

ANGLINTRAD: Towards a purpose specific interpreting corpus

By Michela Bertozzi (Università di Bologna, Italy)

Abstract & Keywords

English:

Corpus-based interpreting methods are effective for analyzing important phenomena that has been neglected in research (Shlesinger 1998), but little attention has been paid to their possible exploitation in interpreter education and to the benefits of corpus-derived insights for trainee interpreters (Bendazzoli 2010a). The aim of this paper is to describe how *Anglintrad*, a purpose-specific intermodal Italian-Spanish corpus, is being built and to suggest some preliminary exploitation criteria for interpreter (and translator) training and practice. This paper focuses on the presence of unmodified English loanwords in Italian political speeches (Marzocchi 2007) and their renditions in simultaneous interpreting and written translation into Spanish. The possibility of comparing the same phenomena (unmodified English loanwords) from two different perspectives (interpreting and translation) represents an unprecedented opportunity entailing possible didactic applications to enhance interpreter and translator training and practice.

Keywords: interpreting, anglicisms, corpora, loanwords, italian, spanish

1. Introduction

Over the last decades, corpus-based and corpus-driven interpreting studies (CIS) have significantly developed from the ground-breaking research on the first manual corpora of courtroom interpreting by Shlesinger (1989), the mid-late nineties with an increasing number of studies on interpreting corpora by Pöchhacker (1994), Kalina (1998), Setton (1997, 1999) and the years between the late nineties and the beginning of the third millennium, characterized by Shlesinger's plea (1998) to make research efforts into the compilation and use of electronic, machine-readable interpreting corpora:

From the standpoint of interpreting research, the compilation of bilingual and parallel corpora is indeed overdue, given the potential to use large, machine-readable corpora to arrive at global inferences about the interpreted text in relation to other forms of oral discourse; and in relation to other forms of translation. (Shlesinger 1998: 2)

This paved the way for a new approach in interpreting research, where the methodology of Corpus Linguistics was applied to the creation and consultation of the first machine-readable interpreting corpora (Cencini and Aston 2002, Wallmach 2002, Bendazzoli *et al.* 2004, Timarova 2005, Shlesinger 2008). Over the last few years, researchers have been channeling their efforts towards open-access electronic corpora (House, Meyer and Schmidt 2012, Monti *et al.* 2005, Bendazzoli and Sandrelli 2005-2007, Sandrelli *et al.* 2010).

However, so far little attention has been paid to the possible exploitation of these corpora for interpreter training and the benefits of corpus-derived insights for didactic purposes. Providing interpreting (and translation) trainees with a user-friendly platform, for instance collecting data on unmodified English loanwords, would be beneficial from a didactic point of view for several reasons: first, raising awareness on the issue of unmodified English loanwords in Italian and how this phenomenon can be managed in interpreting and translation; second, providing students with a set of different strategies applied by professionals in a high-quality, homogenous and comparable setting can bring added value to interpreting and translation training sessions, which would entail the possibility to compare the trainee's renditions (or translations) with the professional ones. As a matter of fact, the speeches, interpretations and translations at European Parliament have already been used for teaching purposes as a source of didactic material and this corpus could be an extra tool to be used both by teachers and students in their training sessions; finally, a platform comparing the same phenomenon from the interpreter and translator's point of view could be exploited to make trainees expand their own perspective on the array of possible strategies that can (or cannot) be used both in interpreting and translation.

1.1 Objectives

The aim of the present study is to present the methodology and contents of *Anglintrad[1]*, a purpose-specific intermodal (interpreting and translation) Italian-Spanish corpus, and to highlight some preliminary didactic implications for future interpreters and translators.

The idea of the *Anglintrad* corpus came from the practical need to shed light on a particularly challenging phenomenon in Italian-Spanish simultaneous interpreting, that is the frequent use of unmodified English loanwords[2] in Italian political speeches (Marzocchi 2007) and the different Spanish mechanisms of loanword integration (Tonin 2010); these phenomena have been widely studied in translation, but little attention has been paid to understanding how they can affect the interpreter's performance. Therefore, *Anglintrad* was specifically designed with a view to selecting a number of oral texts delivered within the same setting (the European Parliament plenary sittings) sharing a common characteristic (the presence of unmodified English loanwords in the original Italian speeches), then comparing them with the corresponding Spanish interpreted speeches and official translations. The fact that the corpus is intermodal (including both interpreted and translated texts) may lead to future comparative studies, as already suggested by Shlesinger's paper on the comparison between written and oral corpora:

Ideally, the notion of comparable corpora in interpreting studies should be extended to cover setting up three separate collections of texts in the same language: interpreted texts, original oral discourses delivered in similar settings, and written translations of such texts. This would allow for the identification of patterns specific to interpreted texts (regardless of their source language) as pieces of oral discourse, in relation to comparable texts in the same language. It would also allow us to identify the patterns which single out interpreted texts as distinct oral translational products in a given language irrespective of their source languages, through comparisons with comparable written translational products. (Shlesinger 1998: 4)

In the light of the above, the ultimate goal of the *Anglintrad* project is a bilingual intermodal corpus to observe a particular phenomenon (the presence of unmodified English loanwords in Italian original speeches delivered in the European Parliament plenary sitting), the way it is managed by simultaneous interpreters into Spanish and by translators into the same target language not only within the same setting (the plenary sitting itself) but within the same original text that is studied from two different perspectives.

1.2 Corpus structure

The *Anglintrad* corpus is divided into two main sub-corpora: oral (1) and written (2) texts (see Figure 1). The former includes original Italian speeches delivered at the European Parliament plenary sitting in the year 2011 (1A) with the related interpreted Spanish versions (1B); the latter is made up of the official revised Spanish translations referred to the same original speeches (2A).

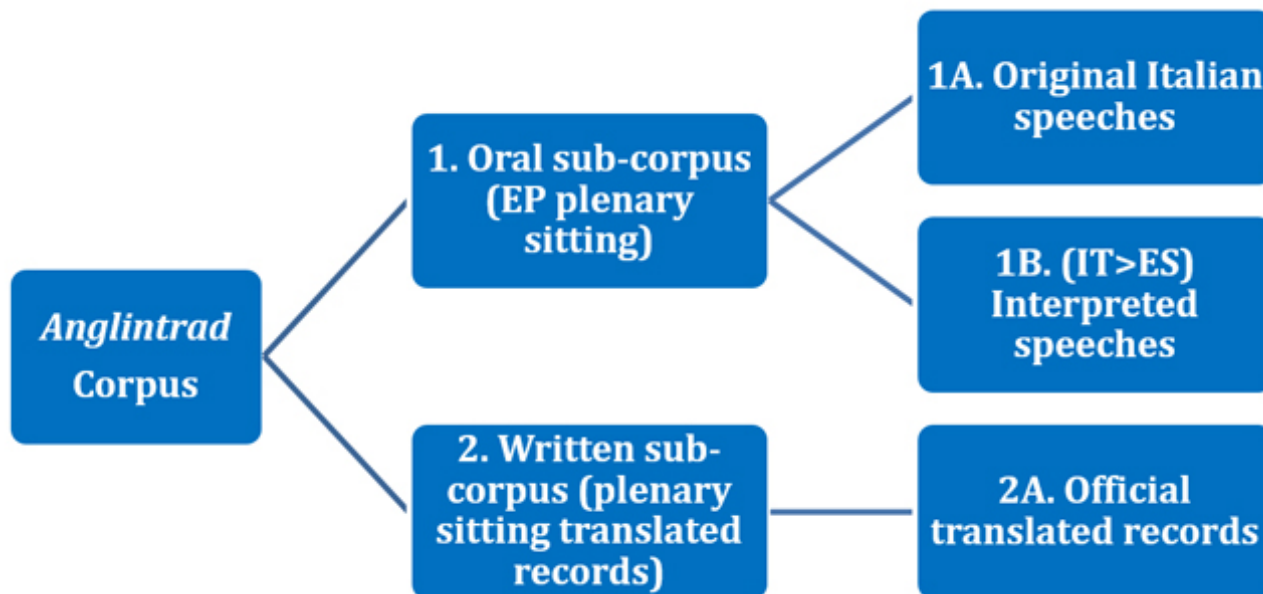


Figure 1. *Anglintrad* structure

2. The corpus

2.1 Methodology

Following the principles underlying the compilation of EPIC[3] and its transcription conventions, the *Anglintrad* corpus was designed to serve a specific purpose: providing a significant amount of data to observe a particular phenomenon, without challenging some basic methodological assumptions which were described by Bendazzoli:

La creazione di EPIC rappresenta uno dei primi tentativi di superare gran parte degli ostacoli descritti [...], in quanto ci si è avvalsi di materiale autentico, omogeneo rispetto a numerose variabili e in quantità sufficienti per essere rappresentativo; lo stesso materiale è stato poi elaborato e reso disponibile in formato elettronico, in modo da poter farne uso a fini di ricerca e didattica attraverso procedure semi-assistite e pertinenti con la linguistica computazionale. [...] La scelta del materiale da includere nello studio è stata guidata da molteplici fattori, quali le fonti disponibili, gli strumenti tecnici più idonei alla raccolta, conservazione ed elaborazione del materiale oggetto di studio e le risorse tecnologiche disponibili al momento dell'attivazione del progetto [...]. (Bendazzoli 2010: 117)

In the light of the need for data accessibility and above all comparability, the European Parliament plenary sitting was selected as the source of all the materials included in *Anglintrad*. This guarantees not only the authenticity of the original material, one of the main methodological challenges in Corpus-based Interpreting Studies (Shlesinger 1998), but also its homogeneity, since oral data coming from different contexts and settings may compromise the basic principles of the study. The selected texts were all delivered in 2011 in 26 plenary sittings where a total number of 241 items (unmodified English loanwords in the original Italian speeches) were identified.

