

ORALITY AND GENDER: A CORPUS-BASED STUDY ON LEXICAL PATTERNS IN SIMULTANEOUS INTERPRETING

Mariachiara Russo

mariachiara.russo@unibo.it
University of Bologna at Forlì

Abstract

Corpus-based Translation and Interpreting Studies so far explored several phenomena pertaining to different dimensions of language production. In particular, Laviosa (1998) investigated lexical variety (LV), i.e. linguistic richness, and lexical density (LD), i.e. prevalence of content words, in English original and translated texts and reported a higher degree of both features in the former vs. the latter. Inspired by these results, another study analysed the European Parliament Interpreting Corpus (EPIC) to verify Laviosa's results comparing English, Italian and Spanish source and target speeches in these three languages; this study obtained less clear-cut and language-dependent results (Russo et al. 2006). LV and LD in either original or interpreted speeches could be not only language- but also gender-dependent. The present quantitative study showed mixed results depending on the language. Yet, some statistically significant trends emerged, among them higher LD in the speeches produced by Italian female vs. male Members of the European Parliament (MEPs), and higher LV in English and Spanish female vs. male interpreted speeches.

RIASSUNTO

I Corpus-based Translation and Interpreting Studies hanno, fino ad ora, esplorato molteplici fenomeni appartenenti a diverse dimensioni della produzione linguistica. In particolare, Laviosa (1998) ha studiato la varietà lessicale (VL), cioè la ricchezza linguistica, e la densità lessicale (DL), cioè la prevalenza di parole contenute, in testi originali e in testi tradotti in lingua inglese, e ha riscontrato una maggiore presenza di entrambe nei testi originali, rispetto a quelli tradotti. Un successivo studio ha analizzato l'European Parliament Interpreting Corpus (EPIC) allo scopo di verificare se i risultati di Laviosa venissero confermati comparando testi originali e testi interpretati simultaneamente in inglese, italiano e spagnolo; in questo caso, i risultati si sono

rivelati non altrettanto univoci e maggiormente dipendenti dalle coppie di lingue analizzate (Russo et al. 2006). VL e DL nei discorsi originali e interpretati potrebbero dipendere non solo dalla lingua, ma anche dal genere dell'oratore. I risultati ottenuti dal presente studio quantitativo differiscono a seconda della lingua analizzata. Emergono, tuttavia, alcune tendenze statisticamente significative, quali una maggiore DL nei discorsi prodotti da Eurodeputati di sesso femminile rispetto a quelli di sesso maschile, e una maggiore VL nei testi interpretati in inglese e in spagnolo da donne rispetto a quelli interpretati da uomini.

Keywords: Lexical density. Lexical Variety. Source speech. Target speech. Gender.

Parole chiave: Densità lessicale. Varietà lessicale. Testo fonte. Testo meta. Genere.

Manuscript received on May 23, 2015 and accepted for publication on September 19, 2015.

1. Introduction

The way in which someone speaks is one of the most revealing factors of their personality, identity, intentions or motivations, among other things. It was only in the last decades, however, that speech behaviour or speaking styles were analysed from a gender perspective. Over the years, this perspective was enriched by the reflection that the sameness of the activities in which men and women are engaged make them a Community of practice (Eckert & McConnell-Ginet 1992) and may also explain some of the differences observed in their linguistic output (see § 1.1).

The present paper explores the possibility that the speeches produced by the simultaneous interpreters working in the European Parliament (EP), a particular community of practice,¹ might indicate trends highlighting differences between males' and females' richness and variety in speaking patterns. These will be investigated by measuring the lexical density (LD) and the lexical variety (LV) of the target speeches contained in the European Parliament Interpreting Corpus (EPIC).² We hypothesise that there might be a gender-related LD and LV difference in favour of female interpreters, on the assumption that the qualitative difference reported by López García & Morant (1991) and further specified (§ 1.1) may be the result of a more content-rich style and greater lexical availability. However, since target speeches are linguistic productions depending on source speeches, this study tries to verify whether a similar pattern emerges also in the EPIC speeches produced by female and male original speakers (Members of the European Parliament, MEPs), who

