A STUDY OF OMISSIONS IN INTERPRETED ENGLISH-ROMANIAN EUROPEAN PARLIAMENT MEETINGS

Ph.D. coordinator
Prof. univ. dr. habil. Anca Greere

Ph.D. candidate
Matei Idu

Cluj-Napoca, 2024

Contents

Cuvinte-cheie

conference interpreting, omissions, European Parliament, English-Romanian, taxonomy, corpus, quality, ethics, strategies, errors

Rezumatul în limba engleză al tezei de doctorat

Introduction and context

Certain phenomena of the contemporary world, such as globalization and migration, imply an increased social mobility of people at a global level. The world is experiencing an unrivaled level of transnational interaction and social change brought about by global institutions and a growingly universal culture, according to several researchers. This implies, in turn, a greater need for professionals who are able to facilitate cross-linguistic and cross-cultural understanding. In this context, conference interpreting is enjoying increasing recognition and institutionalisation nowadays.

This thesis studies the product of interpretation services in the European Parliament (the Committee on Foreign Affairs/AFET Committee) in the English-Romanian language combination, focusing on the issue of omissions. We investigate the output of interpreting activities in the context of simultaneous interpreting services at the European Parliament, relying on a vast corpus and a viable methodological framework, and we aim to scrutinize this output by primarily looking at omissions (types, numbers, frequencies, causes) in order to ascertain what and how much is omitted in interpreting. The thesis touches upon adjacent sub-fields and tackles issues pertaining to interpreting ethics, pedagogy and training, quality, etc. These are not only topics that are arduously debated within the field, but also represent essential sub-domains that any thorough investigation into interpreting must touch upon.

The state of the art in interpreting studies

We start from outlining initial links between the practices of translation and interpretation, as well as connections between the two fields of study. However, we emphasize, in our literature review, that interpreting eventually became a field in its own right. While the initial practice of interpreting (which used to be consecutive and not simultaneous until the 20th century) was limited to a few domains, nowadays, the need for interpreters spans all domains: legal, medical, business, educational, political, governmental, academic, to name just a few. We outline the consolidaton of interpreting studies while emphasizing the specific particularities that separate interpreting from written translation – such as its orality, immediacy, limited directionality, problem triggers, time lag and ephemeral character. This examination is conducted as part of the larger literature review that guides the first part of the thesis and informs it from a theoretical standpoint.

Despite its quite short-lived existence, the field of interpreting studies has already gone through a number of paradigms. It is mandatory to understand the evolution and consolidation of the field by means of these paradigms. Thus, after differentiating between different modes and types of practice in conference and community interpreting (and outlining the specificities of simultaneous interpreting in the context of the study itself, namely the European Parliament), we investigate the main research paradigms that defined the field. We start from the Paris School of Interpreting, one of the most ground-breaking and lasting theories which imposed a theory of sense, according to which interpreters deverbalize a message and relay that message without carrying out a word-for-word translation. The thesis also goes through the stages of interpreting and the mechanisms available to professionals, according to this theory. Cognitive approaches have also subsequently shaped the field, with more interdisciplinarity spurred by interest coming from neuroscience, psycholinguistics and similar related fields of study. Corpus-based interpreting studies are tackled as a potential paradigm and as a dire need for the field at the moment, with some attention being paid to discourse analysis in the context of corpora.

Subsequently, we tackle the transfer of information in interpreting, attempting to provide a comprehensive literature review of the main topics that have troubled interpreting researchers over the years, mainly how information is transferred inter-linguistically. This inevitably leads to a discussion of interpreting quality, one of the over-arching concepts in interpreting that has received plenty of scholarly attention. While it is not the aim of the study to assess interpreting quality, the illustration of the major works in this sub-field is justified as it is closely linked to the issue of omissions. Hence, we provide a thorough display of how interpreting quality has been tackled over the years, how it is understood by users and how its core criteria are perceived by scholars. We relied on an extensive literature review in order to comprehensively discuss the topic that is at the crux of the study, namely this transfer of information and omissions in interpreting. A discussion of quality is mandatory, seen as an intricate topic which fosters little scholarly consensus. Our scientific incursion set out to go beyond this issue and tackle the adjacent realm of interpreting ethics as well. We hint towards a certain connection between interpreting omissions and ethics, in the sense that the loss of information that is integral to omissions can have repercussions which are deeply-rooted in ethics. To that end, we study different codes of ethics and their limitations, along with the theoretical role of the interpreter and the issue of fidelity. Subsequently, we delve into omissions and rely on most of the relevant available literature in this sense in order to provide a comprehensive picture of what they are, how and why they occur, as well as how they have been scrutinized thus far.

