UNIVERZITA MATEJA BELA V BANSKEJ BYSTRICI
FILOZOFICKÁ FAKULTA

STUDENT PERCEPTIONS OF INTERPRETING QUALITY IN SLOVAKIA

Diplomová práca

02cee431-a36d-4f29-a231-3266b3808f3c

Študijný program: Slovenský jazyk a literatúra – anglický jazyk a kultúra

Študijný odbor: slovenský jazyk a literatúra a prekladateľstvo a tlmočníctvo

Pracovisko: Katedra anglistiky a amerikanistiky

Vedúci diplomovej práce: PhDr. Martin Djovčoš, PhD.

Banská Bystrica 2018   **Bc. Andrea Tokárová**

ČESTNÉ VYHLÁSENIE

Vyhlasujem, že predkladanú diplomovú prácu som vypracovala samostatne pod odborným dohľadom vedúceho záverečnej práce a uviedla som všetku použitú literatúru.

V Banskej Bystrici dňa 16. 4. 2018   ....

Andrea Tokárová

"Differences of habit and language are nothing at all if our aims are identical and our hearts are open."

– J.K. Rowling

Acknowledgments:

Firstly, I would like to express my sincere gratitude to my supervisor, Martin Djovčoš, for his continuous support, patience, and guidance, as well as his endless enthusiasm for this project.

I would also like to thank professors Miroslava Melicherčíková, Martin Djovčoš, Vladimír Biloveský, Igor Tyšš, Pavol Šveda, Ester Demjanová, Miroslava Gavurová, Jaroslav Kompán, Mária Šupínová, Peter Fleischer, and Marika Siekelová for their invaluable help with the distribution of my questionnaires. My gratitude also goes to the 1,150 students who were willing to spare a few minutes to participate in my research.

Last but not least, I want to thank Alexander Millington for proofreading this thesis after having had to hear about it for about a year or so.

ABSTRAKT

TOKÁROVÁ, Andrea: *Vnímanie kvality tlmočenia slovenskými študentmi.* [Diplomová práca]. – Univerzita Mateja Bela v Banskej Bystrici. Filozofická fakulta; Katedra anglistiky a amerikanistiky. – Konzultant: PhDr. Martin Djovčoš, PhD. Stupeň odbornej kvalifikácie: magister. Banská Bystrica: FF UMB, 2018. 82 s.

Diplomová práca sa zaoberá vnímaním kvality tlmočenia slovenskými študentmi odboru prekladateľstvo a tlmočníctvo, ako aj študentmi iných odborov. Jej cieľom je zistiť, či medzi študentmi a profesionálmi v daných odboroch existujú zásadné rozdiely a či je možné takéto rozdiely nájsť aj medzi jednotlivými skupinami študentov. Ako prvá práca svojho charakteru na Slovensku zároveň predstavuje veľký prínos do slovenskej vedy o tlmočení a jej výsledky môžu poslúžiť nielen translatológom, ale aj samotným tlmočníkom. Práca sa delí na šesť kapitol. Prvé tri majú teoretickú povahu a autorka sa v nich zaoberá rozličnými náhľadmi na kvalitu tlmočenia, históriou výskumu tohto fenoménu a obmedzeniami vlastného výskumu. V praktickej časti sú najprv uvedené metódy, ktorými sa pri výskume postupovalo. Tie zahŕňajú najmä kvantitatívny dotazníkový prieskum medzi slovenskými univerzitnými študentmi, ako aj hodnotenie nahrávky tlmočenia vybranými zástupcami jednotlivých pozorovaných odborov. Autorka ďalej predstavuje desať hypotéz, ktoré následne testuje v poslednej, šiestej kapitole. Za ňou sa nachádza záver, v ktorom sú uvedené hlavné zistenia výskumu a návrhy na jeho ďalšie rozšírenie. Prácou sa potvrdila existencia rozdielov medzi študentmi a profesionálmi v jednotlivých odboroch, hoci boli zistené aj podobnosti v názoroch na relatívnu dôležitosť vyselektovaných 14 kritérií pre kvalitné tlmočenie. Rozdiely existujú aj medzi študentmi samotnými. Kým faktory ako pohlavie a ročník štúdia nemajú na vnímanie kritérií veľký vplyv, potvrdilo sa, že študijný odbor či skúsenosť s tlmočením môžu do veľkej miery predurčiť názory respondentov. Autorka tiež zistila nezhodu medzi prioritami, ktoré študenti uviedli v dotazníku, a ich hodnotením reálneho tlmočníckeho výkonu.

Kľúčové slová: Tlmočenie. Kvalita. Očakávania. Hodnotenie. Kritériá. Študenti.

ABSTRACT

TOKÁROVÁ, Andrea: *Student Perceptions of Interpreting Quality in Slovakia.* [Diploma thesis]. – Matej Bel University in Banská Bystrica. Faculty of Arts; Department of English and American Studies. – Supervisor: PhDr. Martin Djovčoš, PhD. Qualification level: Master. Banská Bystrica: FF UMB, 2018. 82 p.

The presented diploma thesis examines the perceptions of interpreting quality of Slovak university students of translation and interpreting, as well as students of other academic disciplines. It aims to establish whether there are differences between students and professionals in given fields and whether such differences can also be found among individual groups of students. As the first study of its nature in Slovakia, it is a valuable contribution to Slovak interpreting studies and its results can be an asset not only to interpreting scholars, but to interpreters as well. The thesis is divided into six chapters. In the first three chapters, the author reviews the theoretical background of interpreting quality by looking at the perspectives of different parties involved in the interpreting process, the history of research of this phenomenon, as well as limitations of her own research. The practical part of the thesis begins with a detailed description of the methods used in the research. These include, above all, quantitative survey research among Slovak university students and an assessment of interpreting recording by representatives of the individual academic disciplines. The author then presents a list of ten hypotheses, which are put to a test in the last, sixth chapter. This chapter is followed by a conclusion, in which the major findings of the study are stated alongside suggestions for future research. The thesis proved the existence of differences between students and professionals, although it also discovered several similarities in their perception of the relative importance of the selected 14 criteria for quality interpreting. Differences were also found among various groups of students. While factors such as gender and year of study had little effect on the perception of the criteria, the field of study and experience with interpreting influenced it significantly. The author also discovered discrepancies between the priorities listed by students in the questionnaire and their actual assessment of an interpreter's performance.

Key words: Interpreting. Quality. Expectations. Assessment. Criteria. Students.

Foreword

The diploma thesis *Student Perceptions of Interpreting Quality in Slovakia* places its primary focus on the expectations of students of interpreting as well as other academic disciplines in relation to interpreting quality. It originated from the author's interest in interpreting as a complex process in interlingual and intercultural transfer.

The opinions of the target population collected through quantitative survey research are analysed with respect to the respondents' socio-demographic background and compared to the findings of other researchers interested in the phenomenon of user expectations. The study also includes an assessment part which aims to establish whether the users' conscious expectations are reflected in their evaluation of interpreting performance.

A relatively well-researched phenomenon in foreign literature, user expectations remain a blank space in Slovak translation and interpreting studies. This thesis thus aims at colouring in the first area of this blank space by providing information on user expectations of a specific group – university students. Even though the subjects are not yet real professionals in their field, as students of higher education they have a great potential to achieve this position in the future and, subsequently, to also become users of interpreting services at various conferences. The thesis can therefore be very informative for both interpreters and interpreting scholars and one of its ultimate goals is to complement Slovak research on interpreting and to initiate further investigation into interpreting quality and phenomena related to it.

TABLE OF CONTENTS

List of figures:

List of tables:

List of abbreviations:

\# – ranking; number

AIIC – International Association of Conference Interpreters

A – accurate

avg. – average

CI – conference interpreting

CPU – Constantine the Philosopher University

CU – Comenius University

def. – definitely

exp. – experience

F – female

FM – final mark

$FM_\%$ - final mark in percentage form

H – hypothesis

IcC_{sum} – the sum of average scores for content-related criteria

IdC_{sum} – the sum of average scores for delivery-related criteria

IlC_{sum} – the sum of average scores for language-related criteria

L – lenient

M – male

MA – Master's

MBU – Matej Bel University

No. – number

non-TI – not studying translation and interpreting

OI – overall impression

OIS – overall impression score

PcC_{sum} – the sum of points awarded for content-related criteria

PdC_{sum} – the sum of points awarded for delivery-related criteria

PlC_{sum} – the sum of points awarded for language-related criteria

pref. – prefer; preference

prob. – probably

s – second(s)

S – strict

SI – simultaneous interpreting

SL – slightly lenient

SOSR – The Statistical Office of the Slovak Republic

SS – slightly strict

ST – source text

TI – translation and interpreting

TT – target text

UNIPO – the University of Prešov

UPJS – Pavol Jozef Šafárik University

INTRODUCTION

Research of quality in interpreting went hand in hand with the process of professionalization and has mainly focused on the simultaneous mode. Decades after its emergence and thousands of pages later, interpreting scholars are still not united in their opinions on interpreting quality. This is largely due to the unique and elusive nature of the phenomenon in question.

The motivation for writing this thesis originated in our interest in interpreting as such as well as in a particular line of research concerned with user expectations. Our original research plan consisted of replication of other existing studies and included both real interpreters and users of interpreting services. However, due to factors beyond our control, we were forced to change the target population to Slovak university students. This made our thesis unique in terms of the observed subjects and allowed us to work with more participants than any researcher of user expectations before.

The main goal of our thesis is to open the gates of user expectation research in Slovakia and to, hopefully, promote this phenomenon in a way which attracts future researchers and leads to its further investigation in our country. We want to ascertain whether student subjects are comparable to professionals, be it interpreters, or experts in other fields, and whether there are significant differences among various groups and subgroups of respondents. Lastly, we want to shed some light on the issue of interpreting assessment.

Our thesis consists of two mutually complementary parts – theoretical and practical, and is divided into six chapters. The first chapter looks at basic approaches in the research on interpreting quality, i.e. interpreting as a product, process, and service. Here, we express the opinion that the interpreter should always be given an opportunity to comment on their own performance, so as to avoid unfair assessment, and we also mention the problem of ideal and optimum quality. Lastly, we look at perspectives of various parties involved in the interpreting process, namely the speaker, listeners, client, interpreter, interpreter's booth partner, and the researcher, and talk about the assessment possibilities and limitations of each of them.

The second chapter is a summary of previous research of interpreting quality both abroad and in Slovakia. We begin with psychologists Henri Barik and David Gerver, continue with Hildegund Bühler, Ingrid Kurz, and Peter Moser, and get to researchers such as Franz Pöchhacker and Cornelia Zwischenberger, Ángela Collados

Aís and her colleagues at Granada University, and many others. Due to the limited extent of Slovak research on quality in interpreting, in this part of the chapter, we also mention several studies on other modes and settings of interpreting, e.g. consecutive and court interpreting, which nevertheless provide useful information.

The third chapter is devoted to user expectations as such and discusses certain limitations and shortcomings of this king of research. We look in more detail at problems such as the inability of users to assess the content of the interpreter's output, perceived importance of criteria vs. subconscious assessment, unwillingness of users to cooperate with the researcher, and their varied interest in the target text. Despite these shortcomings, we still consider our research valuable and informative, which we state in the last subchapter.

The fourth chapter consists of a detailed description of the methods used in our research. We discuss the creation of the questionnaires as well as the development of a mathematical formula used for determining the accuracy of assessors in later stages of the research. This chapter also includes information on the collection of the responses for our survey and an analysis of our two samples – students of translation and interpreting and students of other academic fields – as well as methods used for balancing the various sizes of analysed subgroups.

The shortest, fifth chapter is comprised of a list of ten hypotheses which we formulated in accordance with the goal of this thesis.

The sixth chapter summarises the results of our research and is split into subchapters according to the proposed hypotheses. The results, which are of a numerical nature, are presented in transparent and easily understandable tables and graphs. The last subchapter is a discussion of the results, in which we compare them to results of other studies and state our opinions on why a given hypothesis proved to be true or false and mention shortcomings peculiar to our research.

In the conclusion, we summarise the most important findings of our research and comment on the contribution of the thesis, which we believe can provide a valuable insight not only for interpreting scholars, but for interpreters and future interpreters as well. We also suggest several directions of further investigation which an eager translation student or scholar may take in the future.

Lastly, we would like to mention that due to the extent of this thesis, many of its parts have been moved to the appendices, which can be found at its very end.

1 QUALITY

Quality, in its most widespread sense, is not an under-defined concept. A mere look at dictionary definitions of quality gives one a solid idea of what is generally understood under this term – "[t]he standard of something as measured against other things of a similar kind"[1]; a "degree of excellence"[2]; "how good or bad something is"[3].

We live in an age of well-defined quality. Quality standards (often international) can be found for most things of any size and nature, from needles to bridges. Indeed, standards even exist for crucial equipment of simultaneous interpreters – booths. Why then do we still not have a unified definition of quality in interpreting?

1.1. Quality in Interpreting

Discussions about quality in interpreting went hand and hand with the process of professionalization, but even though decades have passed since it first became a topic of research, interpreting scholars are still not united in this matter. The lack of a single, generally recognised definition of quality in interpreting stems from its very nature – mainly its variability and elusiveness. Demands placed on interpreters and their output differ depending on the interpreting mode (e.g. simultaneous, consecutive, signed-language) and setting (conference, community, media, etc.). However, these are not the only determinants and, as will be discussed in Chapter 2, requirements differ not only with various groups of users (whether made up of experts in a particular field, or of people with none or significant experience with interpreting), but also with every individual. What one person perceives as an average or even poor interpreting performance, may be experienced by another as the best interpreting they have ever heard. Furthermore, interpreting cannot be assessed like a toy or chocolate, for it is neither tangible, nor, as was already said, will it ever appear the same to two individuals (admittedly, the latter could be argued for both a toy and chocolate). Some degree of "tangibility" can be achieved through an interpreting recording or transcript, which will be described in more detail in the next subchapter. However, even then, interpreting can never be a truly tangible concept. On the contrary, most interpreters' outputs are not

---

[1] Available at: <https://en.oxforddictionaries.com/definition/quality>. [accessed 2017-08-03]
[2] Available at: <https://www.merriam-webster.com/dictionary/quality>. [accessed 2017-08-03]
[3] Available at: <https://dictionary.cambridge.org/dictionary/english/quality>. [accessed 2017-08-03]

recorded, thus making it a one-time event which cannot be repeated and is therefore highly elusive.

1.1.1 Approaches to Interpreting or What Are We Trying to Assess?

To even begin thinking about setting generally applicable interpreting standards, researchers would first need to agree on what interpreting actually is. Here, two main perspectives are clearly distinguishable – that, which perceives interpreting as a product, and that which sees it as a process[4].

1.1.1.1 Interpreting as a Product

The product of interpreting, i.e. the interpreter's output, was traditionally seen as the sole indicator of quality in interpreting. Product-oriented quality assessment began at the turn of the decades in the late 1960's and early 1970's with the works of Henry Barik, who, based on recordings of an interpreters' output, developed a system of departures from the ST (omissions, additions, substitutions and errors) (Barik, 1971). While these terms are still commonly used in the description of TTs in interpreting, Barik's system has been criticised for disregarding the context in which the interpreting is taking place, as well as the fact that not all departures from the ST are necessarily a feature of low quality interpreting (see 2.1.1 for more details).

Error counts or propositional accuracy scores[5] are indeed a highly efficient and objective way of evaluating the content correspondence between the ST and the TT. However, they simply cannot be used as the sole indicator of interpreting quality, as they are perhaps too objective and assess every departure from the ST as a negative feature of the TT. When using these methods, the evaluator should therefore take into account the overall effect a given departure might have on the understanding of the TT.

A further shortcoming of propositional accuracy scores with regard to assessing interpreting as a product is the fact that this approach does not encompass extra-linguistic features of the TT, which constitute an important part of the interpreter's output (and are perhaps even more important than content-related features, as suggested by the works of Collados Aís (1998, 2007)). Nevertheless, assessment scales which

---

[4] The terminology differs across the works of various authors – in this thesis, we use the term product to refer to the interpreter's output, i.e. the target text (TT), while the term process denotes the wider communicative process of interpreting, including the various influencing factors.

[5] I.e. splitting the ST and TT into small units of meaning and identifying corresponding units in both; used by many authors, e.g. Tommola and Helevä in 1998 (in Gile et al., 2010), Mackintosh in 1983 (in Pöchhacker, 2001), Melicherčíková (2017), to mention but a few.

take form-related features into account do indeed exist and a very elaborate system was recently developed by the Slovak interpreter Lýdia Machová in her dissertation thesis (2016). Her evaluation form, however, also includes questions about the wider context of the interpreting event (e.g. the difficulty of the ST and the mental state of the interpreter on the day of interpreting), which is why we decided to describe it in the following subchapter.

To summarise, we perceive the sole assessment of interpreting as a product as insufficient. This approach disregards important factors which are often beyond the interpreter's control (such as bad working conditions and other input variables) and may in fact evaluate a perfectly satisfactory performance as one with severe deficiencies (or vice versa).

1.1.1.2 Interpreting as a Process

To assess the quality of interpreting as a process means to evaluate the appropriateness of the interpreter's choice of interpreting strategies and the degree to which they managed to cope with the many challenges of their profession. While this approach is more common in the consecutive rather than the simultaneous mode, it has been increasingly used in the latter as well.

A well-known train of thought in this regard are the effort models by David Gile (1997), according to which the process of simultaneous interpreting is made up of four efforts – a listening and analysis effort, a production effort, a memory effort, and a coordination effort, which allocates certain amounts of processing capacity to each of the three other efforts. In an ideal situation, each effort is given enough attention (processing capacity) to cover the requirements needed for a satisfactory interpreting performance. In a more realistic situation, deficiencies happen due to inaccurate distribution of attention. Every interpreter knows that this job rarely happens in ideal conditions and therefore, we should not speak about ideal quality either, but rather about "optimal quality" (Moser-Mercer, 1996, p. 44) or "quality under the circumstances", as Pöchhacker put it (1994, in Kurz et al., 2008, p. 1). David Gile (2009) goes as far as to say that even the combination of ideal conditions and skilled interpreters will necessarily lead to errors in interpreting and that ideal quality is unreachable. While Gile's theories are useful in explaining the difficulties of interpreting, they are not suitable (nor meant) as an evaluation scheme.

An interesting approach to interpreting as a process can be found in the works of Kalina (2002, 2005), who divided it into four partial processes – **a pre-process** (interpreter's skills, teamwork, preparation, etc.), **a peri-process** (the working conditions), an **in-process** (various requirements which need to be met by individual parties at an interpreted event), and **a post-process** (self-evaluation and studying terminology which was lacking during interpreting). While all-encompassing, this system would be rather difficult to use as an evaluation scheme for a third-party researcher, since many of its parts would be invisible to them and they would have to rely on the trustworthiness of the interpreter in question (e.g. with matters such as the extent of their preparation or the post-studying of terminology). However, it very clearly demonstrates how complex and lengthy the interpreting process actually is and could perhaps serve as an educating tool for clients and users.

Although evaluating interpreting as a process may sound fairer to the interpreter, it has a shortcoming similar to that of the sole output (product) evaluation – if we only evaluate the appropriateness of interpreting strategy usage and the overall handling of the interpreting task, we might end up with conflicting results. Imagine, for example, a relatively inexperienced student of interpreting , who does everything to the best of their abilities; due to their inexperience, it is likely that their delivery would be still poor compared to that of a skilled professional, who was perhaps much less diligent in their preparation.

Thus, we think it is necessary to combine these two approaches and work with both the interpreter's output and the wider context surrounding it. This will inevitably include taking into account what the interpreter has to say about the task and will help us avoid awkward and unfair situations, such as accusing the interpreter of omitting entire sentences when, in fact, the equipment was faulty and they were not receiving any signal. That is why our original questionnaire for interpreters (see Appendix D) gives them the opportunity to describe the working conditions they experienced as well as anything else they felt was important to say.

The evaluation form designed by Machová in her dissertation thesis (2016) is a useful self-evaluation tool meant for students of interpreting. It splits quality criteria into three areas – delivery, language, and content. Students use a point system to assess their performance, but are also given the opportunity to verbally express their thoughts. The second part of the evaluation form consists of process-related questions, such as asking students how they felt during interpreting or what they perceived as most

problematic. In our opinion, with slight adjustments, this form would also be suitable for use in real-life interpreting. The assessor could decide on their own criteria (and distribute points depending on the importance of each individual one), while the second part of the form could be completed through an interview with the interpreter after the event or by the interpreter themselves.

1.1.1.3 Interpreting as a Service

As Machová (2016) notes, treating interpreting as a service is scarce among Slovak translation and interpreting scholars to the point where it might actually be seen as degrading to the profession. Western literature on interpreting is slightly more open towards the idea of interpreters as service providers.

Here, however, we must be careful, for it would be easy to classify the approaches of many interpreting scholars as belonging to this category. For example, in her many studies, Kurz was interested in the user perspective, for she felt that users of interpreting as receivers of this service play a vital role in interpreting assessment (it is them we as interpreters are trying to satisfy and we should take into consideration their requirements of interpreting) (1993, 2001). Her research focuses on the criteria perceived as important by various groups of users, thus breaking the interpreter's output into several features and evaluating them as more or less crucial for good quality interpreting. While considering this approach as service evaluation is not entirely wrong, the fact that Kurz focuses on the mentioned criteria links her assessment to that of product assessment.

On the other hand, we think of the service approach as a more holistic one, in the perhaps slightly radical sense presented Jonathan Downie's 2016 book *Being a Successful Interpreter: Adding Value and Delivering Excellence.* Here, Downie describes a conference which he attended as an interpreter and, not knowing his audience, failed to satisfy his client on the very first day, trying to interpret as much information as possible and often missing nuances of what was said. On the second day, after realizing whom he was speaking to and what his listeners actually wanted from his service, he reduced the dense ST significantly, giving them an overview of what was said and making sure the general tone of the TT was the same as that of the original. As he later notes, the interpreting was a success and the client as well as the users were satisfied. (Ibid.) His story could be loosely interpreted as "know your audience and give them what they want", which is the approach we imagine when we think of interpreting

as a service. Nevertheless, some of Downie's stories might make it look as though interpreters are free to do what they want, as long as it makes their clients happy which, if not taken with a grain of salt, might explain the controversy of his book. Nevertheless, it does a great job of explaining our idea of interpreting as a service, however permissive it is.

A similar idea, although perhaps explained in a more academic way, is Pöchhacker's notion of *hypertext* (1995). Hypertext is essentially a conference, which has a particular function. With regards to interpreting, various hypertext types (conference types) place different demands on interpreters (a highly diplomatic event should be interpreted in a different way than a conference on selling cosmetic products). Thus, knowing the hypertext type is like knowing your audience in Downie's book. The better your knowledge of them, the better your interpreting can be.

Because of its holistic nature, evaluation of interpreting as a service may seem like an easy task, with the satisfaction of the client/users being the sole indicator of quality. However, we believe it is much more complex than that, with matters such as price or long-term relationship with the client also playing their role here and making things more complicated[6]. Furthermore, this approach does not help us reach the objective of our thesis, which is why we shall not be taking it.

To conclude, while there are no universal quality standards for interpreting, its assessment does exist. Every interpreting organisation, be it AIIC or small interpreting agencies, has its own admission procedures, which necessarily include interpreting quality assessment. Whether we decide to approach interpreting as a product, process, or service, we need to make sure we are taking into account the various factors beyond the interpreter's control, for, as Kopczyński stated, quality "is not an absolute value, but rather contextually determined" (1994; in Melicherčíková, 2017, p. 67).

1.1.2 Perspectives of Various Parties and Their Assessment Possibilities and Limitations

If we think of the interpreting process as a communicative event, we can clearly see that it encompasses a multitude of parties involved in it to varying degrees. The objective of this subchapter is to discuss these parties' perspectives and look at the possibilities as well as limitations of their assessment of interpreting quality.

---

[6] Would we as clients be satisfied with poor interpreting simply because it was cheap? (Gile, 1991; in Kurz, 2001, p. 405).

We primarily focus on the simultaneous mode (and the conference setting), which presents the assessors with the added challenge of real-time delivery and elusiveness of both the source and the target text. Brief notes on the various perspectives in consecutive interpreting can be found in Appendix A (their former location is indicated in the text by the following symbol: […]).

1.1.2.1 Speaker

The speaker's ability to assess the quality of interpreting is severely limited, given the fact that, in conference settings, the interpreter's output is streamlined straight into the TT listeners' headphones. Even if the speaker had the headphones at hand, it would be virtually impossible for them to listen to the interpreted speech, as they must focus all of their attention to their own output. Thus, we can say that the only giveaway of the interpreting quality for the speaker are the audience's reactions (such as laughter following a joke) and their questions/contributions in the discussion part of the session (provided there is one). There, the speaker can detect any misconceptions seemingly coming from his or her own speech which might in fact be attributed to the interpreter's output. However, one needs to be careful when blaming the interpreter for such instances, as it may also simply be a case of misunderstanding on the part of the listener in question (here, widespread confusion is a safer indicator of misinterpreted ideas).

[…]

1.1.2.2 Listeners

If we approach interpreting as a service, we can then apply more general definitions of service quality to it. This is what Ingrid Kurz did in her paper on quality in conference interpreting, when she stated that the marketing principle "quality must begin with customer needs and end with customer perception" (Kotler and Armstrong, 1994, in Kurz, 2001, p. 394) should also apply to conference interpreting. Several authors seem to agree that the listener perspective is the most important one when assessing the quality of interpreting (among others Seleskovitch, 1986, in Kurz, 1993, p. 314; Kalina, 2005).

End users are indeed the main, if not the only reason why interpreting at conferences takes place. But are they well-suited to carry out the complicated task of evaluating the interpreter's performance? First of all, we can assume that TT listeners have limited abilities in the language used by the speaker, otherwise they would not need interpreting. Furthermore, their intercultural competence will most likely be

limited as well (Machová, 2016). This, as well as the fact that users must inevitably choose between the source and the target text, debilitates their ability to assess the content of the interpreter's output. Of course, there are cases where it is possible for a user to judge even the content of the TT, but these are rather scarce. Moreover, different users have different demands on interpreting, which can lead to significant differences in assessment (these matters are discussed in more detail in Chapter 3).

Nonetheless, what the end users can evaluate fairly accurately are the formal features of the TT, such as intonation, hesitation, grammaticalness, use of correct terminology, synchronicity with the speaker, quality of the interpreter's voice, etc. In this aspect, they are probably the most important assessors of all, because, as we have already said, it is their satisfaction we need to strive to achieve.

Questionnaires seem to be an appropriate way of establishing how pleased the listeners were with interpreting. They can include both closed and open-ended questions. Open-ended questions (such as "How did you find the interpreters' performance?" or "Did you find any aspect of the interpreter's output irritating?") invite perhaps more intuitive and genuine answers than closed ones and might be useful in a qualitative research. However, other assessment methods, such as point scales, would be more appropriate in quantitative research.

[…]

1.1.2.3 Client

Whether a client is able to assess the quality of interpreting depends on several variables, beginning with their presence at the conference. As Machová (2016) points out, clients often secure the services of interpreters but do not themselves participate in the event and even if they do, they may or may not listen to the interpreters' output. Furthermore, their prospects as assessors depend on their interlingual and intercultural knowledge. Their evaluation ability may be similar to that of the end users. However, clients will also inevitably assess other qualities of the interpreter which are invisible to the users, such as the price, flexibility, loyalty, etc. (Moser-Mercer, 1996). Moreover, one should not forget that there are various types of clients – agencies, conference organisers, individuals, to mention a few – and their background (alongside the type of event they are organising) will most certainly have an effect on their expectations, requirements and, ultimately, on their assessment.

1.1.2.4 Interpreter

Against the past conjecture that interpreters do not in fact listen to their own output stands empirical evidence in the form of corrections, "false starts", cohesive devices, etc. We know for a fact that interpreters are indeed aware of their own speech which raises a question of their self-assessment ability. In most cases, an interpreter is capable of evaluating their own performance immediately after their turn. Such an evaluation can be based, for example, on the interpreter's inner feeling or intuition, which, in turn, stems from the amount of difficulty (or ease) they experienced while interpreting. However, without a recording of their own as well as the speaker's output, an interpreter's assessment of quality will mostly be holistic and therefore insufficient.

Self-assessment is an important part of every interpreter's professional development and is often recommended by teachers of interpreting as a good habit to get into. When provided with recordings of both the ST and the TT, the interpreter, as an interlingual and intercultural communicator, is well-suited to evaluate their own performance. Furthermore, their (previous) presence at the particular event allows them to better judge the used interpreting strategies (i.e., they, more than anyone else, know when an omission was necessary, desirable or faulty). Needless to say, such an extensive evaluation process is extremely time-consuming, which is probably the reason for its rare, if any, occurrence in research on interpreting quality.

[…]

Lastly, one should not forget the issue of subjectivity. Of course, no professional in any field wants to threaten their goodwill for the sake of someone's research on service quality. This is a complex issue which needs to be looked at with care. Indeed, it would be interesting to see a statistical comparison of interpreting performance evaluation by the interpreter in question and a qualified third party.

1.1.2.5 Interpreter's Booth Partner

The interpreter's booth partner is, of course, a qualified interpreter with the knowledge of both the languages and cultures in question. Furthermore, as Machová (2016) points out, they are able to listen to both the TT and the ST at the same time, because they are used to splitting their attention. However, even though a layman may think that the other interpreter is enjoying a "time off" after their turn is over, this is not true at all. In fact, they will more often than not be listening to the speaker, so as not to lose track of their speech, as well as helping their colleague with terminology and other

problem triggers, such as names and numbers, which they can write down. The idea that they could focus all of their attention (or even a significant part of it) on evaluating their colleague's performance is simply absurd. Using recordings to evaluate each other's interpreting could decrease the risk of subjectivity in self-assessment, although we think that a sense of professional solidarity would prevent interpreters from harshly criticising their colleague.

1.1.2.6 Researcher

Pöchhacker (2001) in his article *Quality Assessment in Conference and Community Interpreting*, distinguishes between an internal and an external researcher. While an internal researcher investigates concrete interpreting events, an external one is more interested in hypothetical or past ones. The two approaches are both useful in their own way. While an external researcher may, for one thing, strive to establish norms or general views on interpreting, an internal researcher has a good position for evaluating the quality of actual interpreting. This can be done in relation to the interpreter's output or to the overall process of communicative interaction as such (Ibid.). Machová (2016) considers an internal researcher (an interpreting scholar who, at the event, acts as an observer) to be the most qualified person of all to judge the quality of interpreting, provided that they have access to recordings of both the TT and the ST. Considering the fact that there is a much lower risk of subjectivity, we cannot but agree with the author, but we would also like to add that the researcher should be provided with additional details, such as the client's and/or speaker's demands on the interpreter, the interpreter's working conditions, etc.

[…]

1.1.2.7 Other Perspectives and Combinations

A perspective which is rarely mentioned in papers on quality of interpreting is that of a relay interpreter. As a person who does not speak the original speaker's language a relay interpreter would struggle to assess the content side of the first interpreter's output (although knowledge of the field might serve as a good indicator). The issue of the pivot version qualities is briefly mentioned by Kahane (2000), who concludes that not enough research has been carried out to establish whether an ideal interpretation is the same for both the listeners and the "second" (relay) interpreters.

The options for a combination of two or more parties in the evaluation of interpreting quality are bountiful. For example, it might be interesting to see the

assessments of a speaker and a client, an interpreter and their booth partner, or indeed a user, an interpreter, and a researcher, compiled and compared. A researcher could also work with users to assess the quality of the TT's content. They could, for example, test their comprehension of the TT to ascertain whether the information was correctly transferred from SL to TL. The questions would have to be designed with great care to ensure that they were not in fact testing the listeners' memory and the listeners' background knowledge should be on similar levels. A recent example of using comprehension testing is the study by Cheung (2013), in which the author combined comprehension questions with the listeners' evaluation of interpreting[7].

To conclude, there are a multitude of possible combinations which would without a doubt yield interesting and constructive results. A well-designed scheme may lead to a further widening of the topic of quality research in interpreting studies and open up new possibilities for a fairer assessment of interpreting quality.

---

[7] Abstract available at: <https://benjamins.com/#catalog/journals/intp.15.1.02che/details>. [accessed 2018-02-19]

2 HISTORY OF RESEARCH

The main focus of research on interpreting quality from its beginnings at the dawn of the 1960's until the present time has been on simultaneous interpreting, often with a reference to conference interpreting. As with many other areas of interpreting studies, the initial interest came from experts in other fields, namely psychology.

In this chapter, we look at the most important research on interpreting quality to date and discuss the focus, methodology, and possible shortcomings of individual studies.

2.1 Simultaneous/Conference Interpreting

The simultaneous mode of interpreting has enjoyed a steady growth in popularity since its televised "debut" at the Nuremberg Trials. After being adopted by such important international organisations as the United Nations and the European Union, it has largely overshadowed the consecutive mode, especially at multi-lingual events.