-
1. We consider EP interpreters a community of practice because they share the following features: they are highly skilled, rigorously selected and multilingual professionals who work under tight time constraints in the same setting; they produce target speeches (TS) whose contents and topic-comment planning depend on source speeches (SS) which differ in delivery style, input rate, text length, topic etc.; they are familiar with EP discourse dynamics, procedures, topics, speech typologies, speakers (MEPs), etc.
 2. EPIC is an open, machine-readable, on-line resource for simple and advanced queries at <http://sslmitdev-online.sslmit.unibo.it/corpora/corporaproject.php?path=E.P.I.C>. It is also freely distributed by ELRA in its full video, audio and transcript version only for academic purposes at http://catalog.elra.info/product_info.php?products_id=1145

make up another specific community of practice (Wodak 2003, as quoted in Bergvall et al. 2011). Finally, in the light of the results of our previous study in which interpreted speeches generally displayed a slightly higher LD and lower LV vs. original speeches (Russo et al. 2006), we shall also try to ascertain whether the same results are obtained when comparing female speakers and female interpreters in favour of the latter.

1.1. Brief overview of gender-based studies

The first gender-based studies appeared in the Seventies and aimed at detecting socio-psychological aspects conveyed through language, such as stereotypes attached to men and women, and power asymmetries in their roles (Labov 1972, Lakoff 1975). In their wake, studies on language and gender focused on language use in interaction following the Conversational Analysis paradigm (Tannen 1990, 1994) to highlight dominance or cooperative gendered communicative exchanges and differences in discourse conversational cultures.

Subsequently, linguistic and pragmatic studies concentrated on phonological, morpho-syntactical and lexical variants and turn-management strategies used preferably by men or by women (Muchnik 1997; Mikros 2013). These studies covered a wide range of communication forms: oral discourse (Holmes 1990, Labov 1990), informal written texts (Mulac et al. 1990), or texting (Herring 1996), among others. In particular, the following speech patterns were described (Koppel et al. 2002): males used more determiners, cardinal numbers and a more specific and informative style, while women used more first and second personal pronouns (*I, you*). Concerning the use of personal pronouns, Tannen (1990, 1994) reported a wider use of *we* and other markers of courtesy and cooperative spirit such as the discourse marker *you know*, rhetorical questions and question tags (these last features were also reported by García Mouton 1999). The greater use of personal pronouns (among other distinctive features) in women has recently been reported also by Mikros (2013), who followed a corpus-based gender approach to perform automated text categorization and determine author-gender attribution with successful results. A previous study based on the same approach by Shlesinger et al. (2009) had tried to determine translator-gender attribution with unsuccessful results. As far as Spanish is concerned, López García & Morant (1991) reported marked differences in male and female speech patterns with reference to the use of interjections, vocatives, prefixes, suffixes, word truncations, negations, comparisons, euphemisms and preferred sentences.

Some authors went as far as indicating an overall qualitative difference in favour of female speech styles:

Todos los estudios sociolingüísticos llevados a cabo en distintos países del mundo en los últimos 20 años coinciden en observar que el habla de las mujeres es *cualitativamente mejor* que la de los hombres: ya que se trate del español de Bahía Blanca, del inglés de Norwich, o de la situación lingüística inserta en modelos culturales completamente alejados del nuestro, como es la de los indígenas siberianos chukchees, lo cierto es que en iguales condiciones de edad, clase social y nivel educativo, las mujeres tienen un *vocabulario más rico*, una sintáxis más completa, y una pronunciación más cuidada que sus compañeros varones. (López García & Morant 1991:11-12, my italics).

However, it must be noted that traditional investigations on gender-based speech behaviour were confronted with criticism by authors who contested such gender polarization because oral communication is produced by individuals who cannot possibly be reduced to a homogenous binary distinction (Bergvall et al. 1996):

It should be clear that there is nothing about any particular activity or communicative task (or communicative style) that is itself inherently female or male, nor are the speaking patterns of individuals mindless habits that exist as disembodied linguistic reflexes. [...] We can, in addition, document the similarity of the speech of the women and the men [engaged in similar activities] and hypothesise that the setting and related speaking activities themselves were responsible for the sameness in language. (Freed 1996: 68)

The differences are therefore to be traced back to situation-based – rather than gender-based – factors, which seem to determine individuals' speech similarities or otherwise, as speakers make up what Eckert & McConnell-Ginet (1992) call a "Community of practice". Freed further specifies (1996: 69):

We need to embark on a close examination of different communities and settings and of the various individuals who move in and out of these communities as they engage with one another in talk.