Omissions in interpreting

The concept of omission has been subject to a fair share of divergent definitions. For the purposes of this study, the term ‘omission’ is understood as what occurs when information transmitted in the source language with one or more lexical items does not appear in the target language, and therefore potentially alters the meaning, thus borrowing a definition from interpreting research. The study employs the syntagm ‘transfer of information or ‘information transfer’, to denote the main inter-linguistic activity of professionals. We use the term ‘information’, understood as all kinds of linguistic material, bearing a certain oral form when spoken and bearing a certain lexical form when transcribed.

Moreover, it is imperative to emphasize that we are dealing with omissions, namely a loss, as the aforementioned definition also suggests. This loss is not necessarily a negative one, and thus, adopting the alternative concept of a “loss of meaning” would not be appropriate, as it brings up negative connotations. Meaning is certainly and integral part of any spoken event, but it does not represent the main focus of this thesis. Conversely, a loss of lexical information in a certain context can even be beneficial for the interpreting output, for the end user, etc. At the same time, we must emphasize the idea that this is not a study of quality and that interpreting quality is so broad and complex that it cannot be reduced to a factual examination of interpreting omissions, such as the one we are conducting. In doing so, we also aim to shed some light on the different concepts that scholars use when discussing quality – namely equivalence, fidelity, effectiveness, accuracy, competence, completeness, faithfulness etc. The only consensus that we can point to is that measuring quality in interpreting is a difficult research avenue and that as a multi-faceted element, there are numerous dimensions of output quality that could be taken into account. These range from completeness, clarity, discretion, lack of disturbances and comprehensibility to grammatical correctnes, voice quality, discipline, articulation and many more. We conclude that there is, unfortunately, a lack of reliable parameters and frameworks when it comes to assessing and analyzing quality in interpreting. The field of interpreting studies is in need of more research on quality and could benefit from different attempts at breaking it down and measuring it, however fallible, limited and subjective these approaches may be, as they could teach researchers more about this concept. Furthermore, issues of interpreting assessment and quality in general should concern the organizers of admission tests and accreditation exams, along with trainers, scholars, practitioners etc. Our subsequent ethical incursion is two-sided, following a theoretical avenue discussing major pieces of research, but also analyzing more concrete codes of ethics in interpreting.

Tackling omissions, scholars have studied them in a number of ways and the diverse and profound connections between non-renditions and other areas of interpreting are largely acknowledged, on a theoretical level. Omissions can take many forms — or lengths, rather — as they can represent one or more than one lexical item. In other words, interpreters can omit one word or a whole idea, phrase, or larger portion of speech. Regardless, they still fall under the category of omission. Such occurrences are also generally considered significant or insignificant (small or big) in assessment and pedagogy. Significant omissions can be seen, by interpreting trainers, either as removing a larger portion of speech or as eliminating a smaller element, perhaps even one lexical item, but one with considerable importance to the overall message of the speech. In other words, the general term 'omission' denotes a consierably vast range of elements that can be excluded by interpreters. We attempt to the discuss numerous definitions and facets of omissions present in research, outlining their benefits and limitations.