The unrevised verbatim reports of the original Italian speeches were first scanned as in fast reading in order to detect those containing at least one phenomenon to study; then, the selected texts were analysed and transcribed, following the EPIC transcription conventions[4], and the same procedures were applied to the related Spanish interpreted versions. In the last phase, these text segments were aligned to their official revised translations to allow for an immediate comparison between the three texts (original speech – interpreted version – translated version). A summary of the main methodological steps for the compilation of *Anglintrad* is provided in Fig. 2:

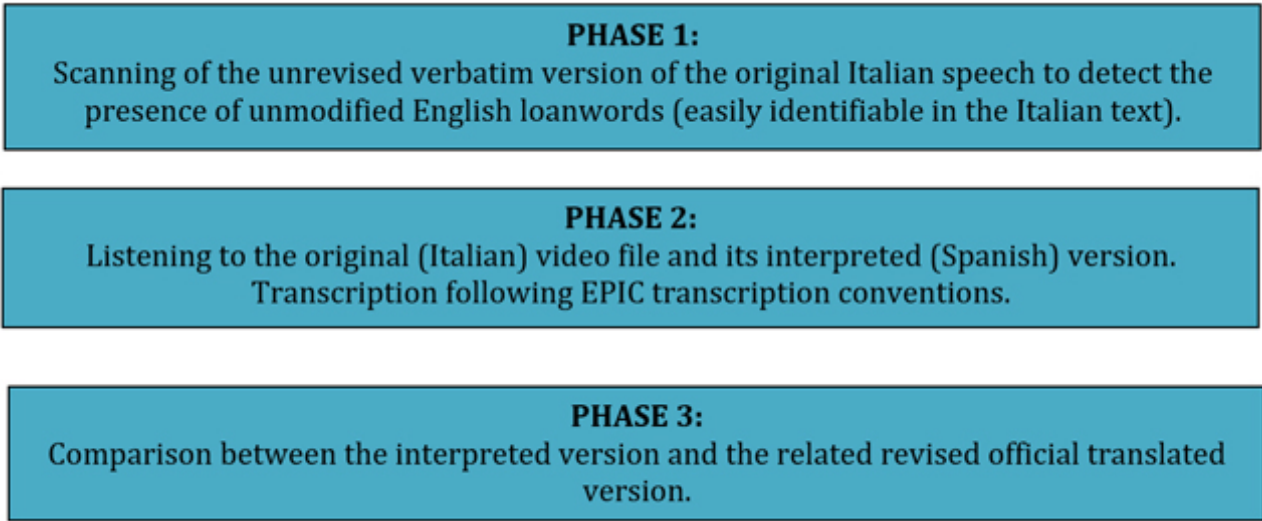


Figure 2. Main methodological steps

2.2 Anglintrad characteristics

The comparison between the three main text segments (original Italian speech – interpreted Spanish version – translated Spanish version) was meant to be as immediate and user-friendly as possible, therefore their structure was organized in a spreadsheet.

Every phenomenon identified in the corpus was classified and matched with a set of metadata on the plenary sitting (for example, “*Dibattito 17_01_11*”), a link to the official translated version (for example, “*Resoconto tradotto*”), a link to the related verbatim unrevised version of the original Italian speech (for example, “*Resoconto*”), the specific topic (for example, “*Dichiarazioni del Parlamento Europeo sulla situazione in Tunisia*”) and the speaker (for example, “*Pier Antonio Panzeri*”).

For a quicker comparison between the three versions of the same phenomenon, the structure was divided into three columns: the first one indicates the transcription of the text segment where the phenomenon was identified; the second one includes the transcription of the same text segment in the interpreted version, while the third column includes the official translated version. This layout allows for an immediate comparison between the phenomena, highlighting possible problems and strategies adopted in simultaneous interpreting/translation. This visualization can be easily exploited for didactic purposes in interpreter and translator training.

Another important element to be considered when using these materials for pedagogical objectives is the composition of the corpus itself. As already mentioned, *Anglintrad* includes 241 unmodified English loanwords detected in the speeches delivered by 46 different speakers (32 men and 14 women) during 26 plenary sittings held in 2011; the total number of occurrences delivered by men is 184 and by women is 57. When a loanword was found, a few words before and after the item were transcribed and included in the corpus in order to preserve the meaning of the sentence. Since it is a purpose-specific corpus and given the main research objective, a full transcription of the whole speeches in which a loanword is present is not provided because the research focus is meant to remain within the analysis of this particular phenomenon and the way it is managed by interpreters and translators.

Further information and metadata on the structure of the corpus itself (distribution of phenomena by topic, type of entry, and type of pronunciation in the original Italian speech) is provided in Figures 3-6 below[5]:

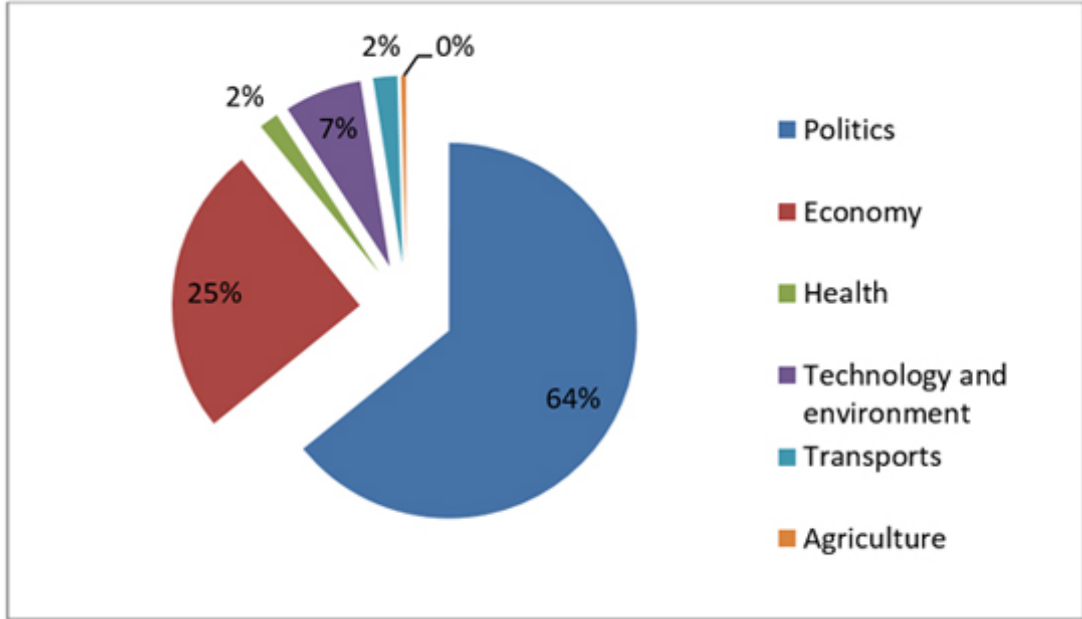


Figure 3. Percentage of loanwords by topic

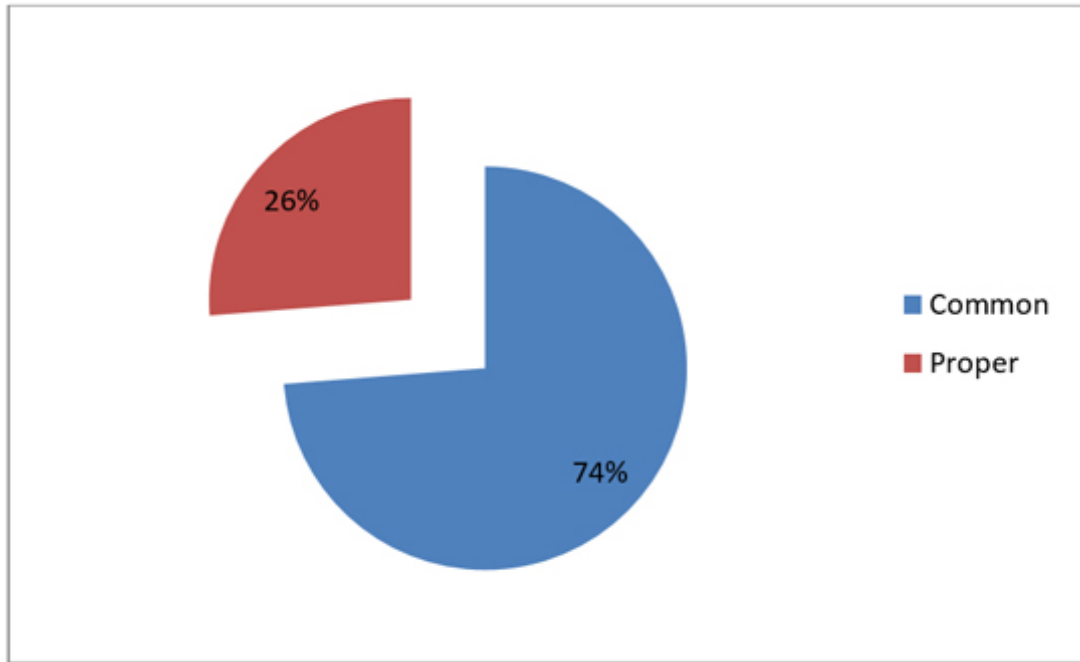


Figure 4. Shares of common and proper items in loanwords

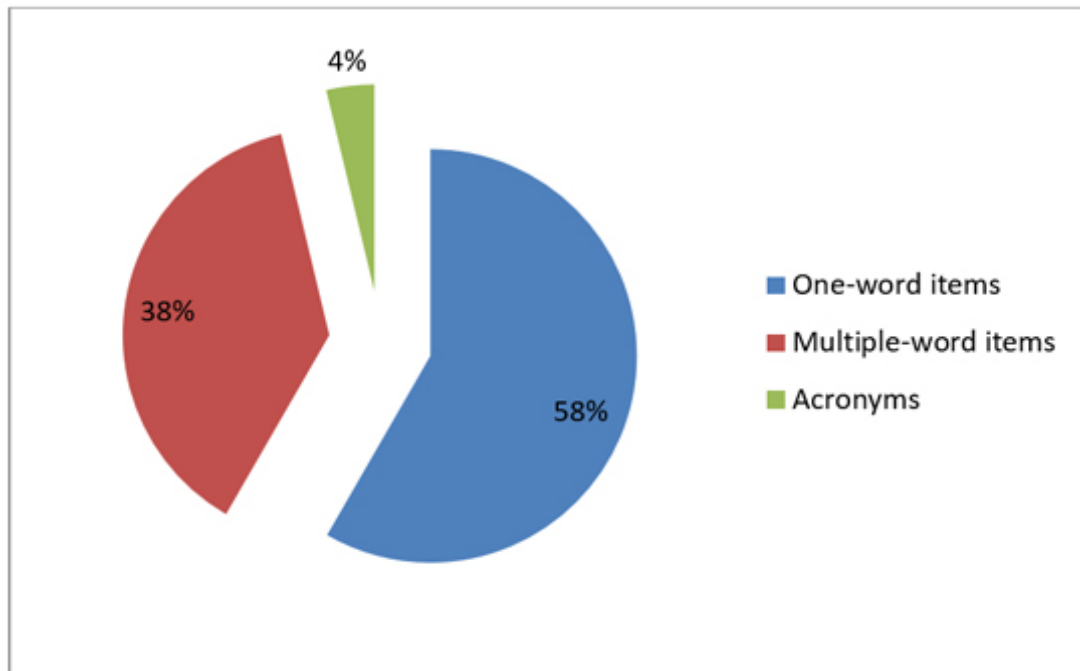


Figure 5. Shares of item types in loanwords

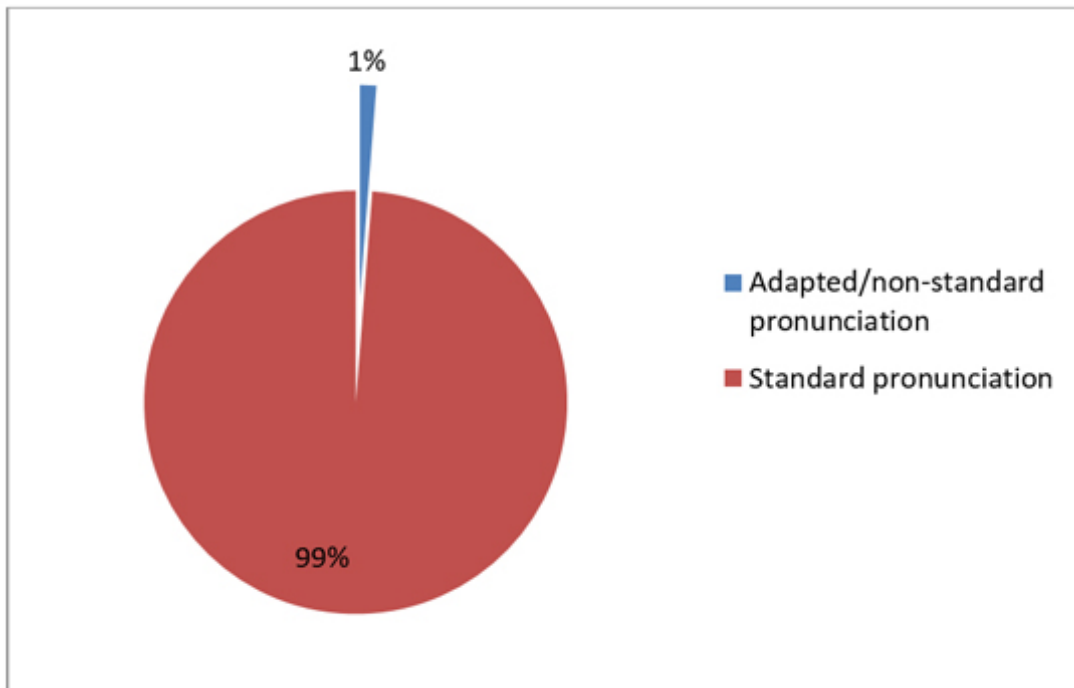


Figure 6. Share of pronunciation type[6] in original Italian speeches

Some weighted percentages[7] were also calculated for the number of speakers by gender and the related weighted percentage of phenomena divided by speaker's gender (see Figures 7-8); the number of speakers by political group (S&D – Social and Democrats, EPP – European People's Party, EFD – Europe of Freedom and Democracy, ALDE – Alliance of Liberals and Democrats for Europe) and the related weighted percentage of phenomena by political group (see Figures 9-10):

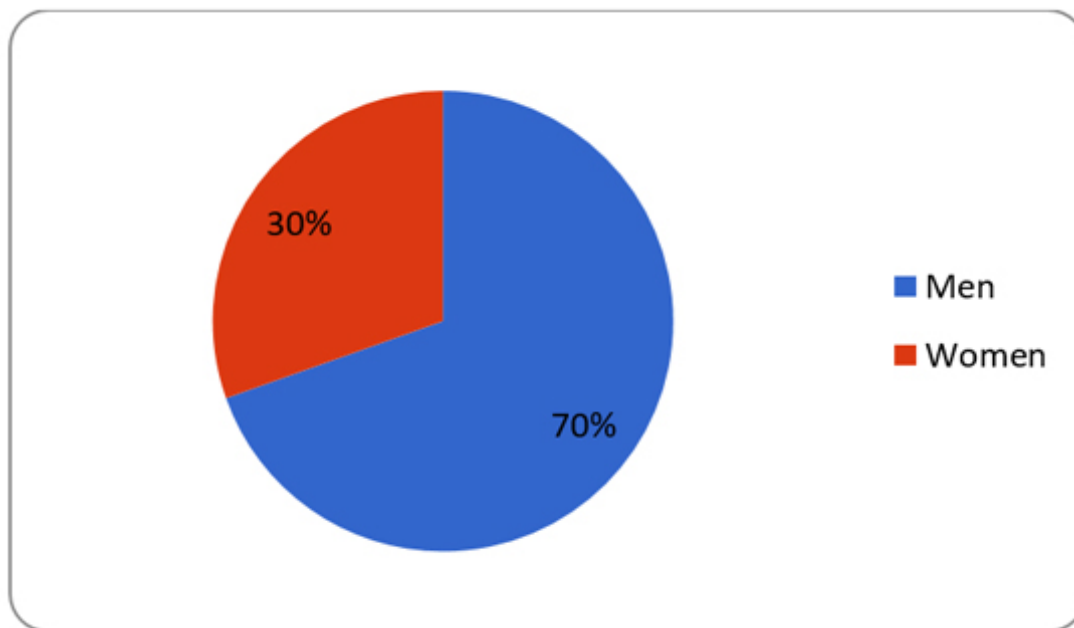


Figure 7. Gender distribution of speakers

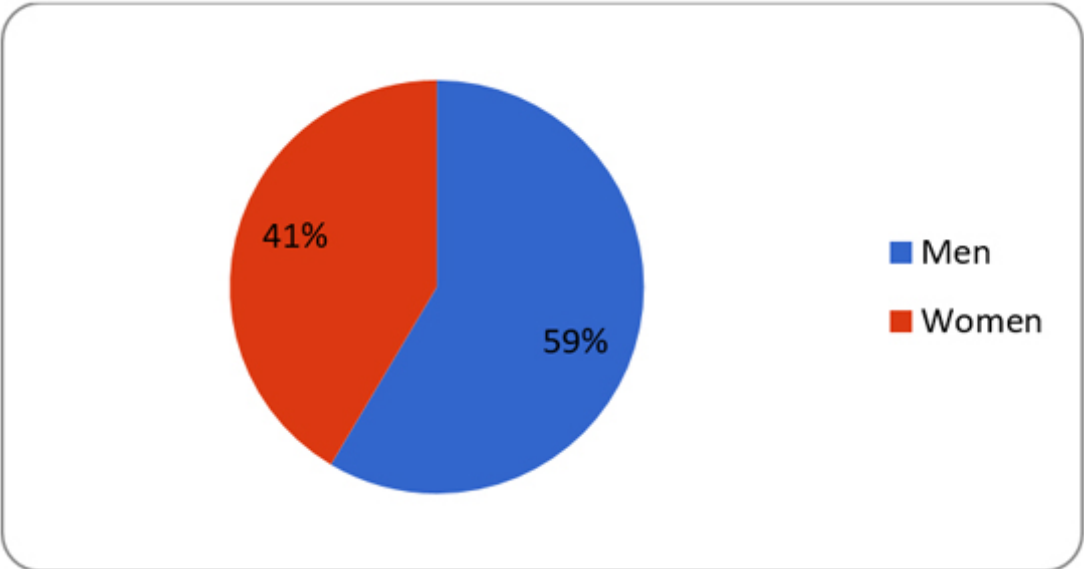


Figure 8. Weighted percentage of loanwords by speaker gender

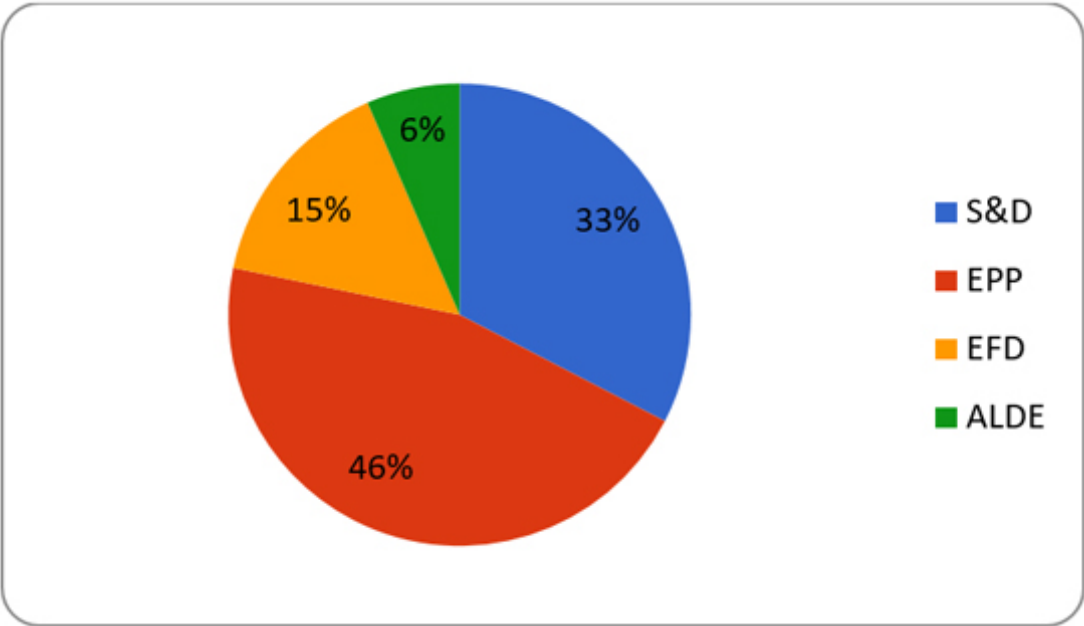


Figure 9. Distribution of speakers over political groups

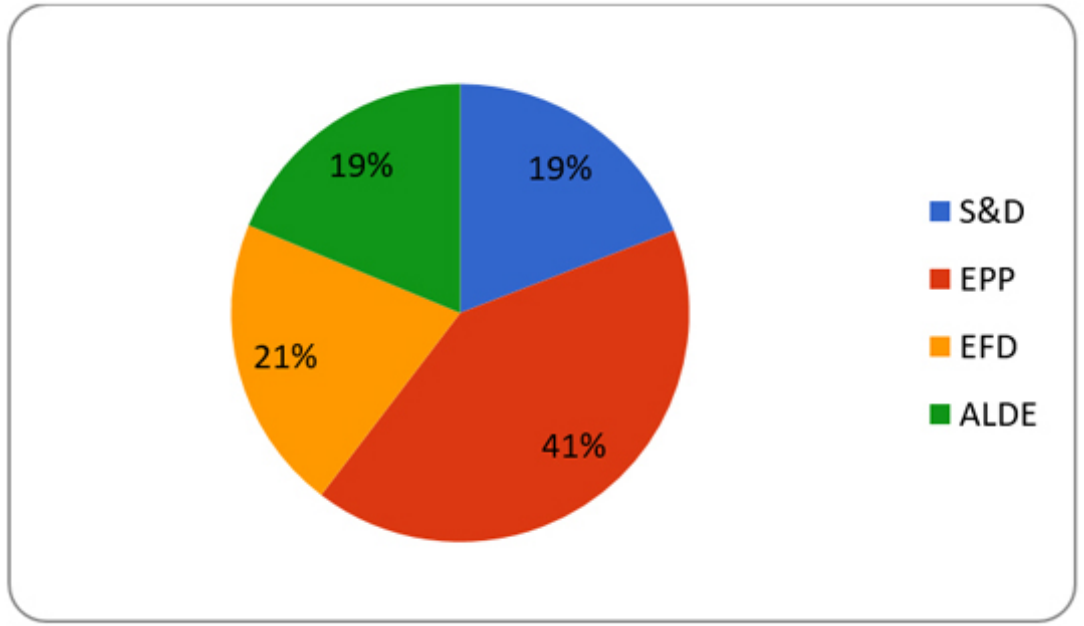


Figure 10. Weighted percentage of loanwords by political group

3. Preliminary results

3.1 Strategies: similarities and differences

The corpus allows for a double-perspective observation of the same phenomenon (unmodified English loanwords in the original Italian speech) because it is intermodal and directly comparable, since the same oral text is aligned to the related interpreted and translated versions. The fact that *Anglintrad* is a purpose-specific corpus (with a particular aim: observing what strategies are applied by simultaneous interpreters and translators to tackle the same potentially challenging lexical item such as unmodified English loanwords in the Italian>Spanish language combination) meets the need to create a user-friendly tool that can be easily exploited for didactic purposes.

