying various language pairs to the issue of intrusive *we* as such as it seems an interesting topic to be explored in future multilingual research projects undertaken in search of interpreting tendencies.

## References

Bernardini, S., Ferraresi, A., Miličević, M., *From EPIC to EPTIC - building and using an intermodal corpus of translated and interpreted texts*. Power Point Presentation. At: New Ways of Analysing Translational Behaviour in Corpus-Based Translation Studies, SLE 2013 - Split, 18 September 2013. Available at: <http://www.vtc.ugent.be/file/88> [retrieved: 15.12.2014].

Beaton-Thome, M., 2010. Negotiating Identities in the European Parliament. The Role of Simultaneous Interpreting. In: M. Baker et al., eds, 2010. *Text and Context. Essays on Translation and Interpreting in Honours of Ian Mason*. Manchester: St. Jerome Publishing, pp. 117–138.

Calzada Pérez, M., 2007. *Transitivity in Translating. The Interdependence of Texture and Context*. Oxford: Peter Lang.

de Cock, B., 2011. Why we can be you: The use of 1st person plural forms with hearer reference in English and Spanish. *Journal of Pragmatics* 43(11), pp. 2762–2775.

Dumara, B., 2014, (in prep.) Discourse analysis for interpreters: the consequences of intrusive "we" in simultaneous interpreting. Post-conference paper. Teaching Translation and Interpreting 4. Łódź. 28–29 November 2014.

Dumara, B., 2014, (in prep.) Polish vs. British we in the European Parliament. Post-conference paper. Political Linguistics III. Warsaw. 8–10 May 2014.

- Gile, D., 2009. The Effort Models of Interpreting, In: D. Gile, 2009. *Basic concepts and models for interpreter and translator training*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 107–120.
- Levý, J., 1967. Translation as a Decision Process. In: L. Venutti, ed., 1995. *The Translation Studies Reader*. London: Routledge, pp. 148–159.
- Kamio, A., 2001. English generic *we*, *you*, and *they*: An analysis in terms of territory of information. *Journal of Pragmatics* 33(7), pp. 1111–1124.
- Koturbasz, B. 2012. *Ideology in the European Parliament: We markers in one-minute speeches and their modifications in translated and interpreted discourse*. MA thesis [unpublished]. University of Gdańsk.
- Petersoo, P., 2007. What does ‘we’ mean? National diexis in the media. *Journal of Language and Politics* 6(3), pp. 419–436.
- Puurtinen, T., 2000. Translating Linguistic Markers of Ideology. In: A. Chesterman, N. Gallardo SanSalvador and Y. Gambier, eds, 2000. *Translation in Context . Selected contributions from the EST Congress, Granada 1998*. Amsterdam/Philadelphia: John Benjamins, pp. 177–186.
- Schlesinger, M., 1998. Corpus-based Interpreting Studies as an Offshoot of Corpus-based Translation Studies. *Meta: Translators' Journal* 43(4), p. 486–493.

# THE TRANSLATION CHALLENGES OF PREMODIFIED NOUN PHRASES IN SIMULTANEOUS INTERPRETING FROM ENGLISH INTO ITALIAN: A CORPUS-BASED STUDY ON EPIC

SERENA GHISELLI, UNIVERSITY OF BOLOGNA

[serena.ghiselli2@studio.unibo.it](mailto:serena.ghiselli2@studio.unibo.it)

This paper examines the handling of complex noun phrases in the simultaneous interpretation into Italian of English speeches of the electronic corpus EPIC. The complex noun phrases analyzed in this study are noun phrases with two or more premodifying items included in the following categories: nouns, adjectives, adverbs, cardinal numbers and genitives.

EPIC contains transcripts of speeches delivered at the European Parliament in February 2004 in English, Italian and Spanish and the interpretation of each speech into the other two languages involved. EPIC corpus is structured into nine sub-corpora. This comparable analysis is based on two sub-corpora: the sub-corpus of English source speeches (SS) and the corresponding sub-corpus of Italian target speeches (TS). The SS analyzed are 81, for a total of 42,705 words.

After a theoretical overview of the research field of corpus-based interpreting studies, this paper deals with the description of adjectives, nouns, adverbs, cardinal numbers and genitives with the function of premodifiers in the English language. The structure of a premodified noun phrase and the genitive structure are described.

The aim of the project is extracting complex noun phrases from a large sample of authentic English speeches and comparing the extracted phrases with their corresponding translation into Italian in order to study the strategies used by interpreters. The hypothesis is that complex noun phrases pose a translation challenge in simultaneous interpreting from English into Italian because the two languages have structural and lexical diversities. Moreover, the high information density of the strings of modifiers brings about specific processing constraints due to memory overload. The expected reaction to this difficult interpreting situation is that interpreters adopt output strategies, such as approximation and generalization, or emergency strategies, such as omission of text segments and parallel reformulation.