2. Materials and method

2.1. Background: lexical density and lexical variety in previous T&I research

Lexical density and lexical variety are two quantitative measures that provide some indications concerning the information density and vocabulary size of the language produced by the speaker. They were typically investigated in corpus-based studies and have already been analysed also in Translation and Interpreting (T&I) research. Laviosa (1998), who published the first corpus-based study on translation from this perspective, provided the following operational definitions. Lexical density is expressed as a percentage and is

calculated by subtracting the number of function words in a text from the number of running words (which gives the number of lexical words) and then dividing the result by the number of running words (Laviosa 1998: 565). Lexical variety is the proportion of high frequency vs. low frequency words and is calculated by producing word frequency lists and then selecting the top 100 words from them (list heads). Then, the overall word count is obtained for the list head, and the percentage of the corpus covered by the list head is calculated. In other words, if the 100 most frequent words in a corpus account for a large part of that corpus, the lexical variety is low, because it means that the same words are used over and over again.

Laviosa contrasted English non-translational texts vs. English translational texts and found that: a) translated texts have a relatively lower percentage of content words vs. grammatical words³ (i.e. their lexical density is lower); b) the proportion of high frequency words vs. low frequency words is relatively higher in translated texts; c) the list head of a corpus of translated texts accounts for a larger area of the corpus (i.e. the most frequent words are repeated more often) (Laviosa 1998: 563).

Following the same methodology, our research group wanted to ascertain whether similar trends emerged when comparing original and interpreted speeches contained in EPIC, a pos-tagged, lemmatised and indexed corpus made up of nine sub-corpora, three sub-corpora of English, Spanish and Italian original speeches and 6 sub-corpora of the corresponding simultaneously interpreted versions in these three languages, namely English>Italian, English>Spanish, Italian>English, Italian>Spanish, Spanish>English and Spanish>Italian (for a detailed description of EPIC, its applications and developments, see Sandrelli et al. 2010 and Russo et al. 2012). Our results differed from Laviosa's ones and shed light on the role of language pair and directionality (Sandrelli & Bendazzoli 2005; Russo et al. 2006).

With respect to lexical density, the effect noted by Laviosa (1998) in translated texts, i.e. a lower lexical density than in texts originally written in English, was not confirmed by our studies where the lexical density tended to be slightly higher in interpreted speeches than in speeches originally delivered in the same language (with only two exceptions for the Spanish into English and the Spanish into Italian sub-corpora). Interestingly, lexical density patterns in EPIC were consistent within individual languages (considering both

3. Content – or lexical – words are nouns, verbs, adjectives and adverbs; function – or grammatical – words are prepositions, determiners, pronouns, conjunctions, particles, numerals, interjections, negatives, etc.

original and interpreted speeches, about 57% in English, 58-59% in Italian and 53-54% in Spanish).

As far as lexical variety is concerned, Laviosa's finding that translated texts reveal a higher proportion of high frequency words vs. low frequency words, was confirmed in our studies for the speeches interpreted into English and into Spanish which featured a higher percentage of high frequency words than source speeches in the same languages (in other words, the TL list heads accounted for a larger area of each respective sub-corpus). This means that both interpreted English and interpreted Spanish in EPIC have a lower lexical variety than original English and original Spanish (in line with Laviosa's findings on translational English). A similar result regarding lower lexical variety in oral interpretations vs. written translations of the same source text was obtained by Shlesinger (2009), who applied a different method – type-to-token ratio – to calculate linguistic richness in her intermodal corpus. She carried out both an intra- and inter-subject evaluation and, according to her results, the only variable determining this finding seemed to be the translational activity.