A first analysis of the corpus based on the simple observation of unmodified English loanwords highlighted some preliminary results that, despite being far from thorough and complete, can provide an interesting initial overview of the similarities and differences in the strategies used by interpreters and translators dealing with the same lexical item, in the same text and within the same context.

The first step of this empirical observation was an attempt to classify the different strategies detected both in the interpreted and in the translated versions (see Table 1).

List of Strategies	Definition	Italian	Spanish	References
1 Cancellation	The phenomenon is not rendered.	/ [...] registro delle lobby che dovrebbe essere formato da qui a breve/	XXX	Bakti 2009
2 Exact rendition	The phenomenon is rendered with no modifications.	/ [...] la commissione si è sempre schierata a favore del Made In/	/ [...] a favor del Made In/	Wadensjö 2001, Schjoldager 1996
3 Generalization	The communicative intention or the basic concept is rendered in a generic way.	/ [...] l'azzeramento dei dazi sui prodotti coreani contro l'innalzamento degli standard ambientali e sociali in Corea/	/ [...] tendrán demasiadas ventajas, muchas más que los productos europeos/	Al-Khanji <i>et al.</i> 2000, Bartłomiejczyk 2006
4 Substitution	The phenomenon is reformulated at a lexical level (use of synonyms) or at a syntactic level.	/ [...] i meccanismi di pubblicità il web e altre modalità efficaci/ /l'Europa deve essere in grado di intervenire con misure comuni...ed efficaci per la sicurezza... dell'approvvigionamento alimentare per evitare le forti asimmetrie a-ancora esistenti relative agli standard di-di sicurezza tra i prodotti UE ed extra UE grazie/	/ [...] todos los medios que existen como...la internet/ /Europa debe intervenir con medidas comunes y eficaces...para que haya un surtido alimentario adecuado evitando...las fuertes asimetrías aún existentes ehm... relativas a las normas de seguridad...entre productos europeos y no europeos/	Wadensjö 2001, Straniero Sergio <i>et al.</i> 2012
5 Translation	The entry is adapted to the morphological and lexical norms of the target language or the lexicalised equivalent in the target language is used.	/ [...] il rimborso del prezzo del biglietto in caso di partenza annullata ritardo superiore alle due ore o overbooking/	/ [...] que se le reembolse el billete en caso de que se le cancele la salida o un retraso de más de dos horas o cuando haya ehm sobreventa/	
6 Expansion	At different levels, the interpreter/translator can make additions to the source text.	/ [...] stress test [...]/	/ [...] pruebas de... aguante / resistencia [...]/	Bartłomiejczyk 2006

Table 1. List of strategies

The first strategy detected in the corpus is **cancellation**, indicating the lack of any type of rendition of the original phenomenon in the target text: at first sight, it may seem that cancellation necessarily entails a partial or complete loss of the original content, but it can also be a strategy activated by interpreters or translators to make the original message clearer, provide better cohesion to the target text or eliminate redundancy if present in the source text (Russo and Rucci 1997). In the case of interpreted texts, this holds even more true since cancellations can:

[...] help guarantee the best possible quality of interpretation under the circumstances. [...] In some cases, omissions are deliberate and aimed at economy of expression, ease of listening for the audience and maximum communication between the speaker and audience. (Jones 1998: 139)

The second strategy is **exact rendition**, meaning that the loanword is simply transposed into the target language without any type of modification. In the light of the different mechanisms that these cognate languages (Italian and Spanish) have to integrate new unmodified loanwords (Tonin 2010), this strategy may be regarded as

potentially complex to be applied correctly. However, the typical features of this setting (the European Parliament plenary sitting) and its microlanguage (Bertozzi 2016) allow for a purpose-specific use of both oral and written language, since all participants share the same knowledge and background: that is why a “non-domestication” strategy may be perfectly acceptable in this setting:

L'unico soggetto che potrebbe discostarsi dal gruppo è lo stesso interprete, in quanto difficilmente avrebbe la possibilità di condividere lo stesso livello di esperienza e preparazione degli altri partecipanti, pur preparandosi adeguatamente all'incarico assegnato. In questo caso l'interprete prediligerebbe il più possibile un uso tecnico e specifico della lingua; eventuali lacune sarebbero generalmente compensate dalla conoscenza degli ascoltatori. (Bendazzoli 2010: 151)

The third strategy detected in the corpus is **generalization**, where the communicative intention or the basic concept of the source text is rendered in a generic way in the target text. This category also includes:

[...] l'utilizzo di acronimi e il ricorso alla deissi, utilizzata in sostituzione di porzioni di testo più lunghe, grazie alle conoscenze che l'interprete condivide con oratore e pubblico. (Voncina 2009: 28)

This technique has always been widely used and studied both in Interpreting and Translation Studies; more specifically, in the case of interpreting, Gile (1995) included generalization among the so-called “preventive and reformulation tactics” consisting of ‘replacing a segment with a superordinate term or a more general speech segment’ (Bartłomiejczyk 2006: 152).

Substitution is the fourth type of strategy identified in the corpus, meaning that the phenomenon is reformulated at a lexical level (in other words, with the use of synonyms) or at a syntactic level. This macro-category includes a set of sub-strategies such as morpho-syntactic transformation, chunking (Seleskovitch and Lederer 1989), permutation or the re-arrangement of elements within the same sentence (Pippa and Russo 2002) and paraphrasing. Restricting the scope to interpreting, this strategy can be particularly demanding in terms of cognitive load and lexical retrieval capacity since, in some cases, this type of rendition is far longer and more complex than the original message, with a subsequent lengthening in the interpreter’s *décalage* and possible carry-over effects in the following segments. That is why it must not come as a surprise that ‘experts did more than twice as much lexical elaboration than novices’ (Setton and Motta 2008: 217).

The fifth strategy is **translation**, where the phenomenon is adapted to the morphological and lexical norms of the target language or where the lexicalised equivalent in the target language is used. If, on the one hand, the Italian language has always tended to integrate unmodified loanwords (possibly modifying only their phonetic level), on the other hand the Spanish language has a more restrictive approach and tends to use target terms more frequently (Tonin 2010). Many examples such as “*budget - presupuesto*”, “*road map – hoja de ruta*” or “*bond – bono*” can be found in the corpus.

The sixth and last strategy is **expansion**, where the interpreter/translator makes additions to the source text at different levels. In interpreting, this phenomenon has also been called “addition”:

Addition is treated as a strategy when the interpreter decides to add, by way of explanation, something the original speaker did not say because the interpreter thinks the interpretation may otherwise not be clear for the audience (e.g. due to discrepancies between the source- and target-language cultures).(Bartłomiejczyk 2006: 160)

This holds true also for translation, where expansion can be used for discourse-planning purposes, or to provide better cohesion to the target text.

After direct comparison between the three versions of the same phenomenon and the identification of a set of strategies (which, far from being exhaustive, can however provide a necessary attempt to classify the different strategies adopted for didactic purposes), the next step entailed the subdivision of the strategies adopted by interpreters and translators into two main macro-categories: same and different strategies, where the first ones include marked similarities (at a lexical, pragmatic level, and so on) between the interpreted and the translated renditions, while the second ones indicate how the interpreter and the translator facing the same linguistic phenomenon can adopt different strategies (cancellation, exact rendition, generalization, substitution, translation, expansion) (see Figures 11-12). This type of classification proved to be the most suitable for didactic purposes, where the need to simplify this structure as much as possible and therefore the need to provide a user-friendly tool for interpreting and translation trainees is crucial:

COMPARISON BETWEEN INTERPRETED AND TRANSLATED RENDITIONS

<p>Italian original speech</p>	<p>/l'Europa deve essere in grado di intervenire con misure comuni...ed efficaci per la sicurezza...dell'approvvigionamento alimentare per evitare le forti asimmetrie a-ancora esistenti relative agli standard di sicurezza tra i prodotti UE ed extra UE grazie/ /</p>
<p>Interpreted Spanish text</p>	<p>/Europa debe intervenir con medidas comunes y eficaces...para que haya un surtido alimentario adecuado evitando...las fuertes asimetrías aún existentes ehm... relativas a las normas de seguridad...entre productos europeos y no europeos/[4 Substitution]</p>
<p>Translated Spanish text</p>	<p>Europa tiene que ser capaz de intervenir con medidas comunes y eficaces para garantizar el suministro de alimentos y evitar las graves desigualdades que todavía existen en relación con las normas de seguridad entre los productos comunitarios y nocomunitarios. [4 Substitution]</p>

Figure 11. Example of identical strategies

The example in Figure 11 shows that the interpreter reformulated the segment containing the loanword (*standard di sicurezza*) at a lexical level (*normas de seguridad*), which can be classified as a substitution (see fig. 12); the use of this strategy in simultaneous interpreting is particularly frequent and reformulation is often associated with the activity of interpreting itself:

L'abitudine alla riformulazione, a una maggiore flessibilità lessicale può trasformarsi in una strategia automatizzata che consente di distribuire al meglio le proprie risorse per prevenire una resa insoddisfacente imputabile a una cattiva suddivisione delle stesse. (Riccardi 1999: 172)

More specifically, with regard to the example above (Figure 11), one can hypothesize that the interpreter tried to retrieve the same word in Spanish and the latter may have not been immediately available in his/her memory (Gran 1992), as one could assume given the presence of a filled pause (*ehm...*) just before this segment (Ahrens 2002); therefore, due to time constraints, the interpreter may have tried to find a possible strategy to render this potentially challenging phenomenon (an unmodified loanword from a third language that is not included in the pair being activated in simultaneous mode) by reformulating the source segment. Interestingly, despite the many obvious differences characterizing translation and interpreting activities, the same strategy (substitution) was activated by the translator as well (*normas de seguridad*). This may suggest that what may seem to be an “emergency strategy” in interpreting (reformulation as a consequence of difficulties in retrieving the right word/segment) can actually be a specifically-targeted strategy as such in translation: as a matter of fact, the segment “*normas de seguridad*” is particularly frequent in Eurlex[8], so this may prove that the translator was provided with specific terminological guidelines in advance (which may also apply to interpreters, but the simultaneous mode does not always allow for an immediate retrieval of single specific terms, even if provided in advance).