EPIC is POS-tagged, so it is possible to extract data automatically by means of specific item sequences in CQP language through the advanced query function. Sequences to retrieve nouns premodified by other nouns, adverbs, adjectives, cardinal numbers or genitives are developed.

The selected noun phrases are compared with the corresponding translation into Italian and a classification of the strategies used by interpreters is drawn up. The percentages of the word strings out of the total word count of every text are calculated. After the collection and comparison of all the data, quantitative and qualitative conclusions on the translation of premodified noun phrases in simultaneous interpreting are drawn.

## References

- Bendazzoli C. (2010) *Corpora e interpretazione simultanea*, Bologna, Asterisco
- Kruger A., Wallmach K. and Munday J. (eds) (2011) *Corpus-based translation studies: research and applications*, London, Continuum
- Quirk R., Greenbaum S., Leech G. e Svartvik J. (1985) *A comprehensive grammar of the English language*, Harlow, Longman
- Riccardi A. (2005) "On the Evolution of Interpreting Strategies in Simultaneous Interpreting", *Meta*, vol. 50, n°2: 753-767
- Russo M., Bendazzoli C., Sandrelli A. and Spinolo, N. (2012) "The European Parliament Interpreting Corpus (EPIC): implementation and developments" in Straniero Sergio F., Falbo C. (eds) (2012) *Breaking Ground in Corpus-based Interpreting Studies*, Bern, Peter Lang
- Sandrelli A., Bendazzoli C. and Russo M. (2010). "European Parliament Interpreting Corpus (EPIC): Methodological issues and preliminary results on lexical patterns in SI", *International Journal of Translation* 22 (1-2), 165-203
- Straniero Sergio F. and Falbo C. (eds) (2012) *Breaking Ground in Corpus-based Interpreting Studies*, Bern, Peter Lang

# INTERPRETING UNIVERSALS: A STUDY ON EXPLICITNESS IN THE INTERMODAL EPTIC CORPUS

NICCOLÒ MORSELLI, UNIVERSITY OF BOLOGNA

[nmorselli@hotmail.it](mailto:nmorselli@hotmail.it)

Explicitation is one of the most popular ‘candidates’ in the search for translation and interpreting universals. The investigation presented in this study, stemming from a master’s degree dissertation, is a first attempt at searching for explicitness in Italian and English texts, both translated and interpreted in the simultaneous mode.

The study set out to test the findings of the quantitative analysis on explicitness in English interpretations carried out by Kajzer-Wietrzny (2012) in the *European Parliament Translation and Interpreting Corpus* (EPTIC). EPTIC (Bernardini *et al.* 2013) is a bilingual, bidirectional and intermodal corpus of EP plenary session speeches in English and Italian, a resource newly developed in Forlì which is so far a unique corpus of its kind as it contains both transcriptions of speeches and their interpreted versions, as well as their written counterparts in the form of minutes and their translations.

Kajzer-Wietrzny’s indicators of explicitness for the investigation of English (linking adverbials, apposition markers and optional that) were matched by comparable indicators for the investigation of Italian and applied to the relevant subcorpora of EPTIC. First, quantitative analysis was carried out, both from a monolingual comparable perspective (comparing speeches and interpretations in the same language), and from an intermodal perspective (comparing interpretations and translations). Second, a parallel qualitative analysis was performed. Some interesting differences according to language directionality emerged, such as the Italian interpreters’ preference to add apposition markers, which they used more frequently than English speakers, or the tendency of English interpreters to leave out linking adverbials more frequently than Italian speakers. Furthermore, parallel concordances allowed to highlight some shortcomings of traditional monolingual comparable analyses, whose results are sometimes disconfirmed by a parallel investigation. These results show the huge untapped potential of bilingual, bidirectional and intermodal corpora like EPTIC, and the need to enlarge the corpus in order to gain further research insights on the nature of translation and interpreting universals.

# GENERATING AND COMPILING AN INTERPRETING CORPUS FOR A PHD DISSERTATION: CHALLENGES OF CORPUS DESIGN, ACCESSIBILITY, DATA COLLECTION AND PROCESSING

DANUTA PRZEPIÓRKOWSKA, GESIS - LEIBNITZ INSTITUTE FOR THE SOCIAL SCIENCES /  
INSTITUTE OF APPLIED LINGUISTICS UNIVERSITY OF WARSAW  
[d.przepiorkowska@uw.edu.pl](mailto:d.przepiorkowska@uw.edu.pl)

Corpus-based PhD dissertations in Translation Studies have mushroomed in recent years thanks to the very wide availability of corpora, whether mono- or multilingual, for research in linguistics and translation. However, a review of available online resources seems to suggest that PhD dissertations in Interpreting Studies which are based on corpora are quite scarce. One important reason is the limited availability of generally accessible interpreting corpora, and available language pairs within them. Another reason is that such doctoral projects are generally time-consuming and work-intensive.