By contrast, the opposite trend was found in EPIC Italian material (Russo et al. 2006): original Italian speeches appeared to have a lower degree of lexical variety than interpreted Italian speeches from English and Spanish. This may suggest that Italian interpreters, irrespective of the language-pair specificity, prefer to use as wide a vocabulary as possible. However, since Italian speeches were only 17, this trend could be disproved by a wider sample size.

2.2 Methodology

Based on the above-mentioned results and following the same methodology, the present study investigates both measures, that is lexical density (LD) and lexical variety (LV), by separately considering the speeches delivered by male and female speakers and interpreters, to gain an insight into their expressive skills and linguistic richness from a gender-based perspective.

A t-test was performed on each pair of results – that is the number of lexical items for LD and the proportion of list-head items for LV – to verify whether the observed gender-based differences reached statistical significance.

2.3 EPIC size and composition by gender

As already anticipated, EPIC is a trilingual corpus consisting of nine sub-corpora. At present it is not equally balanced between its linguistic components as English original speeches and corresponding interpreted speeches into

the other two languages account for the largest part with 81 source speeches (SS) and 162 target speeches (TS), followed by Spanish (21 SS and 42 TS) and Italian (17 SS and 34 TS). The speeches were delivered in three different modalities – read, impromptu and mixed – and covered several topics, but 70% of them concerned health, justice and politics. The original speeches and the interpreters' performances were recorded during the EP Plenary sessions of February and March 2004 (Monti et al. 2005, Sandrelli et al. 2010). Table 1 shows EPIC size and composition by speech gender.

Sub-corpus	Females		Males		F + M	
	N. of speeches	Tot. running words	N. of speeches	Tot. running words	N. of speeches	Tot. running words
org-en	16	6,215	65	39,308	81	45,523
int-en-es	59	31,122	22	9,530	81	40,652
int-en-it	68	28,637	13	9,491	81	38,128
org-it	3	1,551	14	5,550	17	7,101
int-it-en	9	3,585	8	3,617	17	7,202
int-it-es	10	4,514	7	2,927	17	7,441
org-es	7	3,832	14	11,284	21	15,116
int-es-it	21	13,646	0	0	21	13,646
int-es-en	5	5,164	16	8,741	21	13,905
Total	198	98,266	159	90,448	357	188,714

Table 1. EPIC size and composition by gender

Female speakers at the EP are greatly outnumbered by male speakers in our corpus, whereas the opposite occurs with female interpreters. However, since the interpreters' identities are unknown (and the same interpreters may well have interpreted more than one speech), the overwhelming female presence ought to be the impression received by the users of interpreting services in those sessions, based on the voices they heard.

3. Results

We analysed EPIC as a comparable corpus, that is we analysed original speeches in one language and the interpreted speeches in that language (intra-lingual approach), in line with Laviosa's approach. Our first hypothesis was that there might be a gender-related difference in LD in favour of women. The first level of analysis entailed a comparison between female and male original speakers (org-en, org-es, org-it). Our hypothesis has been confirmed by the data for English and Italian, but not for Spanish (see Table 2). However, the higher lexical density found in English female speakers is a trend which does not reach statistical significance. Yet, a word of caution is mandatory since female speakers are over four times less numerous than male speakers for English. This is the case also for Italian female speakers, but not for Spanish where female speakers are only twice less numerous than their male counterparts (§ 2.3). However, despite the difference in sample size, the gender-based differences observed in Italian originals reached statistical significance in both cases (Table 2).

We then proceeded to verify the general trends detected across EPIC, where interpreted speeches were slightly richer in content words than original speeches (Russo et al. 2006, Sandrelli et al. 2010), in order to ascertain whether this could be mainly due to female vs. male speaking styles, the primary objective of our present study.

The emerging trends are divergent across the three languages. Generally, female original speeches show a slightly higher lexical density than female interpreted speeches in the same language, except for Spanish, where it is identical.