Omissions are central to the problem of the transfer of information in simultaneous conference interpreting. Seeing that there is limited consensus on how interpreters transfer information and how this transfer should be quantified, assessed, achieved and taught, it becomes apparent that studying omissions can provide scholars with some insights not necessarily into the transfer of information per se, but on the loss of information. While the general transfer is very complex and difficult to analyze in its entirety, on a practical and empirical level, the process of removing content can be scrutinized more easily. Additionally, by finding out more about what interpreters leave out and how much they remove, we can uncover more about the content that they do focus on — the types of information that never or rarely seem to go missing and that are prioritized in the transfer of information. While linguistic content, meaning and quality are too complex and broad to be empirically analyzed in interpreting, the loss of information can be measured and assessed in a more straightforward manner, and that is why we chose to investigate omissions. Also referred to as 'non-renditions', omissions are inherently linked to accuracy, fidelity (along with more general ethical considerations) and quality, but can be argued to represent a more measurable parameter in interpreting. Research has been quite scarce in this area but seems to point to consecutive interpreting being more accurate compared to the simultaneous mode, due to the lag in delivery and the possibility of taking notes. When looking at the research landscape, it is important to highlight that there is still no general consensus among researchers on how omissions should be approached, as an object of study.

The thesis also illustrates the main ways in which omissions have been approached in research, starting with the error analysis trend of the 70s and 80s, followed by theoretical works arguing in favor of qualitative analyses and more abstract works. At the same time, there is a lot to cover on the strategy-versur-error debate. On one hand, omissions can be regarded as mere interpreting mistakes (and we attempt to showcase potential causes and tendencies). On the other hand, the alternative viewpoint implies regarding omissions as techniques, as deliberately-employed interpreting strategies.

A vast number of studies dealing with omissions on an analytically descriptive or empirical level have been experimental in nature. Such studies present severe limitations by creating an artificial background for the interpreted interaction, a setting that implicitly alters the conduct of practitioners. A shift toward descriptive studies that seek to observe and examine real-life, genuine interpreted situations would be beneficial and we attempt to align to this tendency. While error analysis remains a feasible avenue for research, despite some drawbacks, it is mandatory to also emphasize that omissions are not just errors. Omissions are occurrences that are present in the output of interpreters, either as mistakes or techniques, and it is the duty of research to understand them better. Moreover, there is no condition or obligation to classify omissions as either techniques or strategies from the beginning or as an underlying premise. Depending on the individual instance, an omission can either be a strategical device employed by the interpreter, but it can equally well be an error that comes about due to a comprehension mistake. Comprehension mistakes, for instance, can be generated, in turn, by limited cognitive capacities on behalf of the interpreter, but can also be directly brought about by a number of problem triggers linked to the original discourse, such as unclear and poor oratory/elocution on behalf of the speaker. We will tackle problem triggers shortly in this summary. Moreover, the concept of 'omission' seems to have been negatively charged by scholars for a while, as it was long associated with the idea of interpreting errors exclusively. We believe that including compensatory techniques employed by interpreters in order to skip or not convey certain lexico-semantic elements of the original speech under the more general umbrella of omissions can alleviate the negative connotation of the concept itself and can allow for a better understanding of what is omitted and which omissions do indeed serve discourse functions and which do not. This takes us to an analysis of problem triggers, namely those specificities of the original input or the interpreting setting that can have a direct impact on the interpreting output, potetially leading to omissions. In this sense, we adopt a framework of problem triggers from interpreting research, with the aim of studying potential links between these particularities of the input and the omissions present in the output.

Methodology and research design

Our study employs a representative corpus (of over 94 thousand words, representing 5 different meetings held in the course of one month, with a total of 10 hours of audio and video material) and a viable methodology (based on an innovative classification of omissions) to find out what is omitted, how often certain information is omitted and why omissions occur. The study starts from an examination of the state of the art in interpreting studies, outlining gaps such as the limied corpus-based approaches and the insufficient research on omissions. It then presents a rigorous methodology that can enable the study of omissions by means of a mixed-methods corpus-based approach. Along with illustrating an innovative taxonomy of omissions, the thesis scrutinizes types, numbers and frequencies of omissions, as well as problem triggers and potential implications for the usage of the findings, touching on adjacent sub-fields of interpreting.