The simultaneity of the task has always attracted the attention of psychologists, psycholinguists, neuropsychologists (and others) more than consecutive interpreting, perhaps because many of them saw the newly emerging mode as more difficult and challenging. This view of superiority has also been shared by not a small number of conference interpreters themselves, so much so that Sergio-Viaggio, a UN interpreter and interpreting scholar once referred to them as the "boothed gentry" (1996, in Pöchhacker, 2004, p. 197).

With so much interest from experts not only within the field but also outside it, it is understandable why research on interpreting quality has also focused mainly on this mode of the interlingual and intercultural transfer.

2.1.1 The Late 1960's and the Early 1970's – Psychologists

In 1969, Henri C. Barik completed the very first PhD thesis on simultaneous interpreting. In this thesis, as well as in his subsequent works in 1971 and 1975, he analysed various types of TT departures from the ST. In his experimental layout, Barik had six people participate in his research – two professional interpreters, two fresh graduates of an interpreting programme, and two bilinguals with no special training in interpreting. In each group, one of the participants' dominant language was French

while the other one's was English. He observed the most frequent departures from the ST and divided them into three main categories, namely omissions, additions, and substitutions and errors. These were further split into subcategories according to their severity and/or origin. (Barik, 1971)

It is evident that the focus of Barik's studies was on interpreting as a product rather than a process. However, there is a major shortcoming in his research, of which he is aware, having stated it multiple times in the paper, and that is the insensitivity, as he puts it, to the elegance of the target text, and the sole comparison of the ST and TT correspondence. (Ibid.) For instance, Barik treats every omission as an error. However, as Stenzl points out, not every piece of information is equally important and the target audience might well find "a clear and intelligible text with some information loss […more useful than one which…] aims at completeness at the cost of clarity and intelligibility" (Stenzl, 1983, p. 29-30). This opinion is supported by findings of Moser's study (1995), according to which users of interpreting at conferences often prefer concentration on essential information on the part of the interpreter. A 2012 paper by Korpal, *Omission in simultaneous interpreting as a deliberate act*, explores the pragmatic side of omissions. The author, just like Pym (2008) is of the opinion that omissions do not always constitute an error, nor can they only be the result of a coping strategy, but that they should be looked at within a wider context and with the communicative function and pragmatic aspect of interpreting in mind.

The second shortcoming of Barik's experiment is the low number of participants, which limits the reliability of the presented results of his study. In practice, each of his six types of interpreters is only represented by one person, which really does not suffice to draw conclusions (although it appears plentiful for the creation of a system of departures from the ST). (Barik, 1975)

David Gerver, another psychologist interested in simultaneous interpreting, studied various aspects of this phenomenon, such as the issue of divided attention, information processing and memory performance. As part of his doctoral thesis, he carried out experiments in order to assess the impact of input variables such as an increased presentation rate, noise, and non-standard intonation on the target text. Findings from his partial study on the effect of the ST presentation rate from 1969 confirmed that the quality of the interpreters' output decreases significantly when the source text is presented at faster rates (e.g. 142 w/m as opposed to 120 w/m), while ear-voice span increases in a similarly significant fashion. (Ibid.) Like Barik's, Gerver's

study also suffers from an insufficient number of participants (five interpreters and five "shadowers"), but on the plus side, Gerver is slightly more benevolent when assessing the correctness of interpretation, i.e. he accepts paraphrases as correct renditions of the ST, since "a word-for-word translation was not expected and, indeed, would not have been a good translation from the interpreter's point of view" (Gerver, 1969, in Pöchhacker – Shlesinger, 2002, p. 56).

2.1.2 1986 – Hildegund Bühler

It is quite difficult to find a study on interpreting quality that does not mention Hildegund Bühler and her research conducted in 1986. The reason for this is that Bühler was the one who took the very first step towards establishing quality criteria which AIIC members considered important when sponsoring candidates for membership in their association (Bühler, 1986). As pointed out by Pöchhacker and Zwischenberger, her quality criteria "became something like the backbone of empirical research on quality in conference interpreting" (Pöchhacker – Zwischenberger, 2010, p. 1) and her study "inspired one of the most prolific and coherent lines of research on quality in interpreting research" (Ibid.).

The sixteen criteria to be evaluated on a four-point scale from very important to irrelevant are as follows: native accent, pleasant voice, fluency of delivery, logical cohesion, sense consistency (with original), completeness of interpretation, correct grammatical usage, use of correct terminology, use of appropriate style, thorough preparation of conference documents, endurance, poise, pleasant appearance, reliability, ability to work in a team, positive feedback of delegates (Bühler, 1986, in Melicherčíková, 2016a, p. 63). As can be seen from some of the criteria (e.g. poise), Bühler's study focused not only on the simultaneous, but on the consecutive mode as well.

Out of the nine TT-related criteria, the ones that the 47 AIIC members ranked as most important were sense consistency (with an average rating of 3.957), logical cohesion (3.8), and use of correct terminology (3.489), closely followed by fluency of delivery (3.468) and completeness of interpretation (3.426) (Bühler, 1986).

Having analysed the results of her questionnaire-based survey, Bühler suggested that they corresponded to the requirements of interpreting service users. However, this assumption needed to be put to a test.

2.1.3 Ingrid Kurz

In 1989, the Vienna-based interpreting scholar Ingrid Kurz conducted a bilingual survey among participants at a medical conference, asking them to evaluate eight of Bühler's output-related criteria (native accent, pleasant voice, fluency of delivery, logical cohesion, sense consistency, completeness of interpretation, correct grammatical usage, use of correct terminology) on a four-point scale. She had questioned Bühler's statement about the correspondence of requirements (expectations) of interpreters and users of interpreting services. Kurz discovered that while some of the criteria were attributed similar degrees of importance by the users (namely sense consistency, logical cohesion, and use of correct terminology), others received significantly lower ratings from users of interpreting than from interpreters themselves. In general, users attributed lower degrees of importance to the listed criteria than interpreters. (Kurz, 1989; in Kurz, 2001, p. 398)

Another important hypothesis of Kurz was that different user groups would have different expectations. In order to empirically test this, she carried out two more surveys, one among engineers at an international conference on quality control and the other among participants at a meeting of the Council of Europe. The yielded results proved the hypothesis, with significant differences found among the users' assessment of criteria such as correct grammatical usage, use of correct terminology, completeness of interpretation, and logical cohesion. (Kurz, 1993)

2.1.4 Replication of research

As already mentioned, Bühler's initial step in determining quality criteria has led to a relatively large number of similar studies aiming at establishing the demands various participants of interpreted events (but mainly the users) place on the interpreter and their output. In 2011, Franz Pöchhacker examined some of these efforts in his paper *Replication in Research on Quality in Conference Interpreting*. As he pointed out, replication in research of any kind serves as a way of validating (or, indeed, invalidating) the original study on which it is based (Pöchhacker, 2011a).

From several types of replication, the most commonly employed ones seem to be partial replication and replication with update. This is because the majority of researchers either changed the questionnaire items (e.g. added or removed some of the criteria) or conducted the survey in a different socio-cultural environment. (Ibid.)

Among the most notable studies focusing on empirical research of interpreting quality (apart from Bühler's and Kurz') are those of Gile (1990), Vuorikoski (1993; 1998, in Kurz, 2001, p. 400) Kopczynski (1994, in Kurz, p. 401), Moser (1995), Collados Aís (1998), Chiaro and Nocella (2004), and Pöchhacker and Zwischenberger (2010). Each of them approaches the issue of quality in its own way.

[…][8]

The 1995 survey conducted by Moser and commissioned by AIIC is by far one of the most elaborate in the area of research on quality in interpreting. Just like Kurz and many others, Moser was interested in the expectations of conference interpreting users. The questionnaire he used is a combination of closed and open-ended questions. Apart from content and form-related criteria, the author also inquired about the users' perceptions of the role of interpreters, irritating aspects of their output and other, rather innovative matters[9]. What makes this study particularly interesting is the multifaceted categorization of the 201 responses according to, among others, conference types, respondents' gender, age, nationality, role at the conference, and experience with simultaneous interpreting. The results clearly show that there are significant differences between the preferences of conference "oldtimers" and newcomers, between men and women, participants of different conference types, etc. (Moser, 1995)

A very interesting and important line of research into the assessment of interpreting quality was started by Ángela Collados Aís and complemented by her colleagues at the University of Granada. In 1998, Collados Aís published an article on the importance of nonverbal communication in simultaneous interpreting quality assessment. Her theory was that while interpreting users consciously perceived form-related criteria as inferior to those related to content, their evaluation of actual interpreting output would be subconsciously influenced by those aspects of an interpreter's output which they considered less important. To test this, the author created three recordings of simultaneous interpretation, one presented with a monotonous intonation, but fully consistent with the content of the original message, one with a lively intonation but containing content errors, and one presented with a lively intonation and fully consistent with the ST content. The three recordings were

---

[8] […] indicates the former location of the parts of this subchapter moved to Appendix A.
[9] For example, asking the respondents what they considered interesting/difficult about the profession, how long they thought an interpreter's turn should be, whether they could think of any comparable professions, etc. (Moser, 1995).

assessed by three groups of legal experts (42 participants in total) and the first video was indeed ranked as the one with the lowest interpreting quality. This led Collados Aís to the conclusion that users are in fact "not good judges of quality, simply because they are not in a position to perform this task" (Collados Aís, 1998, in Pöchhacker – Shlesinger, 2002, p. 336). This study was later followed by several more, based on the same principle of manipulating individual parameters of the interpreter's output (e.g. grammar, accent, style, etc.), further proving that deficiencies in just one single parameter can have severely negative effects on the overall assessment of interpretation (Collados Aís *et al.*, 2007).

[…]

Pöchhacker and Zwischenberger (2010) chose the Internet as the medium for their survey among AIIC members and received an impressive amount of responses – 704 in total. The questionnaire consisted of three parts – the first part elicited information about the respondents' socio-demographic background, while the second part replicated (and updated) Bühler's study (respondents were asked to rate the importance of 11 criteria for quality interpreting[10]) and also included an experiment, in which the interpreters were asked to share their thoughts on a short audio sample of simultaneous interpretation. The third part of the questionnaire was devoted to the perceptions of the conference interpreter's role. It is clear from the results that more than 20 years later, interpreters still consider sense consistency with the original to be the most important criterion for the quality of interpreting. Just like in Bühler's study (1986), it is followed by logical cohesion (however, several respondents pointed out that this is sometimes beyond the interpreter's control). Although these two criteria received lower ratings from the "new generation" of interpreters, other ones, such as correct terminology and correct grammar, were given higher ratings by the same group. Almost 45% of the respondents stated that the importance of individual criteria varied depending on the type of the interpreted event. The results of the audio sample experiment are also rather interesting and somewhat contradictory to the findings of Collados Aís. Half of the respondents were given a recording with a monotonous intonation, while the other half listened to interpreting with a lively intonation. The difference in overall evaluation was small (0.16 on a six-point scale). Furthermore, the youngest group of interpreters (aged 30 to 47) rated the two recordings almost

---

[10] The two new criteria were lively intonation and synchronicity (Pöchhacker – Zwischenberger, 2010).

identically. However, as Pöchhacker pointed out, it is likely that the impact of monotonous intonation would increase with the length of the interpretation (the audio sample was only one minute long). In the third part, devoted to the interpreter's role, the authors noticed significant differences between respondents of different gender, age, and working experience, in relation to matters such as intervention into the source text (more experienced professionals were more ready to intervene), loyalty to the speaker (felt as more important by female interpreters), etc. To conclude, the study by Pöchhacker and Zwischenberger provided useful results, whose validity is supported by the large number of responses. (Pöchhacker – Zwischenberger, 2010)

Of course, the studies we have briefly mentioned in this subchapter do not constitute an exhaustive list and there are many more similar efforts to look at, including, but not limited to, Meak (1990), Ng (1992), Mack and Cattaruzza (1995; in Kurz, 2001, p. 396), Garzone (2002), Cheung (2003[11], 2013[12], 2015), Lee (2008[13]), Jolibois (2010, in Pöchhacker, 2011, p. 47-49), as well as some more theoretically oriented papers, such as Moser-Mercer (1996), Shlesinger (1997, in Kalina, 2005), Kahane (2000), Pöchhacker (2001, 2011a), Kalina (2005), and others.

2.2 User Expectation/Interpreting Quality Research in Slovakia

Interpreting research in Slovakia is relatively young and not very extensive. Nevertheless, the 21st century and especially the past decade have seen several publications which, if not devoted to quality of interpreting as such, at least scrape the surface of this issue. In this subchapter, we will present a selection of them in chronological order.

In her book, Makarová perhaps indirectly defines good quality interpreting through a list which she calls the "Ten Commandments for Interpreters" (Makarová, 2004, p. 6). However, this list focuses solely on formal features of the TT and while it is a helpful guide for beginner students of interpreting, we still perceive it as rather limited even in its scope. The same could be said about the list of Ten Commandments found in Müglová's 2009 publication, with the difference that this author also mentions

---

[11] Abstract available at: <https://www.sciencedirect.com/science/article/pii/S0346251X16300537?via%3Dihub>. [accessed 2018-02-18]

[12] Abstract available at: <https://benjamins.com/#catalog/journals/intp.15.1.02che/details>. [accessed 2018-02-19]

[13] Abstract available at: <http://www.tandfonline.com/doi/abs/10.1080/1750399X.2008.10798772>. [accessed 2017-11-20]

several pre-process "commandments", such as the preparation of glossaries and widening one's general knowledge.

A summary of Barik and Kopczyński's classifications of errors in interpreting can be found in Šramková's 2008 paper *Druhy chýb pri tlmočení a ich dopad na komunikáciu*[14]. Despite the fact that this article does not deal with quality of interpreting as such, in the conclusion, Šramková notes that the interpreter's output should be assessed as a whole and that, most importantly, it should show equivalence, coherence, and cohesion. Nevertheless, we cannot forget the fact that this output is predetermined by its oral realisation and as such will always include certain formal deficiencies. The author therefore suggests that we distinguish between severe, unacceptable errors (whether content or form-related) and deficiencies acceptable in the given communicative situation. (Šramková, 2008)

Bohušová, in her book on neutralisation in transcultural communication (2009), just like many foreign authors, mentions two opposing parts of quality assessment in interpreting – (imagined) needs, i.e. expectations of users and their subjective perception of the actual situation. She talks about the difficulty of interpreting quality assessment which she believes exists due to the nature of interpreting – the fact that this service always follows a custom order and is therefore unique and nonrecurring. There exist no prototypes or samples and the interpreter's performance is always dependent on the situation. The author also stresses the need for a stricter, more selective process of university admission. According to her, only talented students with realistic ideas about the translator or interpreter's job should be accepted. (Ibid.)

The year 2012 was rather productive in terms of interpreting-related studies published in Slovakia, although the following one was, admittedly, written by a Czech author. Čeňková (2012) explores three participant perspectives in community interpreting during an asylum process – that of the interpreter, the employee of a state institution (e.g. the police), and the asylum seeker. She notes that their expectations are often very different. While the police officer might wish for the interpreter to be unbiased and interpret exactly what was said, the asylum seeker's only determinant of quality is the achievement of their goal (being granted an asylum). (Ibid.) While Čeňková's study deals with consecutive rather than the simultaneous mode, we believe

---

[14] *Types of errors in interpreting and their impact on communication.*

that the situation with different perspectives and expectations is universal for all modes of interpreting.

Taida Nováková (2012) also ventures into a non-conference setting of interpreting. She looks at interpreting in court and mentions not only the various perspectives of interpreting assessment (event organiser, TT listener, interpreter, interpreter's colleague), but also the three main elements of quality analysis – content, language, and overall impression. (Ibid.) All three of these parts are further described and, with slight adjustments, they could be used in other modes and/or settings of interpreting as well.

An interesting study into interpreting quality was conducted by Stahl (2012). While his paper is only ten pages long, it provides the reader with a surprising amount of information. First of all, Stahl offers various definitions of the term "quality" and subsequently looks at quality in interpreting which, he says, is necessarily determined by the perspective of the assessor. His work also includes a survey among clients (people who order interpreting services, rather than users). He claims that while the TT **should** preserve the criteria of communicative equivalence, it **must** meet the expectations of the client. His survey is very short and includes six criteria for a successful interpreting – fluency, grammatical correctness, logical cohesion, pleasant voice, content equivalence, and other factors – which the respondents (clients) were asked to order from the most to the least important one. The most important factor, according to them, is logical cohesion, while "other factors" were seen as the least important. Unfortunately, Stahl's questions were only answered by six clients, which, in our opinion, limits the validity of the findings (the author also realises this shortcoming). (Ibid.) Nevertheless, his research is interesting, for it targets clients rather than users and, as every interpreter surely knows, it is the clients whom the interpreters must convince of the quality of their services. A satisfied user is just that – a satisfied user. But the ultimate decision maker when it comes to ordering interpreting services will always be the client, and Stahl seems to be very aware of this fact.

The year 2015 saw the publication of the very fruitful collection of studies under the name *Tlmočenie v interdisciplinárnej perspektíve so zameraním na osobnostné charakteristiky a kognitívne aspekty*[15]. Here, several articles reflect on the topic of interpreting quality. In their theoretical study, Šveda and Poláček (2015) explore the

---

[15] *Interpreting in an interdisciplinary perspective with a focus on personality traits and cognitive aspects.*

issue of student admission into interpreting courses at Slovak universities. The authors first look at the situation abroad and mention notable experiments with the selection procedure. They then discuss the situation in Slovakia and conclude that aptitude testing needs to be put into practice to ensure a higher quality of graduates as well as to help students develop the skills they are naturally good at. Šveda and Poláček suggest four scenarios for such a testing and stress the importance of including the Slovak language in it, since they feel the students' excellent knowledge of their mother tongue is usually taken for granted while in fact it often lacks in quality. (Ibid.)

The same publication includes a study by Stanislava Moyšová (2015), who likens the interpreter's performance to that of a professional sportsperson. Although it is not focused on quality as such, the article offers advice for teachers of interpreting on what to teach their students in order to help them improve their skills. She sees elements such as resilience to stress and correct breathing as important parts of the interpreter's skillset. (Ibid.)

Michalčíková's 2015 monograph aims at developing a system for evaluating consecutive interpreting of students. She does so by exploring similar systems used at several universities around the globe, and ultimately arrives at her own model, which includes three main elements – the semantic side (of interpreting), the formal and language side (of interpreting), and the interpreter's extra-lingual skills. (Ibid.) Although her model is based on consecutive interpreting, many of the criteria listed under the three elements could also be used in the evaluation of simultaneous interpreting and, indeed, served as an inspiration to us while we were compiling the criteria for our own questionnaires.

Machová's 2016 dissertation also served as a great source of inspiration. The aim of her thesis was to develop a self-evaluation form for students of interpreting. We got acquainted with this form in our fourth year of study, when we frequently used it in our interpreting classes. We are convinced that it is a great tool which helps students realise their weak points and motivates them to work on their improvement. Furthermore, we have also decided to use Machová's classification of the output-related criteria in our questionnaires by dividing the selected criteria into three categories – delivery, language, and content. We also approve of Machová's interest in the interpreter's working conditions – not an insignificant part of her self-evaluation form is devoted to this aspect of interpreting. (Ibid.)

Lastly, we want to mention Melicherčíková's 2017 book *Kognitívne charakteristiky a tlmočnícky výkon: „Súvisia spolu?"*[16] This very recent publication makes use of knowledge of cognitive sciences in the study of the interpreting phenomenon. The author suggests that the cognitive trait "field-independence" could be an indicator of quality performance in both simultaneous and cognitive interpreting. Furthermore, she found out that students of translation and interpreting excel in terms of their attention (concentration) abilities. An intriguing part of Melicherčíková's research tries to prove the existence of a correlation between the ST and the TT content equivalence (determined through a propositional analysis) and a positive assessment of independent evaluators. (Ibid.) The results indeed prove this correlation, which is contrary to the findings of many other researchers from abroad. However, further discussion with the author revealed that, in general, the students who managed to transfer less information also lacked formal qualities in their speech, while those who struggled less with the content transfer also performed better in terms of formal features. This is why we believe that the method of propositional analysis has significant shortcomings and needs to be supplemented by other methods or at least a verbal description of the interpreter's output.

---

[16] *Cognitive Characteristics and Interpreting Performance: "Are They Connected?".*

3 USER EXPECTATIONS – LIMITATIONS AND SHORTCOMINGS

In her 2001 paper *Conference Interpreting: Quality in the Ears of the User*, Kurz expressed the idea that service[17] quality (SQ and/or customer satisfaction) is what we get after subtracting expected service (ES) from actual service (AS). If we accept this formula (SQ = AS – ES), we must also accept that user expectations are a major determinant of interpreting quality.

For this reason, we decided to carry out a user expectation survey in Slovakia. With no predecessors, it is the first survey of its kind in the whole country and we hope it will help interpreters (and teachers of interpreting) deliver the best service possible in order to satisfy the only party that will undoubtedly be judging their performance – the users. However, as with every research, we must be aware of its possible limitations and shortcomings, listed in this chapter.

3.1 Users' Inability to Assess the Content

The inability to judge the content of the TT and its faithfulness to the ST is a major limitation of user assessment in SI, as discussed in subchapter 1.1.2.2. This is particularly problematic, since the category of sense consistency is rated as very important in all surveys, whether targeted at users or interpreters.

Despite perceiving sense consistency as one of the crucial factors for quality interpreting, the users' ability to assess whether the interpreter is actually fulfilling this criterion, is low; such cases may include a factual error where the users can, for example clearly see a number in a presentation different to the one they have just heard in their headphones, or an error recognizable due to their knowledge of the field, such as an interpreter at a medical conference saying that the humerus is located in the leg, rather than the arm. It is for this reason that the suitability of users to do the job of interpreting assessment is often questioned by interpreting scholars (e.g. Moser-Mercer (1996), Collados Aís et al. (2007)) – how can they judge an interpreter's performance if they are unable to assess the fulfilment of one of the most important criteria? Furthermore,

---

[17] Our view of interpreting as a service, as discussed in 1.1.1.3, differs to that of Kurz, who does not approach it in a holistic sense. When dealing with user expectations, we shall not use the word service in the holistic sense which we believe distinguishes it from interpreting as a product/process.

do they realise this limitation? The latter question might perhaps be answered by our research.

3.2 Perceived Importance of Criteria vs. Subconscious Assessment

While the users may be unable to judge the sense consistency of the ST and the TT, what they can assess are mainly formal criteria such as fluency, intonation, accent, etc. Problems arise when these criteria, despite usually seen as inferior to content-related ones, prevail over criteria such as sense consistency in the overall assessment. Thorough research into this matter, conducted at the University of Granada in Spain, proved that in many instances there are significant differences between our conscious perception and subconscious importance of individual criteria.

In our opinion, interpreting assessment by users only is insufficient and should be combined with that of a researcher with access to the recordings of the ST and the TT, who can evaluate parameters such as sense consistency. Even an evaluation like this might nevertheless be of little use to the users, who could simply reject our objections to their assessment and keep the opinion that the interpreting they have just heard (which was, for example, perfect in terms of sense consistency but poor in terms of confidence in voice) was of low quality. Essentially, the important thing to remember is that the users' expectations and the criteria they see as important should not be accepted without reservations.

3.3 Non-cooperativeness with the Researcher and Inconsistent Interest in the TT

A common problem researchers in any field face is the low return rate of respondents. This is particularly frequent in the written mode (e.g. a uniform questionnaire distributed among conference goers), while an oral interview usually yields a higher percentage of responses (however, it is more time-consuming for both sides, and mainly for the researcher) (Moser-Mercer, 1996). Non-cooperativeness can be a problem, especially if the researcher has to keep a low profile in order not to disturb or bother the respondents. During an interpreted conference, stumbling across rows of chairs and distributing questionnaires would likely be frowned upon by both the audience and the conference organisers. Therefore, the only chance for an interpreter to get in contact with the users is before or after the event, or during the breaks, which, again, might be perceived as a nuisance.

On the other hand, a questionnaire distributed online can reach a higher number of respondents at a time that is suitable for them, although this does not guarantee a particularly high return rate. For example, a survey by Chiaro and Nocella (2004) gained 169 responses out of about 800 successfully delivered invitations (which gives it a 21% return rate). Melicherčíková's 2017 study managed to get up to 31% (five assessors out of 16 addressed in total) through a more personal invitation. A similar response rate was achieved by the duo Pöchhacker & Zwischenberger in their 2010 study among AIIC interpreters (28.5%).

Another problem, this time related to quality assessment rather than user expectations, is the fact that users are not equally interested in the whole interpreted event, or even in the whole speech of one speaker. As a result, an almost perfect interpreting with one poorly rendered part may be judged as low quality overall or vice versa. We therefore propose asking the users how much of the interpreting they actually listened to. While this does not completely solve the above mentioned issue, it does give the researcher at least a vague idea of the type of respondent they are dealing with.

3.4 What Can We Learn from the Results?

Despite the listed shortcomings, we still believe the results of our research will be interesting and informative. While the expectations may not be entirely reflected in the users' actual assessment, the question of what they perceive as crucial for good quality interpreting is no less important, for it gives us an insight into their conscious minds. We also think that making the users think about the various criteria and the role of the interpreter might give them an idea of how complex interpreting actually is and how many aspects of it interpreters have to keep in mind while doing their job. It will be particularly interesting to see the differences in perception of criterion importance by different groups and subgroups of students.

As for the survey targeting students of interpreting and translation, we believe that the opinions of the students who are actually planning on becoming interpreters are particularly relevant, for it is likely that they will soon join their skilled fellows in the field (if they have not done so yet). That is why it is important to ask the respondents from this sample about their future plans as well as past experiences.

4 METHODOLOGY

The research methods for this thesis were devised and refined as soon as the first half of the year 2017. However, due to factors beyond our control, we were forced to change the original research plan in January 2018 and carry on taking an alternative route. This chapter looks at the secondary research methodology, while the original one is described in detail in Appendix B.

4.1 Student-Based Research

The new research design included several necessary changes. Instead of real users (experts in their field) we approached university students of any field other than translation and interpreting ("non-TI students") and instead of interpreters, we worked with students of interpreting ("TI students"). We dare disagree with Moser-Mercer (1996) who claims that the results of a study carried out on students rather than professionals cannot be generalised to a wider population of actual interpreting users. While differences may occur, we are of the opinion that they will not be too significant and that the general pattern of results seen in most major studies on user expectations will be preserved. Furthermore, we might also gain interesting insights into the differences in thinking between TI students in, for example, their first and last year of study, or between future interpreters from different universities. Therefore, we believe that with a sufficient and representative sample for every field of study observed, the results can still be very informative and tell us a lot about the preferences of potential future users of interpreting services (and future interpreters).

4.1.1 Recycling the Questionnaires[18]

Since no longer related to a specific event, the original questionnaires had to be adapted to meet our new needs. Inevitably, their assessment part was left out, as there was nothing to assess. However, we managed to preserve all of the main questions.

4.1.1.1 Questionnaire for Non-TI Students

The questionnaire targeted at Slovak students of any field other than translation and interpreting ("Questionnaire A") began by briefly introducing the aim of our

---

[18] For more information about the original questionnaires, see Appendix B.

research and informing the respondents about the estimated amount of time it would take to fill it in.[19]

There were five socio-demographic questions in total. We wanted to know the students' **gender**, **age**, **field of study**, **year of study**, and **experience with interpreted conferences** (none, some, ample experience).

The first part of Questionnaire A was followed by a short instruction, telling the respondents to imagine that they were attending a conference on matters from their academic discipline where they were using simultaneous interpreting. They were informed that the following parts of the questionnaire would be related to this hypothetical conference. Next, we asked the respondents these questions:

- **"Should the interpreter be a man or a woman?"** with five possible answers:
  - I prefer female interpreters;
  - I prefer male interpreters;
  - I prefer a speaker/interpreter gender match;
  - I do not have a preference;
  - other (specify).

  This question was inspired by Vuorikoski's 1993 research and was asked in order to determine whether users had strong feelings about the gender of the interpreter or not.

- **"What kind of interpreting would you prefer?"**
  - The interpreter interprets everything said by the speaker without adding or omitting anything (we use the terms "everything", "full rendition", or "ghost role" in reference to this answer in the research part);
  - The interpreter gives a faithful rendition of the speech but can add (explain) or omit (e.g. redundant) information ("free");
  - The interpreter only summarises what has been said by the speaker ("summary", "summarising");
  - other (specify) ("other").

  This question was inspired by both Downie's book (2016) and Moser's 1995 survey, in which he asked the respondents what they considered more

---

[19] The beginning also included our contact details (e-mail address), in case the respondents had some questions.

important, "concentration on essentials" or "completeness of rendition" (Ibid., p. 15).

The next step, identical in both questionnaires, consisted of rating **14 output-related criteria** on a four-point scale.[20] The criteria were listed in a random order and they are discussed in more detail subchapter 4.2. To give the respondents a chance to add their own criteria which they perceived as important, we included an optional open-ended question at the end of our list: **"Are there any other criteria that would be important for you?"**.

At the end, we thanked the respondents for their time and cooperation and invited their opinion on the questionnaire.

4.1.1.2 Questionnaire for TI Students

The second questionnaire ("Questionnaire B") was targeted at Slovak students of translation and interpreting[21]. It was derived from the original interpreters' questionnaire. Just like Questionnaire A, it also opened with a brief introduction of the research aim and information about its length.

The socio-demographic data section included seven questions in total. We inquired about the respondents' **gender** and **age**, **university** and **year of study**, **studied languages**[22], **experience with simultaneous interpreting**[23], and **desire to become an interpreter** in the future[24,25].

Before filling in the second part of the questionnaire, the respondents were told to imagine that they were simultaneously interpreting an international conference. They were then asked the following questions:

---

[20] We considered adding an extra point to the original Bühler's (1986) scale, but opted against this for two reasons – firstly, we wanted our research to be compatible with all major studies dealing with criteria importance (namely Bühler (1986), Kurz (1989, in Kurz, 2001; 1993), and Pöchhacker – Zwischenberger (2010)), and, secondly, we thought this would force the users to think in cases where they could not decide, as opposed to simply choosing the "middle ground" often used as the "I don't know./I can't be bothered thinking." option.

[21] In Slovakia, interpreting is always studied together with translation.

[22] We did not ask the respondents to specify whether their languages were A, B, or C languages, since we think that this may still be undecided at this point.

[23] Possible answers: a) I don't have any experience yet; b) My only experience is from lessons of interpreting; c) I have interpreted simultaneously as part of my university's interpreting experience programme; d) I have done simultaneous interpreting a few times outside of my university; e) I often do simultaneous interpreting outside of my university; f) other (specify).

[24] Possible answers: a) Definitely not; b) Probably not; c) Maybe yes; d) Definitely yes.

[25] We believe this is crucial information, since, in our opinion, wanting/not wanting to be an interpreter in real life determines to a great degree one's attitude towards interpreting, be it at school or behind its walls.

- **In your opinion, how does a good interpreter interpret?** (offered the same answers as the question on the interpreter's role in Questionnaire A);
- **How important would the following criteria be for providing quality interpreting?**;
- **Are there any other criteria that would be important for you?** (optional);
- **Do you think the importance of these criteria would change depending on the topic of the conference? If so, briefly describe how.** (optional).

After the last question, we thanked the respondents for their time and invited their opinion on the questionnaire.

4.2 The Criteria

The total number of criteria for the respondents to assess was 14. This subchapter describes how and why they were chosen for our survey.

4.2.1 Kurz' Criteria[26]

Eight of our 14 criteria were the ones chosen by Ingrid Kurz for her user expectation studies – native accent[27], pleasant voice, fluency (of delivery), logical cohesion, sense consistency, completeness of interpreting[28], correct grammar, correct terminology (Kurz, 1993).

4.2.2 Other Criteria

We added six more criteria on top of those used by Kurz. These criteria came from various sources. In the following paragraphs, we state the reasons for their addition on the list and say which study inspired it.

4.2.2.1 Lively Intonation

Proven by Collados Aís (1998) to be of extreme (unconscious) importance, we were particularly curious to see the scores this criterion would receive from both TI and non-TI students since, as the author discovered, it is often perceived as not very

---

[26] Although we refer to them as "Kurz' criteria", they are merely the eight criteria she chose from Bühler's 1986 study.

[27] Even though the criterion of native accent is not very relevant in Slovakia (the cases in which a non-native speaker interprets into Slovak are much less common than, for example, the cases of non-native interpreters interpreting into English), we decided to keep it, simply for a better consistency of our research with that of Kurz' and others.