When interpreted speeches are compared by gender, our hypothesis was confirmed only for females' speeches interpreted into Italian, but not for those interpreted into Spanish (Table 2). In other words, only female Italian interpreters appear to be responsible for the slightly higher lexical density reported in our previous research. As far as English is concerned, two opposing trends emerged: in the speeches from Italian into English, lexical density is higher for females and the reverse occurred in the speeches interpreted from Spanish into English. This time the difference, albeit small in percentage terms, is highly statistically significant (Table 2).

Sub-corpus	Total running words		Lexical words		Function words		Lexical density	
	Females	Males	Females	Males	Females	Males	Females	Males
org-en	6,215	39,308	3,286	39,308	2,929	19,184	53	51
int-it-en	3,585	3,617	1,851	1,835	1,734	1,782	52	51
int-es-en	5,164	8,741	2,599	4,476	2,565	4,265	50*	51*
org-es	3,832	11,284	1,930	5,757	1,902	5,527	51	53
int-it-es	4,514	2,927	2,297	1,551	2,217	1,376	51	53
int-en-es	31,122	9,530	15,891	4,834	15,231	4,696	51	51
org-it	1,551	5,550	864	3,056	687	2,494	56**	55**
int-es-it	13,646	0	7,298	0	6,348	0	54	0
int-en-it	28,636	9,491	15,614	5,132	13,022	4,359	55	54

Table 2. Lexical density by gender across EPIC. *p = 0.0003 **p = 0.02

As far as lexical variety is concerned, let us start by considering the first level of analysis, that is the comparison between original female speakers vs. their male counterparts (org-en, org-es and org-it; Table 3). Our hypothesis was confirmed only for Italian female speakers.

Sub-corpus	List Head Word Count		% of Sub-corpus	
	Females	Males	Females	Males
org-en	1,159	6,087	35	30
int-it-en	778	781	42*	43*
int-es-en	979	1,743	38	39
org-es	765	1,731	40	30
int-it-es	904	634	39**	41**
int-en-es	5,1901	1,698	33	35
org-it	200	987	23	32
int-es-it	2,324	0	32	0
int-en-it	4,905	1,767	31	34

Table 3. Lexical variety by gender across EPIC. *p = 0.002 **p = 0.007

The further levels of analysis included a comparison between interpreted vs. original speeches, and then female vs. male interpreted speeches. Generally, interpreted speeches show a lower degree of lexical variety vs. originals because the proportion of high frequency words is larger, except for Spanish female target speeches. Finally, the comparison between female and male interpreted speeches indicates that female speeches for all languages have a richer vocabulary than the male interpreted speeches because the proportion of high frequency words is always smaller. In two cases, the difference is highly significant (interpreted speeches from Italian into English and into Spanish) (Table 3).

4. Discussion and conclusions

The conditions under which both MEPs and simultaneous interpreters work in the specific setting of EP Plenary sessions are extremely taxing: speaking time is strictly regulated and kept to a minimum, topics vary greatly as do speakers' modes and speed of presentation. Therefore, we had expected that possible differences between both MEPs and interpreters' speech styles would be determined more by those variables than by gender, and yet some gender-based trends did emerge.

Before discussing the results of our small study, let us consider a general feature characterising EPIC (Russo et al. 2006), which provides an insight into EP speech types. Interestingly, lexical density patterns in EPIC are consistent within individual languages (considering both original and interpreted speeches, about 57% in English, 58-59% in Italian and 53-54% in Spanish). This further confirms that speeches typically interpreted at the EP are closer in nature to written texts⁴ along the written-to-oral continuum. Indeed, source speeches are often read or presented in a mixed mode (i.e., read and impromptu), mostly at breakneck speed: here lies the greatest challenge for interpreters.