In other words, we present the main underpinnings of our research project, tackling objectives and investigative questions, anticipated outcomes, target audience, envisioned impact and limitations, as well as research design. We set out to find the right methodological tools and framework in order to reach our objectives. In our case, these represented a corpus-based approach, whose design, procedures and limitations were exhibited accordingly. While the transcribing conventions were essential to the generation of the corpus, aspects related to the selection of material, scope, representativeness and replicability were also considered crucial. The initial taxonomy of interpreting omissions was presented under the same methodological incursion, before defining the main methodological tool we employed, namely discourse analysis. Replicability was one of the main objectives we took into consideration when designing the methodology and we are confident that the framework can indeed be replicated across different modes, types, fields and settings of interpreting, as well as different language combinations.

Existing taxonomies of omissions are discussed at length before presenting the route that we took, namely that of creating an innovative taxonomy, taking into account the limitations of previous ones. Omissions are complex and their study, over time, has gone through as many twists and turns as has interpreting as a field. Some methodological consensus with respect to omissions is desirable and, in that sense, we believe that being aware of the limitations of earlier works can provide valuable insights into the potential directions of future research. By being aware of what has already been researched and what that research lacked, we can enhance the quality of future projects and further our quest to understand omissions better. Moreover, in the case of the present study, examining these limitations was a crucial stage before designing our own updated taxonomy of omissions, a taxonomy that has managed to bypass or directly address some of the drawbacks present in earlier research. Thus, following a discussion on corpus design, technical procedures, limitations, transcription and selection of material, we tackle the representativeness and replicability of this corpus-based approach.

This is then linked precisely to our proposed taxonomy of omissions, exposing the methodology behind it and illustrating our intention of designing a classification which would be able cover a high number of possible types of omissions that interpreters can make, and to make it potentially replicable — in the sense that it could provide a tool for analyzing non-renditions that can be applied to other studies and in other contexts, including other language combinations and other modes or contexts of interpreting. We reached this classification by conducting a relatively limited assessment on a short sample of written interpreted material — namely around the first five thousand words of the original corpus. We choose to address the crux of the matter and scrutinize the available corpus in order to find out which types of informational content are omitted in simultaneous interpretation. Thus, we differentiate between ten categories of speech elements that tend not to be rendered at times and these are omissions: of agent, of example/part of an enumeration, of small talk, of links between ideas, of entire ideas or larger portions of speech, of redundancies, of implicit speech, of names, of numbers and of attributes. This taxonomy appears slightly revised and improved towards the end of the thesis, after presenting our findings. Apart from the traditional discourse analysis that we employ, it is noteworthy to mention that we propose a mixed-methods investigation into the corpus, thus presenting both qualitative and quantitative results. The discourse analysis that we conduct on the corpus is primarily bottom-up in nature, in that we allowed the data to guide the research and reveal what we should be investigating, in terms of patterns, particular tendencies and special features, as opposed to approaching the corpus with pre-established, specific hypotheses to confirm and with certain characteristics to identify. The corpus, in turn, is a parallel and billingual one, adhering to strict transcribing conventions established beforehand.

Analysis, findings and implications

The qualitative side of the investigation starts by exploring introductory speeches in the meetings under scrutiny and then takes omissions by category, linking certain types in order to showcase regularities of patterns. The section on findings also recurrently factors in problem triggers and references previous scholarly findings, relying on the literature review we initially conducted. The quantitative results show numbers and frequencies of omissions for each category, therefore demonstrating that some types of information from the input indeed go missing more frequently than others and that some omissions seem to be mere interpreting errors, while others appear to be strategic tools employed by interpreters.

Our corpus-based data analysis shows a significant number of omissions in all the meetings we analyzed. We conclude that we can observe omissions of all kinds and lengths and we identified recurring patterns that pointed to the existence of problem triggers and contextual factors that often influence the interpretation. Some categories of the taxonomy proved to be inter-dependent, in the sense of observing overlap, and we explain how this overlap is both unavoidable and desirable, in fact, as it allows researchers and practitioners to scrutinize inextricable connections between types of omissions and their respective elements of linguistic information. We also discuss the findings in light of the available literature and conclude that there are, indeed, certain types of information that go missing more frequently than others and that these findings reveal consistent interpreting patterns. We found that the interpretation employs a wide range of strategical tools in handling problem triggers and that some non-renditions managed to increase the coherence of the output, serving clear purposes. One final illustration of the omission taxonomy is included, in light of the findings and in order to facilitate further replications of the methodological framework, before relating the findings to existing research and displaying validity and limitations.