[28] Here, we slightly changed the criterion and instead used "completeness of delivery", in accordance with Machová's dissertation thesis (2016).

important while the influence it actually has on the evaluation of interpreting is significant. Furthermore, lively intonation was also included in Pöchhacker and Zwischenberger's 2010 study and, in a way, in Moser's (1995) study.

4.2.2.2 No Filler Words & Hesitation Noises

Although filler words (such as "like" in English or "teda/vlastne/proste" in Slovak) are not mentioned in any of the listed "mainstream" studies, more than 74% of Moser's respondents rated "ums and ahs" as very or fairly irritating (1995, p. 29). Machová (2016) also made hesitations part of her self-evaluation form. In our questionnaire, we decided to combine filler words and hesitation noises because we believe they have a very similar, if not the same origin, i.e. they mainly occur when the interpreter is "stuck" and feels the need to say at least something to break the silence. We also think the irritation in both cases is caused by a frequent repetition of these phenomena and we treat both of them as deficiencies in delivery.

4.2.2.3 No Booth Noises

Booth noises such as coughing or rustling papers are warned against by many, e.g. Gile (2009), Taylor-Bouladon (2011), the AIIC website (2005)[29], to mention but a few sources. Almost 14% of Moser's respondents also spontaneously mentioned "poor microphone discipline" as one the major irritants (1995, p. 28), which further proves that avoiding booth noises is very important, not only for the sake of the listeners, who will undoubtedly hear any noise very loudly, but for our booth partner's comfort as well.

4.2.2.4 Synchronicity with the Speaker

Maintaining synchronicity with the speaker is a criterion which appears in Pöchhacker and Zwischenberger's 2010 study and is rated as important by 52% of all respondents. While at the same time spontaneously mentioned by several users in Moser's study (1995), one cannot but agree that it is a criterion difficult to assess by the listeners. However, just like with sense consistency, there are certain hints that can give away the interpreter's lagging behind the speaker.[30]

4.2.2.5 Clear Articulation

---

[29] Available at: <https://aiic.net/page/1676/>. [accessed 2018-01-22]

[30] For example, when the speaker tells a joke, the people in the audience who do not require interpreting will laugh immediately while the users have to wait for their share of fun – which, in case of jokes in particular, may come significantly later, if at all. Another, perhaps more common case, is the speaker switching the slide in their presentation before the interpreter has finished talking about it. It is situations like these where lagging can be perceived as a nuisance.

Interpreters are often likened to orators and it is common knowledge that orators should excel in articulation. We do not need to look far to find clear articulation mentioned as a crucial element of the interpreter's output. It is extensively dealt with even in Slovak literature. For example, Makarová (2004) lists clear articulation as one of her ten commandments for interpreters, while Michalčíková includes it in her thorough model of interpreting assessment (2015, p. 30). The very existence of the book *Tlmočník ako rečník*[31] (Vertanová, et al., 2015) is a further proof that the interpreter's oratory skills (which definitely include clear articulation) are not taken lightly. Palčeková's 2018 master's thesis deals with rhetoric as part of the interpreting course at Slovak universities and her work will definitely shed more light on this matter. It is therefore curious that none of the major studies dealing with user expectations lists clear articulation among the criteria for good quality interpreting[32].

4.2.2.6 Confident voice

A brilliant demonstration of the difference a confident vs. unconfident voice makes can be found in Machová's speech at the 2015 Polyglot Gathering – *The Pleasures and Pains of Working as a Conference Interpreter*[33]. She also includes this criterion in her self-evaluation form, where it falls under the category of delivery (2016). Stahl (2012) notes that an interpreter speaking with confidence will easily gain the trust of their listener. It is, again, curious that confident voice as a separate criterion is missing in all major studies on user expectations, when it is often considered to be one of the most influential factors in user evaluation of interpreting (however subjective its perception might be). It may, nevertheless, be similar to intonation in that the users might consciously perceive it as a relatively unimportant criterion when, in fact, it significantly influences their evaluation.

4.2.3 Three Categories of Criteria

Inspired by the self-evaluation form developed by Machová (2016), we decided to put the 14 criteria in three categories, as follows:

- **delivery:** fluency, native accent, completeness of delivery, pleasant voice, lively intonation, no filler words & hesitation noises, no booth

---

[31] *Interpreter as a Speaker.*
[32] Perhaps apart from Moser (1995), who lists "[speaking] very quietly" as an irritant (p. 29) while his respondents mention "clear enunciation" as an important criterion (Ibid.).
[33] Available at: <https://www.youtube.com/watch?v=mMLXYXOEHk0&t=5s>. [accessed 2018-01-23]

noises, synchronicity with the speaker, clear articulation, confident voice;

- **language:** correct terminology, correct grammar;
- **content:** logical cohesion, sense consistency.

Delivery and language-related criteria make up over 85% of our list, which corresponds to the popular opinion that content-related criteria are not user evaluation friendly and users should therefore not assess them at all. However, we decided to keep both logical cohesion (which, we believe, can be quite well assessed by the users, although it does not always depend on the interpreter) and sense consistency simply because they are usually seen as the most important criteria of all (alongside correct terminology and fluency of delivery) and we expect the scores they get to be similar to those in other user expectation surveys.

4.3 Collecting Responses & Representativeness of Samples

The first questionnaire to be initiated was Questionnaire B. On 25 January 2018, it was shared for the first time on the social network Facebook, in groups for TI students. We also sent a direct URL link to it to our acquaintances and professors with contacts for TI students at other universities and received help with its distribution from professors at our own university as well.

Questionnaire A was published just a day later, on 26 January 2018, and was also shared on the same social network, this time in university groups, as well as with our personal acquaintances. After approximately 200 responses, the number of respondents per day decreased rapidly, which is why, in the second half of February, we started going to universities and personally asking students to fill in the questionnaire. We visited five towns (Banská Bystrica, Zvolen, Žilina, Martin, and Bratislava) and approximately 15 faculties. Luckily, students were very cooperative and we also received invaluable help from several professors who allowed us to distribute the questionnaire in their classes.

On March 5, we stopped collecting responses to Questionnaire B, after the number of respondents had stopped at 250. Questionnaire A was stopped on March 13 after gaining exactly 900 responses.

4.3.1 TI Students' Socio-Demographic Structure

In total, we collected 250 responses from students of translation and interpreting. 205 (82%) of them were women and 45 (18%) men. Compared to data we acquired from the Statistical Office of the Slovak Republic (SOSR, 2018), it may seem that women are slightly underrepresented (there, women make up as much as 89% of all TI graduates). However, we think the 7% difference was mainly caused by the fact that our sample included first and second year students as well. Therefore, we consider this sample to be representative in terms of gender.
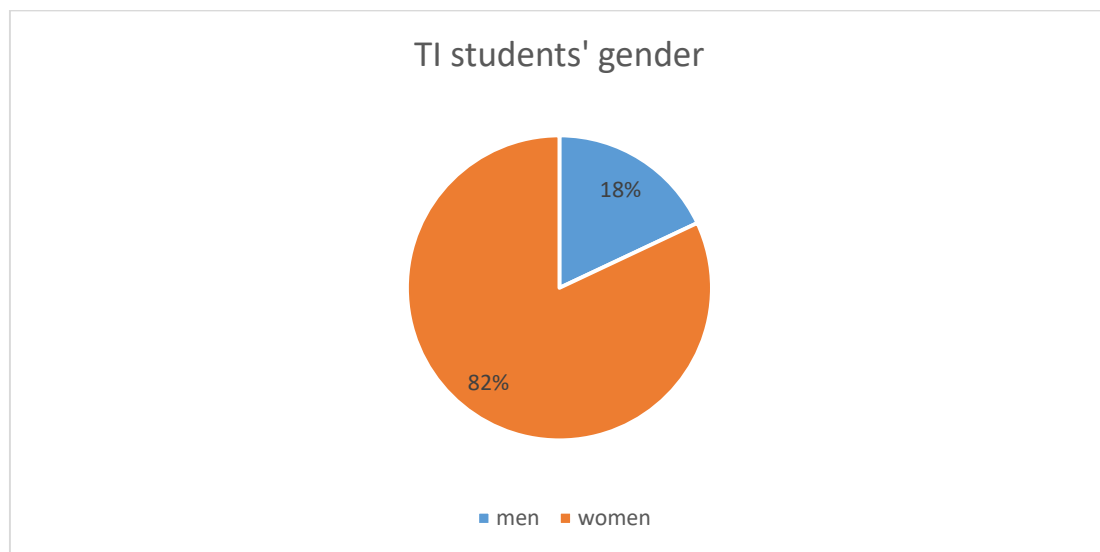
Figure 1: TI sample – gender

The following is the structure of our sample in terms of students' universities:

- Matej Bel University – 139 students; 55.6% [35%][34];
- Constantine the Philosopher University – 48 students; 19.2% [18%];
- Comenius University – 25 students; 10% [29%];
- the University of Prešov – 19 students; 7.6% [11%];
- Pavol Jozef Šafárik University – 19 students; 7.6% [6%].

---

[34] The percentage given in square brackets is the ratio of TI students who graduated from each of the five universities in 2014 (SOSR, 2018).
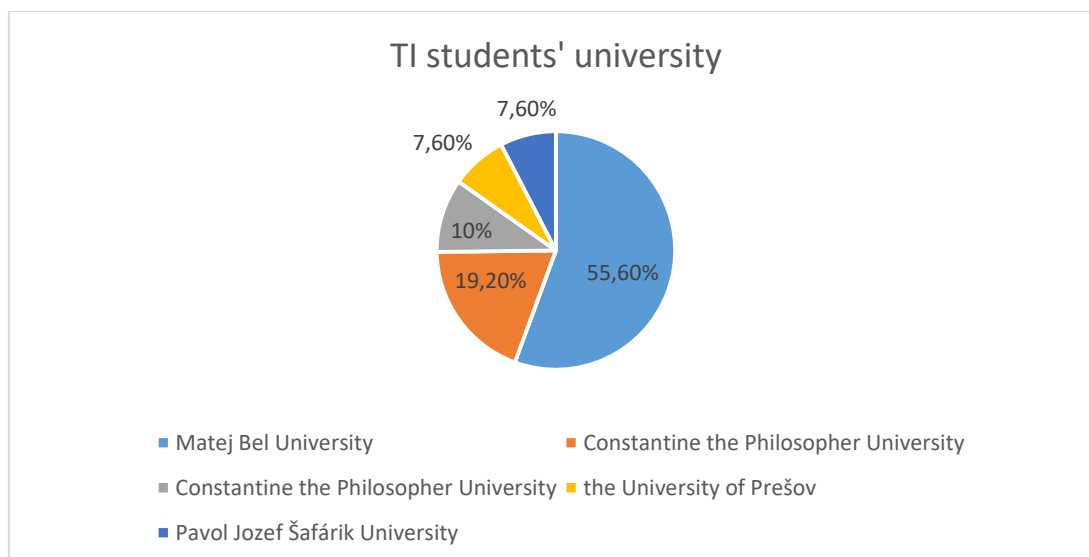
Figure 2: TI sample – university

As we can see, while MBU students are overrepresented by approximately 20%, students of CU are underrepresented by almost as much. The ratio for the remaining three universities is more accurate.

The distribution of students in different years of study looks like this:

- 1st year: 60 (24%);
- 2nd year: 35 (14%);
- 3rd year: 41 (16.4%);
- 4th year: 62 (24.8%);
- 5th year: 49 (19.6%);
- PhD level: 3 (1.2%).
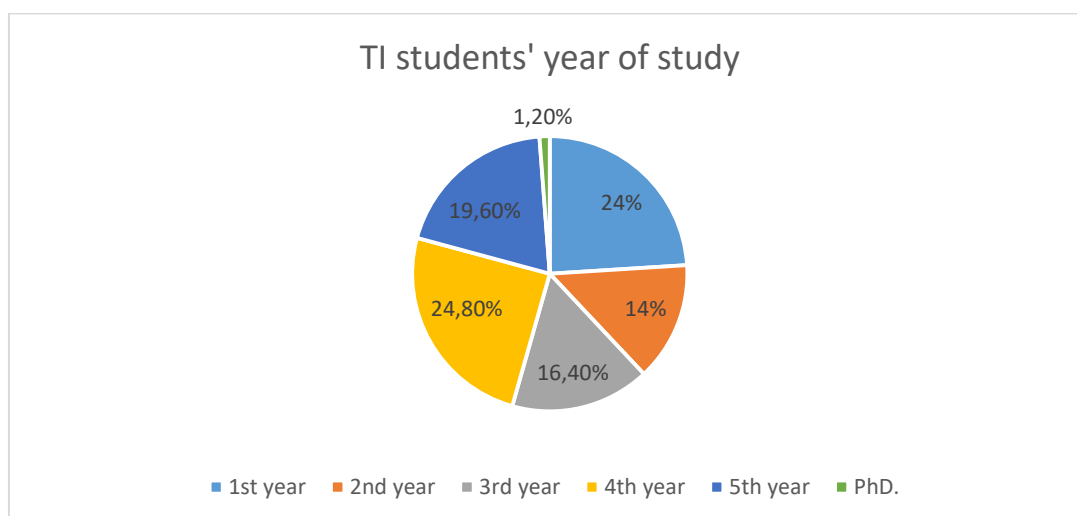
Figure 3: TI sample – year of study

While not having statistical data on the number of TI students in each year of study, we realise that our sample is not quite representative enough in this regard. However, we believe we have collected enough responses from each group (apart from PhD students) to be able to generalise our findings in an isolated manner (separately for each year). When referring to the entire sample of TI students, we shall be cautious, provided that we find significant differences between these groups.

When asked whether they would like to become interpreters in the future, TI students answered as follows:
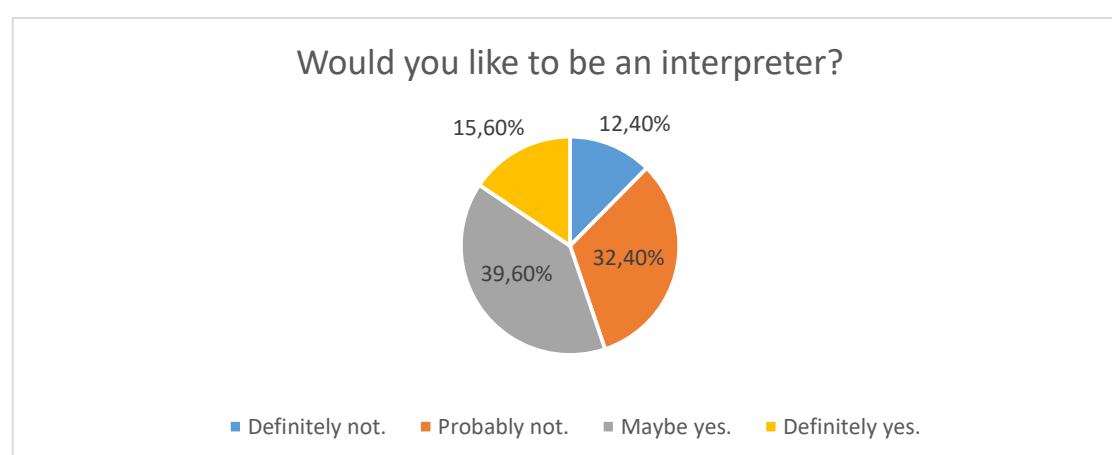
Figure 4: TI sample – desire to work as an interpreter

The next graph shows that most of our respondents have either none or lesson only experience with simultaneous interpreting, while only 39 have done out-of-lesson SI.[35]
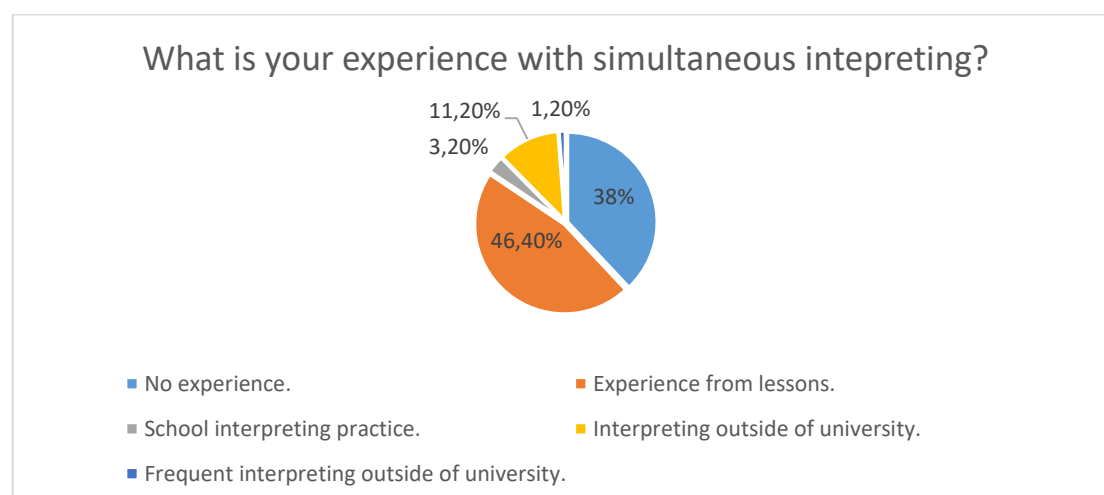
Figure 5: TI sample – simultaneous interpreting experience

---

[35] A few respondents used the "other" option to say that they had interpreted for their family/friends while on holiday. However, as this was most likely not simultaneous interpreting, we decided to add them to the "no experience" category.

4.3.2 Non-TI Students' Socio-Demographic Data

The total number of responses collected from non-TI students was exactly 900, out of which 575 (63.89%) were women and 325 (36.11%) were men. This means that there was a slight overrepresentation of women in the whole sample[36].
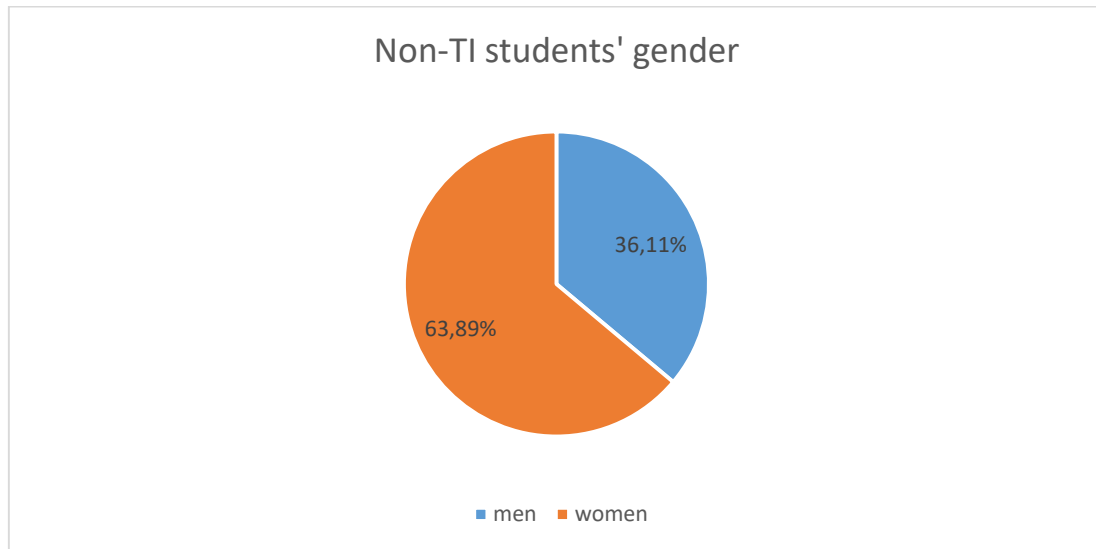
Figure 6: Non-TI sample – gender

The following graph shows the structure of our sample in terms of the students' year of study:
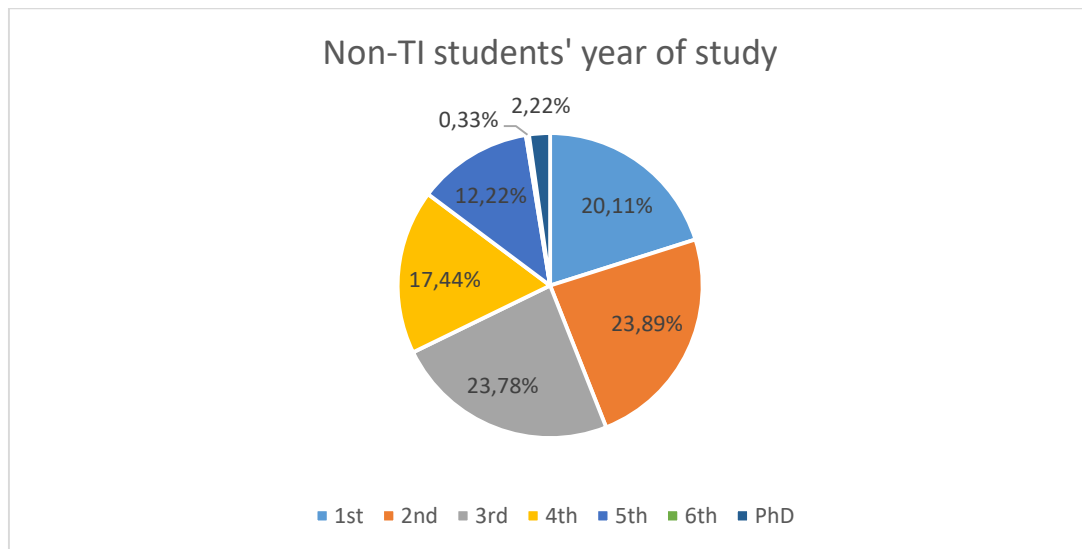
Figure 7: Non-TI sample – year of study

Each group out of the most common first five years is represented by at least 100 respondents. While the 6th year group may seem underrepresented, it is important

---

[36] According to the *Statistical Yearbook of the Slovak Republic 2016* (2016), in October 2015, 58.88% of all Slovak full-time university students were female.

to note that only three study fields in Slovakia require six years of study – medicine, dentistry, and veterinary medicine.

In Questionnaire A, we asked the respondents whether they had ever attended a conference at which they were using simultaneous interpreting. As expected, the majority of our sample has not had such experience:
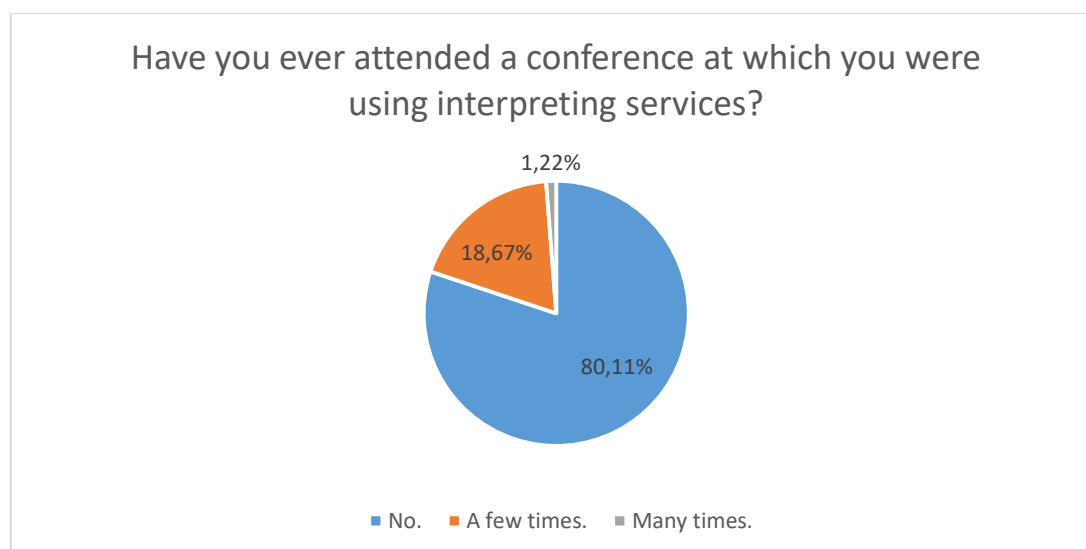
Figure 8: Non-TI sample – experience with interpreting

The most important differentiating factor was the non-TI students' field of study. After analysing the responses, we split the students into several categories depending on their field of study. We tried to create as many fields as possible in order to arrive at as accurate answers as possible. The groups (in alphabetical order) are as follows:

Field of Study	Abbreviation(s)	N (total)	n (men)	n (women)
art/aesthetics	art	33	6	27
civil engineering	civil engin.	31	16	15
finance/economy/management	finance	80	26	54
foreign languages	foreign lang.	32	1	31
forestry/agriculture/wildlife management	forestry	43	33	10
healthcare	healthc.	89	23	66
international relations	intern. rel.	46	19	27
information technology	IT	40	24	16
journalism/media	journal.	23	5	18
law	law	52	18	34

medicine	med.; medic.	69	22	47
natural sciences	nat. sc.	27	12	15
physical education/coaching	PE	33	29	4
pharmacy	pharm.	34	7	27
political science	polit.	30	16	14
public administration	public admin.	27	6	21
Slovak language	Slovak	33	2	31
social work	soc. w.	42	4	28
special pedagogy	special. pedag.	20	3	17
technical engineering	techn. engin.; tech. en.	39	30	9
tourism	tourism	33	10	23
transport/logistics/postal services	transport	24	8	16

Table 1: Non-TI sample – fields of study[37]

In most categories, women are slightly overrepresented, according to the data found in the *Statistical Yearbook of the Slovak Republic 2016* (2016)[38]. Grey-coloured fields indicate overrepresentation by 10% and more.

4.3.3 Assessing the Significance of Differences in Criterion Importance

After discussing the nature of our data with several statisticians, we established that the best way to balance the various sizes of individual groups and subgroups of respondents was to develop a scale, which we would use as a tool for deciding what we would treat as a significant difference in the perception of criterion importance in the context of individual sample sizes. This scale was created after carefully analysing the collected responses and was designed to reduce the risk of a small number of responses significantly influencing the mean value of the whole group. Therefore, we will only consider the following differences on the four-point importance scale to be significant:

- 0.2 point or more in case of groups of 45 or more respondents[39];

---

[37] An overview of study programmes included in each group can be found in Appendix B.
[38] However, this comparison is only approximate, since the publication does not split students into categories of study fields equivalent to ours.
[39] While many of the groups are much larger (e.g. all the female respondents of Questionnaire A), we consider a difference smaller than 0.2 point to be too insignificant in practice, whether or not it may be statistically significant.

- 0.25 point or more in case of groups of 30 to 44 respondents[40];
- 0.3 point or more in case of groups of 20 to 29 respondents;
- 0.4 point or more in case of groups of 10 to 19 respondents.

These ranges apply to both Questionnaire A and Questionnaire B. Furthermore, we decided not to analyse subgroups of fewer than 10 respondents, as we do not consider them to be representative. We shall also be cautious when analyzing the results of groups of 10 to 19 respondents[41].

4.4 Interpreting Assessment

Since our original research also included interpreting assessment, we did not want to leave this part out in the student-based research. We could not ask all of the respondents to assess a recording of interpreting, as it would have led to significantly lower response rates. Therefore, we chose 10 TI students and 10 non-TI students (one from each of the most represented study groups) to evaluate an interpreting recording of a fourth-year (female) student of MBU.

The topic of the original speech was the placebo effect. The short speech (**4:54 minutes**) was delivered with a neutral **British accent** at a speed of **183 syllables per minute** and it contained a minimal amount of terminology (we provided the student-interpreters with a short glossary containing **eight terms** in total[42]; their context was explained to the students beforehand).

After listening to 14 recordings made, we chose the one we considered the best. This recording was then sent to the 20 assessors together with an evaluation form, which can be found in Appendix E. The assessors were asked to give the interpreter one to five points for each of the 14 criteria found in our questionnaires, as well as a score of one to ten points for overall impression (OI). They could also add a comment if they felt they needed to do so.

Subsequently, the filled-in assessment forms were collected and analysed. We used a slightly altered formula developed for the original research to ascertain whether the OI score matched the score the interpreter should have been given according to the

---

[40] Although the range usually changes after 10 respondents, we decided to be stricter in this case and chose the range to finish with 44 rather than 39 respondents, simply because most of our study groups fall into this category and we wanted to be more cautious when analysing their responses.

[41] This applies to subgroups and sub-subgroups only (e.g. men within the political science group). For groups of students of different academic disciplines ("study groups") as a whole, we required at least 20 members.

[42] The glossary and the transcript of both the ST and the TT can be found in Appendix F.

points given for the 14 criteria and the perception of their study group of the criteria's importance. The formula looked like this:

$$\frac{\left(\dfrac{IdC_{sum}}{10} \ast \dfrac{PdC_{sum}}{10}\right) + \left(\dfrac{IlC_{sum}}{2} \ast \dfrac{PlC_{sum}}{2}\right) + \left(\dfrac{IcC_{sum}}{2} \ast \dfrac{PcC_{sum}}{2}\right)}{\left(\dfrac{IdC_{sum} \ast 5}{10}\right) + \left(\dfrac{IlC_{sum} \ast 5}{2}\right) + \left(\dfrac{IcC_{sum} \ast 5}{2}\right)} = FM_{\%}$$

Figure 9: Formula for counting the final mark

While it may look complicated, the formula is actually rather simple. The 14 criteria (C) are divided into three categories – delivery, language, and content (d, l, c). The "weight" or importance (I) of each category is calculated by adding up the mean scores of all the criteria in the given category (IxC_{sum}, where the letter in place of the "x" represents one of the three categories (d, l, c)) and dividing it by the number of criteria in that category (e.g. $IdC_{sum}/10$ is the "weight" of the delivery category, which includes ten criteria).

To calculate "weighted points" (i.e. points incorporating the importance of each category), we simply multiply the weight of the given category by the average score or points (P) assigned to all the criteria in it (e.g. PdC_{sum} for the category of delivery divided by 10, because there are 10 criteria in this category).

We repeat the process for all three categories and add up the numbers, to get the total amount of "weighted points" the interpreter was awarded – this is the value calculated in the numerator of the complex fraction.

The denominator of the fraction calculates the maximum amount of weighted points the interpreter could have been awarded for each of the three categories by multiplying the weight of the categories by five (the maximum amount of non-weighted or raw points). If we add these three values up, we arrive at the overall maximum of awarded (weighted) points.

By dividing the awarded weighted points by the maximum possible amount of weighted points, we get the $FM_{\%}$ value, which gives us the percentage "success" of the interpreter. We have decided to round this value to two decimal places (e.g. 80.49%).

When comparing the given and the appropriate overall impression scores (OIS), we worked with the $FM_{\%}$ value divided by 10 (FM – which puts it on a 10-point scale)

and allowed the assessors a certain amount of leeway by establishing the following levels of accuracy:

- **accurate**, if the value of FM differed from OIS by no more than 0.499 point;
- **slightly strict**, if the value of FM was higher than OIS by 0.500 to 0.749 point;
- **strict**, if the value of FM was higher than OIS by more than 0.749 point;
- **slightly lenient**, if the value of FM was lower than OIS by 0.500 to 0.749 point;
- **lenient**, if the value of FM was lower than OIS by more than 0.749 point.

In case of inaccuracy, the average marks given for each category of criteria can tell us which of the three categories likely influenced the assessor more than they thought it would. This formula therefore helps us decide whether we can truly rely on user expectations only.

We will also make use of the independent samples t-test to determine whether the differences between the two groups (TI and non-TI) are statistically significant or not.

4.4.1 Choosing the Assessors

The 20 assessors were chosen once we had determined the structures of our two samples. For the assessors chosen from the TI group, our aim was to faithfully represent all five universities according to the number of respondents from each of them. We also attempted to make the group representative in terms of gender and include students from various years of study (although we were primarily be interested in older students).

The non-TI assessors were chosen from the ten highest-represented study groups. At the same time, we also tried to make the sample representative in terms of gender. However, this group consisted of students in higher years of study only, as we were interested in the opinions of the people who are most likely to actually stay in their field and who, at the same time, already have a deeper understanding of it.

Thus, our TI assessors' group had the following structure:

- eight women and two men;
- five students from MBU, two from CPU, one from CU, UNIPO, and UPJS;
- five 5th year students, two 3rd year students[43], one 4th, 2nd, and 1st year student.

The structure of the non-TI assessors' group was as follows:

- six women, four men;

---

[43] There were many 3rd year students in the CPU group.

55

- one student of each of the following fields: healthcare, finance, medicine, law, international relations, forestry, social work, IT, technical engineering, and pharmacy;
- eight 5th year students, one 3rd year student (healthcare), and one PhD student (social work).

5 HYPOTHESES

The goal of our thesis is to ascertain whether TI and non-TI students differ from interpreters and users of interpreting (respectively) or not, as well as to determine whether there are differences among various subgroups of students. We have therefore formed the following hypotheses which shall be put to a test in Chapter 6:

H1: TI students view the 14 criteria similarly to Bühler's (1986) and Pöchhacker and Zwischenberger's (2010) interpreters. Non-TI students of specific disciplines have similar opinions on the importance of the criteria to Kurz' (1993) professionals in those disciplines.