As anticipated, trends about gender-based differences emerged both for lexical density (LD) and lexical variety (LV). This places our previous results into a new perspective. Considering LD, that is a larger presence of content vs. functional words, as an indication of greater information density, we could assume that female MEPs, who display higher LD vs. their male colleagues, probably try to communicate as much as they can in the limited time allotted

4. According to Halliday (1993, as quoted in Castello 2004), lexical density is a distinctive measure of the difference between spoken and written language since it is likely to be twice as high in the written mode.

to them. This is the case for English and Italian MEPs, and not for Spanish MEPs (Table 3). As already mentioned, female English and Italian MEPs are much less numerous than male MEPs (Table 1), which may lead us to hypothesise that their being a minority speaking to a male majority could play a role in this observation. A qualitative analysis of their speeches is of course indispensable to prove this assumption. Nevertheless, it seems to be an indication of a speech style that appears to be in line with what sociolinguist Wodak (quoted in Bergvall et al. 1996: 421) observed, namely:

that female EU parliamentarians in order to be successful in 'doing politics', develop practices that construct them as 'assertive activists', 'experts', 'being different in a positive way (special bird)', or combinations of this habitus, and that these types of female gender role constructions of successful women are quite different from those found in other institutional or working settings.

Furthermore, the emerging trends for LD differ across the three languages (Table 2). Female source speeches generally tend to have a slightly higher LD than female interpreted speeches in that same language (except for Spanish, where it is identical). This finding differs from the results of our previous study (Sandrelli & Bendazzoli 2005, Russo et al. 2006) and seems more in line with Laviosa's observation on translational vs. non-translational lexical patterns. The present result could be due to the different sizes of the sub-corpora. Another possible explanation may be linked to the interpreters' need to ease processing and output efforts due to the speakers' input rate or the delivery mode: they may be required to carry out considerable input chunking operations leading to syntactical restructuring and lexical deletions.

In contrasting female and male interpreting outputs, once again results differ across languages and directions (Table 2). Female interpreted speeches between English and Italian (in both directions) display more content words. From a current qualitative study on EPIC (Russo forthcoming), we observed that additions of content words due to self-repairs, explicitations or synonymic pairs and paraphrastic reformulations of the same concepts are more present in female than male target speeches. Interestingly, our present findings seem to indicate that these two languages, which belong to two different linguistic stocks, do in fact require greater lexical expansion and rearrangement during simultaneous interpreting in this language combination.

As far as lexical variety (LV) is concerned, a measure of linguistic richness and flexibility, heterogeneous trends emerged (Table 3). In comparing female and male EU parliamentarians, only female Italian MEPs seem to have a wider range of lexical selection, hence expressive resources.

The same occurred when comparing source and target speeches in the same language: only female Spanish interpreters seem to pay greater attention to their vocabulary and expressive versatility.

Finally, the only consistent trend in favour of female target speeches emerged when female and male interpreted performances were contrasted. The comparison between female and male interpreted speeches indicates that female speeches for all languages display a richer vocabulary than the speeches interpreted by males, because the sub-corpus list head proportion is always smaller (Table 3). Considering that linguistic variety is generally one of the attributes of an elegant speaking style, female simultaneous interpreters are probably perceived as being stylistically more refined by their clients.

To conclude, our study highlighted some unexpected gender-based trends concerning information density (lexical density, LD) and expressive creativity (lexical variety, LV) in source and target speeches contained in EPIC. However, a better balance in the size of the 9 sub-corpora and a qualitative study on a larger set of performances are necessary to test the soundness of these preliminary insightful observations.

References

- BERGVALL, Victoria; Janet M. Bing & Alice F. Freed (eds.) (1996) *Rethinking Language and Gender Research. Theory and Practice*. New York: Addison Wesley Longman.
- BERGVALL, Victoria; Janet M. Bing & Alice F. Freed. (2011) "Rethinking Language and Gender Research. Theory and Practice." In: Wodak, Ruth; Barbara Johnstone & Paul E. Kerswill (eds.) *The Sage Handbook of Sociolinguistics*. London: Sage, pp. 411-423.
- CASTELLO, Erik. (2004) "Calcolo della densità lessicale e dell'intricatezza grammaticale di corpora linguistici." In: Taylor Torsello, Carol; Maria Grazia Busà & Sara Gesuato (eds.) *Lingua Inglese e Mediazione Interlinguistica. Ricerca e didattica con supporto telematico*. Padua: Unipress, pp. 131-151.
- ECKERT, Penelope & Sally McConnell-Ginet. (1992) "Communities of practice: where language gender and power all alive." In: Hall, Kira; Mary Buchholz & Birch Moonwoman (eds.) *Locating power: proceedings of the second Berkeley women and language Conference*. Berkeley: Berkeley Women and Language Group, University of California, pp. 89-99.
- FREED, Alice F. (1996) "Language and gender research in an experimental setting." In: Bergvall, Victoria et al. (eds.) *Rethinking Language and Gender Research. Theory and Practice*. New York: Addison Wesley Longman, pp. 54-76.
- GARCÍA MOUTON, Pilar. (1999) *Cómo hablan las mujeres*. Madrid: Arco/Libros.