Doctoral candidates who embark on an empirical PhD project in Interpreting Studies usually have to choose from two options: 1) compile their own corpus based on available resources which have not been structured into a corpus, or 2) generate their own corpora using methods applied in empirical research.

One example of the former approach is the doctoral dissertation of Morven Beaton, based on a corpus of German-to-English simultaneous interpretations from the European Parliament, which she downloaded from the EP website and processed for her own needs (a total of 7.5 hours of recording).

The example of the second approach is described in this paper. The empirical PhD project in Interpreting Studies was an offshoot of a cross-national project in social sciences and focused on interpreters' agency in social research (more specifically: interpreting focus group discussions). Permission was obtained to use a focus group guide from the cross-national project.

After formulating her research questions, the author prepared a preliminary research design, consulted it with two eminent IS professors and then set out to compile a corpus. As the design was based on ten interpreters interpreting the same event (a focus group discussion), the first step was to organise and record the event in the source language and then to arrange for ten different interpreters to interpret the same discussion into the target language.

The paper describes challenges faced in the process: finding a research agency to conduct a professional focus group, briefing the moderator who facilitated the discussion, checking the agency's technical capabilities, recruiting interpreters with experience in this type of interpretation, making appointments, preparing the set-up and recording the material.

The paper also raises confidentiality issues, ethical considerations and the technical challenges that emerged in the process, as well as the problems faced in transcription, approach to data analysis (based on the qualitative paradigm) and the costs involved in compiling the corpus (the funding was provided from the social sciences project).

The final corpus compiled for the project consists of 2 hours of source material (in Polish) and 20 hours of interpreted material (in English).



Hopefully, the experience of this case study may prove useful for other small-scale research projects with a constrained budget.

## **References**

Beaton, Morven (2007) 'Interpreted Ideologies in Institutional Discourse. The Case of the European Parliament', *The Translator* 13(2):271–296.

Shlesinger, M. (1998). "Corpus-based Interpreting Studies as an Offshoot of Corpus-based Translation Studies". *Meta*, 43 (4), 486–493.

Silverman, David (2012) *Interpreting Qualitative Data. 4th Edition*. SAGE Publications Ltd.

Underlying project:

<http://www.cmi.no/research/project/?11118=legal-cultures-in-transition-the-impact-of-eu>

# INTERPRETING TEACHING IN LIBYA: CHALLENGES AND FUTURE INSIGHTS FROM A CORPUS-BASED PERSPECTIVE

MOHAMMED ZAGOOD, EL-MERGIB UNIVERSITY, KHOMS

[zagood81@yahoo.co.uk](mailto:zagood81@yahoo.co.uk)

Interpreting is a mediate oral translation of oral discourse (Jones, 1998). Interpreting is of different modes and types: consecutive, simultaneous, liaison, and whispered. Interpreting is being taught in Libya as two courses to the master degree students at the Libyan Academy in Tripoli, Libya. One course is named *Interpreting I* in which students study interpreting theory and methodology and practice interpreting from English into Arabic. *Interpreting II* in which students study interpreting theory and methodology and practice interpreting from Arabic into English. The main course book taught is *Conference Interpreting Explained* by Rodrick Jones (1998) in addition to other articles/essays from other sources. Moreover, most of the time is given to students to practice the four modes of interpreting. These two courses were taught by the same lecturer. This paper is limited to recap and evaluate the teaching process of both courses: *Interpreting I* (English-Arabic Interpreting) and *Interpreting II* (Arabic-English Interpreting). Therefore, this paper is an attempt to evaluate the course structure, materials taught, speeches interpreted and time spent during the teaching of these courses. This evaluation is done via a questionnaire which consists of a number of items followed by three open questions given to the students who attended the courses. The answers are analyzed to recognize the challenges and difficulties encountered and to suggest solutions for the sake of improving the teaching process of interpreting in future.

# WHAT HAVE WE LEARNED FROM CSI AND WHAT CAN OTHERS LEARN?

JANA ZIDAR FORTE, UNIVERSITY OF LJUBLJANA

[jana.ZidarForte@ff.uni-lj.si](mailto:jana.ZidarForte@ff.uni-lj.si)

Despite various technical and practical obstacles that may arise in the process of gathering interpreted texts, their processing and corpus compilation, which also make it an especially time consuming task, corpus-based interpreting studies (CIS) have proved invaluable in the development of interpreting research. After considerably lagging behind, CIS are nowadays gradually gaining a foothold in the general sphere of corpus studies. However, the majority of research is conducted by individual researchers or as part of final theses, which invariably limits their scope and use. On the other hand, collective approach, e.g. as the one adopted for the EPIC corpus and the subsequent EPTIC corpus, may yield better results, notably with respect to corpus size, linguistic combinations, and observed parameters, to name a few.