An example of different strategies activated by interpreters and translators in the corpus is provided in figure 12:

Italian original speech	/a tale proposito ho chiesto ai miei servizi/ che ringrazio per il contributo che danno sempre... all'attivià legislativa della Commissione/ di prepara- di preparare una road map sulla...implementazione che ho intenzione di inviarmi...nonappena... sarà possibile/
Interpreted Spanish text	/por ello he pedido... a mis servicios a los que doy las gracias por la contribución que siempre... s-ofrecen a la actividad legislativa de la Comisión/ preparar una hoja de ruta para la aplicación...con la intención de enviarla ehm lo más rápidamente posible/[5 Translation]
Translated Spanish text	A tal fin he pedido a mis servicios —a los que agradezco la aportación que siempre hacen a la labor legislativa de la Comisión— que elaboren un plan de trabajo para su aplicación, que enviaré a Sus Señorías lo antes posible.[4 Substitution]

Figure 12. Example of different strategies

In this case, the original speaker is making use of an unmodified loanword (*road map*) that is becoming more and more common in the Italian language, especially in the press and in the political domain. Given the importance of an in-depth analysis for each type of loanword, its main characteristics and use in modern Italian, every phenomenon identified in the Italian sub-corpus was provided with a specific terminological sheet (an example is provided in Table 2) indicating its grammatical category, gender and number, the related original word in English, its definition taken from main modern Italian dictionaries, the use of the linguistic phenomenon in context (from the Lexis Nexis Database[9]), the year of first appearance in dictionaries (where reported), its further productivity in Italian (if any), any indications on pronunciation and some information on the history of the loanword in Italian (whether it is a neologism, it is reported as “anglicism” in the dictionaries, it is present in previous editions of the same dictionary or it is part of a sectoral language):

Lessema	ROAD MAP
Categoria grammaticale	lessema ingl. (propr. «carta stradale»), usato in ital. come sost. femm.
Genere	femm.
Numero	invar. (Gabrielli); Treccani ammette il plur. road maps <... mäps>.
Derivazione inglese (Oed)	noun; 1A map, especially one designed for motorists, showing the roads of a country or area. 2A plan or strategy intended to achieve a particular goal: "a road map for peace in the region".
Fonti lessicografiche /terminologiche italiane	VOCABOLARIO TRECCANI: 1. Spec. nel linguaggio giornalistico, piano diplomatico e strategico accuratamente programmato, e da realizzarsi in diverse tappe, in vista del raggiungimento di uno specifico obiettivo, spec. con riferimento al conflitto tra israeliani e palestinesi. 2. estens. Tabella di marcia, programma di lavoro e sim.: attenersi scrupolosamente alla road map fissata. DIZIONARIO GABRIELLI: Piano, progetto dettagliato, scandito a tappe come una tabella di marcia, in vista di un obiettivo da perseguire.
Contesti	Il piano di pace del Quartetto Usa-Ue-Onu-Russia, la cosiddetta 'road map', e' stato ufficialmente presentato questo pomeriggio al nuovo premier palestinese Mahmud

Abbas a Ramallah (Ansa 2003 - Database Lexis Nexis). Toccherà al Quartetto, cioè ai quattro mediatori internazionali (Stati Uniti, Russia, Unione europea e Nazioni Unite) che hanno redatto la *road map*, valutare i progressi nell'attuazione del piano (La Stampa 2003 - Database Lexis Nexis).

Il nuovo capo dell'Anp Abu Mazen ha detto oggi che i palestinesi sono pronti a attuare gli impegni assunti nella *Road Map*, il percorso di pace delineato due anni fa dal Quartetto Usa-Onu-Ue-Russia (Ansa 2005 - Database Lexis Nexis).

Se la divisione destra/sinistra ha ancora un senso, e si rimprovera alla *road map* di Monti di aver seguito politiche sbilanciate nell'una o nell'altra direzione, in una coalizione destra-sinistra rimproveri del genere non sono evitabili e segnalano che la *road map* funzionerebbe meglio se avesse alle sue spalle una maggioranza politicamente coerente (Corriere della Sera 2012 - Database Lexis Nexis).

In quella sede, sono emerse, nei tavoli di lavoro, le varie proposte della *Road Map* in ausilio ed in funzione della legge di stabilità 2016 (Italia Oggi 2015 - Database Lexis Nexis).

Anno	2003 (Treccani).
Produttività del lessema/ulteriori apporti dall'inglese	La locuzione nasce in un contesto ben specifico, quello del conflitto israelo-palestinese (piano diplomatico e strategico accuratamente programmato, e da realizzarsi in diverse tappe, in vista del raggiungimento di uno specifico obiettivo, spec. con riferimento al conflitto tra israeliani e palestinesi) e si estende in seguito alla seconda accezione (Treccani), ad oggi molto frequente: tabella di marcia, programma di lavoro.
Indicazione di pronuncia	<róud mäp>, road maps <... mäps> (Treccani). Non indicata in Gabrielli.
Riferimenti	http://www.treccani.it/vocabolario/road-map/ (19/02/16) http://dizionari.repubblica.it/Italiano/R/roadmap.php (19/02/16) http://www.oxforddictionaries.com/definition/english/road-map?q=road+map (19/02/16) https://it.wiktionary.org/wiki/road_map (19/02/16).
Note	prestito linguistico dall'inglese <i>road map</i> , entrato nel linguaggio italiano inizialmente tramite il gergo giornalistico per riferirsi al processo di pace israelo-palestinese (Treccani, Wikizionario).
Carattere neologico	1) PRESENZA NEI DIZIONARI DI LINGUA GENERALE: sì (Treccani 2003, Gabrielli). Non indicato da De Mauro né Sabatini Coletti. Il Dizionario De Agostini 1995 e lo Zingarelli 1970 non lo riportano. 2) SEGNALATO COME ANGLICISMO: sì, da Treccani. Non segnalato da Gabrielli. 3) PRESENZA INDICAZIONE DI PRONUNCIA: solo Treccani la riporta. 4) LINGUAGGIO SETTORIALE/LINGUA GENERALE: il lessema scaturisce dal linguaggio giornalistico, con particolare riferimento al conflitto israelo-palestinese e solo successivamente si estende al linguaggio generale nella sua accezione più ampia di tabella di marcia, programma di lavoro (Treccani).

Table 2. Example of terminological sheet

In the specific case illustrated in Figure 12, the use of the loanword “*road map*”, which entered the Italian vocabulary through the journalistic language with reference to the Israeli-Palestinian conflict, is becoming more and more frequent, thus potentially affecting the way interpreters may deal with this phenomenon: as a matter of fact, interpreters tend to rely on automatic mechanisms to render the most frequent linguistic features (such as loanwords). This may be one of the reasons why the interpreter did not hesitate in using a translation strategy in this case (“*hoja de ruta*” is the semantic equivalent of “*road map*”); the translator did not opt for the same strategy, relying on a substitution (“*plan de trabajo*”), which may seem to be more frequent in interpreting (due to time constraints and the difficulties in retrieving the exact word, thus potentially leading to a reformulation). In this case, the translator’s aim was making the target text more recipient-oriented and clearer from a linguistic point of view.

Another example of same strategies used by interpreters and translators in the corpus is provided in Figure 13:

Italian original speech	/alcuni diritti positivi vengono tutelati come quelli dei portatori di handicap /
Interpreted Spanish text	/hay derechos positivos bien plasmados y tutelados y amparados sobre todo entonces los discapacitados o personas con movilidad reducida / [6 Expansion]
Translated Spanish text	Algunos derechos están protegidos, como los de personas con discapacidad o con movilidad reducida . [6 Expansion]

Figure 13. Example of same strategies

The loanword *handicap* (and the related expression *portatori di handicap*) has a long tradition in the Italian vocabulary (the first occurrences in the main Italian dictionaries date back to the late 19th century); the same applies for the Spanish language, but with a difference: the entry in the *Diccionario de la Real Academia Española*[10] is *hándicap* (with acute accent) and, in the *Diccionario Clave*[11], *handicap* is in italics, since it is classified as a foreign word. The *Diccionario Panhispánico de Dudas*[12] suggests the use of *discapacitado* or *minusválido* instead of the unnecessary anglicism *handicapado*. The choice made by the interpreter and the translator in this case (Figure 13) is particularly interesting because they both rely on an expansion, a strategy requiring additional efforts, especially in the simultaneous mode (Bartłomiejczyk 2006). It appears quite clearly that both interpreters and translators are particularly sensitive to the “politically correctness” issues inherent in language and one could assume that, within the European institutions, they are provided with guidelines on how to render/translate these potentially challenging phenomena: this could be one of the reasons why they both opted for an expansion of the original text, even if there was no need to further clarify the source message and despite the fact that the previous segment might have been particularly difficult to render in simultaneous mode.

4. Conclusions

In this paper the methodological steps undertaken to create a bilingual (Italian > Spanish) intermodal (simultaneous interpreting and written translation) corpus for pedagogical purposes have been presented. The *Anglintrad* corpus is being built to explore the strategies used by interpreters and translators when dealing with unmodified English loanwords in the Italian source text. An easy-to-use classification of interpreting/translation strategies along with convenient parallel display of both source and target texts have been designed and can be exploited in interpreter and translation training.

A thorough analysis of the whole corpus was still beyond the scope of the present work. However, a preliminary data observation highlighted that in some cases translation and simultaneous interpreting are much closer than could be expected in terms of strategies applied to face the same problem within the same context and setting (see Figures 11 and 13). This may be due to the fact that both interpreters and translators within the European Parliament share a similar background, a demanding specialized training and the use of standardized terminology for certain terms is highly recommended by the DG Translation and Interpreting.

In other cases observed so far (see Figure 12), the strategies adopted by interpreters and translators can vary considerably due to a number of factors that are not only linked to the different time constraints but also to the different purposes and recipients of the interpreted and the translated renditions.

This dual approach in the observation of the same linguistic phenomenon provides a valuable opportunity entailing resourceful teaching applications for interpreter and translator training and practice. More specifically, in addition to the corpus under construction, the *Anglintrad* project includes the creation of a platform containing useful material for didactic purposes; this platform is currently being developed and will soon make the following material available to interpreting and translation trainees and teachers: first, the bilingual intermodal corpus (Italian original text – Spanish interpreted rendition – Spanish translated version) as described in Section 2, containing additional information on the speaker (name and surname, political group, sex), the type of source text (topic, delivery speed, type of delivery – impromptu or read) and the type of phenomenon detected in the source text (one-word or multiple-word anglicism, proper or common item); second, a terminological sheet for each phenomenon in the Italian sub-corpus (see Table 2), containing an in-depth analysis of the loanword and its history/use in the Italian language (frequency of use, definitions, contexts, specific domains, and so on); finally, a user-friendly classification of the strategies adopted by interpreters and translators (see Table 1), allowing for direct comparison between the two (same/different strategies).

This twofold (interpreting vs translation) perspective has already been hypothesized by some scholars. It is in particular worth mentioning the article by Viezzi (1993) in this context, in which written translation and simultaneous interpreting are contrasted in a case study and the work by Padilla Benítez et al. (1999), who apply the principles of cognitive theory to the two disciplines. However, the same approach has never been used to study a specific linguistic feature: that is the reason why the creation of an open-access platform for the analysis and comparison of the strategies adopted by interpreters and translators, as well as the main challenges involved, can be particularly beneficial for didactic purposes and can provide new insights based on different perspectives and strategies, bearing in mind that each discipline can learn something from the other.