- HERRING, Susan C. (1996) "Two variants of an electronic message schema." In: Herring, Susan C. (ed.) *Computer-Mediated Communication: Linguistic, Social & Cross-Cultural Perspectives*. Amsterdam: John Benjamins, pp. 81-106.
- HOLMES, Janet. (1990) "Hedges and boosters in women's and men's speech." *Language and Communication* 10:3, pp. 185-205.
- KOPPEL, Moshe; Shlomo Argamon & Anat Rachel Shimoni. (2002) "Automatically categorizing written texts by author gender." *Literary and Linguistic Computing* 17:4, pp. 401-412.
- LABOV, William. (1972) *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- LABOV, William. (1990) "The intersection of sex and social class in the course of linguistic change." *Language Variation and Change* 2, pp. 205-254.
- LAKOFF, Robin. (1975) *Language and woman's place*. New York: Harper & Row.
- LAVIOUSA, Sara. (1998) "Core patterns of lexical use in a comparable corpus of English narrative prose." *Meta* 43:4, pp. 557-570.
- LÓPEZ GARCÍA, Ángel & Ricardo Morant. (1991) *Grámatica femenina*. Madrid: Cátedra
- MIKROS, George. (2013) "Systematic stylometric differences in men and women authors: a corpus-based study." In: Köhler, Reinhard & Gabriel Altmann (eds.) *Issues in Quantitative Linguistics 3. Dedicated to Karl-Heinz Best on the occasion of his 70th birthday*. Lüdenscheid: RAM, pp. 206-223. Electronic version: <http://users.uoa.gr/~gmikros/Pdf/Systematic%20stylometric%20differences%20in%20men%20and%20women%20authors.pdf>
- MONTI, Cristina; Claudio Bendazzoli; Annalisa Sandrelli & Mariachiara Russo. (2005) "Studying directionality in simultaneous interpreting through an electronic corpus: EPIC (European Parliament Interpreting Corpus)." *Meta* 50:4. Electronic version: <http://www.erudit.org/revue/meta/2005/v/n4/019850ar.html>
- MUCHNIK, Malka. (1997) "Men vs. women – Different communication patterns." *Hebrew Linguistics* 41-42, pp. 79-86.
- MULAC, Anthony; Lisa B. Studley & Sheridan Blau. (1990) "The gender-linked language effect in primary and secondary students' impromptu essays." *Sex Roles* 23:9-10, pp. 439-469.
- RUSSO, Mariachiara; Claudio Bendazzoli & Annalisa Sandrelli. (2006) "Looking for lexical patterns in a trilingual corpus of source and interpreted speeches: extended analysis of EPIC (European Parliament Interpreting Corpus)." *FORUM, International journal of interpretation and translation*, 4:1, pp. 221-254.
- RUSSO, Mariachiara; Claudio Bendazzoli; Annalisa Sandrelli & Nicoletta Spinolo. (2012) "The European Parliament Interpreting Corpus (EPIC): implementation and developments." In: Straniero Sergio, Francesco & Caterina Falbo