The findings thus show a significant number of omissions and a higher frequency of omissions for certain types of linguistic information, such as elements of enumerations, attributives, implicit items, redundant items, but also entire portions of speech. Moreover, a link between omissions and certain particularities of the original speech, such as speed, density or language errors primarily, becomes evident. The thesis also discusses certain strategies employed by professional interpreters during the inter-linguistic transfer of information, but emphasizes that, from a strictly quantitative point of view, 16.8 words go missing in the interpreting output for every 100 words uttered by the original speaker and that the highest frequency of omissions is that of entire ideas or larger portions of speech, usually linked to imposing problem triggers. Moreover, qualitative findings illustrated for every category of omission present in the taxonomy showcase tendencies and patterns within that category, frequently relying on corpus-based examples and evoking problem triggers.

These findings also display certain strategic tools that interpreters frequently employ and provide examples for functional omissions, where the non-rendition of information present in the original actually enhances the coherence of the interpreting output. The final interpreting product often seems to represent a deverbalization of the original input, with professionals striving to mainain a sense consistency with that original, clearly conveying or attempting to convey the message of the speaker and not the mere words. Finally, the outliers of the analysis are also tackled, with some specific patterns serving to improve the initial omission taxonomy. We show during our qualitative analysis that sometimes there is a certain degree of overlap between categories and that there are inter-dependencies between the different types of non-renditions and their respective information — attributives can be placed in enumerations, names can become redundant or implicit and agents are often expressed by means of names. However, in the case of non-renditions that did see overlap, we did not count them twice in this quantitative analysis, and we always selected one category that seemed to overshadow the other one. In almost all such instances, one category of non-renditions seemed dominant to the detriment of the other one that could be considered. Finally, judging by the number of outliers that we managed to find in these two meetings, it becomes apparent that additional categories can be considered. The addition of the category of personal references would leave very few instances of omissions that can truly be considered outliers. This would improve the thoroughness and stability of the proposed taxonomy.

All in all, it is evident that enumerations pose problems to simultaneous interpreters and that series of elements enhance cognitive load and processing ability so much that this phenomenon generates several non-renditions, for instance. We also expected to see a high number of omissions of implicit and redundant information, as they are part of the deliberately employed interpreting technique of ‘cleaning up’ the incoming message, simplifying it and converting it into a coherent output in the target language that still conveys the original message. Names and numbers proved to be a challenge for our interpreters, as expected, but arguably to a much more limited extent than expected.

We manage to establish, by means of this brief quantitative outlook on the corpus that the simultaneous interpretation in Romanian does omit information and that a significant amount of information is not rendered into the target language. Furthermore, we were able to pinpoint those types of information that are more frequently omitted than others and this leads to the creation of clear patterns that were upheld throughout the corpus. It is important to know not just what is omitted, but also how much. We would like to point out the importance and potential of quantitative approaches in simultaneous interpreting. After the earlier decades of the field, where scholars were adamant to count interpreting errors and limit the realm of interpreting quality to these figures, research has largely steered away from quantitative considerations, mostly disregarding their validity and potential to show a complete picture. However, research in interpreting studies could benefit from returning to incorporating quantitative considerations and conducting mixed-methods investigations into the intricate realities of simultaneous conference interpreting. Such observations and examinations in research projects can provide insights into matters that have long stayed under-studied and despite the fact that figures cannot illustrate the entire truth behind what interpreters do in their daily practice, quantitative investigations can complement qualitative ones really well.