H2: TI students attribute more importance to the 14 criteria than students of other fields, similarly to the interpreters and users in Kurz' 1993 study.

H3: There are significant differences in the perception of criterion importance between students of interpreting in lower years (first and second) and their older fellows.

H4: There are significant differences in the perception of criterion importance between students of interpreting who are considering a career in interpreting and those who are not. The former are stricter in their perception of criterion importance than the latter.

H5: No significant differences in the perception of criterion importance or the interpreter's role exist among students of the five Slovak universities offering the translation and interpreting study programme.

H6: A large majority of TI students prefer free interpreting to both the ghost role and summarising. Free interpreting is more popular with TI than non-TI students.

H7: There are significant differences among students of various fields, similarly to professionals in Kurz' 1993 study. Some criteria receive more varied answers than other ones.

H8: Non-TI students who have previously experienced interpreting at a conference are more likely to prefer a free rather than a full rendition of the ST, and are more likely to have no gender preference for the interpreter.[44]

H9: TI students are stricter assessors of interpreting quality than non-TI students.

H10: The majority of the 20 chosen assessors will give the interpreter an inaccurate overall impression score (significantly different from the calculated FM value).

---

[44] The first part of this hypothesis challenges the finding of Moser (1995), who discovered that "oldtimers" (users with ample experience with interpreting) at large conferences preferred completeness of rendition. On the contrary, we believe that it is inexperienced non-TI students who prefer a full rendition.

6 RESULTS AND DISCUSSION

In this chapter, we present the results of our research in an order correspondent to our hypotheses. We also try to analyse the findings and suggest reasons why a given hypothesis was accepted or rejected.

Due to the extent of the research, many of its parts had to be excluded from this chapter. However, shall the reader be interested in further comparisons of various subgroups, they can find the remainder of the results in Appendix C.

6.1 Hypothesis 1

H1: TI students view the 14 criteria similarly to Bühler's (1986) and Pöchhacker and Zwischenberger's (2010) interpreters. Non-TI students of specific disciplines have similar opinions on the importance of the criteria to Kurz' (1993) professionals in those disciplines.

First, let us compare our results with those of Bühler (1986) and Pöchhacker and Zwischenberger (2010). TI students' scores are written in the middle column for an easier visual comparison.

	Bühler N = 47	TI students N = 250	Pöchhacker & Zwischenberger N = 675-704
fluency of delivery	*3.468*	3.488	***3.7***
native accent	***2.9***	2.26	***2.662***
logical cohesion	3.8	3.756	3.744
correct terminology	3.489	3.556	3.6
completeness of delivery[45]	3.426	3.344	3.408
correct grammar	3.38	3.208	**3.489**
sense consistency	**3.957**	3.716	3.877
pleasant voice	**3.085**	2.52	**3.123**
lively intonation	-	2.752	**3.148**
synchronicity with the speaker	-	2.884	2.799

Table 2: Criteria – TI students and interpreters

It is clearly visible from the table that differences between TI students and interpreters do exist. Figures written in boldface indicate instances in which the scores

---

[45] In Bühler's study, this criterion is written as **completeness of interpreting**, while Pöchhacker and Zwischenberger work with **completeness** only.

of TI students differ significantly from those of actual interpreters. However, such differences are sometimes found only with respect to the scores taken from one of the two mentioned studies, while they are similar to the other one (the two criteria in which the students differ from both groups of interpreters are indicated in boldface). Furthermore, Bühler's and Pöchhacker and Zwischenberger's interpreters are also not completely united in their opinions, namely in two cases – fluency of delivery and native accent, indicated by the italicised figures.

However, what we find very interesting is the fact that while the scores themselves differ significantly in many cases, TI students and Bühler's interpreters were **consistent in the ranking** of the eight criteria with only one difference – the students perceived logical cohesion as marginally more important than sense consistency (by 0.05 p.). Pöchhacker and Zwischenberger's interpreters' ranking differed slightly more from both of the other groups.

Nevertheless, we must conclude that differences between TI students and interpreters, while perhaps not seemingly major, but statistically important according to our own established limits, do indeed exist and we therefore **reject** the first part of our hypothesis.

Next, we will have a look at how non-TI students' scores compare to those collected by Ingrid Kurz (1993) at three conferences. To match her respondents, we have chosen the following groups:

- attendants of an international conference on general medicine (MDS): **medicine**;
- attendants of an international conference on quality control (Eng.): **technical** and **civil engineering**[46] (engin.);
- attendants of a Council of Europe meeting (CE): **political science**.

	MDS N = 47	med. n = 69	Eng. N = 27	engin. n = 70	CE N = 48	polit. n = 30	**avg.** Kurz	**avg.** Tokár.
native accent	2.3	**1.986**	2.2	2.138	2.08	2.233	2.193	2.119
pleasant voice	2.6	2.507	2.4	**2.707**	2.396	2.4	2.465	2.536
fluency of delivery	2.9	**3.638**	2.966	**3.386**	3.208	**3.533**	3.025	**3.519**
logical cohesion	3.6	3.696	3.1	3.316	3.3	3.333	3.523	3.448

---

[46] Using the average scores of both groups.

sense consistency	3.6	3.652	3.655	**3.315**	3.6	3.633	3.618	3.533
completeness of interpretation[47]	3	**3.435**	2.9	3.17	3.458	3.3	3.119	3.302
correct grammar	2.4	**2.986**	2.03	**2.692**	2.688	2.9	2.372	**2.859**
correct terminology	3.4	3.594	3.138	**3.482**	3.729	3.7	3.422	3.592
average value	3	3.187	2.8	3.026	3.06	3.129	2.967	3.114

Table 3: Criteria – non-TI students and users

The table clearly shows that students of **political science** and attendants of a **Council of Europe meeting** gave the eight criteria very similar scores, apart from the criterion of fluency of delivery. Furthermore, their rankings of the criteria were exactly the same.

On the contrary, we found many significant differences in the two remaining pairs in both scores and rankings. The only criterion which received similar scores from all of the groups was **logical cohesion**, while **fluency of delivery** received significantly different scores from all of them. Average scores for all eight criteria were similar in all cases. Nevertheless, due to the considerable differences found among the first two pairs, we also **reject** the second part of Hypothesis 1.

To conclude, despite finding certain similarities in both cases (such as the rankings of TI students and Bühler's interpreters, or the scores and rankings of political science students and EU politicians), we **reject** Hypothesis 1. Further research would have to be conducted in order to determine whether students in general differ from professionals or whether the differences observed could perhaps be attributed to the nature of the Slovak language or the time that has passed since both Bühler's and Kurz' studies were published. Lastly, we found it interesting that while TI students were generally more lenient than interpreters, non-TI students, on the contrary, tended to be stricter in their perception of the criterion importance.

6.2 Hypothesis 2

H2: TI students attribute more importance to the 14 criteria than students of other fields, similarly to the interpreters and users in Kurz' 1993 study.

The following table shows criterion importance as seen by TI and non-TI students:

---

[47] Our criterion was called "completeness of delivery", which may account for the differences.

	TI all N = 250	non-TI all N = 900
fluency of delivery	3.488	3.434
native accent	2.26	2.123
logical cohesion	**3.756**	3.487
correct terminology	3.556	3.43
completeness of delivery	3.344	3.327
correct grammar	**3.208**	2.922
sense consistency	3.716	3.528
pleasant voice	*2.52*	2.731
lively intonation	2.752	2.666
no filler words & hesitation noises	**3.108**	2.699
no booth noises	3.144	2.962
synchronicity with the speaker	*2.884*	2.946
clear articulation	3.516	3.397
confident voice	3.32	3.226
average value	3.184	3.063

Table 4: Criteria – TI & non-TI students

TI students gave the 14 criteria an average score of 3.184, while the average score from non-TI students was only 3.063. TI students gave higher scores to as many as 12 criteria. The only two criteria which were seen as more important by non-TI students were **pleasant voice** (significant difference) and synchronicity with the speaker (insignificant difference). We therefore **accept** Hypothesis 2.

6.3 Hypothesis 3

H3: There are significant differences in the perception of criterion importance between students of interpreting in lower years (first and second) and their older fellows.

The following table shows TI students' average scores according to their year of study:

	TI all N = 250	1st year n = 60	2nd year n = 35	3rd year n = 41	4th year n = 62	5th year n = 49	PhD. n = 3
fluency of delivery	3.488	3.367	3.457	3.512	3.452	3.673	3.667
native accent	2.26	**1.983**	2.371	2.488	2.194	2.408	2.333
logical cohesion	3.756	3.6	3.743	3.829	3.71	3.939	4
correct terminology	3.556	3.567	3.686	3.659	3.5	3.49	2.667
completeness of delivery	3.344	3.267	3.314	3.341	3.323	3.469	3.667

correct grammar	3.208	3.3	3.314	3.439	**2.919**	3.245	2.333
sense consistency	3.716	3.733	3.629	3.707	3.694	3.816	3.333
pleasant voice	2.52	2.433	2.543	2.537	2.339	**2.837**	2.333
lively intonation	2.752	2.6	2.571	2.951	2.726	**2.959**	2.333
no filler words & hesitation noises	3.108	3	3.257	3.049	3.081	3.245	2.667
no booth noises	3.144	3.25	3.2	3.171	3.129	3	2.667
synchronicity with the speaker	2.884	3.067	**3.143**	3.024	2.71	**2.653**	1.667
clear articulation	3.516	3.617	3.657	3.683	**3.258**	3.51	3
confident voice	3.32	3.283	3.543	3.22	3.113	**3.531**	3.667
average value	3.184	3.148	3.245	3.258	3.082	3.27	2.881

Table 5: Criteria – TI students in different years of study

In general, we can say that TI students in various years of study do not differ significantly from their younger or older fellows[48]. The differences between the mean values for all criteria together do not go over our established borders of significance, even though some criteria received significantly different scores from certain groups (figures in boldface).

Overall, the difference between the mean scores for the 14 criteria given by 1st and 2nd year students (the ones most likely to have no experience with SI) and those given by 3rd, 4th, 5th year and PhD students is almost non-existent (3.183 and 3.184, respectively) and we therefore **reject** Hypothesis 3.

6.4 Hypothesis 4

H4: There are significant differences in the perception of criterion importance between students of interpreting who are considering a career in interpreting and those who are not. The former are stricter in their perception of criterion importance than the latter.

The next table shows the scores of TI students split into four groups, based on their answer to the question: "Would you like to become an interpreter in the future?":

---

[48] This is also true for the sample of non-TI students, whose scores were remarkably similar (see Appendix C).

	TI all N = 250	Definitely not. n = 31	Probably not. n = 81	Maybe. n = 99	Definitely yes. n = 39
fluency of delivery	3.488	3.516	3.457	3.434	3.667
native accent	2.26	2.258	2.309	2.303	2.051
logical cohesion	3.756	3.742	3.79	3.727	3.769
correct terminology	3.556	3.484	3.519	3.545	3.718
completeness of delivery	3.344	3.226	3.309	3.343	3.513
correct grammar	3.208	3.065	3.123	3.303	3.256
sense consistency	3.716	3.645	3.753	3.697	3.744
pleasant voice	2.52	2.71	2.42	2.545	2.513
lively intonation	2.752	2.71	2.654	2.778	2.923
no filler words & hesitation noises	3.108	3.29	**2.901**	3.182	3.205
no booth noises	3.144	2.903	3.148	3.202	3.179
synchronicity with the speaker	2.884	3	2.889	2.889	2.769
clear articulation	3.516	3.419	3.469	3.535	3.641
confident voice	3.32	3.323	3.21	3.333	3.513
average value	3.184	3.164	3.139	3.201	3.247

Table 6: Criteria – TI students with different outlooks on their future interpreting careers

The scores of these groups show a balanced situation with only one significant difference – students who stated that they will probably not consider an interpreting career, saw the criterion of **no filler words & hesitation noises** as less important than the whole sample.

We can clearly see that the groups considering the career of an interpreter are stricter than the other ones (the combined mean scores are 3.214 vs. 3.146, respectively), although this difference is not significant enough to prove our hypothesis, which we therefore **reject**.

6.5 Hypothesis 5

H5: No significant differences in the perception of criterion importance or the interpreter's role exist among students of the five Slovak universities offering the translation and interpreting study programme.

This is how TI students from different universities perceive the 14 criteria:

	TI all N = 250	MBU n =139	CPU n = 48	CU n = 25	UNIPO n = 19	UPJS n = 19
fluency of delivery	3.488	3.403	3.604	3.68	3.421	3.632
native accent	2.26	2.122	**2.542**	2.16	2.579	2.368
logical cohesion	3.756	3.719	3.875	3.92	3.579	3.684
correct terminology	3.556	3.554	3.5	3.68	3.579	3.526
completeness of delivery	3.344	3.338	3.375	3.4	3.316	3.263
correct grammar	3.208	3.144	3.354	3.28	3.316	3.105
sense consistency	3.716	3.734	3.729	3.8	3.632	3.526
pleasant voice	2.52	2.504	2.583	2.48	2.368	2.684
lively intonation	2.752	2.719	2.833	2.84	2.579	2.842
no filler words & hesitation noises	3.108	3.036	3.146	3.24	3.105	3.368
no booth noises	3.144	3.137	3.125	3	3.316	3.263
synchronicity with the speaker	2.884	2.763	**3.167**	2.72	3.053	3.105
clear articulation	3.516	3.496	3.688	3.28	3.526	3.526
confident voice	3.32	3.273	3.313	3.32	3.526	3.474
average value	3.184	3.139	3.274	3.2	3.207	3.240

Table 7: Criteria – TI students from different universities

Significantly different scores can be found only with respect to two criteria – **native accent** and **synchronicity with the speaker** – which are seen as more important by students from CPU, who are also the strictest group of all. Overall differences are negligible. We will now look at how these groups see the role of the interpreter.
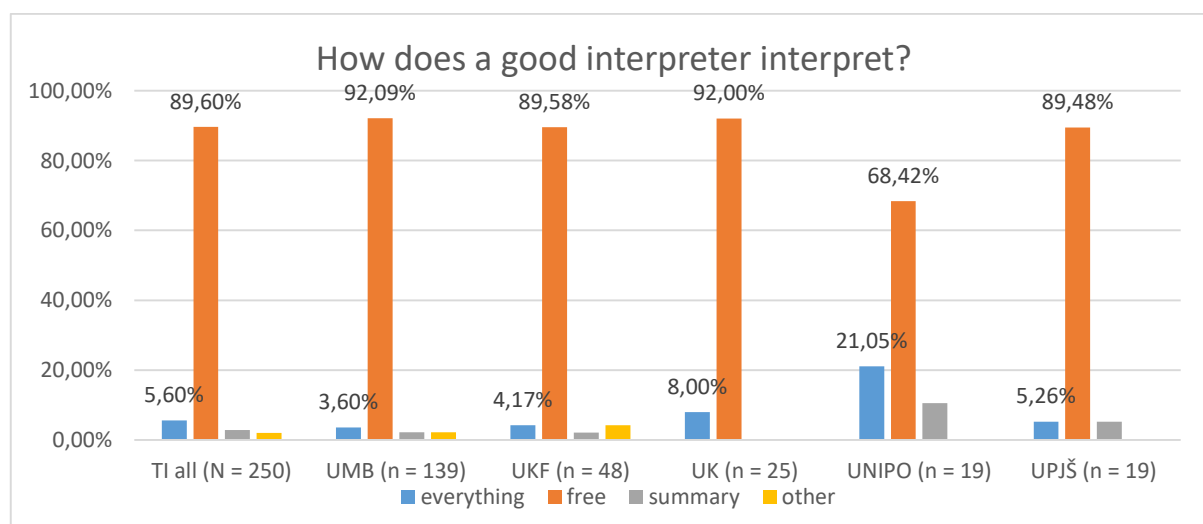
Figure 10: Interpreting styles – TI students from different universities

The percentages are similar in all cases but one – over 20% of UNIPO TI students chose "everything" over the most common "free" answer. Although this number may seem strikingly high, in reality, it represents only four students. Similarly, while 8% for "everything" (CU) may also seem like a substantial difference, this answer was only chosen by two students from this university. Thus, we do not consider these findings statistically significant due to the low representation of both UNIPO and CU students in our sample and we **accept** Hypothesis 5.

6.6 Hypothesis 6

H6: A large majority of TI students prefer free interpreting to both the ghost role and summarising. Free interpreting is more popular with TI than non-TI students.

We could already see the high preferences for free interpreting in the previous graph. Let us now compare TI and non-TI students' answers:
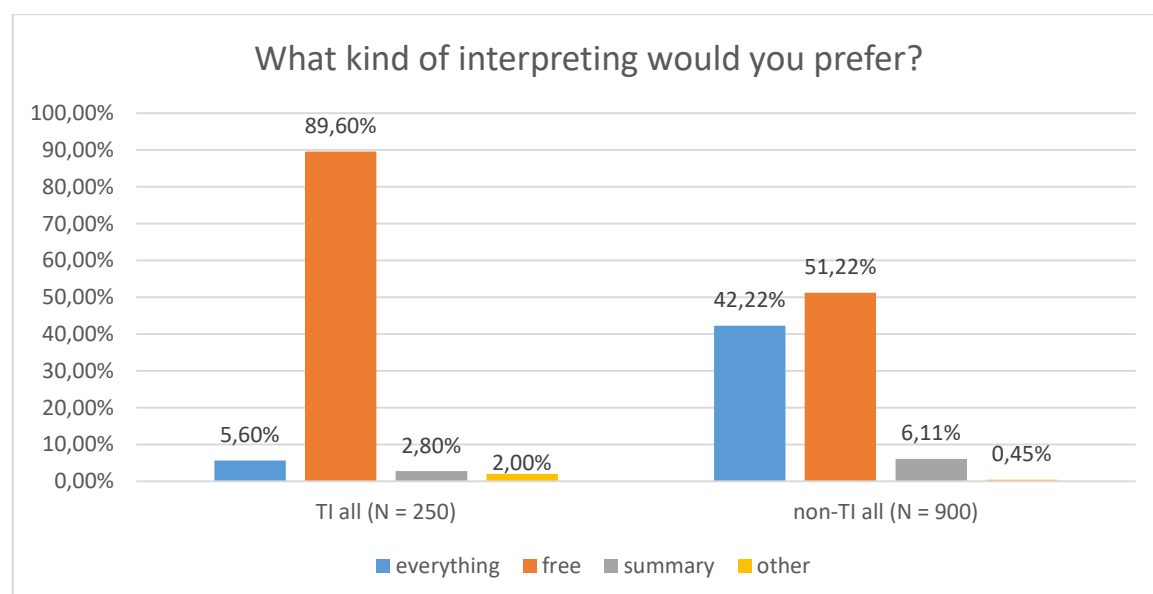
Figure 11: Interpreting styles – TI & non-TI students[49]

The percentages very clearly show the TI students' preference for free interpreting to both the ghost role and summarising (89.60%, 5.60%, and 2.80%, respectively)[50]. Furthermore, while free interpreting is the preferred style of 89.60% of

---

[49] The question for TI students offer the same answers, but it was phrased differently: "How does a good interpreter interpret?".

[50] There were also five "other" answers (TI), four of which were concerned with an accurate transmission of the meaning or all of the important information, while allowing omissions and additions. One respondent (R235, female) also stated that the interpreter should "tone down" (neutralise) expressive words. Four non-TI students chose the "other" answer, explaining that the style of interpreting depended on the type of event and topic while one respondent said she would allow omissions of repeated information, but no additions (R245).

TI students, this is only true for 51.22% of non-TI students. Based on these figures, we **accept** Hypothesis 6.

6.7 Hypothesis 7

H7: There are significant differences among students of various fields, similarly to professionals in Kurz' 1993 study. Some criteria receive more varied answers than other ones.

On the following two pages, we look at the 22 groups of non-TI students created on the basis of their field of study. The values are presented in two tables which use colours to help the reader notice significant differences among them – the red colour indicates a significantly stricter (higher) value, while the green colour indicates a significantly more lenient (lower) value.

Although we only present the most important findings in this subchapter, we also invite the reader to see Appendix C, where each individual criterion is discussed separately and in great detail, just like in Ingrid Kurz' 1993 study.

	non-TI all N = 900	art n = 33	civil engin. n = 31	finance n = 80	foreign lang. n = 32	forestry n = 43	healthc. n = 89	intern. relations n = 46	IT n = 40	journal. n = 23	law n = 52	medic. n = 69
fluency of delivery	3.434	3.424	3.516	3.475	3.563	3.209	3.315	3.587	3.3	3.565	3.519	3.638
native accent	2.123	2.152	2.097	1.938	2.438	2.116	2.135	2.109	2.325	2.043	1.692	1.986
logical cohesion	3.487	3.242	3.323	3.463	3.719	3.163	3.427	3.652	3.325	3.652	3.769	3.696
correct terminology	3.43	3.061	3.323	3.325	3.563	3.395	3.315	3.587	3.225	3.478	3.538	3.594
completeness of delivery	3.327	3.242	3.161	3.35	3.188	3	3.36	3.37	3.375	3.174	3.596	3.435
correct grammar	2.922	3	2.742	2.763	3.219	2.93	2.944	3.065	2.675	3.130	3	2.986
sense consistency	3.528	3.576	2.935	3.513	3.844	3.326	3.281	3.63	3.425	3.739	3.769	3.652
pleasant voice	2.731	2.758	2.645	2.85	2.625	2.651	2.775	2.67	2.9	2.913	2.462	2.507
lively intonation	2.666	2.606	2.419	2.563	2.844	2.605	2.809	2.67	2.675	2.87	2.654	2.594
no filler words & hesitation noises	2.699	2.485	2.581	2.588	2.813	2.674	2.742	2.891	2.5	2.609	2.385	2.768
no booth noises	2.962	2.727	3.226	2.763	2.813	3.023	3.056	3.087	2.575	3.043	2.962	2.913
synchronicity with the speaker	2.946	2.333	3	2.95	2.563	2.977	2.944	3.304	2.8	2.957	3.038	3.130
clear articulation	3.397	3.212	3.194	3.375	3.406	3.209	3.404	3.522	3.4	3.522	3.462	3.464
confident voice	3.226	2.939	3.065	3.213	3.094	3.047	3.337	3.326	3.2	3.391	3.269	3.159
average value	3.063	2.911	2.945	3.009	3.121	2.952	3.060	3.176	2.979	3.149	3.08	3.109

Table 8: Criteria – non-TI students of different study fields, part 1

	non-TI all N = 900	natural sciences n = 27	PE n = 33	pharmacy n = 34	political science n = 30	public admin. n = 27	Slovak n = 33	social work n = 42	special pedag. n = 20	techn. engin. n = 39	tourism n = 33	transport n = 24
fluency of delivery	3.434	3.333	3.364	3.5	3.533	3.148	3.545	3.452	3.55	3.256	3.545	3.208
native accent	2.123	2.074	2.424	2.147	2.233	2.037	2.091	2.523	2.05	2.179	2.303	2.167
logical cohesion	3.487	3.63	3.182	3.647	3.333	3.259	3.636	3.595	3.45	3.308	3.727	3.25
correct terminology	3.43	3.407	3.394	3.5	3.7	3.148	3.485	3.476	3.3	3.641	3.515	3.375
completeness of delivery	3.327	3.296	3.182	3.5	3.3	3	3.394	3.381	3.6	3.179	3.515	3.208
correct grammar	2.922	2.63	2.939	2.853	2.9	2.704	3.242	2.905	3.1	2.641	3.061	2.958
sense consistency	3.528	3.481	3.515	3.765	3.633	3.444	3.576	3.595	3.25	3.692	3.515	3.416
pleasant voice	2.731	2.519	2.97	2.676	2.4	2.63	2.697	3.31	2.75	2.769	2.636	2.917
lively intonation	2.666	2.593	2.879	2.706	2.433	2.259	2.758	2.905	2.45	2.59	2.727	2.667
no filler words & hesitation noises	2.699	2.778	2.818	2.912	2.867	2.074	2.97	2.833	3.1	2.436	3.03	2.5
no booth noises	2.962	2.63	3.273	2.941	3.133	2.481	3.212	3.214	3	2.846	3.394	3
synchronicity with the speaker	2.946	2.963	3.030	2.971	2.733	2.519	3.091	3.214	3.05	2.795	3.212	2.708
clear articulation	3.397	3.185	3.364	3.382	3.567	3.148	3.606	3.429	3.65	3.308	3.545	3.417
confident voice	3.226	3.037	3.455	3.147	3.633	2.815	3.273	3.405	3.6	3.256	3.152	3
average value	3.063	2.968	3.128	3.118	3.1	2.762	3.184	3.231	3.136	2.993	3.206	2.985

Table 9: Criteria – non-TI students of different study fields, part 2

Tables 8 and 9 clearly demonstrate that there are indeed significant differences among students of individual fields. In fact, there were only four groups whose scores did not differ significantly for any of the criteria – *finance/economy/management*, *journalism/media*, *pharmacy*, and *transport/logistics/postal services* students. By far the lowest demands on the interpreter's output were those of **public administration** students; their average importance score given to the 14 criteria was only **2.762**, which is also the only significantly different mean value of all of the 22 groups. On the contrary, the most demanding groups were those studying social work (3.231), tourism (3.206), and the Slovak language (3.184, an average value perfectly matching the average value given by TI students).

Furthermore, some criteria received many more "conflicting" views than others – the criteria with the highest level of disagreement were: **no booth noises** (8 significantly different answers), **no filler words & hesitation noises** (6), **and synchronicity with the speaker** (6). On the other hand, the criterion of **clear articulation** was seen similarly by all study fields. Since both of our presumptions proved to be right, we **accept** Hypothesis 7.

6.8 Hypothesis 8

H8: Non-TI students who have previously experienced interpreting at a conference are more likely to prefer a free rather than a full rendition of the ST, and are more likely to have no gender preference for the interpreter.

On the following page, we present two graphs, both showing the preferences of three groups of non-TI students – those with no, some, or ample experience with interpreted conferences – in relation to the interpreter's role and gender.
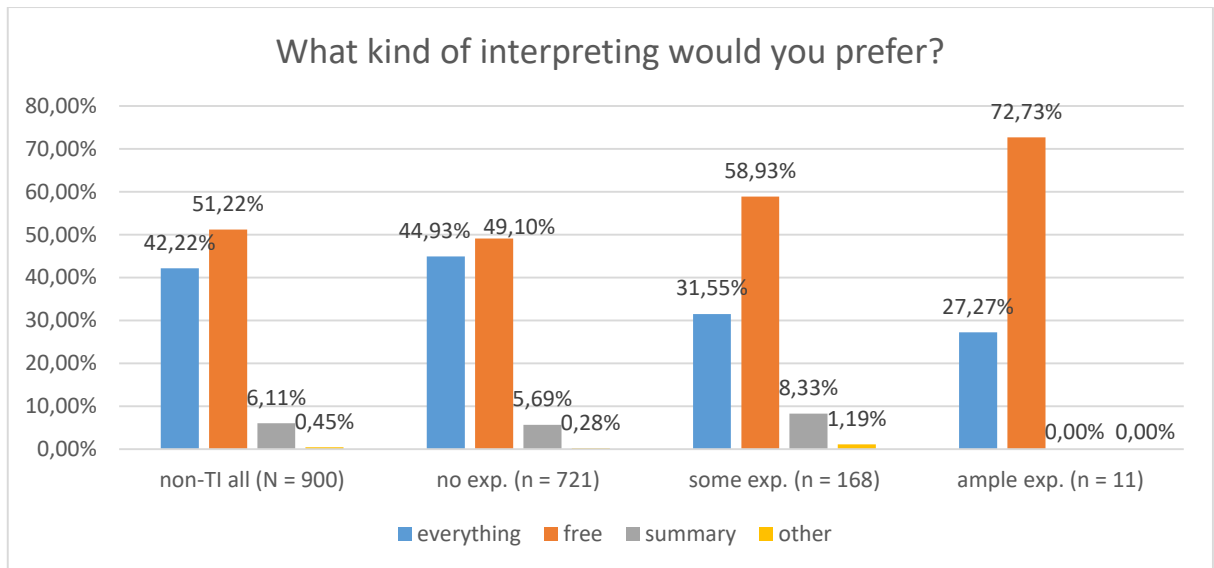
Figure 12: Interpreting styles – non-TI students with various CI experience
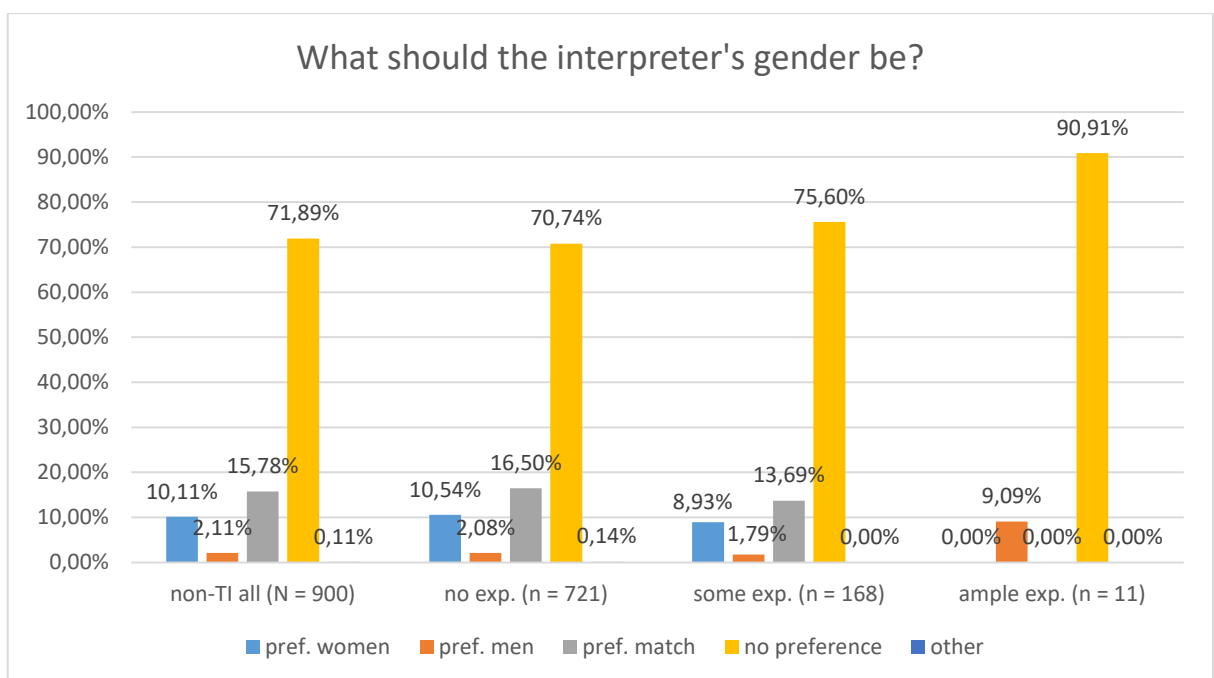
Figure 13: Gender preferences – non-TI students with various CI experience

Figure 12 clearly demonstrates the decreasing preference for the ghost role and the increasing preference for free interpreting with increasing experience with interpreted conferences.

Similarly, the yellow bar in Figure 13, representing no existing gender preference, shows an increasing tendency with more experienced students. While it may look like there is a relatively high preference for male interpreters in the most experienced group (9.09%), this in fact only represents one respondent and it can be

assumed that the ration would decrease significantly with more respondents in this group. Again, we **accept** Hypothesis 8.

When analysing the results, we also noticed another interesting tendency – while the option of the speaker/interpreter gender match was almost equally popular with both men and women, as many as 23.08% of the men stated that they preferred female interpreters (the opposite was only true for 2.78% of women).

6.9 Hypotheses 9 and 10

H9: TI students are stricter assessors of interpreting quality than non-TI students.
H10: The majority of the 20 chosen assessors will give the interpreter an inaccurate overall impression score (significantly different from the calculated FM value).

We have decided to look at Hypotheses 9 and 10 together in one subchapter, because to test them, we will use the same data.

The table on the following page compiles the scores of all 20 assessors, and gives the following information about them:

- study field;
- gender;
- university (TI students only);
- year of study;
- average scores given for delivery-, language-, and content-related criteria;
- overall impression (OI) score;
- $FM_{\%}$ value;
- accuracy of assessment.