- (eds.) *Breaking Ground in Corpus-Based Interpreting Studies*. Bern: Peter Lang, pp. 53-90.
- RUSSO, Mariachiara. (forthcoming) "Speech patterns and gender: trends emerging from Epic." Paper presented at *The Forlì First international workshop Corpus-based interpreting studies: the State of the Art*, Forlì 7-8 May 2015.
- SANDRELLI, Annalisa & 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). Electronic version: <http://www.birmingham.ac.uk/research/activity/corpus/publications/conference-archives/2005-conf-e-journal.aspx>
- SANDRELLI, Annalisa; Claudio Bendazzoli & Mariachiara Russo. (2010) "European Parliament Interpreting Corpus (EPIC): Methodological issues and preliminary results on lexical patterns in SI." *International Journal of Translation* 22:1-2, pp. 165-203.
- SHLESINGER, Miriam. (2009) "Towards a definition of Interpretese: An intermodal, corpus-based study." In: Hansen, Gyde; Andrew Chesterman & Heidrun Gerzymisch-Arbogast (eds.) *Efforts and Models in Interpreting and Translation Research. A tribute to Daniel Gile*. Amsterdam & New York: John Benjamins, pp. 237-253.
- SHLESINGER, Miriam; Moshe Koppel; Noam Ordan & Brenda Malkiel. (2009) "Markers of translator gender: do they really matter?" *Copenhagen Studies in Language* 38, pp. 138-198. Electronic version: http://u.cs.biu.ac.il/~koppel/papers/Shlesinger%20et%20al_183-198.pdf
- TANNEN, Deborah. (1990) *You just don't understand. Women and Men in conversation*. New York: William Morrow.
- TANNEN, Deborah. (1994) *Gender and Discourse*. New York & Oxford: Oxford University Press.

BIONOTE / NOTA BIOGRAFICA

MARIACHIARA RUSSO graduated in Conference interpreting from the Advanced School of Modern Languages for Interpreters and Translators (SSLMIT) of the University of Trieste in 1987 and has been a freelance conference interpreter ever since. In 1993 she became Associate Professor at the SSLMIT of Trieste, where she taught simultaneous and consecutive interpreting from Spanish into Italian. In 2001 she moved to the SSLMIT (now Department of Interpretation and Translation, DIT) of the University of Bologna at Forlì, where she also teaches Interpreting Theory. In 2005-2012 she directed the Post Graduate Degree Program in Conference Interpreting. She coordinated

the EPIC (European Parliament Interpreting Corpus) project, which is an on-line resource freely available at <http://sslmitdev-online.sslmit.unibo.it/corpora/corpora.php>. She is on the Scientific Board of *Puentes* (Univ. Granada), *Trans* (Univ. Malaga) and *Translation and Translanguaging in Multilingual Contexts* (John Benjamins), and is a referee for many T&I Journals. Main research fields: aptitude testing for simultaneous interpreting; corpus-based interpreting studies; public service interpreting; effects of morpho-syntactic asymmetries from Spanish into Italian; simultaneous interpretation of films. In 2014 she received the national habilitation for Full Professorship.

MARIACHIARA RUSSO si è laureata in Interpretazione di Conferenza presso la Scuola Superiore di Lingue Moderne per Interpreti e Traduttori (SSLMIT) dell'Università di Trieste nel 1987, e da allora svolge la professione di interprete di conferenza. Nel 1993 è diventata Professore Associato alla SSLMIT di Trieste, dove ha insegnato interpretazione simultanea e consecutiva dallo spagnolo all'italiano. Nel 2001 si è trasferita alla SSLMIT (ora Dipartimento di Interpretazione e Traduzione, DIT) dell'Università di Bologna, dove insegna inoltre Teoria dell'Interpretazione. Dal 2005 al 2012 ha diretto il corso di Laurea Magistrale in Interpretazione di Conferenza. Ha coordinato il progetto EPIC (*European Parliament Interpreting Corpus*), risorsa online disponibile alla pagina <http://sslmitdev-online.sslmit.unibo.it/corpora/corpora.php>. Fa parte del Comitato Scientifico di *Puentes* (Univ. Granada), *Trans* (Univ. Malaga), *Translation and Translanguaging in Multilingual Contexts* (John Benjamins) ed è *referee* per molte riviste di Traduzione e Interpretazione. Principali interessi di ricerca: attitudine all'interpretazione simultanea; *corpus-based interpreting studies*; interpretazione per i servizi pubblici; effetti delle dissimmetrie morfosintattiche dallo spagnolo all'italiano; interpretazione simultanea di film. Nel 2014 ha ricevuto l'abilitazione nazionale per la prima fascia.