According to Setton (2011), the prevailing research orientations and topics in CIS include features of the interpreting product or 'interpretese': collocations, part-of-speech distributions and frequencies as well as lexical richness. Several studies within CIS focused on the use of lexical patterns in interpreting (Laviosa 1998, Bendazzoli and Sandrelli 2010, Shlesinger 2009, Shlesinger and Ordan 2012, Bernardini et al. 2013). Other prevalent orientation is determining key factors in performance and quality (Prada Macias 2009, Bendazzoli 2011). Authentic corpora also prove indispensable in understanding cognitive processes and strategies in SI, and, following the recent socio-cultural turn in translation studies, research also focuses on social and ideological aspect. There were also studies regarding specific interpreting settings, i.e. television interpreting (Cencini 2000; Falbo 2012), interpreting in football press conferences (Sandrelli 2012); interpreting in court hearings (Biagini 2012).

Regarding practical pedagogical applications, corpora with professional interpretations invariably present an important source of authentic material that can be used as a reference or as practice material for interpreting trainees. Interpreter trainers, on the other hand, can benefit from CIS findings as they offer detailed insight into the process and even more specifically the result of interpreting (sometimes on its own, or compared to translation). Thus, they enable interpreter trainers to complement (or even move beyond) anecdotal and purely experience-based assumptions and arguments that have long been prevalent in interpreter training, and replace them with evidence-backed findings. However, interpreter trainers are more often than not professional interpreters not closely linked to the research world, which makes it hard for them to learn about the recent developments in CIS, let alone apply them to teaching practice in the classroom where time is limited and rather spent on interpreting exercises.

This is why at least part of the future efforts in CIS should also focus on knowledge transfer, especially through existing platforms within the university network EMCI or the professional association AIIC and similar national or regional associations. Besides widening the scope of corpus-based interpreting research to other aspects and even relating to other disciplines, the CSI community should strive to get closer to actual (and future) interpreters

## References

- Bendazzoli, Claudio and Annalisa Sandrelli (2009) Corpus-Based Interpreting Studies: Early Work and Future Prospects, *Tradumatica 7: L'applicazione del corpus linguistico a la traduzione*. Available online at <http://www.fti.uab.es/tradumatica/revista/num7/articles/08/08central.htm> (accessed 9 January 2015)
- Bendazzoli, Claudio (2011) Disfluencies in Simultaneous Interpreting: A Corpus-Based Analysis. In: Kruger, Alet et al. (ed.) *Corpus-Based Translation Studies*. London, Bloomsboory. 282-306
- Bernardini, S. et al. (2013) From EPIC to EPTIC - building and using an intermodal corpus of translated and interpreted texts. Split, ESL.
- Biagini, Marta (2012) Data collection in the courtroom: challenges and perspectives for the researcher. In: Straniero Sergio, Francesco and Caterina Falbo (ed) *Breaking Ground in Corpus-based Interpreting Studies*. Bern, Peter Lang. 231-252.
- Cencini, M. (2000). Il Television Interpreting Corpus (TIC). Proposta di codifica conforme alle norme TEI per trascrizioni di eventi di interpretazione in televisione. Unpublished dissertation. Advanced School for Translators and Interpreters (SSLMIT), University of Bologna at Forlì
- Falbo, Caterina (2012) Topical Coherence in television interpreting: question/answer rendition. in: Straniero Sergio, Francesco and Caterina Falbo (ed) *Breaking Ground in Corpus-based Interpreting Studies*. Bern, Peter Lang. 155-186.
- Laviosa, S. (1998). The Corpus-based approach: a new paradigm in Translation Studies, *Meta*, 43/4. Prada Macias 2009
- Sandrelli, Annalisa (2012) Introducing FOOTIE (Football in Europe): simultaneous interpreting in football press conferences. in: Straniero Sergio, Francesco and Caterina Falbo (ed) *Breaking Ground in Corpus-based Interpreting Studies*. Bern, Peter Lang. 119-154.
- Setton, Robin (2011) Corpus-Based Interpreting Studies (CIS): Overview and Prospects. in: Kruger, Alet et al. (ed.) *Corpus-Based Translation Studies*. London, Bloomsboory. 33-75.
- Shlesinger, Miriam (2008). Towards a definition of Interpretese. An intermodal, corpus-based study, in G. Hansen, Chesterman, A. & H. Gerzymisch-Arbogast (eds.) (2008), *Efforts and Models in Interpreting and Translation Research. A tribute to Daniel Gile*. Amsterdam / Philadelphia: John Benjamins.
- Shlesinger, Miriam and N. Ordan (2012). More spoken or more translated? Exploring a known unknown of simultaneous interpreting. *Target* 24, 43-60.