No. of assessor (#)	study field	university	gender	year of study	avg. delivery score (1-5)	avg. language score (1-5)	avg. content score (1-5)	overall impression score (1-10)	FM% value (%)	accuracy of assessment
1	TI	MBU	F	5	4.5	3	4	7	76.33	SS
2	TI	MBU	F	4	**4.5**	5	5	9	97.01	SS
3	TI	MBU	F	3	3.7	4	4.5	8	81.89	A
4	TI	MBU	F	2	3.5	3.5	4	6	73.68	S
5	TI	MBU	M	1	4.3	4.5	4	8	85.12	SS
6	TI	CPU	F	5	4.8	5	4	9	91.44	A
7	TI	CPU	F	3	4.4	4.5	4	9	85.72	A
8	TI	CU	M	5	3.6	3	4	8	70.95	L
9	TI	UNIPO	F	5	4.6	3.5	4	9	80.25	L
10	TI	UPJS	F	5	3.9	4.5	4.5	8	86.41	SS

No. of assessor (#)	study field	gender	year of study	avg. delivery score (1-5)	avg. language score (1-5)	avg. content score (1-5)	overall impression score (1-10)	FM% value (%)	accuracy of assessment
11	healthc.	F	3	4.3	4.5	5	8	92.28	S
12	finance	F	5	3.8	3.5	3	7	68.15	A
13	medicine	M	5	3.7	3.5	3	6	67.49	SS
14	law	F	5	3.6	4.5	3	6	73.37	S
15	inter. rel.	F	5	3.4	5	4.5	7	86.61	S
16	forestry	M	5	3.7	3.4	4	8	74.05	SL
17	soc. w.	F	PhD	**4.6**	5	5	9	97.45	SS
18	IT	M	5	**4.5**	5	5	9	96.85	SS
19	tech. en.	M	5	3.9	4	4	8	79.4	A
20	pharm.	F	5	3.3	2.5	2.5	3	54.84	S

Table 10: Assessment of interpreting by TI and non-TI students[51]

[51] Accuracy of assessment abbreviations and colours: A = accurate, SS = slightly strict, S = strict, SL = slightly lenient, L = lenient.

Firstly, we would like to give our own brief evaluation of the interpreting[52]. The interpreter whose recording we chose for the last part of the research, despite omitting substantial amounts of the ST, managed to sustain the logical cohesion of the TT with some terminological inaccuracies[53] and avoided both frequent filler words and hesitation noises. Perhaps her biggest problem was fluency and synchronicity with the speaker, as her *décalage* was relatively long (often, there were pauses longer than three seconds and sometimes as long as eight). Apart from this, we think her output was very satisfactory.

Contrary to our expectations, TI students were actually more lenient than students of other fields. The average OI scores given by these groups were 8.1 and 7.1 and the average $FM_{\%}$ values 82.88% and 79.049%, respectively.

However, although the difference between the OI scores may seem quite large (given that OI was given on a 10-point scale), it is nevertheless statistically insignificant. This was determined through an independent samples t-test with the following data, in which the p-value was found out to be 0.1452[54].

Assessors	N	Mean OI score	Standard Deviation
TI	10	8.1	0.9944
non-TI	10	7.1	1.792

Table 11: Data used in calculating the p-value

Furthermore, if we run the $FM_{\%}$ values through the same test, the p-value is actually much higher (0.4668). Thus, we cannot say that there is a significant difference between the assessments of TI and non-TI students.

Nevertheless, if we combine the facts that although statistically insignificant, the TI students' marks were more favourable, many of them left a comment saying they enjoyed the interpreting, and that there were also two lenient assessors in the TI group, we can say that there seems to be at the very least a tendency for more appreciative marks within the TI sample and we therefore **reject** Hypothesis 9.

---

[52] This is by no means meant as some kind of "ultimate assessment"; rather, we want to enable the reader to get a better idea of the nature of the recording.
[53] See Appendix F for the comparison of ST and TT transcripts.
[54] For a hypothesis to be accepted, the p-value should be 0.05 or less.

Table 10 also gives us interesting results in terms of the assessors' accuracy. Overall, there were only five (25%) accurate assessors[55] – three in the TI and two in the non-TI group. Students tended to give a stricter assessment than their partial scores (for individual criteria) suggested. This was the case of 12 assessors (60%), out of which seven gave a slightly strict OI score (35%) and five gave a strict OI score (25%). On the other hand, there were also three students (15%) who gave the interpreter a favourable overall score which did not quite correspond to their partial scores. Two of them, both lenient, were from the TI group, while one slightly lenient student was also found in the non-TI group. Based on this analysis, we **accept** Hypothesis 10.

In subchapter 1.1.1.2, we expressed our opinion that the interpreter should always be given an opportunity to comment on their own performance, so as to avoid unfavourable assessment in cases where, for example, factors beyond their control influenced their performance.

Therefore, we asked the interpreter in question to comment on her recording. Although we did not require her to give herself marks for the individual criteria, she decided to do so and was subsequently found to be a strict assessor – her average scores for delivery, language, and content were 4.1, 4, and 4, respectively, while her OI score was only 6. According to our formula, her $FM_{\%}$ value was 80.33%, which means that her OI score should have been 8. She also added verbal assessment, which can be found below (as translated by us):

"The interpreting went quite well, I felt rested and, surprisingly, despite not expecting to be interpreting this speech, I did not feel stressed. Perhaps it was because I had already done some warm-up, as we had been interpreting for some time.

"I'm not satisfied with my décalage. I think it was too long, because I did not know much about the topic and I was too afraid to anticipate. This décalage meant that my delivery was not fluent. Sometimes, it is possible to hear agitation in my voice (especially at points where I was lagging too much behind the speaker and I was trying to catch up).

---

[55] I.e. assessors whose OI scores did not differ from their FM values by more than 0.499 point.

"I am satisfied with the low amount of hesitation noises (this used to be a big problem for me).

"I gave myself a score of 6/10, because while being satisfied with the content of the output, its non-fluent character severely affects the overall impression."

6.10 Discussion

In Chapter 5, we formulated ten hypotheses, which were put to a test in this chapter. Now, we would like to comment on the results and try to explain why we think the hypotheses were accepted or rejected. We will also discuss several limitations peculiar to our research.

Our very **first hypothesis** was **rejected**. In many instances, TI students' scores differed significantly from at least one of the groups of interpreters we compared them to, and in two cases (native accent and pleasant voice) they differed from both.

We believe we can readily explain at least one of these differences – that in the perception of native accent. Although native accent is quite a widely discussed concept in the English-speaking world, its significance for Slovak speakers is relatively small. Instances of a non-native Slovak speaker simultaneously interpreting into this language are scarce. Furthermore, it is hard to imagine a Slovak interpreter with a regional accent so strong that it would actually bother the listeners. We think that these facts result in the perception of a native accent in interpreting as a relatively unimportant criterion.

In spite of the differences discovered, we would like to restate the fact that our results confirmed Bühler's (1986) findings in terms of the relative importance of the criteria – TI students ranked them almost identically to her interpreters, with an insignificant difference in the first two criteria. Both of these groups slightly differed from Pöchhacker and Zwischenberger's (2010) interpreters.

When we compared three study groups of non-TI students to Kurz' (1993) doctors, engineers, and politicians, we found similarities only between the last group and students of political science. The other groups differed significantly in both scoring and ranking of the criteria. However, there were almost no significant differences in the perception of the two content-related criteria. Logical cohesion received similar scores from all six groups and the only different score for sense consistency was the one from our group of engineering students. It is important to mention that we do not know what kind of engineers Kurz worked with, and thus we combined our civil and technical

engineering students. However, civil engineering students stood out in our sample for their extremely low score for sense consistency (2.935) and had we considered students of technical engineering only, our score for this criterion would be very similar to Kurz' (3.692 and 3.655, respectively).

The **second hypothesis** was **accepted**, because TI students indeed attributed more importance to the 14 criteria than non-TI students. This finding is hardly surprising, since TI students have to go through years of training during which they are constantly told to work on improving their output and see most of the criteria as important or even very important. Non-TI students only saw two criteria as more important than TI students – pleasant voice and synchronicity with the speaker.

The **third hypothesis** concerned TI students of lower (1st and 2nd) and higher years of study and it was **rejected** because we did not find significant differences between these groups. While several individual criteria received higher or lower scores from some of the groups, their scoring was consistent for the most part. We think the reason behind this is the fact that even 1st and 2nd year students usually have at least some theoretically oriented interpreting lessons, where they learn the basics of good interpreting (such as the "ten commandments" mentioned in previous chapters).

We also had to **reject** the **fourth hypothesis**, because we found no significant differences among students with different outlooks on their future interpreting careers. We believe that the similarity of the scores is caused by the fact that all TI students, whether wanting to be interpreters or not, have to undergo the same basic training, where they learn the same principles of interpreting, which seemingly "stick" with them throughout the whole university at least. Nevertheless, the students considering a career in interpreting were indeed slightly stricter than the other group.

The **fifth hypothesis** was **accepted** because we found no significant differences between TI students from the five Slovak universities with respect to either the criteria or the interpreter's role. Although it may seem as if many (21.05%) UNIPO students preferred the ghost role of the interpreter, this group was very small (19 members) and we assume the ration would decrease with more respondents. We think TI programmes are relatively similar at all Slovak universities which offer them and thus we did not find this result surprising.

We also **accepted** the **sixth hypothesis**, because our figures proved that while the majority of TI students (89.60%) prefer free interpreting, this is only true for 51.22% of non-TI students. The difference between the two samples is remarkable, but not

surprising. We think it stems from the fact that while TI students spend all of their university years learning about the importance of not trying to transfer everything from the ST into the TT in interpreting, but rather focusing on essential information, non-TI students are likely to want a full rendition and see the option of interpreting with additions and omissions as giving the interpreter a free hand in deciding what is necessary for them and what is not.

Here, we arrived at similar results than Pöchhacker and Zwischenberger (2010) in that the readiness to intervene in the ST increased with more interpreting experience (free interpreting was preferred by 88.99% of students with none or lesson-only experience with SI, and 97.62% of more experienced students).

Interestingly, we also discovered that men are more prone to wanting a full rendition of the ST than women, but also more likely to "only" request a summary of what was said (this was also true for TI students).

The **seventh hypothesis** was **accepted**, since we found many differences among non-TI students of various academic disciplines. While the overall values were relatively similar (except in the case of public administration students), partial scores showed significant differences between most of the groups. Furthermore, some criteria (e.g. no booth noises) were more "conflicting" than other ones (e.g. clear articulation). We think that this is a natural result and that the reason behind it is simply the specific nature of each individual study field. It is only logical that different fields have different demands and that some criteria may be more or less generally considered (un)important than others.

Another hypothesis to be **accepted** was the **eight hypothesis**. Our graphs clearly showed that both the preference for the ghost role and any specific gender preference decrease with an increasing experience with interpreted conferences. In our opinion, this stems from the fact that while inexperienced non-TI students may lack the trust in the interpreter's judgment of important information and think that what they need is a full rendition (without realising it is actually impossible to transfer all information), more experienced students have possibly listened to both a ghost interpreter, trying hard to transfer as much information as possible, and a free interpreter, and realised how much more pleasant (and probably more informative) it is to listen to the latter one. Similarly, while inexperienced students may think that the gender of the interpreter may be an important factor (from our own experience, the opinion that mismatching genders of the speaker and the interpreter would be distracting is popular even among

inexperienced TI students), we believe that students with experience realise that its importance in comparison with other factors, is low.

In his user expectation survey, Moser (1995) discovered that users with little to no experience with interpreting preferred concentration on essentials to a full rendition. Our hypothesis thus challenged this finding and, indeed, we arrived at adverse results. However, for a more informed opinion, we would need to know at least the topic of the conferences at which he conducted his research, for in our sample, we also found significant differences in the perception of the interpreter's role among various study groups.

The **ninth hypothesis** was **rejected**. Contrary to our expectations, TI students' assessments were more favourable than those of non-TI students. We think this was probably caused by one of the two following reasons – either it was solidarity of TI students, or the fact that they realise how difficult interpreting actually is and therefore they praised the interpreter for her strong points, rather than criticised her for the weak ones (it could also be a combination of these two reasons). Even though the differences proved to be statistically insignificant, the combined tendencies to give higher scores, be more lenient, and leave a positive comment made us reject Hypothesis 9. It would be very interesting to see how a professional interpreter would evaluate the recording, without knowing the subject was a student interpreter.

On the contrary, we **accepted** the last, **tenth hypothesis**. The majority (75%) of our assessors were indeed inaccurate and most of them gave the interpreter a stricter overall impression score than their partial scores suggested. There is, of course, a possibility that, by coincidence, we managed to choose 15 assessors (75%) which were not good representatives of their study group. Nevertheless, we accept this hypothesis on the basis of our findings, although we agree that it needs more validation to be applicable to a wider population.

Although it would be bold to confidently say that inaccurate assessors were mainly influenced by delivery-related criteria, we believe we can at the very least assume so in the case of the following assessors: #2, #9, #10, #11, #15, #17, and #18, marked with an underscore in Table 10. For the scores written in boldface in the same table, we can actually claim that delivery-related criteria were more important than previously perceived (in the questionnaires), as the interpreter received full marks for the other two categories of criteria.

Interestingly, only three TI (#4, #6, #7) and one non-TI (#12) assessors pointed out that they cannot assess the criterion of sense consistency with the original. Assessors #6 and #7 did not give a score for this category and we therefore counted their average score for content-related criteria on the basis of the only other criterion in this category, logical cohesion. Overall, seven TI students and four non-TI students left a comment on the evaluation form[56].

6.10.1 Research Limitations and Shortcomings

As with every research design, we need to be aware of as well as prepared to admit possible shortcomings of our research. Thus, we shall mention them in this subchapter.

6.10.1.1 Sample Sizes

While both our non-TI and TI samples may seem large at the first look, one must not forget the fact the non-TI sample, when split into 22 smaller samples according to the study field of its members, leaves us with five groups of fewer than 30 students. In the original research, we did not expect to get responses from more than 20 interpreting users per conference on average (and even that aim was perhaps ambitious), but it would nevertheless be preferable to have at least 30, and ideally 40 students in each group, since the findings of this research are already less generalisable to the whole population due to the observed subject being students.

To balance the disproportions, we established several "borders of significance", which proved to be a good idea, judging by the fact that we found neither too many, nor too few significant differences between individual groups. We realise that, apart from the assessors' scores, the data was not put through a statistical test, although we would like to justify that with a reference to other researchers of user expectations, who also did not reach for statistical testing, and also with the fact that we discussed the nature of our research with several statisticians, all of whom advised us to simply work with levels of importance established by ourselves.

6.10.1.2 Criterion of "Completeness of Delivery"

When we included the criterion of completeness of delivery on our list, what we had in mind was completeness in the sense presented by Machová in her 2016 dissertation thesis (i.e. finishing one's sentences). However, after we started personally

---

[56] See Appendix G for the full assessments.

giving out questionnaires, we occasionally overheard groups of students discussing the criteria and some of the conversations suggested that they considered this to be more of a content-related criterion than anything else. Completeness of delivery was most commonly seen as an important criterion (3.344 from TI and 3.327 from non-TI students), but we cannot truly know whether the individual respondents thought of it in the same way as we did. Nonetheless, we decided to treat it as a delivery criterion in the whole thesis.

6.10.1.3 The Choice of Assessors

In subchapter 6.9, we looked at 20 assessments of interpreting and discovered that the majority of our assessors were inaccurate in the context of their study group's previously expressed opinions on the importance of the 14 criteria. However, it is important to note that while all of the assessors had also filled in Questionnaire A or B, we did not know their individual answers, since we had decided to treat them as representatives of their respective study fields. It is entirely possible that some of them may have been exceptions within their group, which is why their $FM_{\%}$ values and their OI scores differed in so many cases. We concluded that users tended to be strict in their assessment, although we do not present this result as generalisable to a wider population without further validation of this finding.

6.10.1.4 The Respondent's Age

Although not a shortcoming, we would like to admit that asking for the respondents' age was unnecessary in the context of this research. As can be seen, for example, in Figure 15 in Appendix B, we had very few non-TI respondents over the age of 24 and the data on the students' year of study was a sufficient substitute. Nevertheless, we do not think that the age question led to large number of uncompleted questionnaires, which means that including it was not harmful for the research.

CONCLUSION

The original aim of this diploma thesis was to replicate user expectation surveys in Slovakia. However, due to factors beyond our control, we had to change the target population and instead of real users and interpreters, we worked with Slovak university students. Although we perceived this as a disappointing turn of events at first, we soon came to the realisation that the student subjects actually made our research unique[57].

Our work on this study included extensive reading of previously published papers on quality in interpreting, user expectations, quality assessment, etc. Through this lengthy preparation phase, we were able to devise the questionnaires used in the research to the best of our abilities and we found them to be a very useful tool.

The cornerstone of our research were the studies of Hildegund Bühler (1986), Ingrid Kurz (1993), Peter Moser (1995), Ángela Collados Aís (1998), and Franz Pöchhacker and Cornelia Zwischenberger (2010). Many of our findings confirm those of these researchers.

Even though our TI students differed from Bühler's (1986) and Pöchhacker and Zwischenberger's (2010) interpreters in terms of the scores they assigned to the individual criteria for quality interpreting, we noticed similar tendencies in the relative importance of these criteria (ranking).

Students of other academic disciplines were similar to Kurz' users (1993) in that they differed from one another in the perception of the criteria's importance. Even though we only found similar scores and rankings between one of the three compared pairs, namely politicians and students of political science, we confirmed Kurz' finding that users in general give lower scores to the criteria than interpreters.

On the contrary, we arrived at adverse results to those of Peter Moser (1995), when we ascertained that the more experience non-TI students had with interpreted conferences, the more likely they were to prefer free interpreting to a full rendition.

Assessment of interpreting constituted only a minor part of our research, but it nevertheless confirmed what Collados Aís discovered as soon as in 1998 – that our conscious perception of quality interpreting and our subconscious or indeed even conscious assessment were perhaps not as connected as we might think. As many as

---

[57] Of course, we cannot claim this with 100% confidence, as other similar studies may have been published for example in languages which we do not understand.

75% of our assessors gave the interpreter an inaccurate score, which is a striking result even in such a small sample (20 members).

We can also confirm the finding of Pöchhacker and Zwischenberger (2010) that more experienced interpreters are also more ready to intervene in the ST, since our results displayed the same situation among students of translation and interpreting.

We would like to mention a few other findings which we find particularly interesting. While in both samples, there were small overall differences in the perception of criterion importance, this was not true in the case of the interpreter's role. We discovered that men are more likely to prefer the ghost role of the interpreter as well as a summary of the ST than women. Furthermore, they are also more likely to have a gender preference for the interpreter.

The differences between students in individual years of study are insignificant, which applies for both samples (in the non-TI sample, the differences are almost non-existent). Furthermore, TI students from different universities were also relatively consistent in their opinions and their desire to become an interpreter in the future also did not result in many differences in their opinions. These findings made our manifold subgroup results applicable to the whole samples of TI or non-TI students.

Non-TI students, although having lower demands on interpreting than TI students, were actually stricter in the assessment of the recording with which we provided them. Although the differences between the two groups proved to be statistically insignificant, we could see a clear tendency for giving a more favourable assessment in the TI students' group.

While the core of our research consisted of ten hypotheses only, our analyses went much further and we would like to remind the reader that there is much more to see in the Appendices, should they be interested.

Although we realise that some of our findings may require further validation, we think that we have nevertheless managed to uncover some very interesting opinions of Slovak students, who could one day become our listeners at various conferences. Therefore, we believe that our results could be valuable to future researchers and we would like to suggest a few ideas for further research. We also think the findings are relevant to anyone interested in the topic of user expectations in Slovakia, for at the moment, this thesis is the only study of its kind in the country.

Suggestions for Future Research

We hope that our thesis will serve as an inspiration for other TI students deciding on the topic of their final theses, or indeed, even for translation scholars.

As the reader probably realises, this research was very extensive and in fact, we think it could be split into two or even three separate, but connected studies – one on TI students' opinions on quality in interpreting, one on non-TI students' expectations and preferences, and one on assessment of interpreting quality.

It could, for example, be useful to extend any part of the research in terms of the sample sizes. Should someone wish to do so, we would be more than happy to provide them with our raw data, which they could add to, or perhaps compare with their own data.

Furthermore, the target groups could be changed to include real Slovak interpreters and Slovak professionals in various fields, which would definitely provide us with further informative data. The formula for determining the accuracy of assessors developed for this thesis could be used in further analyses of whether the respondents stay true to their conscious opinions, or whether they are subconsciously influenced by criteria they see as inferior. It can be easily adjusted to a different number or structure of criteria.

Lastly, during the analysis of our data, we discovered that there are seemingly significant differences between men and women in terms of their preference for either the gender of the interpreter, or the interpreter's role. Therefore, we believe a gender study in this direction could lead to many interesting findings.

The topic of interpreting quality is extremely wide and the possibilities for research are endless. Thus, whatever the eager researcher chooses to observe, they will definitely not struggle with a lack of material or phenomena.

RESUMÉ

Téma kvality rezonuje vo vede o tlmočení už niekoľko desaťročí, no napriek nemalému záujmu translatológov i odborníkov z iných disciplín dodnes neexistuje všeobecne uznávaná definícia tohto fenoménu. Zdá sa, že rôznorodosť názorov pritom vplýva zo samotnej podstaty tlmočenia ako neopakovateľnej a nepostihnuteľnej interlingválnej a interkultúrnej činnosti.

Diplomová práca *Vnímanie kvality tlmočenia slovenskými študentmi* sa zaoberá názormi slovenských vysokoškolských študentov odboru prekladateľstvo a tlmočníctvo ako aj študentov iných odborov na kvalitu v tlmočení. Primárnou motiváciou na spracovanie tejto témy bol náš záujem o ňu, ako i fakt, že podobná štúdia stále absentuje v slovenskej vede o tlmočení.

Cieľom našej práce bolo zistiť názory študentov slovenských univerzít na kvalitné tlmočenie, analyzovať ich s ohľadom na rôzne socio-demografické údaje respondentov, ale i porovnať naše zistenia so zisteniami zahraničných štúdií zaoberajúcich sa fenoménom užívateľských očakávaní v tlmočení. Keďže prekladaná diplomová práca je prvá svojho charakteru na Slovensku, dúfame, že ňou otvoríme brány výskumu kvality tlmočenia, ktorému sa u nás doposiaľ venovala len malá pozornosť.

Práca je rozdelená na šesť kapitol. V prvej kapitole vymedzujeme hlavné prístupy k tlmočeniu, a síce tlmočenie ako produkt, proces a služba. Domnievame sa, že každý z nich má svoje pozitíva a negatíva – kým hodnotenie tlmočenia ako produktu sa zaoberá len cieľovým textom a ignoruje neverbálne prvky, ako aj okolnosti tlmočenia, náhľad tlmočenie ako na proces môže viesť k nedokonalému posúdeniu napríklad významovej zhody medzi pôvodným a pretlmočeným prejavom. V práci preto navrhujeme spojenie týchto dvoch prístupov, ktoré vedie ku komplexnému hodnoteniu tlmočníckeho výkonu v spolupráci so samotným tlmočníkom. Tlmočenie ako služba sa nám javí ako zaujímavý, hoci v slovenskom kontexte zatiaľ neveľmi rozšírený prístup. Pre jeho holistickú povahu ho však vo svojom výskume nevyužívame.

Prvá kapitola tiež približuje čitateľovi šesť hlavných perspektív hodnotiteľov tlmočenia – perspektívu rečníka, používateľov, klienta, tlmočníka, tlmočníkovho kolegu a výskumníka. Pri každej z nich sa uvádzajú možnosti i obmedzenia hodnotenia (napríklad problémy zhodnotenia významovej zhody používateľmi tlmočenia, ktorí počujú len pretlmočený prejav). V poslednej podkapitole preto navrhujeme pracovať

s viacerými perspektívami naraz, pričom zdôrazňujeme, že tlmočníkovi by vždy malo byť umožnené vyjadriť sa k vlastnému výkonu a pracovným podmienkam, aby sa predišlo nespravodlivému hodnoteniu v prípadoch, keď tlmočník nemohol za zhoršenú kvalitu svojej práce.

Druhá kapitola je zhrnutím zahraničného i domáceho výskumu kvality tlmočenia. Ten siaha do obdobia na prelome šesťdesiatych a sedemdesiatych rokov minulého storočia a začína štúdiami psychológa Henriho Barika, ktorý na základe šiestich nahrávok tlmočenia vytvoril schému odchýlok od pôvodného textu – omisií, adícií, substitúcií a chýb – ďalej rozčlenených podľa pôvodu a stupňa závažnosti (1971). Fakt, že autor klasifikoval každú vynechanú informáciu ako negatívny posun v cieľovom texte, považujeme za nedostatok Barikovej štúdie. Ako totiž poznamenala Stenzlová (1983), tlmočený text s niekoľkými (nezávažnými) omisiami môže byť pre poslucháča užitočnejší než taký, ktorý sa snaží o úplnosť na úkor zrozumiteľnosti.

Štúdia Hildegund Bühlerovej z roku 1986 je základným kameňom výskumu očakávaní používateľov tlmočenia, napriek tomu, že Bühlerová sa v skutočnosti zaoberala názormi tlmočníkov. Jej zoznam 16 kritérií dobrého tlmočenia hodnotilo 47 tlmočníkov z Medzinárodnej asociácie konferenčných tlmočníkov (AIIC) na štvorbodovej škále (irelevantné – veľmi dôležité). Autorka zistila, že za najdôležitejšie boli považované kritériá významovej zhody (3,957), logickej kohézie (3,8) a správnej terminológie (3,489) a vyjadrila názor, že nároky tlmočníkov korešpondujú s nárokmi používateľov tlmočenia. (Bühlerová, 1986)

Toto tvrdenie sa o niekoľko rokov neskôr rozhodla podrobiť skúške Ingrid Kurzová, ktorá napísala niekoľko štúdií týkajúcich sa používateľských očakávaní. V našej práci sme pracovali najmä s jej štúdiou z roku 1993, v ktorej nielen dokázala, že nároky používateľov sú nižšie ako nároky tlmočníkov, ale zistila i to, že používatelia na rôznych konferenciách (lekárskej, technickej a politickej) sa od seba líšili i navzájom. (Kurzová, 1993)

Podobné výskumy uskutočnili aj Peter Moser (1995, používatelia) a tandem Franz Pöchhacker a Cornelia Zwischenbergerová (2010, tlmočníci). Experimentálne štúdie Ángely Collados Aís (1998) ako aj jej kolegýň z Granadskej univerzity (2007) však spochybňujú relevantnosť kritérií, ktoré používatelia považujú za dôležité. Autorky totiž zistili, že nesplnenie kritérií označených za menej dôležité (napr. živá intonácia) viedlo k výrazne horšiemu hodnoteniu inak bezchybného tlmočenia (Collados Aís a kol., 2007).

Slovenský výskum kvality tlmočenia nie je zatiaľ veľmi obsiahly, a preto pri jeho sumarizácii spomíname aj štúdie týkajúce sa konzekutívneho a komunitárneho tlmočenia. Za najrelevantnejšiu pre náš výskum považujeme dizertačnú prácu Lýdie Machovej (2016), ktorej cieľom bolo vytvoriť hodnotiaci formulár pre tlmočníkov ako nástroj sebareflexie. Z Machovej práce je pre nás dôležité najmä jej členenie kritérií do troch kategórií – prednes, jazyk a obsah. S týmto členením neskôr pracujeme vo výskumnej časti našej diplomovej práce. Na formulári oceňujeme i to, že dáva tlmočníkom príležitosť vyjadriť sa k procesuálnym záležitostiam, akými sú napríklad pracovné podmienky či ich vlastný psychický a fyzický stav v čase tlmočenia. (Machová, 2016)

Tretia kapitola našej práce je užšie zameraná na očakávania používateľov a venujeme sa v nej najmä obmedzeniam tohto typu výskumu. Bližšie hovoríme o už spomínanom probléme hodnotenia významovej zhody používateľmi, ako aj o otázke nezhody očakávaní a hodnotenia, ktorej sa venovali výskumníčky z univerzity v Granade (Collados Aís a kol., 2007). Posledným problémom je neochota používateľov spolupracovať s výskumníkom (dotazníkové prieskumy sú známe nízkym počtom odpovedí) a fakt, že používatelia často nemajú záujem o celý tlmočený prejav, no zaujíma ich, napríklad, len určitá jeho časť, na základe ktorej potom hodnotia celkový výkon tlmočníka.

Vo štvrtej kapitole podrobne opisujeme metódy, ktoré sme použili vo výskumnej časti práce. Išlo najmä o kvantitatívny dotazníkový prieskum, pre ktorý sme vytvorili dva dotazníky – jeden pre študentov odboru prekladateľstvo a tlmočníctvo a druhý pre študentov iných odborov (netlmočníkov). Dotazníky sa od seba líšili len niekoľkými otázkami (v závislosti od skupiny nás zaujímali iné socio-demografické údaje, netlmočníkov sme sa pýtali na preferencie tlmočníkovho pohlavia a študentov tlmočníctva na závislosť spomínaných kritérií od typu tlmočenej konferencie).

Hlavné boli pre nás dve časti oboch dotazníkov – názory študentov na „štýl tlmočenia" (prenesenie všetkých informácií, voľné tlmočenie s možnými adíciami a omisiami, sumarizácia) a zhodnotenie dôležitosti 14 kritérií: významová zhoda a logickosť (kategória obsahu), správna terminológia a správna gramatika (kategória jazyka), príjemný hlas, plynulosť prejavu, úplnosť prejavu, príjemný hlas, živá intonácia, nepoužívanie výplnkových slov a hezitačných zvukov, nerušenie tlmočenia zvukmi z kabíny, synchrónnosť s rečníkom, čistá artikulácia a istota v hlase (kategória prednesu). Prvých osem vymenovaných kritérií pochádza zo štúdie Ingrid Kurzovej

(1993), kým zvyšných šesť sme pridali na základe naštudovanej literatúry. Respondenti tieto kritériá hodnotili na štvorbodovej škále, čím sme zaistili kompatibilitu našich výsledkov s predošlými štúdiami.

Dotazníky boli primárne distribuované cez internet, pričom s výskumom nám pomohli aj mnohí vyučujúci zo slovenských univerzít. Najviac odpovedí sme získali osobne, návštevou niekoľkých univerzít v Banskej Bystrici, Zvolene, Martine, Žiline a Bratislave. Do výskumu sa celkovo zapojilo 250 študentov prekladateľstva a tlmočníctva a 900 študentov iných odborov.

Obe vzorky sa vyznačovali relatívne dobrou reprezentatívnosťou s ohľadom na pohlavie respondentov. Hoci to isté nemožno povedať o zastúpení jednotlivých ročníkov, môžeme konštatovať, že všetky ročníky boli zastúpené dostatočným počtom respondentov, a keďže sme medzi jednotlivými ročníkmi nenašli významné rozdiely, môžeme výsledky generalizovať na celú vzorku. Študentov-netlmočníkov sme ďalej rozdelili do 22 skupín podľa ich študijného odboru, pričom ako spodnú hranicu analyzovateľnosti výsledkov sme určili 20 študentov. Určili sme si tiež hranice významnosti výsledkov ako kompenzáciu rozličných veľkostí analyzovaných skupín.

Súčasťou nášho výskumu bolo aj hodnotenie nahrávky tlmočenia, ktorú sme získali od študentky-tlmočníčky z Univerzity Mateja Bela. Na túto časť výskumu sme vybrali desať študentov odboru prekladateľstvo a tlmočníctvo a desať študentov-netlmočníkov, po jednom z desiatich najvyššie zastúpených študijných odborov. Na analýzu výsledkov sme vyvinuli matematickú rovnicu, ktorá nám umožnila posúdiť, či sa názory študentov o dôležitosti kritérií premietli do ich hodnotenia.

Piata kapitola je zoznamom desiatich hypotéz, ktoré sme sformulovali v súlade s cieľom diplomovej práce a testovali v poslednej, šiestej kapitole. Výsledky nášho výskumu prezentujeme v nasledujúcich odstavcoch.

Názory študentov tlmočníctva a prekladateľstva sa líšia od názorov tlmočníkov v Bühlerovej štúdii (1986), ako aj v štúdii Pöchhackera a Zwischenbergerovej (2010). Hoci sme zistili podobné tendencie v hodnotení pomernej dôležitosti kritérií u všetkých troch skupín, študenti vnímali uvedených kritériá ako omnoho menej dôležité než tlmočníci. Opačná situácia nastala v skupine študentov-netlmočníkov, ktorí v porovnaní s Kurzovej (1993) skupinami vnímali dané kritériá ako dôležitejšie. Zhodu názorov sme zistili len medzi skupinou politikov a študentov politológie. Potvrdili sme však Kurzovej zistenie, že používatelia tlmočenia vo všeobecnosti hodnotia kritériá kvalitného tlmočenia ako menej dôležité než tlmočníci.

Neočakávaným výsledkom boli konzistentné názory skupín študentov-tlmočníkov z rôznych ročníkov či stupňov štúdia, ako aj tých, ktorí uviedli, že by v budúcnosti radi či neradi robili prácu tlmočníka. Žiadne významné rozdiely neboli zistené ani medzi študentmi-tlmočníkmi z rôznych slovenských univerzít vyučujúcich tento odbor.