A preliminary analysis of the corpus suggests that the same strategies are used more often than one might expect: this can serve as a starting point for a new approach in translation and interpreting training, providing a platform that collects a number of genuine examples from a real setting and some useful tools (such as the terminological sheets) to raise awareness among trainees on the issue of unmodified English loanwords in Italian, bearing in mind that a good target text (regardless of the translation or interpreting mode) is intrinsically linked to a deep knowledge of the most important emerging trends and features of the source language.

Finally, the same purpose-specific approach can also be applied to the study of other particularly challenging linguistic items and other language combinations as well, paving the way for future research projects and applications.

Appendix

For each chart included in the paper, the related raw frequencies are reported in the frequency bar charts below:

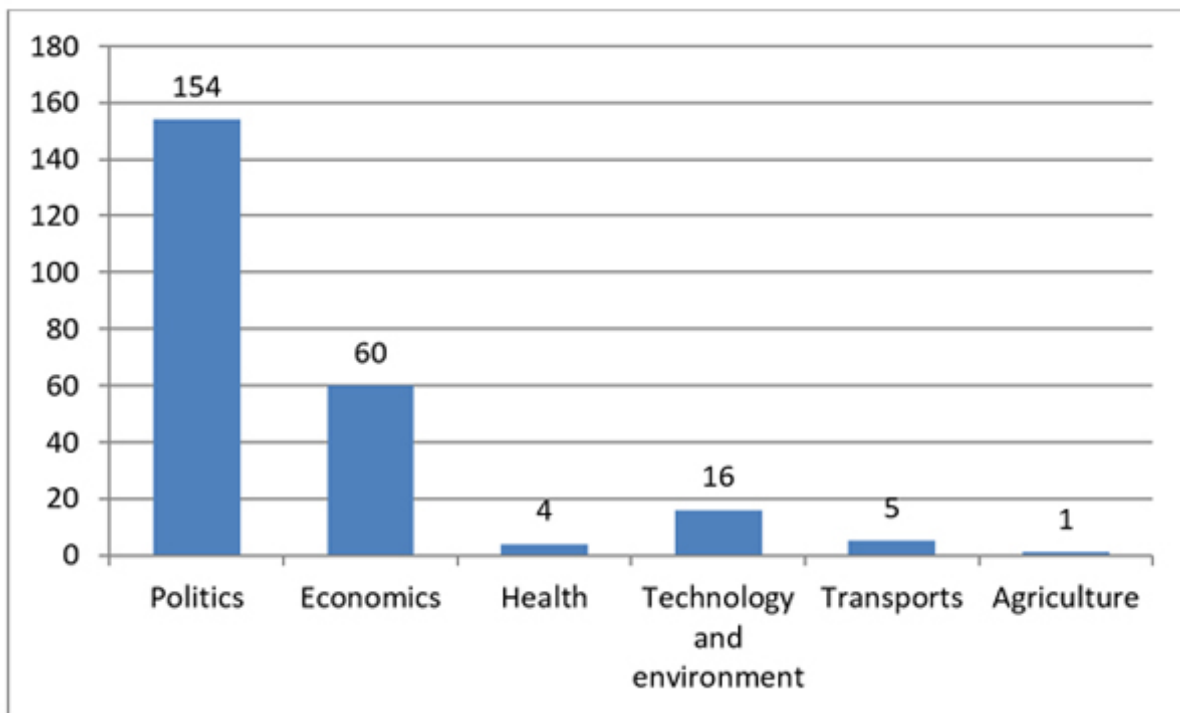


Figure A1. Frequencies of loanwords per topic

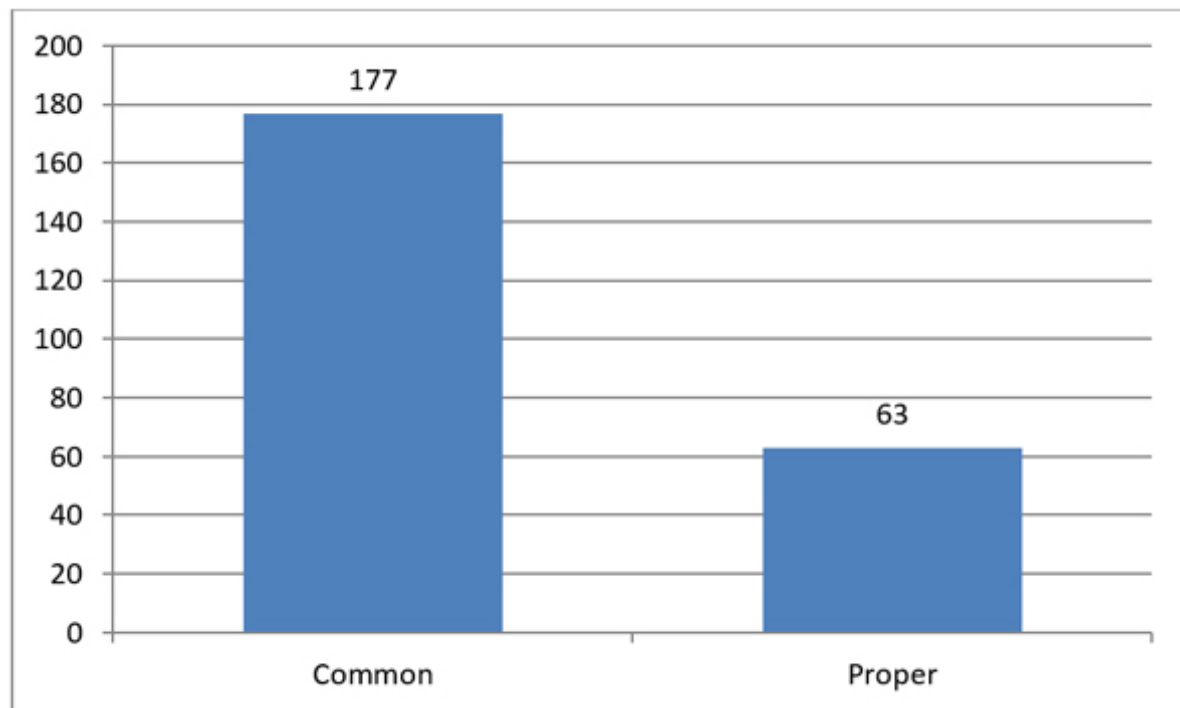


Figure A2. Frequencies of loanwords: common and proper items

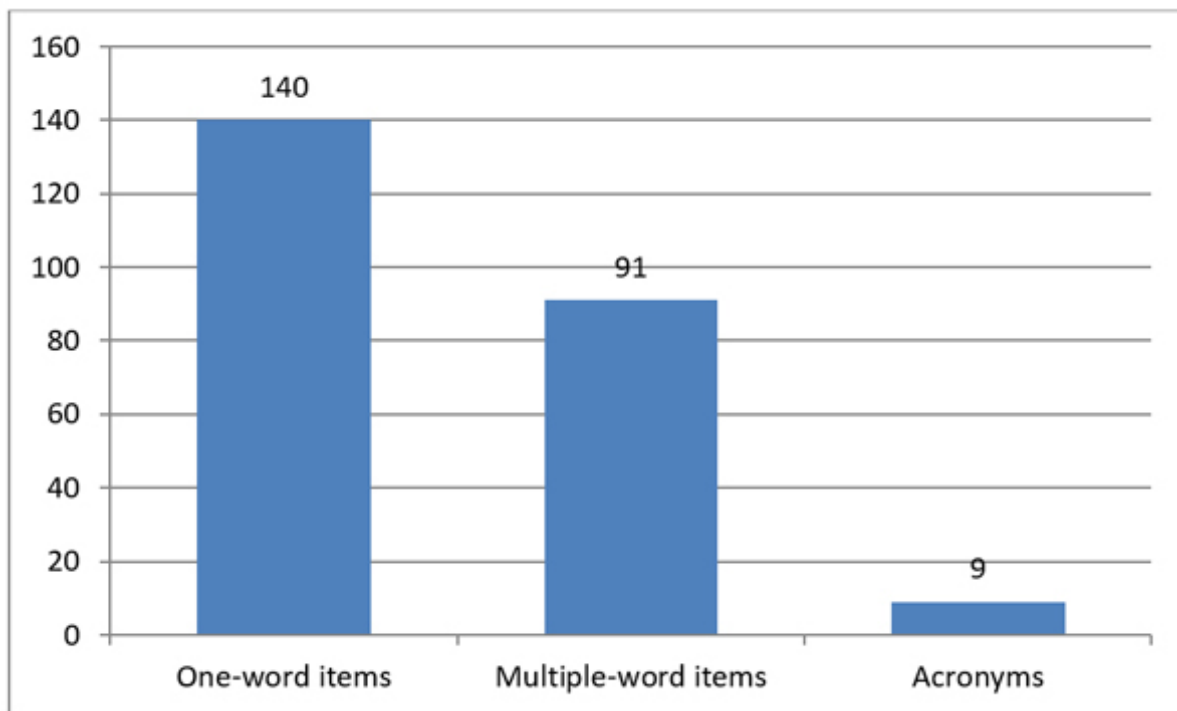


Figure A3. Frequencies of loanwords per type of item

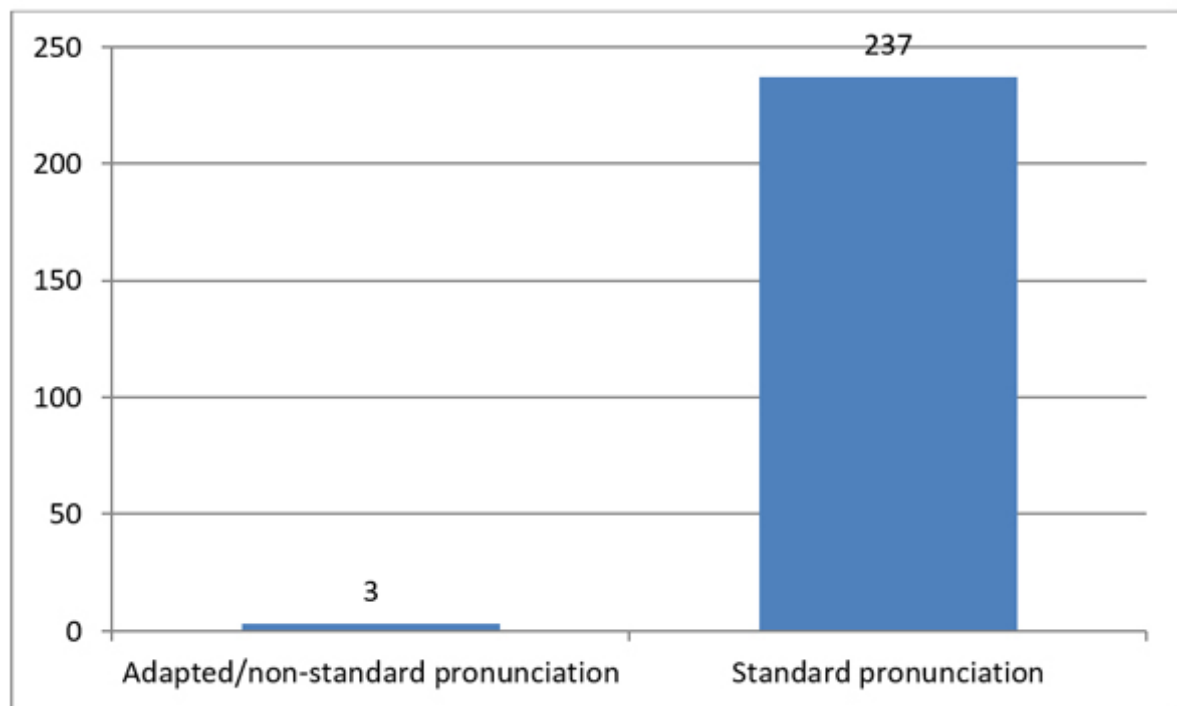


Figure A4. Type of loanword pronunciation in the original Italian speeches

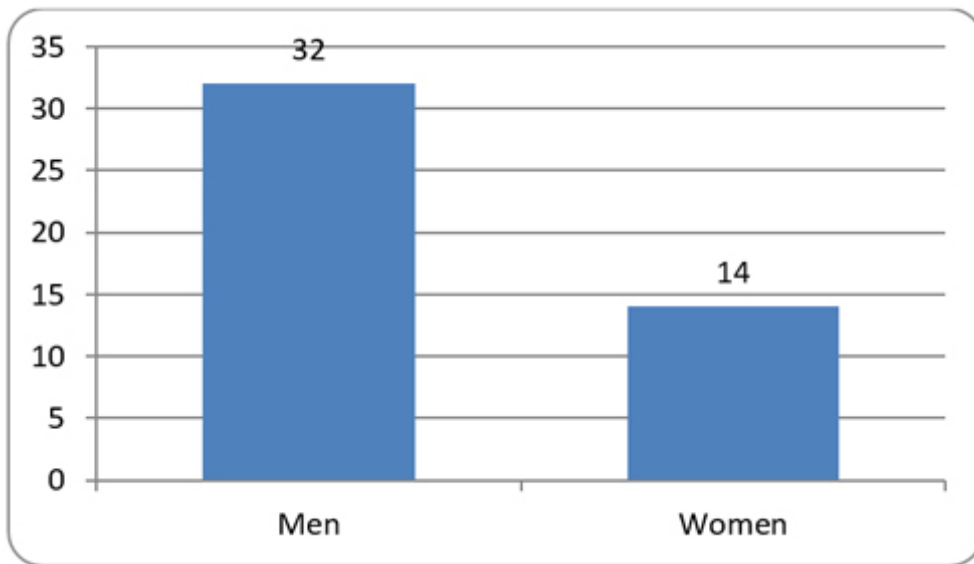


Figure A5. Number of speakers per gender

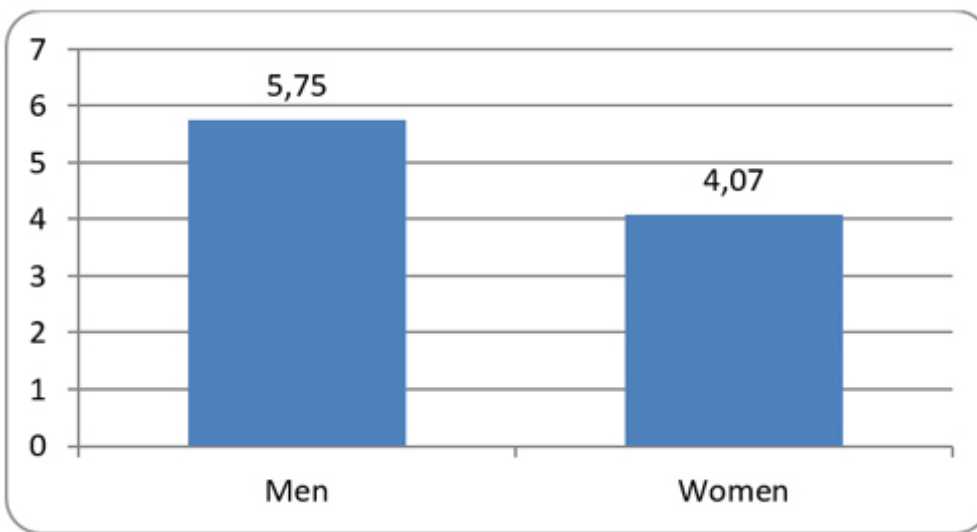


Figure A6. Weighted percentage of loanwords per speaker gender

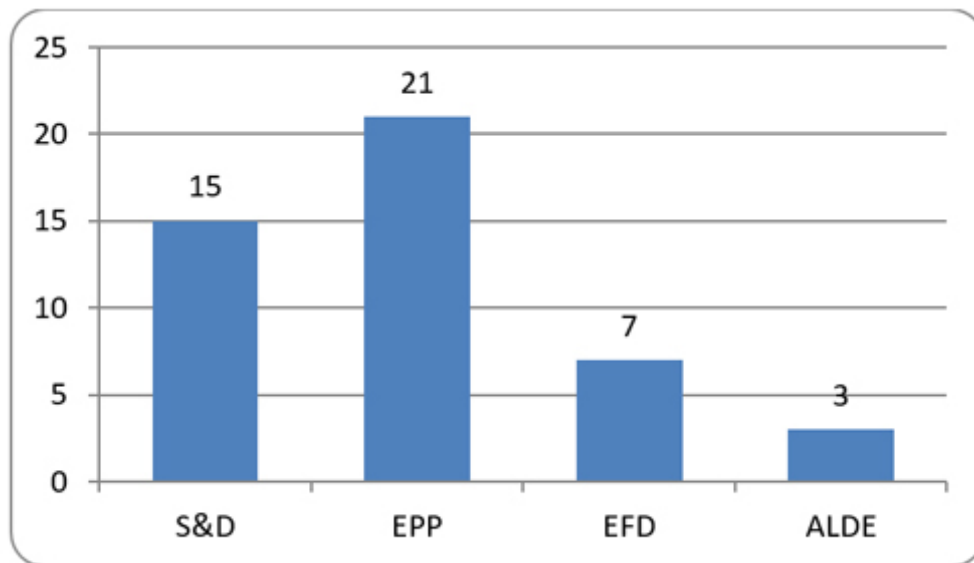


Figure A7. Number of speakers per political group

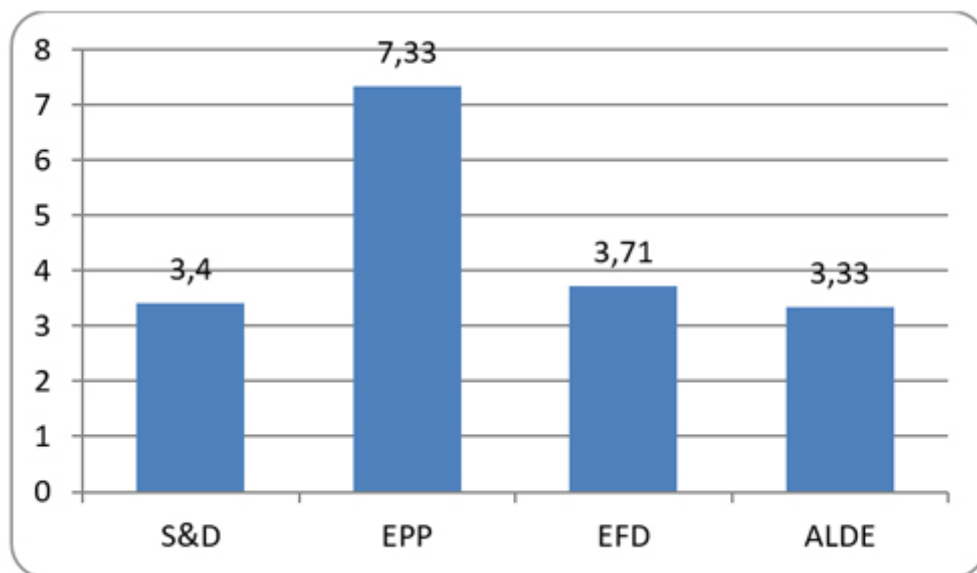


Figure A8. Weighted percentage of loanwords per political group

References

- Ahrens, Barbara (2002) "The Correlation between Verbal and Nonverbal Elements in SI" in *Perspectives on Interpreting*, Garzone, Giuliana, Peter Mead and Maurizio Viezzi (eds), Bologna, CLUEB: 37–46.
- Al-Khanji, Raja, Said El-Shiyab and Riyadh Hussein (2000) "On the Use of Compensatory Strategies in Simultaneous Interpretation", *Meta* 45, no. 3: 548–57.
- Baker, Mona (1993) "Corpus Linguistics and Translation Studies: Implications and Applications" in *Text and Technology: In honour of John Sinclair*, Mona Baker, Gill Francis and Elena Tognini Bonelli (eds), Amsterdam, John Benjamins: 233–50.
- Bakti, Maria (2009) "Speech Disfluencies in Simultaneous Interpretation" in *Selected Papers of the CETRA Research Seminar in Translation Studies 2008*, Dries de Crom (ed.), <http://www.arts.kuleuven.be/cetra/papers> (accessed 20 September 2016).
- Bartłomiejczyk, Magdalena (2006) "Strategies of Simultaneous Interpreting and Directionality", *Interpreting* 8, no. 2: 149–74.