One of our aims was also that of creating a link between the practical corpus-based findings that we outline and relevant theoretical precepts that make up the available scholarly literature in interpreting studies. To that end, we discuss what these findings mean for interpreting practice. We start with an illustration of frequencies of omissions and problem triggers, before revisiting the omission taxonomy and tackling a range of concepts and debates that have already been covered in our literature review. Finally, we tackle the validity and limitations of the present research project. We believe that in relating our findings to existing research, it is clear that our results align with prior studies in demonstrating the importance of examining omissions. However, our study brings a further contribution by highlighting nuanced variations in this regard, by pointing to problem triggers and contextual factors, conducting a mixed-methods investigation, relying on authentic interpreted material and a sound methodological framework.

We believe our study points to certain patterns that can be explored further and that have a clear impact on the daily practice of interpreting. We found that omissions are commonplace in simultaneous interpreting, that certain types of information are omitted more frequently (some have a specific function, others tend to be mere errors) and that while a large portion of non-renditions are linked to problem triggers, the most common factors that influence the output are fast pace, dense and complex language, as well as speech errors. In a nutshell, we have showed that the Romanian interpretation implied a loss of linguistic information when compared to the English original.

Our findings can be relevant for the academic pedagogy and the training of interpreters, as well as for the actual practice of professionals. Descriptive and analytical studies conducted on real-life interpreted material, which have a mixed-methods outlook and seek to scrutinize omissions should be of interest to trainers and trainees, users, professional interpreters and accreditation bodies, along with any other stakeholder who is engaged in the employment of interpreting services. Furthermore, the findings and patterns we displayed throughout this thesis can have a meaningful impact on the manner in which interpreting is taught and practiced. We also outline some directions of future research in interpreting studies that foster potential, seeking to illustrate potential research avenues for the years to come. These include observational corpus-based studies, research on our methodology, ethics and quality, assessment, remote interpreting, pragmatics, comparative research, etc.

Most importantly, existing research has been missing a taxonomy of omissions that is replicable across languages, domains and modes of interpreting, and a taxonomy that is mindful, at the same time, of problem triggers and context. The other scholarly frameworks that we presented throughout the study are limited when it comes to their applicability and replicability, let alone their narrow approach of omissions in general. We firmly believe that the replicability of our framework is significantly improved compared to previous approaches. Our frame of reference can represent a means of replications across languages, domains and modes of interpreting. Evidently, it can also suffer changes and additions, and we discuss potential directions as well. Only a sound and stable taxonomy can allow for comprehensive investigations fostering the potential of making the complex realm of interpreting omissions more transparent.

Additionally, we believe our findings to be valid for a range of interpreted interactions and we regard our corpus as representative for most of the formal contexts that require the services of professional conference interpreters in the simultaneous mode. We managed to design a corpus that fosters diversity within its boundaries, with respect to speakers, topics, procedures, problem triggers, interpreters, types of speeches, durations, etc. Starting from the premise that the corpus is representative for the phenomenon being scrutinized, we believe our findings to also be representative and foster validity within the proposed scope. This validity can and should be enhanced by means of replications that go beyond this language combination (English-Romanian) and beyond this specific domain that we purposefully limited our analysis to (foreign affairs). As a result of our overall scrutiny, we managed to observe not just that interpreters omit information, but that they omit some types of information more than others. We believe that that alone is an important finding in itself, as it clearly relates to an interpreting pattern that has not been studied sufficiently. Following this initial conclusion, we managed to shed a light on what exactly they omit and how much is not rendered in the final output compared to the original input.

All in all, we believe that we managed to accomplish what we initially set out to do. This thesis represents a multi-lateral investigation into a rare source of authentic interpreted material (for which we have received an approval from the European Parliament in order to use the recordings for research purposes), it adopted a sound methodological avenue and consequently revealed intriguing conclusions about the realities of simultaneous interpreting. In other words, these findings can be said to have a paramount relevance for the under-researched realities of interpreting nowadays. The study brings up over 370 important scientific publications that have shaped the field of interpreting over the years, showcasing connections between these and the results of the study illustrated in the present thesis. The results can have a profound and far-reaching impact on several sub-fields, such as the daily practice of professionals, quality in interpreting, ethics, pedagogy, assessment, accreditation, etc., with particular relevance for professional interpreters, students and trainers, users of interpreting services, accreditation bodies and the general public.