Naša hypotéza, že študenti-tlmočníci uprednostňujú voľné tlmočenie (teda tlmočenie s možnými omisiami a adíciami) sa potvrdila – túto možnosť zvolilo takmer 90 % študentov. Vo vzorke študentov-netlmočníkov však voľné tlmočenie preferovalo len 51,22 % študentov. Zistili sme tiež, že u mužov je vyššia pravdepodobnosť než u žien, že budú požadovať buď úplné pretlmočenie pôvodného prejavu (bez omisií a adícií), alebo len jeho sumarizáciu.

Skúmali sme i názory študentov-netlmočníkov s rozličnými skúsenosťami s tlmočenými konferenciami. Zistili sme, že čím väčšie množstvo skúseností títo študenti mali, tým vyššia bola ich preferencia voľného tlmočenia a znižoval sa i počet respondentov s vyhradenou preferenciou tlmočníkovho pohlavia. Zaujímavým zistením bol i fakt, že omnoho viac mužov malo existujúce preferencie pohlavia tlmočníka – až vyše 23 % mužov v dotazníku uviedlo, že uprednostňuje tlmočníčky, kým u žien boli takéto preferencie zanedbateľné (menej ako 3 % v prípade tlmočníkov i tlmočníčok).

Posledné dve hypotézy sa týkali hodnotenia nahrávky tlmočenia. Predpokladali sme, že študenti tlmočníctva a prekladateľstva budú prísnejšími hodnotiteľmi než študenti-netlmočníci. Táto hypotéza sa však nepotvrdila. Priemerný počet bodov v prvom prípade bol 8,1/10, kým druhá skupina dala tlmočníčke v priemere o jeden bod menej (7,1). Hoci štatistický test nepotvrdil relevantnosť tohto výsledku, študentov-tlmočníkov považujeme za miernejších hodnotiteľov aj na základe ich často kladných komentárov pridaných k hodnotiacemu formuláru.

Ako sme predpokladali, väčšina hodnotiacich sa nepridŕžala dôležitosti kritérií uvedenej v dotazníku. Až 12 študentov hodnotilo nahrávku príliš prísne, traja študenti boli zhovievaví a len päť hodnotení zodpovedalo spomenutej dôležitosti kritérií. Naša posledná hypotéza sa teda potvrdila. Uvedomujeme si však možnosť, že sme zvolili nesprávnych zástupcov jednotlivých odborov. Pri výpočtoch v tejto časti sme totiž počítali s priemernými hodnotami kritérií daných jednotlivými študijnými skupinami, ktorých hodnotitelia zastupovali, a nie hodnotiteľmi samotnými. Tento výsledok si teda nutne vyžaduje ďalšiu validáciu.

Hoci náš pôvodný výskum počítal s reálnymi tlmočníkmi a odborníkmi v jednotlivých odboroch, z dôvodov, ktoré sme nemohli ovplyvniť, sme napokon museli spolupracovať s vysokoškolskými študentmi. Nemyslíme si však, že to znižuje hodnotu tejto práce, ktorá je napriek zmenenému subjektom naďalej jediná svojho druhu na Slovensku. Čo viac, práca so študentmi nám umožnila získať omnoho viac odpovedí než s koľkými počítal pôvodný výskum, a jej výsledky môžu byť užitočné nielen pre budúcich bádateľov, ale aj pre tlmočníkov samotných, pretože poskytujú prehľad názorov na kvalitu tlmočenia, aký sme v slovenskom kontexte doposiaľ nemali. Študenti, ktorí sa na našom výskume podieľali, sa možno o niekoľko rokov zaradia medzi našich poslucháčov na rozličných konferenciách, a je preto užitočné vedieť, čo od nás, jazykových a kultúrnych sprostredkovateľov, očakávajú.

Výskum kvality v tlmočení je širokým spektrom mnohých fenoménov, čo dokazuje aj táto práca. V jej závere preto navrhujeme smery, ktorými by sa mohli uberať budúci výskumníci. Domnievame sa, že by bolo užitočné rozšíriť našu vzorku alebo uskutočniť podobný prieskum medzi reálnymi tlmočníkmi a odborníkmi z rozličných odborov. Prínosným by bol i výskum hodnotenia kvality tlmočenia rozšírený o viacerých respondentov. Ako sme spomenuli na predchádzajúcich stranách, počas analýzy našich vlastných výsledkov sme zistili niekoľko zaujímavých rozdielov medzi mužmi a ženami, ktoré by bolo vhodné preskúmať vo väčšom detaile. Nech už sa však budúci bádatelia vyberú akýmkoľvek smerom, sme si istí, že ich práca bude veľkým prínosom, pretože bielych miest na mape slovenského výskumu kvality tlmočenia je naozaj mnoho.

BIBLIOGRAPHY

AMINI, M. – AYOB, L. – IBRAHIM-GONZÁLEZ, N. 2013. Quality of Interpreting from Users' Perspectives. [online]. In *International Journal of English and Education*. ISSN 2278-4012, vol. 2, iss. 1, p. 89-98. [accessed 2017-07-23]. Available at: <https://www.academia.edu/21207271/QUALITY_OF_INTERPRETING_FROM_U SERS_PERSPECTIVES>

ANDERSON, R. B. W. 1976. Perspectives on The Role of Interpreter. In *The Interpreting Studies Reader*. London : Routledge, 2002, p. 208-217. ISBN 0-415-22477-2.

BARIK, H. C. 1971. *A Description of Various Types of Additions and Errors of Translation Encountered in Simultaneous Interpretation.* [online]. ISSN 1492-1421, 1971, vol. 16, iss. 4, 199-210. [accessed 2017-08-08]. Available at: <https://www.erudit.org/fr/revues/meta/1971-v16-n4-meta254/001972ar/>.

BARIK, H. C. 1975. Simultaneous Interpretation: Qualitative and Linguistic Data. In *The Interpreting Studies Reader*. London : Routledge, 2002, p. 78-91. ISBN 0-415-22477-2.

BERK-SELIGSON, S. 1988. The Impact of Politeness in Witness Testimony: The Influence of The Court Interpreter. In *The Interpreting Studies Reader*. London : Routledge, 2002, p. 278-292. ISBN 0-415-22477-2.

BITTNEROVÁ, Z. 2012. Verbálne protokoly v svetle interdisciplinárneho prístupu vo výskume simultánneho tlmočenia. In *Kontexty súdneho prekladu a tlmočenia I*. Bratislava : Vydavateľstvo UK, 2012, 12-27. ISBN 978-80-223-3293-4.

BOHUŠOVÁ, Z. 2009. *Neutralizácia ako kognitívna stratégia v transkultúrnej komunikácii. Lingvistické analógie.* Banská Bystrica : DALI-BB, 2009. 146 p. ISBN 798-80-89090-55-6.

BOHUŠOVÁ, Z. 2015. Mocenské asymetrie pri tlmočení. In *Tlmočenie v interdisciplinárnej perspektíve so zameraním na osobnostné charakteristiky a kognitívne aspekty*. Banská Bystrica : Belianum, 2015, p. 120-130. ISBN 978-80-557-0916-1.

BÜHLER, H. 1986. Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters. In *Multilingua*. 1986, vol. 5, iss. 4, p. 231-235.

CHESTERMAN, W. – WILLIAMS, J. 2002. *The Map: A Beginner's Guide to Doing Research in Translation Studies.* Manchester : St. Jerome Publishing, 2002. 149 p. ISBN 1-900650-54-1.

CHEUNG, A. K. F. 2015. Scapegoating the interpreter for listeners' dissatisfaction with their level of understanding. [online]. *In Interpreting*. ISSN 1384-6647, 2015, vol. 17, iss. 1, p. 46-63. [accessed 2018-01-20]. Available at: <https://benjamins.com/#catalog/journals/intp.17.1.03che/fulltext>.

CHIARO, D. – NOCELLA, G. 2004. Interpreters' Perception of Linguistic and Non-Linguistic Factors Affecting Quality: A Survey through the World Wide Web. [online]. In *Meta*. ISSN 1492-1421, 2004, vol. 49, iss. 2, p. 278-293. [accessed 2017-08-19]. Available at: <https://www.erudit.org/en/journals/meta/2004-v49-n2-meta770/009351ar/>.

CLIFFORD, A. 2001. Discourse Theory and Performance-Based Assessment: Two Tools for Professional Interpreting. [online]. In *Meta*. ISSN 1492-1421, 2001, vol. 46, iss. 2, p. 365-378. [accessed 2017-08-02]. Available at: <https://www.erudit.org/en/journals/meta/2001-v46-n2-meta159/002345ar.pdf>

COLLADOS AÍS, Á. 1998. Quality Assessment in Simultaneous Interpreting: The Importance of Nonverbal Communication [Translated by Pamela Faber]. In *The Interpreting Studies Reader*. London : Routledge, 2002, p. 326-336. ISBN 0-415-22477-2.

COLLADOS AÍS, Á. – GARCÍA BECERRA, O. – PRADAS MACÍAS, E. M. – STÉVAUX, E. 2007. *La Evaluación de la calidad en interpretación simultánea: Parámetros de incidencia.* Granada : Interlingua, 2007. 279 p. ISBN 978-8498360578.

COLLADOS AÍS, Á. – GARCÍA BECERRA, O. 2015. Quality. In *The Routledge Handbook of Interpreting*. London : Routledge, 2015, p. 368-383. ISBN 978-0-415-81166-8.

ČEŇKOVÁ, I. 2012. Kdo určuje kvality komunitního tlumočení? In *Preklad a kultúra 4*. Nitra : Univerzita Konštantína Filozofa, Filozofická fakulta, 2012, p. 63-75. ISBN 978-80-558-0143-8.

DE WIT, M. – SLUIS, I. 2014. Sign language interpreter quality: the perspective of deaf sign language users in the Netherlands. [online]. In *The Interpreters' Newsletter*. ISBN 1591-4127, 2014, vol. 19, p. 63-85. [accessed 2017-11-12]. Available at: <https://www.openstarts.units.it/bitstream/10077/10650/1/deWit-Sluis_19.pdf>

DJOVČOŠ, M. – ŠELENGOVÁ, D. 2015. Potenciálny vzťah medzi tlmočníckym výkonom a závislosťou/nezávislosťou od poľa. In *Tlmočenie v interdisciplinárnej perspektíve so zameraním na osobnostné charakteristiky a kognitívne aspekty*. Banská Bystrica : Belianum, 2015, p. 72-92. ISBN 978-80-557-0916-1.

DJOVČOŠ, M. – ŠVEDA, P. 2017. *Mýty a fakty o preklade a tlmočení na Slovensku*. Bratislava : Veda, 2017. 206 p. ISBN 978-80-224-1566-8.

DIRIKER, E. 2015. Conference Interpreting. In *The Routledge Handbook of Interpreting*. London : Routledge, 2015, p. 171-185. ISBN 978-0-415-81166-8.

DOWNIE, J. 2016. *Being A Successful Interpreter: Adding Value and Delivering Excellence*. London : Routledge, 2016. 109 p. ISBN 978-1-138-11969-7.

GARZONE, G. 2002. Quality and Norms in Interpretation. [online]. In *Interpreting in the 21st Century: Challenges and Opportunities*. Amsterdam/Philadephia : John Benjamins Publishing Company, 2002, p. 107-119. ISBN 9789027216496. [accessed 2017-07-18]. Available on the Internet: <https://air.unimi.it/bitstream/2434/197078/2/Garzone.doc>

GERVER, D. 1969. The Effects of Source Language Presentation Rate on The Performance of Simultaneous Conference Interpreters. In *The Interpreting Studies Reader*. London : Routledge, 2002, p. 52-67. ISBN 0-415-22477-2.

GILE, D. 1990. L'évaluation de la qualité de l'interprétation pa les délégués: une étude de cas. [online]. In *The Interpreters' Newsletter*. ISSN 1591-4127, 1990, vol. 3, p. 66-71. [accessed 2017-08-22]. Available at: <https://www.researchgate.net/publication/277828816_L'evaluation_de_la_qualite_de_l'interpretation_pa_les_delegues_une_etude_de_cas>

GILE, D. 1997. Conference Interpreting as A Cognitive Management Problem. In *The Interpreting Studies Reader*. London : Routledge, 2002, p. 162-177. ISBN 0-415-22477-2.

GILE, D. 2009. *Basic Concepts and Models for Interpreter and Translator Training (Revised edition)*. Amsterdam/Philadelphia : John Benjamins Publishing Company, 2009. 283 p. ISBN 978 90 272 2433 0. [accessed 2018-02-12]. Available at: <https://books.google.sk/books?id=YfHEXethAN4C&source=gbs_navlinks_s>.

GILE, D. – HANSEN, G. – POKORN, N. K. (eds.) 2010. *Why Translation Studies Matters*. [online] Amsterdam; Philadelphia : John Benjamins Publishing Company, 2010. 269 p. ISBN 978 90 272 8864 6. [accessed 2017-08-07]. Available at: <https://books.google.sk/books/about/Why_Translation_Studies_Matters.html?id=6e GPANEvSmUC&redir_esc=y>.

GULDANOVÁ, Z. 2012. Zadávateľ a jeho postavenie v translačnom procese z pohľadu súdneho tlmočenia/prekladu. In *Kontexty súdneho prekladu a tlmočenia I*. Bratislava : Vydavateľstvo UK, 2012, 51-59. ISBN 978-80-223-3293-4.

HEINZOVÁ, Z. – NÁBĚLKOVÁ, E. 2015. Tlmočnícky výkon v kontexte vybraných psychologických charakteristík tlmočníka. In *Tlmočenie v interdisciplinárnej perspektíve so zameraním na osobnostné charakteristiky a kognitívne aspekty*. Banská Bystrica : Belianum, 2015, p. 26-40. ISBN 978-80-557-0916-1.

KAHANE, E. 2000. *Thoughts on The Quality of Interpretation.* [online]. AIIC.NET, 2000. [accessed 2017-06-29]. Available at: <https://aiic.net/page/197/>.

KALINA, S. 2002. Quality in interpreting and its prerequisites: A framework for a comprehensive view. [online]. In *Interpreting in the 21st Century: Challenges and Opportunities*. Amsterdam/Philadelphia : John Benjamins Publishing Company, 2002. ISBN 9789027216496, p. 121-130. [accessed 2017-08-18]. Available at: <https://books.google.sk/books/about/Interpreting_in_the_21st_Century.html?id=Ugs bX4OqPzkC&redir_esc=y>.

KALINA, S. 2005. Quality Assurance for Interpreting Processes. [online]. In *Meta*. ISSN 1492-1421, 2005, vol. 50, iss. 2, p. 768-784. [accessed 2017-07-12]. Available at: <https://www.erudit.org/fr/revues/meta/2005-v50-n2-meta881/011017ar.pdf>.

KORPAL, P. 2012. *Omissions in simultaneous interpreting as a deliberate act.* [online]. Adam Mickiewicz University, 2012. [accessed 2017-08-24]. Available at: <https://www.researchgate.net/publication/297474647_Omission_in_simultaneous_interpreting_as_a_deliberate_act>.

KURZ, I. 1993. Conference Interpretation: Expectations of Different User Groups. In *The Interpreting Studies Reader*. London : Routledge, 2002, p. 312-325. ISBN 0-415-22477-2.

KURZ, I. 2001. Conference Interpreting: Quality in the Ears of the User. [online]. In *Meta*. ISSN 1492-1421, 2001, vol. 46, iss. 2, p. 394-409. [accessed 2017-08-16]. Available at: <https://www.erudit.org/fr/revues/meta/2001-v46-n2-meta159/003364ar.pdf>.

KURZ, I. - PÖCHHACKER, F. – ZWISCHENBERGER, C. 2008. *Quality and role: the professionals' view*. [online]. AIIC.NET, 2008. [accessed 2017-08-18]. Available at: <https://aiic.net/page/3044/quality-and-role-the-professionals-view/lang/1>

MACHOVÁ, L. 2016. *Hodnotiaci formulár ako nástroj tlmočníckej sebareflexie u študentov: didaktické východiská* (Doctoral dissertation). Comenius University in Bratislava, 2016. 168 p.

MAKAROVÁ, V. 2004. *Tlmočenie: Hraničná oblasť medzi vedou, skúsenosťou a umením možného.* Bratislava : STIMUL, 2004. 197 p. ISBN 80-88982-81-2.

MARRONE, S. 1993. Quality: A Shared Objective. [online]. In *The Interpreters' Newsletter*. ISSN 1591-4127, 1993, vol. 5, p. 35-41. [accessed 2017-09-03]. Available at: <https://www.openstarts.units.it/bitstream/10077/14664/1/MarroneIN5.pdf>

MARTÍNEZ-GÓMEZ, A. 2015. Non-Professional Interpreters. In *The Routledge Handbook of Interpreting*. London : Routledge, 2015, p. 417-431. ISBN 978-0-415-81166-8.

MEAK, L. 1990. Interprétation simultanée et congrès médical: attentes et commentaires. [online]. In *The Interpreters' Newsletter*. ISSN 1591-4127, 1990, vol. 3, p. 9-13. [accessed 2018-02-10]. Available at: <https://www.openstarts.units.it/bitstream/10077/2148/1/03.pdf>.

MELICHERČÍKOVÁ, M. 2016a. *Súvislosti medzi tlmočníckym výkonom študentov a vybranými kognitívnymi charakteristikami* (Doctoral dissertation). Matej Bel University in Banská Bystrica, 2016. 163 p. 9be84fb6-e3af-42a3-84e8-a07527f6d551.

MELICHERČÍKOVÁ, M. 2016b. Obsahuje profil začínajúcich študentov OPT znaky hybridnosti? In *Preklad a tlmočenie 12: Hybridita a kreolizácia v preklade a translatológii*. Banská Bystrica : Belianum, 2016, p. 283-298. ISBN 978-80-557-1204-8.

MELICHERČÍKOVÁ, M. 2017. *Kognitívne charakteristiky a tlmočnícky výkon: „Súvisia spolu?"*. Banská Bystrica : Belianum, 2017. 186 p. ISBN 978-80-557-1290-1.

MICHALČÍKOVÁ, J. 2015. *K vybraným aspektom konzekutívneho tlmočenia s notáciou: Metodologické aspekty evaluácie výstupov v dimenzii konzekutívneho tlmočenia.* Banská Bystrica : Belianum, 2015. 120 p. ISBN 978-80-557-0894-2.

MOSER, P. 1995. *Survey on Expectations of Users of Conference Interpretation.* [online]. AIIC.NET, 1995. 50 p. [accessed 2017-08-13]. Available at: <https://aiic.net/page/attachment/1044>.

MOSER-MERCER, B. 1996. Quality in Interpreting: Some Methodological Issues. [online]. In *The Interpreters' Newsletter*. ISSN 1591-4127, 1996, vol 7, p. 43-55. [accessed 2017-08-17]. Available at: <http://hdl.handle.net/10077/8990>.

MOYŠOVÁ, S. 2015. Tlmočenie ako vrcholový šport. In *Tlmočenie v interdisciplinárnej perspektíve so zameraním na osobnostné charakteristiky a kognitívne aspekty*. Banská Bystrica : Belianum, 2015, p. 131-137. ISBN 978-80-557-0916-1.

MÜGLOVÁ, D. 2009. *Komunikácia, tlmočenie, preklad alebo Prečo spadla babylonská veža?* Bratislava : Enigma, 2009. 323 p. ISBN 978-80-89132-82-9.

MÜGLOVÁ, D. 2012. Typológia komunikačných situácií v tlmočení na základe výrazovej sústavy Františka Mika. In *Preklad a kultúra 4*. Nitra : Univerzita Konštantína Filozofa, Filozofická fakulta, 2012, p. 212-219. ISBN 978-80-558-0143-8.

NANPON, H. – OLÉRON, P. 1965. Research into Simultaneous Translation. In *The Interpreting Studies Reader*. London : Routledge, 2002, p. 42-51. ISBN 0-415-22477-2.

NG, B. Ch. 1992. End Users' Subjective Reaction to the Performance of Student Interpreters. [online]. In *The Interpreters' Newsletter*. ISSN 1591-4127, 1992, special iss. 1, p. 35-41. [accessed 2018-01-12]. Available at: <https://www.openstarts. units.it/bitstream/10077/2173/1/04.pdf>.

NOVÁKOVÁ, T. 2012. Tlmočenie cez prizmu hodnotenia jeho kvality. In *Kontexty súdneho prekladu a tlmočenia I*. Bratislava : Vydavateľstvo UK, 2012, 85-97. ISBN 978-80-223-3293-4.

OZOLINS, U. 2015. Ethics and The Role of The Interpreter. In *The Routledge Handbook of Interpreting*. London : Routledge, 2015, p. 319-336. ISBN 978-0-415-81166-8.

PALČEKOVÁ, N. 2018. *Rétorická príprava budúcich tlmočníkov v rámci štúdia prekladateľstva a tlmočníctva na Slovensku* (unpublished MA thesis). Matej Bel Univesity in Banská Bystrica, 2018.

PANETH, E. 1957. An Investigation into Conference Interpreting. In *The Interpreting Studies Reader*. London : Routledge, 2002, p. 30-41. ISBN 0-415-22477-2.

PÖCHHACKER, F. 1995. Simultaneous Interpreting: A Functionalist Perspective. [online]. In *Hermes. Journal of Linguistics*. ISSN 1903-1785, 1994, vol. 14, p. 31-53. [accessed 2017-08-17]. Available at: <https://tidsskrift.dk/her/article/view/25094 /22017>.

PÖCHHACKER, F. 2001. Quality Assessment in Conference and Community Interpreting. [online]. In *Meta*. ISSN 1492-1421, 2001, vol. 46, iss. 2, p. 410-425. [accessed 2017-06-30]. Available at: <https://www.erudit.org/en/journals/meta/2001-v46-n2-meta159/003847ar.pdf>.

PÖCHHACKER, F. 2004. *Introducing Interpreting Studies*. London : Routledge, 2004. 251 p. ISBN 0-415-26886-9.

PÖCHHACKER, F. – ZWISCHENBERGER, C. 2010. *Survey on Quality and Role: Conference Interpreters' Expectations and Self-Perceptions*. [online]. AIIC.NET, 2010. [accessed 2017-08-11]. Available at: <https://aiic.net/page/pdf/3405>.

PÖCHHACKER, F. 2011a. Replication in Research on Quality in Conference Interpreting. [online]. In *T & I Review*. ISSN 2233-9221, vol. 1, p. 35-57. [accessed

2017-07-24]. Available at: <http://cms.ewha.ac.kr/user/erits/download/review_1/02_ Franz%20Pochhacker.pdf>.

PÖCHHACKER, F. 2011b. Researching TV interpreting: selected studies of US presidential material. [online]. In *The Interpreters' Newsletter*. ISSN 1591-4127, 2011, vol. 16, p. 21-36. [accessed 2017-08-22]. Available at: <https://www.openstarts. units.it/bitstream/10077/8252/1/NL_16_4.pdf>

PYM, A. 2008. On omission in simultaneous interpreting. Risk analysis of a hidden effort. [online]. In *Efforts and Models in Interpreting and Translation Research*. Amsterdam/Philadelphia : John Benjamins Publishing Company, 2008, p. 83-105. ISBN 9789027291080. [accessed 2018-01-28]. Available at: <https://pdfs.semantic scholar.org/fdcf/a73aa75ce03f04eef8cea0369574705b8a81.pdf>

ROBERTS, R. P. 2000. *The Critical Link 2: Interpreters in the Community : Selected Papers from the Second International Conference on Interpreting in Legal, Health, and Social Service Settings, Vancouver, BC, Canada, 19-23 May 1998*. [online]. Amsterdam; Philadelphia : John Benjamins Publishing Company, 2000. 316 p. ISBN 9789027216366. [accessed 2017-08-27]. Available at: <https://books.google.sk/ books/about/The_Critical_Link_2.html?id=8do5AAAAQBAJ&redir_esc=y>

ROY, C. B. 1993. The Problem with Definitions, Descriptions, and The Role Metaphors of Interpreters. In *The Interpreting Studies Reader*. London : Routledge, 2002, p. 344-353. ISBN 0-415-22477-2.

SCHJOLDAGER, A. 1995. An Exploratory Study of Translational Norms in Simultaneous Interpreting: Methodological Reflections. In *The Interpreting Studies Reader*. London : Routledge, 2002, p. 300-311. ISBN 0-415-22477-2.

SCHOFIELD, M. – MAPSON, R. 2014. Dynamics in interpreted interactions: An insight into the perceptions of healthcare professionals. [online]. *In Journal of Interpretation*, 2014, vol. 23, iss. 1, article 3. [accessed 2017-12-27]. Available at: <https://digitalcommons.unf.edu/cgi/viewcontent.cgi?article=1032&context=joi>.

SEEBER, K. G. 2015. Simultaneous Interpreting. In *The Routledge Handbook of Interpreting*. London : Routledge, 2015, p. 79-95. ISBN 978-0-415-81166-8.

SNELL-HORNBY, M. – JETTMAROVÁ, Z. – KAINDL, K. (eds.). 1997. *Translation as Intercultural Communication: Selected Papers from the EST Congress, Prague 1995*.

[online]. Amsterdam/Philadelphia : John Benjamins Publishing Company, 1997. 354 p. ISBN 9789027216212. [accessed 2017-10-23]. Available at: <https://books.google.sk/ books/about/Translation_as_Intercultural_Communicati.html?id=XyD_CeM3BYC&r edir_esc=y>.

STAHL, J. 2012. Kvalita tlmočenia v teórii a praxi. In *Kontexty súdneho prekladu a tlmočenia I.* Bratislava : Vydavateľstvo UK, 2012, 112-121. ISBN 978-80-223-3293-4.

STENZL, C. 1983. *Simultaneous Interpretation: Groundwork Towards a Comprehensive Model* (unpublished MA thesis). [online]. Birbeck College, University of London, 1983. [accessed 2017-08-24]. Available at: <https://www. academia.edu/389677/Simultaneous_Interpretation_Groundwork_Towards_a_Compr ehensive_Model>.

STRANIERO SERGIO, F. 2003. Norms and quality in media interpreting: the case of Formula One press conferences. [online]. *In The Interpreters' Newsletter*. ISSN 1591-4127, 2003, vol. 12, p. 135-174. [accessed 2018-01-05]. Available at: <https://www.openstarts.units.it/bitstream/10077/2480/1/06.pdf>.

ŠRAMKOVÁ, M. 2008. Druhy chýb pri tlmočení a ich dopad na komunikáciu. In *Od textu k prekladu*. Praha : Jednota tlumočníků a překladatelů, 2008, p. 152-165. ISBN 978-80-7374-047-4.

Štatistický úrad Slovenskej republiky. 2016. *Štatistická ročenka Slovenskej republiky/Statistical Yearbook of the Slovak Republic 2016.* Bratislava : Veda, 2016. 688 p. ISBN 978-80-224-1541-5.

Štatistický úrad Slovenskej republiky. (2018). *Vs_abs_skoly*. Unpublished internal document.

ŠVEDA, P. – POLÁČEK, I. 2015. Trendy v posudzovaní talentových predpokladov pre štúdium tlmočenia. In *Tlmočenie v interdisciplinárnej perspektíve so zameraním na osobnostné charakteristiky a kognitívne aspekty*. Banská Bystrica : Belianum, 2015, p. 60-70. ISBN 978-80-557-0916-1.

TAYLOR-BOULADON, V. 2011. *Conference Interpreting: Principles and Practice*. 3rd edition. [online]. Adelaide : Crawford House Publishing, 2011. 289 p. ISBN 1-4196-6069-1. [accessed 2018-01-22]. Available at: <https://www.scribd.com/ document/149722079/Conference-Interpreting-Taylor-Bouladon-Valerie>.

TISELIUS, E. 2012. *Research on Quality in Interpreting: We All Want A Quality Performance. But How Do You Define That?* [online]. AIIC.NET, 2012. [accessed 2017-07-02]. Available on the Internet: <https://aiic.net/page/print/6318>.

VERTANOVÁ, S. *et al.* 2015. *Tlmočník ako rečník: Učebnica pre študentov tlmočníctva.* Bratislava : Vydavateľstvo UK, 2015. 231 p. ISBN 978-80-223-4030-4.

VILÍMEK, V. 2012. Neverbální komunikace při mezikulturní komunikaci jako součást curricula filologických a translatologických oborů. In *Preklad a kultúra 4*. Nitra : Univerzita Konštantína Filozofa, Filozofická fakulta, 2012, p. 381-388. ISBN 978-80-558-0143-8.

VUORIKOSKI, A. 1993. Simultaneous interpretation – User experience and expectation. [online]. In *Translation – the Vital Link. Proceedings of the XIIIth FIT World Congress*. London, Institute of Translation and interpreting, 1993, vol. 1, p. 389-397. [accessed 2018-01-22]. Available at: <https://www.academia.edu/3352578/ Simultaneous_interpretation_User_ experience_and_expectation>.

Internet Sources:

Does accent matter? Investigating the relationship between accent and identity in English as a lingua franca communication – ScienceDirect. [online], [accessed 2018-02-18]. Available at: <https://www.sciencedirect.com/science/article/pii/S0346251 X16300537?via%3Dihub>.

Non-native accents and simultaneous interpreting quality perceptions. [online], [accessed 2018-02-19]. Available at: <https://benjamins.com/#catalog/journals /intp.15.1.02che/details>.

Rating Scales for Interpreting Performance Assessment: The Interpreter and Translator Trainer: Vol 2, No 2. [online], [accessed 2017-11-20]. Available at: <http://www. tandfonline.com/doi/abs/10.1080/1750399X.2008.10798772>.

Sign language interpreting on Chinese TV: a survey on user perspectives: Perspectives: Vol 21, No 1. [online], [accessed 2018-02-17]. Available at: <http://www.tandfonline. com/doi/abs/10.1080/0907676X.2011.632690>.

The Hospital Cleaner as Healthcare Interpreter: A Case Study: The Translator: Vol 5, No 2. [online], [accessed 2018-01-20]. Available at: <http://www.tandfonline.com/ doi/abs/10.1080/13556509.1999.10799039?journalCode=rtrn20>.

The Pleasures and Pains of Working as an Interpreter - Lýdia Machová at the Polyglot Gathering 2015 - Youtube. [online], [accessed 2018-01-23]. Available at: <https://www.youtube.com/watch?v=mMLXYXOEHk0&t=5s>.

Quality | Definition of quality in English from by Oxford Dictionaries. [online], [accessed 2017-08-03]. Available at: <https://en.oxforddictionaries.com/definition/ quality>.

Quality | Definition of Quality by Merriam-Webster. [online], [accessed 2017-08-03]. Available at: <https://www.merriam-webster.com/dictionary/quality>.

Quality Meaning in the Cambridge English Dictionary. [online], [accessed 2017-08-03]. Available at: <https://dictionary.cambridge.org/dictionary/english/quality>.

Understanding Booth Manners. [online], [accessed 2018-01-22]. Available at: <https:// aiic.net/page/1676/>.

What are working languages to a conference interpreter? [online], [accessed 2017-10-02]. Available at: <https://aiic.net/page/4004>.

APPENDICES

Appendix A. Sections omitted from Chapters 1 and 2.

1.1.2.1 Speaker

The situation is different in consecutive interpreting, where the interaction between speakers and receivers is often more immediate and the speaker can actually hear all of the interpreted speech. Although this might be of little use to them, provided they have zero comprehension of the target language, they can still (up to a point) assess the prosodic features of the interpreter's output, which are of varying importance depending on, for example, the type of speech. Furthermore, even basic knowledge of the target language can sometimes give away certain changes, such as a change in register, neutralisation (e.g. "toning down" the speaker's language), etc.[58]

1.1.2.2 Listeners

In consecutive interpreting, listeners with a high level of knowledge of the source language may evaluate the content of the TT. The better their language skills in the given language, the more accurate their assessment may be. Nevertheless, we cannot forget that the listeners are most probably not trained interpreters, which can result in unfair statements about the interpreter's output (for example, they may not realise that the interpreter uses strategies such as condensation, to make the TT shorter by omitting unnecessary information, and perceive this as a mistake). Their assessment thus needs to be taken with a grain of salt.

1.1.2.4 Interpreter

In consecutive interpreting, self-assessment (without recording) is just as difficult, with the receptive and productive phases providing the interpreter with little or no time to reflect on their output in real time. To make matters worse, there is usually only one consecutive interpreter, which means that the entire communicative event relies on their continuous interpreting effort, by the end of which they are likely to have forgotten the deficiencies of the previous turns.

---

[58] As previously experienced by us.

1.1.2.6 Researcher

In consecutive interpreting, the situation is rather similar to SI and we stand by the opinion that in order to thoroughly evaluate the interpreting performance, the researcher would need a recording of the event.