- Bendazzoli, Claudio (2010a) *Il corpus DIRSI: creazione e sviluppo di un corpus elettronico per lo studio della direzionalità in interpretazione simultanea*. PhD diss., University of Bologna.
- (2010b) *Corpora e interpretazione simultanea*, Bologna, Asterisco. http://amsacta.unibo.it/view/series/Alma-DL=2E_Saggi.html (accessed 20 September 2016).
- Bendazzoli, Claudio, and Annalisa Sandrelli (2005/2007) "An Approach to Corpus-based Interpreting Studies: Developing EPIC (European Parliament Interpreting Corpus)" in *Proceedings of the Marie Curie Euroconferences MuTra: Challenges of Multidimensional Translation - Saarbrücken 2-6 May 2005*, Heidrun Gerzymisch-Arbogast and Sandra Nauert (eds), http://www.euroconferences.info/proceedings/2005_Proceedings/2005_proceedings.html (accessed 20 September 2016).
- Bendazzoli, Claudio, Cristina Monti, Annalisa Sandrelli, Mariachiara Russo, Marco Baroni, Silvia Bernardini, Gabriele Mack, Elio Ballardini, and Peter Mead (2004) "Towards the Creation of an Electronic Corpus to Study Directionality in Simultaneous Interpreting" in *Compiling and Processing Spoken Language Corpora. LREC 2004 Satellite Workshop IV International Conference on Language Resources and Evaluation*, Nelleke Oostdijk, Gjiert Kristoffersen, and Geoffrey Sampson (eds), Paris, ELRA: 33–9.
- Bertozzi, Michela (2016) "Distinctive Features of Orality in a Microlanguage: The Italian Language in the Plenary Sessions of the European Parliament. Some Preliminary Observations", *MonTI Special Issue*, no. 3: 339–66.
- Bombi, Raffaella (2005) *La linguistica del contatto. Tipologie di anglicismi nell'italiano contemporaneo e riflessi metalinguistici* [Contact linguistics. Typology of anglicisms in current Italian and some metalinguistic considerations], Rome, Il Calamo.
- Cencini, Marco, and Guy Aston (2002) "Resurrecting the Corp(us/se): Towards an Encoding Standard for Interpreting Data" in *Interpreting in the 21st century: Challenges and Opportunities*, Giuliana Garzone and Maurizio Viezzi (eds), Amsterdam, John Benjamins: 47–62.
- Furiassi, Cristiano (2010) *False Anglicisms in Italian*, Monza, Polimetrica.
- Gile, Daniel (1995) *Basic Concepts and Models for Interpreter and Translator Training*, Amsterdam, John Benjamins.
- Gran, Laura (1992) *Aspetti dell'organizzazione cerebrale del linguaggio: dal monolinguisimo all'interpretazione simultanea*, Udine, Campanotto.
- Gusmani, Roberto (1981) *Saggi sull'interferenza linguistica*, Firenze, Le Lettere.
- House, Juliane, Bernd Meyer, and Thomas Schmidt (2012) "CoSi – A Corpus of Consecutive and Simultaneous Interpreting" in *Multilingual Corpora and Multilingual Corpus Analysis*, Thomas Schmidt, and Kai Worner (eds), Amsterdam, John Benjamins: 295–304.