1.1.3 Ideal vs. "Optimum Quality" or "Quality under the Circumstances"

A distinction must be made between the notions of ideal and "optimum quality" (Moser-Mercer, 1996, p. 44), or "quality under the circumstances" as Pöchhacker (1994; in Kurz et al., 2008, p. 1) put it. Ideal quality is "the hypothetical perfect interpretation" (Altman, 1994; in Collados Aís – García Becerra, 2015, p. 369), which should theoretically occur during an unconditioned interpreting, where everything is perfect – the interpreter has had enough time and resources to prepare for the event, the equipment works without the slightest issue, the speaker speaks slowly about a familiar topic, the audience is homogenous in terms of their demands on the interpreter, etc. This situation is indeed ideal and those who believe in it are idealistic.

In practice, interpreters do not work under ideal conditions. Their experience as well as their performance is usually subject to factors beyond their control known as input variables. These include, among others, issues such as input rate[59], foreign accent of the speaker, length of turns, quality of equipment, materials and prior preparation, physical environment, and much more. Researchers therefore agree that to demand perfect interpreting with no errors, whether content- or form-related, would be simply absurd. The terms "optimum quality" (Moser-Mercer, 1996, p. 44) and "quality under the circumstances" (Pöchhacker, 1994; In Kurz et al., 2008, p. 1) thus refer to the quality interpreters can deliver under the various circumstances they have to face. It should therefore be expected that an interpreter's performance will not be perfect, since it is always subject to external factors.

An interesting theory here is that of Daniel Gile (2009), who claims that performance problems and are not limited to difficult STs or unskilled interpreters. On the contrary, they occur even in the work of experienced professionals, interpreting in ideal conditions. We cannot but agree with his opinion, for humans are not machines which work without failures, provided that the conditions are just right. If we accept

---

[59] For the effect of increased input rate on interpreting, see e.g. Gerver (1969).

this theory, the ideal quality then becomes the Holy Grail of interpreting – dreamt about by many, yet far beyond anyone's reach.

2.1.4 Replication of research

[…]

In his 1990 case study, Gile steered away from user expectations and instead collected data on user responses after an interpreted event (a medical conference). The 23 collected questionnaires displayed similar results, with slightly more favourable scores given by American as opposed to French delegates. Gile concluded that while delivery aspects were considered inferior to other quality components, this did not necessarily influence the general quality assessment. (Gile, 1990)

In 1993, Vuorikoski conducted a similar survey, asking 173 participants of five different seminars to express their attitudes towards using interpreting services (for instance, she wanted to know why they listened or did not listen to the interpretation), as well as to assess certain aspects of the interpretation they received (informedness, coherence/easiness to follow, accuracy, pleasant sentence flow and rhythm, fluency, correct use of terminology). Her questionnaire also included personal data questions (among other things, Vuorikoski was interested in the respondents' previous experience with simultaneous interpreting). Those willing to further participate in the study were invited to take part in a telephone interview. The study yielded fruitful results – the author concluded that users' expectations and satisfaction were not necessarily group-determined, but rather individual-specific. Moreover, they tended to change throughout an interpreted event. She also discovered that users sometimes complained about matters beyond the interpreter's control (e.g. uncomfortable headsets, mismatching gender of the speaker and the interpreter) and suggested that there is a need for more collaboration between the various parties involved, if clients are to improve the quality of the offered interpreting services. (Vuorikoski, 1993)

Kopczyński surveyed 54 users of interpreting (speakers and listeners from three different fields – humanities, science and technology, diplomacy) to find out about their expectations of good interpreting in general (i.e. not related to a particular event). Content was considered more important than form by all groups, although some differences were found among the groups with respect to individual criteria and their importance. Kopczyński also inquired about the preferred role of the interpreter by

asking users whether an interpreter should assume the ghost role or whether they should intrude the speaker's output when necessary. The responses clearly pointed towards a preference for the ghost role, although some intrusions were favoured by the respondents. (Kopczyński, 1994; in Kurz, 2001, p. 401)

[…]

Delia Chiaro and Giuseppe Nocella were the first researchers of interpreting quality to use the Internet in 2004 for the distribution of their questionnaire to approximately 1,000 professional interpreters (the total number of responses was 286). The authors asked the respondents to rank (rather than rate) two sets of criteria – linguistic and extra-linguistic, taken from Bühler's 1986 survey – in the order of their importance for high-quality interpreting. (Chiaro – Nocella, 2004) As Pöchhacker (2011a) noted, Chiaro and Nocella's study has one major shortcoming, namely the fact that the target population of their survey is not sufficiently defined. Although the respondents are described in terms of their gender, age, nationality, years of experience, etc., their professional base (organisation/association) is not specified at all, which opens up the possibility that non-conference interpreters were also reached by this web-based survey. Furthermore, the fact that that some of Bühler's criteria were slightly altered in combination with the changed system of their evaluation in terms of importance makes it difficult to compare the results with Bühler's original study.

2.2x Community/Liaison and other types of interpreting

While Bühler's 1986 study looked at both consecutive and simultaneous interpreting (which can be seen from certain interpreter-related criteria such as poise or pleasant appearance), it was not until later that research on quality in consecutive interpreting received more substantial interest from experts in the field. Although fewer in number, studies focusing on non-conference interpreting do exist and they touch various settings and topics. As this type of interpreting is not in the centre of our attention, this subchapter provides but a brief overview of research done in this area.

2.2x.1 Community/Liaison interpreting

The nature of quality-oriented research in community interpreting is rather different to that in conference interpreting. While the latter often focuses on the interpreter's output as such, the former mostly aims at establishing professional standards, dealing with matters such as the interpreter's role and tasks. As observed by

many, e.g. Donk (1994, in Pöchhacker, 2004, p. 152), Scheffer (1997, in Pöchhacker, 2004, p. 152), Davidson (1998 and 2002, in Pöchhacker, 2004, p. 152), Pöchhacker and Kadric (1999[60]), Bolden (2000, in Pöchhacker, 2004, p. 152), to mention a few, interpreters in the community setting often step out of their role and assume the so-called "helper role" (Pöchhacker, 2004, p. 152), acting as a "pre-diagnostic agent", a "co-therapist" (in medical settings), or as a "deputy officer" (in interpreted police interrogation) (Ibid.).

Surveys on client/user expectations and/or client/user satisfaction have also been carried out for community-based domains of interpreting, among others by researchers such as Marrone (1993), Garber and Mauffette-Leenders (1997, in Pöchhacker, 2004, p. 155), Mesa (1997, in Schofield – Mapson, 2014, p. 1; 2000, in Pöchhacker, 2004, p. 154-156), and Kadric (2000, in Roberts, 2000, p. 153-164), who looked at interpreting in legal settings. Results of Mesa's study indicate that a more active role (rather than the ghost role) is preferred by users of interpreting, with 92% of them rating "pointing out a client's lack of understanding" as "very important" (Mesa, 2000, in Pöchhacker, 2004, p. 154).

In her 1988 study, *The Impact of Politeness in Witness Testimony: The Influence of The Court Interpreter*, Susan Berk-Seligson discovered that the interpreter is a powerful mediating agent, for he or she can change the way judges perceive a witness with respect to such qualities as convincingness or trustworthiness. A polite interpretation yielded significantly higher ratings than an impolite one (the original witness' speech was polite in both cases). This leads to an intriguing question of whether a legal interpreter should improve the register of an individual whose utterances they are rendering, or assume the role of an interpreting machine.

2.2x.2 Media interpreting

Although largely underexplored, media interpreting is perhaps the only type of interpreting most people have experienced (possibly without even realising it). Research on quality of interpreting for the media is scarce, given the great potential for unobtrusive observation of interpreters' natural performance – among the few studies dealing with the topic, perhaps the most notable ones are those by Kurz and Pöchhacker (1995, in Pöchhacker, 2011b, p. 30), Kurz (1997, in Snell-Hornby *et al.*, 1997, p. 195-

---

[60] Abstract available at: <http://www.tandfonline.com/doi/abs/10.1080/13556509.1999.10799039?journalCode=rtrn20> [accessed 2018-01-20]

206), Straniero Sergio (2003), Märzluft (2010, in Pöchhacker, 2011b, p. 27-29), and Schwarnthorer (2010, in Pöchhacker, 2011b, p. 29-31), which is interesting to us because of its similarity to our research (users assessing the importance of various criteria).

As one may expect, quality criteria in TV interpreting differ from those in other domains of interpreting. Paralinguistic criteria often seen as the least important ones in conference interpreting (such as pleasant voice, intonation, accent, etc.) rank high in interpreting for the media, as has been proved by some of the above mentioned studies (e.g. the category of pleasant voice ranked fifth, with 4.05/6p. in Schwarnthorer's study (2010, in Pöchhacker, 2011b, p. 29-31)).

2.2x.3 Signed-language interpreting

Research on the quality of signed-language interpreting, which is usually practised in the simultaneous mode, often focuses on the role of the interpreter and users' expectations and perspectives. Furthermore, evaluation-oriented studies have also been written, such as the one by Xiao and Feiyan (2013[61]), which combines the topic of signed-language interpreting with that of interpreting for the media, as it looks at SgL interpreting on Chinese television, or the 1992 study by Strong and Rudser, in which respondents were asked for an overall as well as partial evaluation of SgL interpreters' performance (in Pöchhacker, 2004, p. 156). De Wit and Sluis (2014) found out that much like in other domains of interpreting, users' quality criteria vary depending on the specific setting.

---

[61] Abstract available at: <http://www.tandfonline.com/doi/abs/10.1080/0907676X.2011.632690>. [accessed 2018-02-17]

Appendix B. Sections omitted from Chapter 4.

4.0 The Original Methodology

The original idea of our survey was to let three parties participate in the evaluation of interpreting at real-life interpreted events (conferences). Those parties were the users, the interpreters, and the internal[62] researcher – us. The process was made up of four main stages – collecting responses about the importance of selected criteria and about the perceived quality of the interpreting provided from both the users and the interpreters – these first two stages completed through a questionnaire –, adding our own evaluation and, finally, analysing the results and calculating how well the alleged importance of the criteria corresponded with the final user evaluation.

4.0.1 The Questionnaires – Stage I and II

Two questionnaires were created for the original research – one meant for users and the other one for interpreters. Both questionnaires were in Slovak and were available in two forms – a printed form and an online form (accessible through the website www.survio.com). The only difference between the printed and the online questionnaire was the extra question asking the respondents to briefly identify the conference they had recently attended (by writing its name or date), so that we could match their response with a particular event.

4.0.1.1 Socio-demographic data

In order to be able to better define our sample, the respondents were first asked a few socio-demographic questions. As this was the part with the most significant differences in the two questionnaires, we will deal with both questionnaires separately.

4.0.1.1.1 User Questionnaire Demographics

The user data we were interested in consisted of the user's gender, age, previous experience with interpreting, understanding of the language spoken by the foreign speaker, and the manner in which they used interpreting services at the given conference. The age groups listed corresponded with those given in Peter Moser's 1995 survey (0-29, 30-45, 45-60, over 60), as did the three groups with different experience ("newcomers" using interpreting for the first time, users with some experience with

---

[62] In Pöchhacker's (2001) sense.

interpreted events, and "oldtimers" with ample experience with such events) (Ibid.). We also asked the users to what extent they could understand the speaker's language (not at all, a little, very well) and how much they used interpreting services at the conference (not at all, sometimes, during the whole speech of the foreign speaker). It was the researcher's task to note other data relevant to the research, such as the topic and size of the conference, so as to make the questionnaire as simple and quick as possible for both of the other parties.

4.0.1.1.2 Interpreter Questionnaire Demographics

The interpreters' questionnaire also began by asking about the respondents' gender and age (split in the same age groups as the users' questionnaire). We then asked them about their active interpreting experience (less than 5 years, 5-10 years, 10-20 years, more than 20 years), and their working languages (they were asked to list all of their A, B, and C languages, according to the AIIC classification[63]). Lastly, we wanted to know whether the interpreters specialised in a particular topic(s) or field(s).

4.0.1.2 Preferences and criteria

Before presenting the respondents with our list of criteria, we asked them two questions related to the interpreter themselves. The first question "Should the interpreter be a man or a woman" had four possible answers (and an extra answer for those who did not feel particularly affiliated with any of the given ones): a) I prefer female interpreters; b) I prefer male interpreters; c) the speaker's and the interpreter's genders should match; d) I do not mind either. This questioned was asked to determine whether an average user had strong feelings about the gender of the interpreter or not. A similar, but open-ended question was given to the interpreters: "Do you think the gender of the interpreter is important? If so, briefly describe how."

Next, we asked both users and interpreters about the task of the interpreter – "What kind of interpreting do you prefer?" (users) and "How should an interpreter interpret?" (interpreters). Despite the slight difference in the questions, the answers were the same, with the possibility of a custom answer: a) the interpreter interprets everything said by the speaker without adding or omitting anything; b) the interpreter gives a faithful rendition of the speech but can add (explain) or omit (e.g. redundant) information; c) the interpreter summarises what has been said by the speaker. This

---

[63] Available at: <https://aiic.net/page/4004/>. [accessed 2017-10-02]

question was inspired by both Downie's book (2016) and Moser's 1995 survey, in which he tackled the same issue by asking the respondents what they considered more important, "concentration on essentials" or "completeness of rendition" (Ibid., p. 15).

The next step, identical in both questionnaires, consisted of rating 14 output-related criteria on a four-point scale. We considered adding an extra point to the original Bühler's (1986) scale, but opted against this for two reasons – firstly, we wanted our research to be compatible with all major studies dealing with criteria importance (namely Bühler (1986), Kurz (1993), Moser (1995), Pöchhacker and Zwischenberger (2010)) and, secondly, we thought this would force the users to think in cases where they could not decide, as opposed to simply choosing the "middle ground" often used as the "I don't know/I can't be bothered thinking" option. The criteria were listed in a random order, i.e. delivery, language, and content-related criteria were all mixed together.

[The list of the criteria is the same as the one found in subchapter 4.2.]

4.0.1.2.4 Users and Interpreters' Own Criteria and the Dependency of Criteria

At the end of the criterion list, users and interpreters were given the opportunity to write any other criteria they felt were important in order to provide high quality interpreting.

We also asked the interpreters to tell us whether they thought the importance of the criteria was dependent on the type (topic, size, etc.) of conference or speech. This was an optional open-ended question.

4.0.2 Stage III – Assessment

The next stage consisted of interpreting assessment. Here, the questions differed perhaps most significantly and we shall describe them separately.

4.0.2.1 User Assessment

In the assessment part of the questionnaire, users were asked to evaluate the interpreting they had heard on a five-point scale (in this case, we thought the "middle ground" option was appropriate, as this scale also reminds Slovak users of the assessment system in Slovak primary and secondary education), where one point meant

that they thought the interpreting was very bad and five points meant they thought it was very good[64].

Users were first asked to evaluate the overall impression of the interpreting, and only then did they evaluate the three categories of criteria – delivery, language, and content. The reason for the overall impression being the first one on the list was that we thought it might decrease the risk of bias – if users rated the three categories first, they might then think back to the importance they gave each of the criteria and consciously, rather than intuitively mark their overall impression.

An important point to mention here is the fact that the users were supposed to rate the interpreters as a tandem, rather than individually. This was decided after a thorough discussion with our supervisor, a seasoned interpreter, who is of the opinion that interpreters are always praised or criticised together as a booth. Furthermore, at the end of the questionnaire, users were given the opportunity to express their feelings about the interpreting services received or about the questionnaire itself. In case they felt that the performance of the two interpreters was so incomparable that it deserved to be mentioned, they could potentially express that feeling in this part.

4.0.2.2 Interpreter Assessment

The interpreters' self-assessment was done through an open-ended question. However, interpreters were not only asked to evaluate their own performance, but also to evaluate their working conditions on that particular day. By giving them the chance to list any unfavourable conditions, we wanted to avoid unfair judgment while also letting the interpreters know that we realised their work was dependent on external factors and inviting any complaints they might have about the event and its organisation.

Just like at the end of the user questionnaire, the interpreters were also given the opportunity to mention anything they felt was important in relation to the interpreting or the questionnaire as such.

4.0.3 Analysis of Responses – Stage IV

The analysis of responses in the original research plan had three main parts. The first part was a sole analysis of users and interpreters' opinions regarding the interpreter's gender and role, as well as the criteria importance, the second one was an

---

[64] Having mentioned the Slovak education assessment system, we realise that this scale, which is the opposite (the higher the number the better the interpreting), may be confusing, which is why emoticons were inserted next to the lowest and the highest point options).

analysis of user assessment, and the third one consisted of adding the interpreters' as well as our own insight to this assessment.

4.0.3.1 Preferences and criteria

The collected responses to the questions about the interpreter's gender and role, as well as the ones about the various criteria importance were to be analysed with respect to the socio-demographic data provided by both users and interpreters. In both cases, we were mainly interested in differences between the two genders. We also expected to find significantly different opinions between users with little as opposed to ample experience with interpreted conferences. Similarly, we were curious to see the responses of interpreters with various amounts of experience, as well as interpreters with different working languages.

Furthermore, responses from individual conferences were also to be dealt with as separate units (unless there were more conferences with exactly the same topic), which would result in a comparison not unlike the one in Kurz' 1993 study *Conference Interpretation: Expectations of Different User Groups*. Our own observations about the conferences, regarding their type, size, languages spoken, mode of interaction, etc. would also allow us to group together participants of events similar in at least one of these aspects.

4.0.3.2 User assessment analysis

The first, easier stage of analysing user assessment consisted of collecting all the responses and simply arriving at an average mark for each of the three categories as well as the mark of the overall impression that the users at a particular conference gave the two interpreters.

The more complex analysis consisted of calculating the overall importance of each of the three categories (delivery, language, content), based on the arithmetic mean of all criteria in each individual category. The formula looked like this:

- arithmetic mean for delivery-related criteria (dcM) = (dc1 + dc2 + … dc10) ∕ 10
- arithmetic mean for language-related criteria (lcM) = (lc1 + lc2) / 2
- arithmetic mean for content-related criteria (ccM) = (cc1 + cc2) / 2

These numbers represent the "weight" of each category. To get the total amount of points available for each category (DCmax; LCmax; CCmax), we need to multiply them by 5 (the maximum amount of points that could have been awarded for each of the three categories):

- dcM * 5 = DCmax
- lcM * 5 = LCmax
- ccM * 5 = CCmax

After adding these numbers together, we get the maximum amount of points that the interpreters could have been awarded for their performance, i.e. for the overall impression (OImax):

- DCmax + LCmax + CCmax = OImax

We then multiply the raw points the users actually awarded for each of the category (dcP; lcP; ccP) by their category's weight to arrive at the representation they should get in the final mark (overall impression). For transparency, we will call them weighted points (WP):

- dcP * dcM = dcWP
- lcP * lcM = lcWP
- ccP * ccM = ccWP

Finally, adding the weighted points together and diving the number by the maximum points for overall impression (OImax) gives us the final mark interpreters should have been given for their performance in a percentage form ($FM_{\%}$):

- **(dcWP + lcWP + ccWP) / OImax = FM$_\%$**

The final mark allows us to check whether the users stuck to the alleged criterion importance during their assessment. Thus, for example, if their mark for overall impression is 3, but $FM_{\%}$ is 80%, we can safely say that they were influenced by categories they perceived as less important, or vice versa. Although this formula may look difficult, it is in fact very simple, especially with the use of MS Excel, and the process can even be made fully automatic if the questionnaire results are also exported into this programme.

The adapted formula for calculating the assessors' accuracy works on the same principle as its original version, with a few changes:

- *the respondents were asked to evaluate each of the 14 criteria separately on a five-point scale*
- *overall impression was rated on a ten-point scale*

4.0.3.3 Adding Insights

In subchapter 1.1.1.2, we expressed our opinion that interpreters should always be given the opportunity to justify their decisions during interpreting and/or defend the deficiencies noticed by the listeners. For this reason we wanted to add their insight to the assessment given by users. This insight would consist of their own self-assessment, as well as anything they felt was important to mention in relation to their working conditions.

Furthermore, the third part of the assessment analysis also included adding our own opinion. The original idea was that our role would be that of an internal researcher who not only distributes questionnaires among users and listeners, but also listens to the interpreters' output, while simultaneously listening to the speaker (provided that the conference language was English) and arrives at their own assessment for each criteria on the list (not just the three categories). Because we did not get the permission to record the speaker or the interpreters, our assessment of, for example, sense consistency, would have to be taken with a grain of salt, but would probably be more accurate than that of the users. However, user assessment of correct terminology (provided we asked them to evaluate this criterion separately) would definitely be more accurate, since they were the real experts in the field, not us. Therefore, our assessment was only of a supplementary character and definitely did not serve as an ultimate judgment of the interpreters' performance.

4.0.4 The Preparations and the Failure

The original research idea was discussed with our supervisor as soon as in June 2016. As a professional in the field, he contacted several interpreting agencies and introduced our research idea to them, trying to make their representatives interested in what we were planning to do. By August 2017, our questionnaires were ready and so were we to seize every opportunity available to start the research.

However, negotiations took much longer than expected and by the time we finally got permission to do our research from a well-known Slovak translation and interpreting agency, it was the end of October 2017. The first conference we were allowed to attend was to take place in mid-November. We were supposed to meet the head of the agency's interpreting division and discuss the final details, but for the most part, everything was clear. At the conference, we would act as an employee of the agency's quality assurance department, positioned by the technicians' corner, and give

out questionnaires to users who came to borrow a headset. At the end of the conference, when they were returning it, we would collect their responses. They could also choose to fill in the questionnaire on the Internet, for which we were going to prepare special leaflets with a QR code (as well as a URL address) on them. This code (or URL address) would get them to the questionnaire website and they could fill it in from the comfort of their own home.

Ready to go, we suddenly received an e-mail from the agency saying we could not proceed with the research, because the client (the conference organiser or the person in charge of securing interpreting services) did not wish for any research to be carried out at the conference. Unfortunately, this remained the case for all other conferences and, in the end, despite contacting the agency multiple times, we were never allowed to go to a single one of them.

In January 2018, we started to draft a fallback. By the end of this month, we knew it was too late for the original research idea to work, and we decided to abandon it and initiated the student-based research.

4.4.1 TI Students' Socio-Demographic Data

[…]

Lastly, we wanted to know which languages the students study. While English is studied by a majority of the respondents (232), a large number of them also study other languages. The following chart shows the number of students studying each of the given languages:
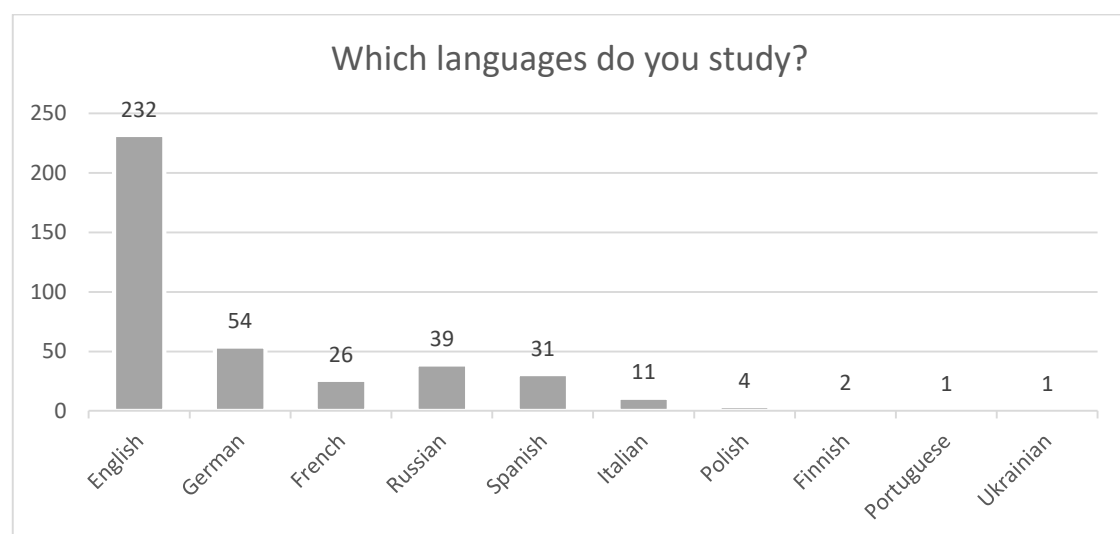
Figure 14: TI sample – studied language(s)

4.4.1 Non-TI Students' Socio-Demographic Data

In terms of age, 185 of our respondents were 21 (20.56%), 184 of them 20 (20.44%) and 172 in total were 22 years old (19.11%). The following graph shows ratios of students of different ages.
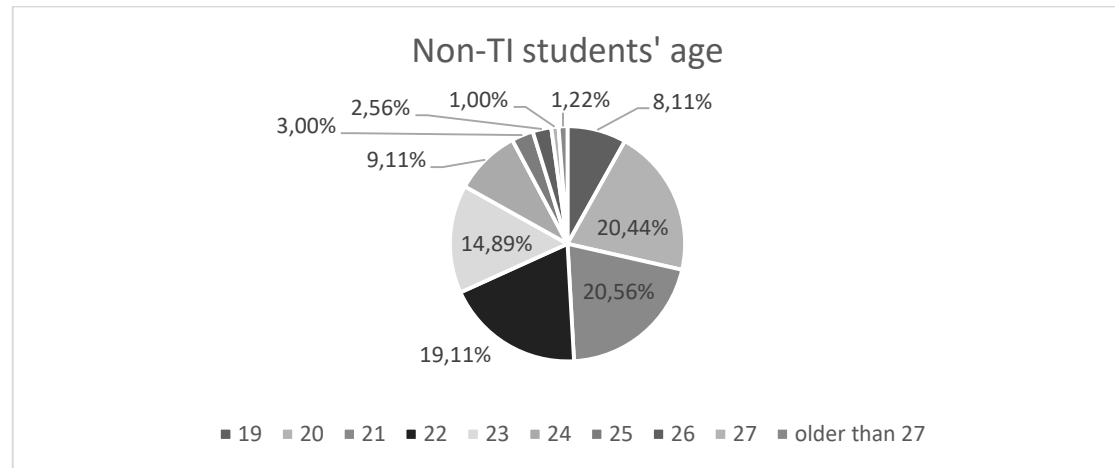
Figure 15: Non-TI sample – age

These age groups are reflected quite well in the structure of our sample in terms of the respondents' year of study. In fact, we shall not analyse the results based on the respondents age, but will rather look at the year of study, simply because a (for example) 20-year-old student is likely to be in one of at least three years of study – first, second, and third (depending on the date of their birth and their age at the time of high-school graduation).

Non-TI students according to their field of study:
- **art/aesthetics:**
  - 33 students (6 men, 27 women);
  - includes students of "history of art" and "aesthetics";
- **civil engineering:**
  - 31 students (16 men, 15 women);
  - includes students of "civil engineering", "civil engineering structures", "bearing structures of buildings", and "roadway engineering";
- **finance/economy/management:**
  - 80 students (26 men, 54 women);

- includes (among others) students of "finance, banking, and investment", "business economics and management", "management", "business management", and "accounting";
- we decided to put both economy and management in one category because we discovered that they were frequently studied together

- **foreign languages:**
  - 32 students (1 man, 31 women);
  - includes students of any foreign language in non-teaching as well as teaching programmes (in the latter case, any possible combinations were also accepted – e.g. Slovak and English, or English and geography);

- **forestry/agriculture/wildlife management:**
  - 43 students (33 men, 10 women);
  - includes students of "forestry", "applied forestry", "applied zoology and wildlife management", and "special animal husbandry";
  - although we included all other management programmes in the finance/economy/management category, we decided not to do so with the applied zoology and wildlife management programme, because its Slovak name does not suggest that this programme is of a managerial nature (aplikovaná zoológia a poľovníctvo);

- **healthcare:**
  - 89 students (23 men, 66 women);
  - includes (among others) students of "physiotherapy", "nursing", "midwifery", "public health";

- **international relations:**
  - 46 students (19 men, 27 women);
  - includes students of "international relations", "international relations and diplomacy", and "security studies";

- **information technology (IT):**
  - 40 students (24 men, 16 women);
  - includes (among others) students of "informatics", "applied informatics", and "telecommunications";

- **journalism/media:**
  - 23 students (5 men, 18 women);

- o includes students of "journalism" and "mass media communication";
- **law:**
  - o 52 students (18 men, 34 women);
  - o includes students of "law";
- **medicine:**
  - o 69 students (22 men, 47 women);
  - o includes students of "general medicine" and PhD students specialising in "oncology", "gynaecology and obstetrics", "clinical biochemistry", and "pathological anatomy";
- **natural sciences:**
  - o 27 students (12 men, 15 women);
  - o includes students of natural sciences, most commonly biology, chemistry, geography, and mathematics, in both teaching and non-teaching programmes in any combination except for combinations with Slovak or any foreign language;
- **physical education:**
  - o 33 students (29 men, 4 women);
  - o includes students of "physical education and coaching" and "physical education";
- **pharmacy:**
  - o 34 students (7 men, 27 women);
  - o includes students of "pharmacy";
- **political science:**
  - o 30 students (16 men, 14 women);
  - o includes students of "political science";
- **public administration:**
  - o 27 students (6 men, 21 women);
  - o includes students of "public administration" and "regional development";
- **Slovak language:**
  - o 33 students (2 men, 31 women);

- includes students of Slovak language (and literature) in both teaching and non-teaching programmes in any combination except for combinations with a foreign language;

- **social work:**
  - 42 students (4 men, 28 women);
  - includes students of "social work";

- **special pedagogy:**
  - 20 students (3 men, 17 women);
  - includes students of "preschool and primary pedagogy", "teaching of professional subjects and practical preparation", and "special pedagogy";

- **technical engineering:**
  - 39 students (30 men, 9 women);
  - includes (among others) students of "automation", "autotronics", "electrical engineering", "biotechnology", and "production technologies";
  - although we considered putting technical engineers and civil engineers in one category, in the end we decided not to, after discovering several considerable differences in their opinions;

- **tourism:**
  - 33 students (10 men, 23 women);
  - includes students of "tourism";

- **transport/logistics/postal services:**
  - 24 students (8 men, 16 women);
  - includes students of "transport", "railway transport", "forwarding and logistics", and "postal technologies and services".

Apart from these categories, there are also two groups of students which we shall not analyse, since they are underrepresented. These are students of history (13 students, 5 men and 8 women) and psychology (7 students, all of them women).

Appendix C. Sections omitted from Chapter 6.

1. TI students & gender:

	TI all N = 250	TI men n = 45	TI women n = 205
average value	3.184	3.127	3.196
fluency of delivery	3.488	3.622	3.459
native accent	2.26	2.422	2.224
logical cohesion	3.756	3.667	3.776
correct terminology	3.556	3.4	3.590
completeness of delivery	3.344	3.378	3.337
correct grammar	3.208	3.022	3.249
sense consistency	3.716	3.622	3.737
pleasant voice	2.52	2.356	2.556
lively intonation	2.752	2.689	2.766
no filler words & hesitation noises	3.108	3.067	3.117
no booth noises	3.144	3	3.176
synchronicity with the speaker	2.884	2.8	2.902
clear articulation	3.516	3.467	3.527
confident voice	3.32	3.267	3.332
average value	3.184	3.127	3.196

Table 12: *Criteria – TI students; men & women*

As can be seen from the table, the two highest ranking criteria both belong to the category of content, **logical cohesion** being perceived as the most important criterion with an average value of 3.756 and followed by **sense consistency** (3.716). **Correct terminology** is seen as the third most important criterion (3.556), closely followed by **clear articulation** (3.516), a criterion not mentioned by a single author of the "mainstream" user expectation surveys. The fifth rank is taken by **fluency of delivery** (3.488).

On the other hand, the criterion of **native accent** is seen as the least important one (2.26), although male respondents placed it before the criterion of pleasant voice (2.422 vs. 2.356). Only three other criteria ranked below three points on average – **pleasant voice** (2.52), **lively intonation** (2.752), and **synchronicity with the speaker** (2.884).