- Iglesias Fernández, Emilia (2007) *La didáctica de la interpretación de conferencias. Teoría y práctica*, Granada, Comares.
- Jones, Roderik (1998) *Conference Interpreting Explained*, Manchester, St. Jerome.
- Kalina, Sylvia (1998) *Strategische Prozesse beim Dolmetschen: Theoretische Grundlagen, empirische Fallstudien, didaktische Konsequenzen* [Strategic processes in interpreting: Theoretical principles, empirical field studies and their implications for teaching], Tübingen, Gunther Narr.
- Marzocchi, Carlo (2007) "Translation — Transcript — Interpretation. Notes on the European Parliament Verbatim Report of Proceedings", *Across Languages and Cultures* 8, no. 2: 249–54.
- Monti, Cristina, Claudio Bendazzoli, Annalisa Sandrelli, and Mariachiara Russo (2005) "Studying Directionality in Simultaneous Interpreting through an Electronic Corpus: EPIC (European Parliament Interpreting Corpus)", *Meta* 50, no. 4. <http://id.erudit.org/iderudit/019850ar> (accessed 20 September 2016).
- Padilla Benitez, Presentación, Maria Teresa Bajo and Francisca Padilla Adamuz (1999) "Proposal for a Cognitive Theory of Translation and Interpreting. A Methodology for Future Empirical Research", *The Interpreters' Newsletter* 9: 61–78.
- Petite, Christelle (2005) "Evidence of Repair Mechanisms in Simultaneous Interpreting: A Corpus-based Analysis", *Interpreting* 7, no. 1: 27–49.
- Pippa, Salvador and Russo Maria Chiara (2002) "Aptitude for Conference Interpreting: A Proposal for a Testing Methodology Based on Paraphrase", in *Interpreting in the 21st Century. Challenges and Opportunities*, Garzone Giuliana and Maurizio Viezzi (eds), Amsterdam, John Benjamins: 245–56.
- Pöschhacker, Franz (1994) "Quality Assurance in Simultaneous Interpreting", in *Teaching Translation and Interpreting 2: Insights, Aims and Visions. Papers from the Second Language International Conference, Elsinore, Denmark 4-6 June 1993*, Cay Dollerup and Annette Lindegaard (eds), Amsterdam, John Benjamins: 232–42.
- Riccardi, Alessandra (1999) "Interpretazione simultanea: strategie generali e specifiche", in *Interpretazione simultanea e consecutiva: problemi teorici e metodologie didattiche*, Falbo, Caterina, Maria Chiara Russo and Francesco Straniero Sergio (eds), Milan: Hoepli: 161–74.
- Russo, Maria Chiara, and Rucci Marco (1997) "Verso una classificazione degli errori nella simultanea dallo spagnolo in italiano" [Towards a classification of errors in Spanish-Italian simultaneous interpreting], in *Nuovi orientamenti negli studi sull'interpretazione*, Gran Laura and Alessandra Riccardi (eds), Padova, Università degli Studi di Trieste: 179–99.
- Russo, Maria Chiara, Claudio Bendazzoli, Annalisa Sandrelli and Nicoletta Spinolo (2012) "The European Parliament Interpreting Corpus (EPIC): Implementation and Developments" in *Breaking Ground in Corpus-Based Interpreting Studies*, Francesco Straniero Sergio and Caterina Falbo (eds), Bern, Peter Lang: 35–90.
- Sandrelli, Annalisa, Claudio Bendazzoli, and Mariachiara Russo (2010) "European Parliament Interpreting Corpus (EPIC): Methodological Issues and Preliminary Results on Lexical Patterns in Simultaneous Interpreting", *International Journal of Translation* 22, no. 1–2: 165–203.
- Schjoldager, Anne (1996) *Simultaneous Interpreting: Empirical Investigations into Target-Text/Source Text Relations*, Aarhus, Aarhus School of Business.
- Seleskovitch, Danica, and Lederer Marianne (1989) *A Systematic Approach to Teaching Interpretation*, Paris, Didier.
- Setton, Robin (1997) "A Relevance-theoretic Account of Simultaneous Interpretation", *Tsuuyaku kenkyuu - Interpreting Research* 13, no. 2: 33–6.
- (1999) *Simultaneous Interpretation: A Cognitive-pragmatic Analysis*, Amsterdam, John Benjamins.
- Setton, Robin and Motta Manuela (2008) "Syntacrobatics: Quality and Reformulation in Simultaneous-with-Text", *Interpreting* 9, no. 2: 199–230.
- Shlesinger, Miriam (1989) "Monitoring the Courtroom Interpreter", *Cahiers de l'Ecole de Traduction et d'Interpretation* 11: 29–36.
- (1998) "Corpus-based Interpreting Studies as an Offshoot of Corpus-Based Translation Studies", *Meta* 43, no. 4: 486–93.
- (2008) "Towards a Definition of Interpretese. An Intermodal, Corpus-based Study", in *Efforts and Models in Interpreting and Translation Research: A Tribute to Daniel Gile*, Gyde Hansen, Andrew Chesterman, and Heidrun Gerzymisch-Arbogast (eds): 237–53.
- Spinolo, Nicoletta (2014) *La resa del linguaggio figurato in interpretazione simultanea: Una sperimentazione didattica*, PhD diss., University of Bologna.
- Straniero Sergio, Francesco, and Caterina Falbo (eds) (2012) *Breaking Ground in Corpus-based Interpreting Studies*, Bern, Peter Lang.
- Timarová, Šárka (2005) "Corpus Linguistics Methods in Interpreting Research: A Case Study", *The Interpreters' Newsletter* 13: 65–70.
- Tonin, Raffaella (2010) *El vaivén de las palabras. Los anglicismos en español y en la traducción al italiano*, Roma, Aracne.
- Viezzi, Maurizio (1993) "Written Translation and Simultaneous Interpretation Compared and Contrasted: A Case Study", *The Interpreters' Newsletter* 5: 94–100.
- Voncina, Katja (2009) *L'interpretazione simultanea al Parlamento europeo sull'esempio delle cabine tedesca, italiana e slovena*, MA diss., SSLMIT, Università degli Studi di Trieste.
- Wadensjo, Cecilia (2001) "Approaching Interpreting through Discourse Analysis", in *Getting Started in Interpreting Research. Methodological Reflections, Personal Accounts and Advice for Beginners*, Gile Daniel,

Wallmach, Kim (2002) "Using Parallel Corpora to Determine Interpreting Strategies for Languages of Limited Diffusion in South Africa", in *Proceedings of the Łódź Session of the 3rd International Maastricht-Łódź Duo Colloquium on "Translation and Meaning*, Barbara Lewandowska-Tomaszczyk, and Marcel Thelen (eds), Maastricht, Hogeschool Zuyd: 503–9.

Notes

[1] The corpus is currently being compiled by the author as part of a PhD project at the University of Bologna at Forlì, Dipartimento di Interpretazione e Traduzione (DIT). The corpus is built for a specific research purpose, though it falls within the context of the EPIC project (see footnote 4).

[2] By "unmodified English loanword" we make reference to Bombi (2005) and Furiassi's (2010) categorisation of anglicisms in Italian, where the lexical borrowing undergoes no modifications in the target language from the morphological and phonetic point of view. This kind of anglicism is often referred to as "*integrale*" since it is the most evident and the least adapted to the rules of the "importing" language.

[3] EPIC, the European Parliament Interpreting Corpus, is a trilingual (English-Spanish-Italian) machine-readable corpus developed at the University of Bologna at Forlì, under prof. Mariachiara Russo's supervision. It consists of online transcripts of original speeches delivered at the European Parliament and of the audio recordings of the related interpreted versions. The corpus is indexed, lemmatised and POS-tagged to make the retrieval of specific features easier and to make online consultation quicker (Sandrelli *et al.* 2010, Russo *et al.* 2012).

[4] For a detailed description of these transcription norms, see Bendazzoli (2010: 126).

[5] For each chart included in this paper, a frequency table is provided in the appendix.

[6] In some cases, the Italian speaker mispronounced the loanword completely, altering the British-American pronunciation (taken as a reference for "standard") and even making it difficult for the recipient to recognize the anglicism as such.

[7] Weighted percentages were calculated by balancing the number of male/female speakers in the first case (Figure 8) and the number of speakers per political group in the second case (Figure 10): this way, the frequency of phenomena is not affected by the highest number of male speakers nor by the most represented political group and a balanced average of phenomena is provided for these two categories.

[8] Eurllex (<http://eur-lex.europa.eu/homepage.html>) is an online database available in 24 languages covering many types of texts produced mostly by the institutions of the European Union, but also by Member States, EFTA, and so on.

[9] Lexis Nexis is an online database of full-text documents from over 17,000 authoritative sources of information (mainly newspapers and press releases) in multiple languages from the early nineties to date (<https://www.lexisnexis.com/en-us/products/lexisnexis-academic.page> accessed 24/02/17).

[10] Real Academia Española. (2014). *Diccionario de la lengua española* (23.º ed.). <http://www.rae.es/rae.html> (accessed 22/02/17).

[11] Clave. (2014). *Diccionario de uso del español actual*. <http://www.smdiccionarios.com/home.php> (accessed 22/02/17).

[12] Real Academia Española y Asociación de Academias de la Lengua Española (2005). *Diccionario panhispánico de dudas*. <http://www.rae.es/obras-academicas/diccionarios/diccionario-panhispanico-de-dudas> (accessed 22/02/17).

©inTRAlinea & Michela Bertozzi (2018).

"ANGLINTRAD: Towards a purpose specific interpreting corpus", *inTRAlinea* Special Issue: New Findings in Corpus-based Interpreting Studies.

Stable URL: <http://www.intraline.org/specials/article/2317>