There are **no significant differences between men and women** when we compare their average scores to the whole group of TI students. However, when compared with each other, a difference of 0.227 point can be seen between the scores

for **correct grammar** and a difference of 0.2 point between the scores given to **pleasant voice**, both of these criteria perceived as more important by women. It is also interesting to see that all four criteria belonging to the categories of content and language are ranked higher by women than men. However, overall, the differences between the two genders are small.
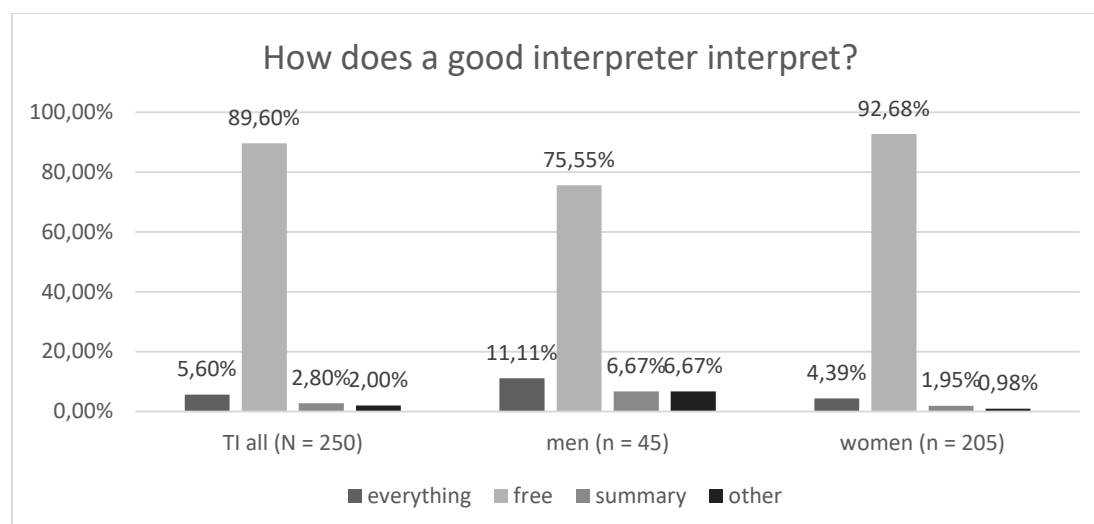
Figure 16: Interpreting styles – TI students; men & women

We can see that almost nine out of ten TI students allow the interpreter to add/omit information. This option is more popular with women than men. Men, on the other hand, are more likely to prefer both a full rendition and a summary. The five "other" answers are all concerned with a correct transmission of the meaning or all of the important information, while allowing for omissions and additions. One respondent (R235, female) also wrote that the interpreter should "tone down" (neutralise) expressive words. […]

2. TI students & their future interpreting careers:
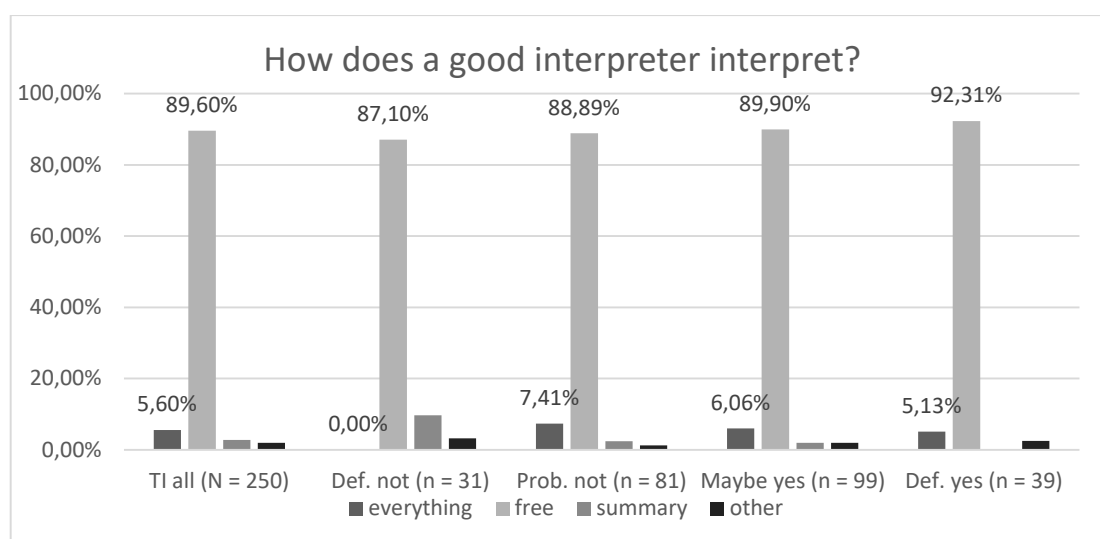
Figure 17: Interpreting styles – TI students with different outlooks on their future interpreting careers

Only marginal differences can be found among groups with different outlooks on their future interpreting careers. Nevertheless, we cannot but point out the fact that the highest percentages for the "free" answer were given by TI students who are considering the job of an interpreter.

3. TI students & SI experience:

When we split Questionnaire B respondents into groups according to their previous experience (exp.) with simultaneous interpreting, we can see significant differences with respect to only three criteria.

	TI all N = 250	no exp. n = 95	lesson exp. only n = 116	out of lesson school exp. n = 8	some out of school exp. n = 28	good out of school exp. n = 3
fluency of delivery	3.488	3.421	3.491	3.375	3.679	4
native accent	2.26	2.179	2.362	2.25	2.071	2.667
logical cohesion	3.756	3.674	3.793	3.75	3.928	3.333
correct terminology	3.556	3.568	3.534	3.625	3.607	3.333
completeness of delivery	3.344	3.242	3.371	3.375	3.571	3.333
correct grammar	3.208	3.284	3.172	3.375	3.107	2.667

sense consistency	3.716	3.726	3.707	3.875	3.714	3.333
pleasant voice	2.52	2.442	2.534	2.875	2.571	3
lively intonation	2.752	2.653	2.828	2.875	2.75	2.667
no filler words & hesitation noises	3.108	3.147	3.164	3.25	**2.75**	2.667
no booth noises	3.144	3.274	3.155	3.75	**2.5**	3
synchronicity with the speaker	2.884	**3.084**	2.836	2.625	**2.464**	3
clear articulation	3.516	3.663	3.431	3.75	3.357	3
confident voice	3.32	3.316	3.302	3.5	3.357	3.333
average value	3.184	3.184	3.191	3.191	3.304	3.102

Table 13: Criteria – TI students with different SI experience
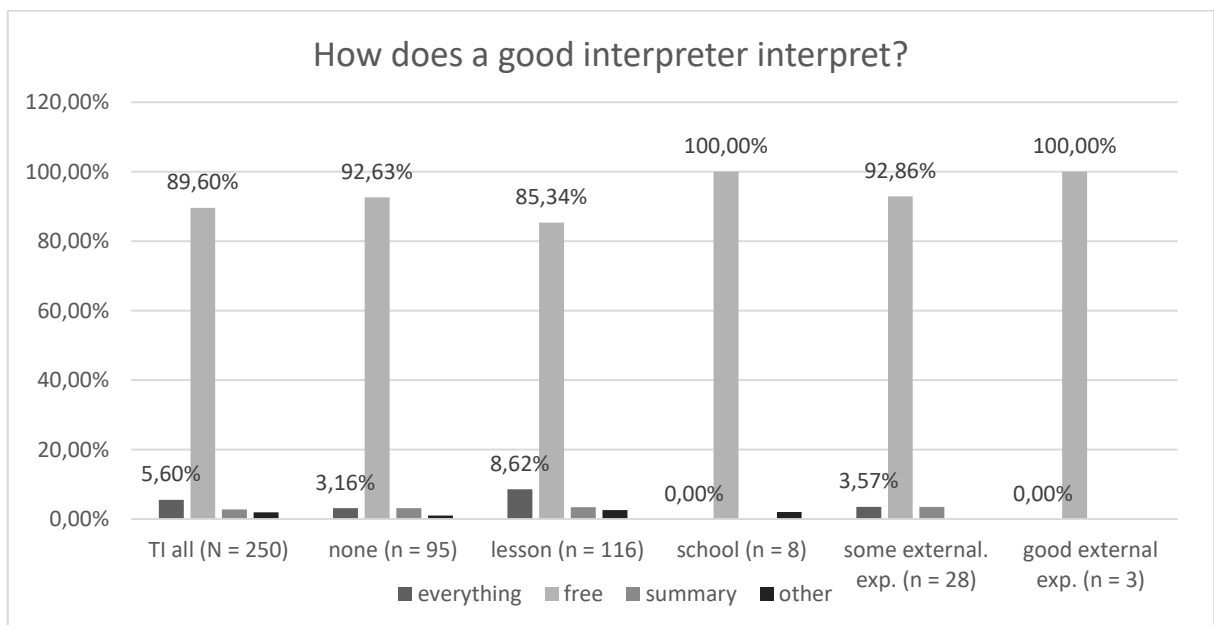
Figure 18: Interpreting styles – TI students with different SI experience

Most "free" answers were given by students who have interpreted simultaneously outside of lessons. Nevertheless, they represent a relatively small group (39 members) and we shall not generalise this result to other experienced students. However, we can also see another interesting result here, if we compare the group with no experience to the group with lesson experience. If we take into account the fact that they are both rather well represented, it is indeed curious that more students with at least some SI experience think that the interpreter should strive for a full rendition.

4. TI students & languages:

Only one significant difference in scores can be found among TI students of different languages – those, who study Russian (in possible combinations with other languages), attributed more importance to the criterion of **no booth noises** (3.41 vs. the overall mean score of 3.144). TI students of Russian language were also the strictest one of all analysed language groups, although the overall differences were not significant.

	TI all N = 250	EN n = 232	DE n = 54	FR n = 26	RU n = 39	ES n = 31	IT n = 11	PL n = 4	other[65] n = 4
fluency of delivery	3.488	3.478	3.481	3.269	3.59	3.548	3.818	4	3.5
native accent	2.26	2.22	2.259	2.038	2.385	2.419	2.182	3.25	2.25
logical cohesion	3.756	3.75	3.704	3.846	3.821	3.71	3.727	4	4
correct terminology	3.556	3.547	3.556	3.615	3.462	3.645	3.545	3.5	3.5
completeness of delivery	3.344	3.349	3.352	3.308	3.487	3.161	3.091	4	3.25
correct grammar	3.208	3.211	3.315	3.115	3.205	3.355	3.364	3	3.5
sense consistency	3.716	3.72	3.685	3.808	3.769	3.71	4	3.5	3.75
pleasant voice	2.52	2.53	2.463	2.538	2.615	2.613	2.636	2	2
lively intonation	2.752	2.733	2.741	2.731	2.795	2.871	2.636	3.5	2.75
no filler words & hesitation noises	3.108	3.086	3.167	3.385	3.103	2.968	3	3.5	3.25
no booth noises	3.144	3.155	3	3.385	**3.410**	3.032	3.273	2.25	2.75
synchronicity with the speaker	2.884	2.849	3.037	2.731	2.718	2.903	2.545	3.5	3.25
clear articulation	3.516	3.526	3.407	3.577	3.615	3.516	3.545	3.5	3
confident voice	3.32	3.319	3.278	3.385	3.385	3.387	3	3	3.5
average value	3.184	3.177	3.175	3.195	3.24	3.203	3.097	3.321	3.161

Table 14: Criteria – TI students of different languages

---

[65] The "other" group includes two students of Finnish, one student of Portuguese, and one student of Ukrainian.
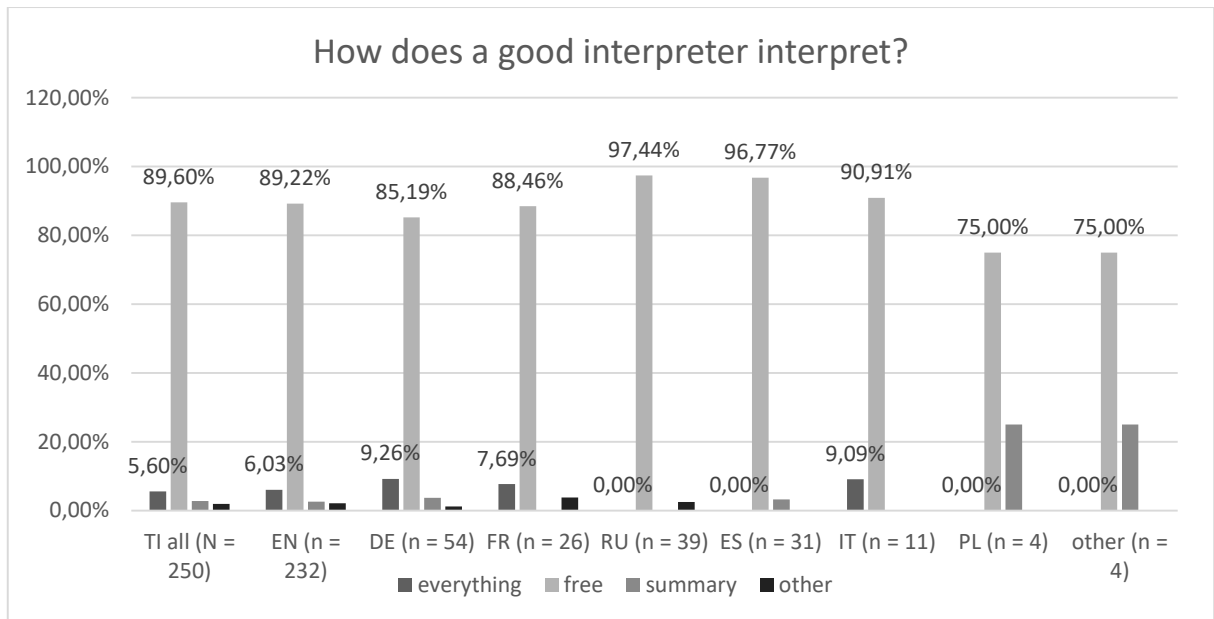
Figure 19: Interpreting styles – TI students of different languages

When split into groups by their studied languages, TI students gave relatively consistent answers, although it is interesting to see that none of the 39 students of Russian and 31 students of Spanish language think that an interpreter should transfer everything. The strictest (analysable) group in this respect is the group of 54 students of German, where over 9 percent allow no omissions/additions to the ST by the interpreter.

5. TI students & year of study:
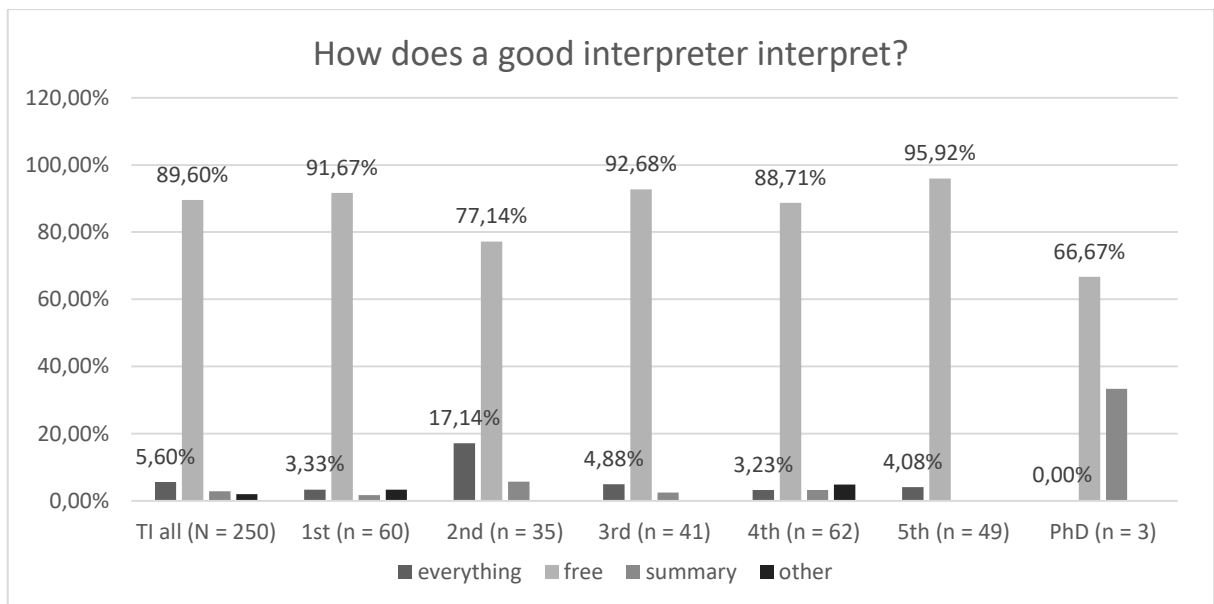
Figure 20: Interpreting styles – TI students in different years of study

From this table, we can say that the group of TI students pushing the percentage of the first answer up, is the group of second year students. After further investigation as to why this group chose this answer much more often than all the other ones, we found out that while only three members of this group are men (i.e. the gender with a higher tendency to choose "everything"), 18 members study at the University of Prešov, which itself stands out in the perception of the role of interpreter.

6. TI students with SI experience – Comparison with other researchers:

An important thing to mention is that our sample of 250 students is made up of 211 students who have never experienced simultaneous interpreting outside of class and of which 95 have never experienced it at all. Therefore, we have decided to try to look at the scores given by the remaining 39 students only, again in comparison with those of Bühler's and Pöchhacker and Zwischenberger's interpreters. The new table then looks like this:

	Bühler N = 47	TI students with experience in SI n = 39	Pöchhacker & Zwischenberger N = 675-704
fluency of delivery	*3.468*	3.684	*3.7*
native accent	***2.9***	2.329	***2.662***
logical cohesion	3.8	3.670	3.744
correct terminology	3.489	3.521	3.6
completeness of delivery[66]	3.426	3.426	3.408
correct grammar	**3.38**	3.05	**3.489**
sense consistency	**3.957**	3.641	3.877
pleasant voice	**3.085**	2.815	**3.123**
lively intonation	-	2.764	**3.148**
synchronicity with the speaker	-	2.696	2.799

Table 15: Criteria – TI students with SI experience and interpreters

After this experiment, we conclude that even with the increased border of significance (0.25, since we are now comparing a group of fewer than 45 members), the differences, except for the criterion of fluency of delivery, do not disappear, while

---

[66] In Bühler's study, this criterion is stated as **completeness of interpreting**, while Pöchhacker and Zwischenberger work with **completeness** only. This may account for differences.

our students now fall significantly behind Bühler's interpreters in terms of their perceived importance of correct grammar.

7. TI students - other criteria and dependency of criteria:

Other Criteria

In total, 18 respondents added their own criteria to the list. Several people mentioned that the interpreter should *dress appropriately* for the event they are interpreting or that they should have *professional behaviour*. Another popular answer was that the interpreter should be *confident* and not let their audience know when they are nervous. Two respondents mentioned that the interpreter should *prepare* for the event to the best of their abilities and should have at least basic *knowledge of the interpreted* topic and three wrote about criteria beyond the interpreter's control, namely *good working conditions* and *working equipment*. Respondent 126 pointed out the *speed* of the interpreter's speech and the *consistency* of their output. Respondent 31 said the interpreter's *intonation* should be appropriate – she added she did not mind a slightly monotonous intonation, but was bothered by an exaggerated one. Appropriateness of style was mentioned by one respondent. Lastly, R216 suggested that rules or norms imposed on interpreters make them feel nervous and that the interpreters should *"be themselves"*.

Dependency of Criteria

58 respondents thought that the listed criteria were not constant, but rather dependent on the conference itself. By far the most commonly mentioned criterion was *correct terminology*. 28 TI students expressed in one way or another that the need for correct terminology was highly dependent on the expertise level of the conference, many of them claiming this criterion would be most important at a medical conference. Interestingly, nine respondents were of the opinion that the criteria were dependent on the working conditions or the physical/mental state of the interpreter, rather than the conference topic. Several respondents thought some of the formal criteria (such as lively intonation, no hesitation noises or fluency of delivery) depended on the level of formality of the conference or on the target audience, while a few others also wrote about the conference type and topic, one respondent even warning us to keep the formal features in mind when interpreting *Hviezdoslavov Kubín*, a popular Slovak tell-tale

contest for primary school children. Other respondents thought the 14 criteria were dependent, but did not specify how.

8. Non-TI students & gender:

	TI all N = 250	non-TI all N = 900	non-TI men n = 325	non-TI women n = 575
fluency of delivery	3.488	3.434	3.372	3.47
native accent	2.26	2.123	2.129	2.12
logical cohesion	**3.756**	3.487	3.385	3.544
correct terminology	3.556	3.43	3.406	3.443
completeness of delivery	3.344	3.327	3.178	3.410
correct grammar	**3.208**	2.922	**2.717**	3.038
sense consistency	3.716	3.528	3.483	3.553
pleasant voice	*2.52*	2.731	2.723	2.736
lively intonation	2.752	2.666	2.548	2.732
no filler words & hesitation noises	**3.108**	2.699	2.538	2.79
no booth noises	3.144	2.962	2.858	3.021
synchronicity with the speaker	*2.884*	2.946	2.889	2.977
clear articulation	3.516	3.397	3.36	3.417
confident voice	3.32	3.226	3.2	3.24
average value	3.184	3.063	2.985	3.178

Table 16: Criteria – TI/non-TI students; non-TI men & women

The average value given to the criteria by non-TI students was 3.063, which is 0.121 lower that the average value given by TI students (3.184). For non-TI students, the highest ranking criterion is **sense consistency** (3.528; TI #2), closely followed by **logical cohesion** (3.487; TI #1), **fluency of delivery** (3.434; TI #5), **correct terminology** (3.43; TI #3) and **clear articulation** (3.397; TI #4). By far the lowest ranking criterion is **native accent** (2.123; TI #14), with an average value lower by as much as 0.543 than the next lowest ranking criterion, **lively intonation** (2.666; TI #12). There are four criteria with significantly different values given to them by TI and non-TI students – **logical cohesion** (3.756 vs. 3.487), **correct grammar** (3.208 vs. 2.922), **pleasant voice** (2.52 vs. 2.731), and **no filler words & hesitation noises** (3.108 vs. 2.699). Interestingly, there are two criteria which are perceived as more important by non-TI students than those studying translation and interpreting – pleasant voice (2.731 vs. 2.52) and synchronicity with the speaker (2.946 vs. 2.884).

Only one significant difference can be found between men and women if compared to the whole non-TI group – men see **correct grammar** as less important than women (2.717 vs. 3.038). However, two more significant differences occur if we compare the two genders to each other – completeness of delivery and no filler words & hesitation noises are both seen as less important by men than by women (3.178 vs. 3.410; 2.538 vs. 2.79).
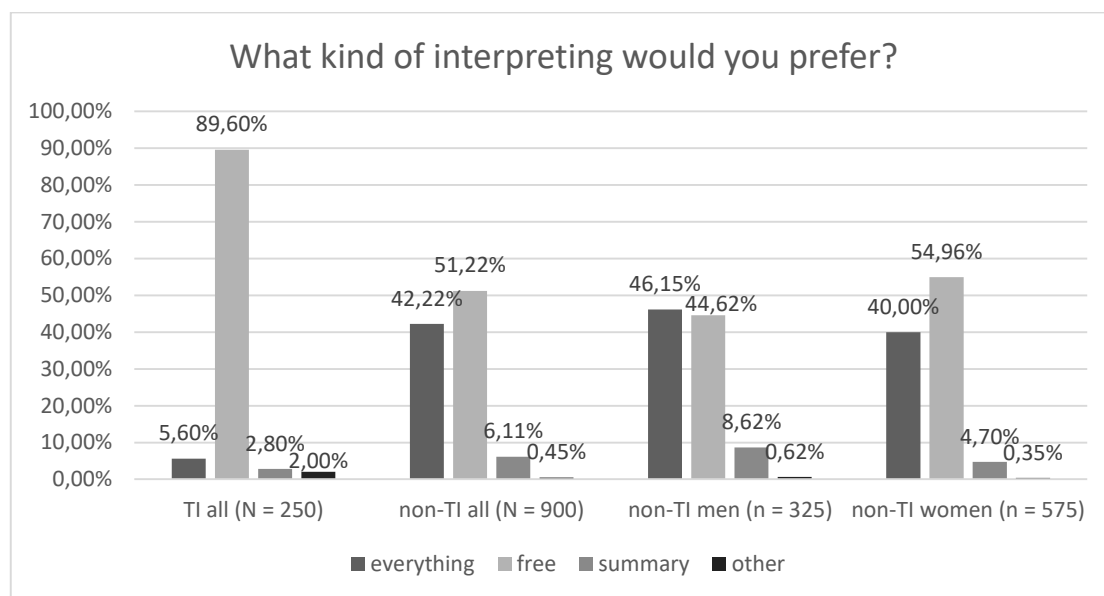
Figure 21: Interpreting styles – TI & non-TI students; men & women[67]

The difference between TI and non-TI students is remarkable, but not unexpected. Furthermore, we found out that men are more prone to wanting a full rendition of the ST than women, but also more likely to "only" request a summary of what was said (this was also true for TI students).

Four respondents chose the "other" option. Two men and one woman said the style of interpreting depends on the type of event and topic (R4, R559, R405) and one woman stated she would allow omissions of repeated information, but no additions (R245).

---

[67] The question for TI students was phrased slightly differently: "How does a good interpreter interpret?".
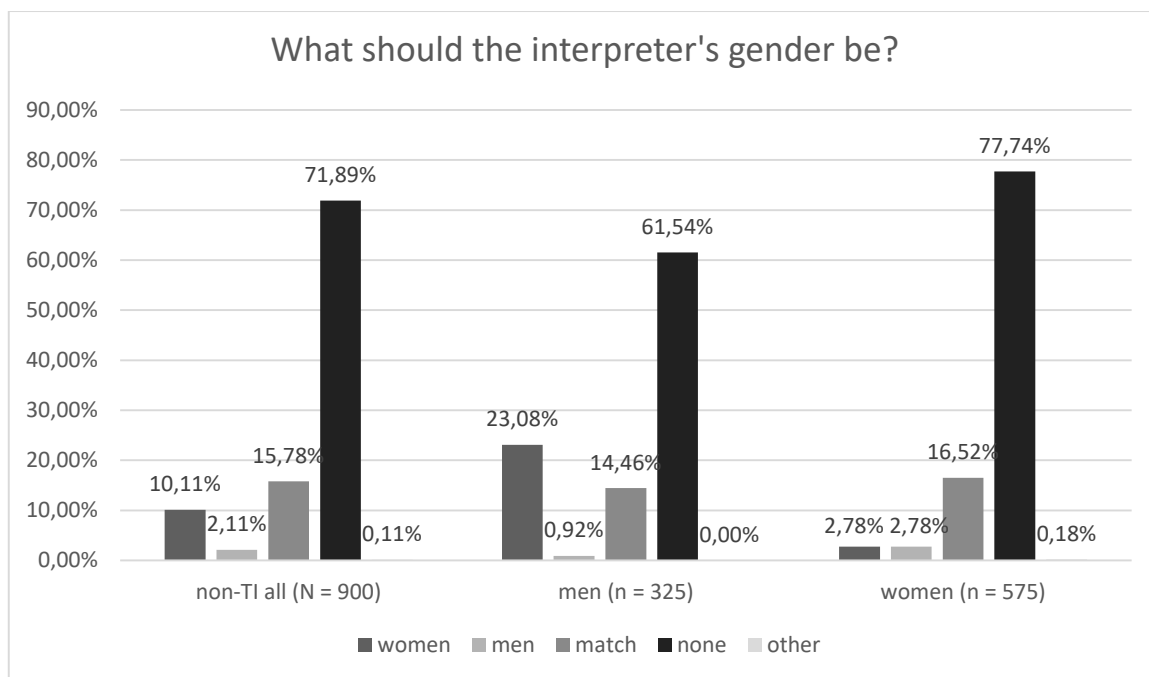
Figure 22: Gender preferences – non-TI students; men & women

This graph clearly demonstrates that men are more likely to have a gender preference (38.42%) than women (22.26%) and that they often prefer female interpreters (23.08%), while the opposite is not true for women (their preference for men is only marginally higher than men's). The speaker/interpreter gender match is also a popular option and the men/women ratios here are very similar. In the whole sample, there was only one "other" answer: R726 (woman, pharmacy, 5th year) stated that her preference depended on the voice of the interpreter.

9. Non-TI students & year of study

	non-TI all N = 900	1st year N = 181	2nd year N = 215	3rd year N = 214	4th year N = 157	5th year N = 110	PhD N = 20
fluency of delivery	3.434	3.442	3.414	3.416	3.446	3.436	3.7
native accent	2.123	2.149	2.028	2.206	2.153	2.055	2.2
logical cohesion	3.487	3.497	3.484	3.435	3.433	3.609	3.65
correct terminology	3.43	3.381	3.344	3.43	3.535	3.482	3.6
completeness of delivery	3.327	3.365	3.358	3.238	3.268	3.445	3.3

correct grammar	2.922	2.994	2.851	2.972	2.854	2.936	2.95
sense consistency	3.528	3.475	3.516	3.495	3.554	3.627	3.65
pleasant voice	2.731	2.702	2.758	2.841	2.624	2.664	2.65
lively intonation	2.666	2.696	2.730	2.598	2.701	2.6	2.45
no filler words & hesitation noises	2.699	2.669	2.647	2.762	2.682	2.8	2.5
no booth noises	2.962	2.812	2.963	2.995	3.006	3.045	3.05
synchronicity with the speaker	2.946	2.912	2.972	2.855	2.943	3.109	3
clear articulation	3.397	3.431	3.4	3.407	3.338	3.345	3.6
confident voice	3.226	3.282	3.256	3.22	3.185	3.164	3.5
average value	3.063	3.058	3.052	3.062	3.052	3.094	3.129

Table 17: Criteria – non-TI students in different years of study[68]

When split into groups by their year of study, no statistically or practically significant differences can be found among non-TI students. On the contrary, the groups' score are remarkably similar. The average values of the groups are as follows:

- 1st year: 3.058
- 2nd year: 3.052
- 3rd year: 3.062
- 4th year: 3.052
- 5th year: 3.094
- PhD: 3.129

As we can see, the biggest difference can be found between the scores of PhD students and the rest of the groups. However, as there are only 20 PhD students in our sample, this does not come as a surprise. Sixth-year students' responses are not analysed for the reason that they are not representative (there are only three members in this subgroup).

10. Non-TI students & experience with interpreted conferences:

The following table shows the values of three groups of non-TI students – those, who have never experienced interpreting at a conference (no exp.), those who have had experience with interpreted conferences (some exp.), and those who have had ample experience with interpreted conferences (ample exp.).

---

[68] Responses of students in year 6 are not analysed because this group has only three members.

	non-TI all N = 900	no exp. n = 721	some exp. n = 168	ample exp. n = 11
fluency of delivery	3.434	3.426	3.44	**3.909**
native accent	2.123	2.119	2.143	2.091
logical cohesion	3.487	3.465	3.554	**3.909**
correct terminology	3.43	3.422	3.446	3.727
completeness of delivery	3.327	3.330	3.321	3.182
correct grammar	2.922	2.936	2.857	3
sense consistency	3.528	3.521	3.542	3.727
pleasant voice	2.731	2.728	2.75	2.636
lively intonation	2.666	2.660	2.673	2.909
no filler words & hesitation noises	2.699	2.678	2.762	3.091
no booth noises	2.962	2.947	3.024	3
synchronicity with the speaker	2.946	2.938	2.952	**3.364**
clear articulation	3.397	3.404	3.357	3.545
confident voice	3.226	3.209	3.3	3.182
average value	3.063	3.056	3.080	3.234

Table 18: Criteria – non-TI students with different CI experience

From this table, we could say that strictness comes with experience, when we talk about expectations of simultaneous interpreting. However, it is also clear that the differences between the three groups, when compared to the whole sample of non-TI students, are insignificant. In fact, the only significant differences were found when we compared the values given by students with considerable experience with interpreted conferences to the whole sample (values with a difference of 0.4 point and more are shown in boldface). Nevertheless, this group is very small (11 members) and the set value for a significant difference for groups of 10-19 members (0.4 ≤) might perhaps be too lenient. Thus, we shall not consider this result generalisable and we conclude that while there seems to be a tendency for strictness with a growing experience with interpreted conferences, the differences are, indeed, insignificant.

11. Non-TI students with/without a preference for the interpreter's gender:

	non-TI all N = 900	existing preference n = 253	no preference n = 647
fluency of delivery	3.434	3.47	3.42
native accent	2.123	2.198	2.094
logical cohesion	3.487	3.435	3.507

correct terminology	3.43	3.403	3.44
completeness of delivery	3.327	3.3	3.337
correct grammar	2.922	2.893	2.934
sense consistency	3.528	3.522	3.53
pleasant voice	2.731	**2.949**	2.646
lively intonation	2.666	2.704	2.651
no filler words & hesitation noises	2.699	2.672	2.709
no booth noises	2.962	2.98	2.955
synchronicity with the speaker	2.946	2.897	2.964
clear articulation	3.397	3.372	3.406
confident voice	3.226	3.182	3.243
average value	3.063	3.07	3.06

Table 19: Criteria – non-TI students with (no) gender preferences

As the table clearly demonstrates, there is only one significant difference between these two groups, but it is nevertheless a very interesting one. The group with an existing preference for the interpreter's gender (in which 142 respondents said they want the gender of the interpreter and the speaker to match, and 91 stated they prefer female interpreters) sees the criterion of **pleasant voice** as more important than the group with no gender preference (2.949 vs. 2.646). The significance of this difference is plausible because the category of voice should theoretically be the only one changing with the gender of the interpreter (provided that their professional skills are equal).
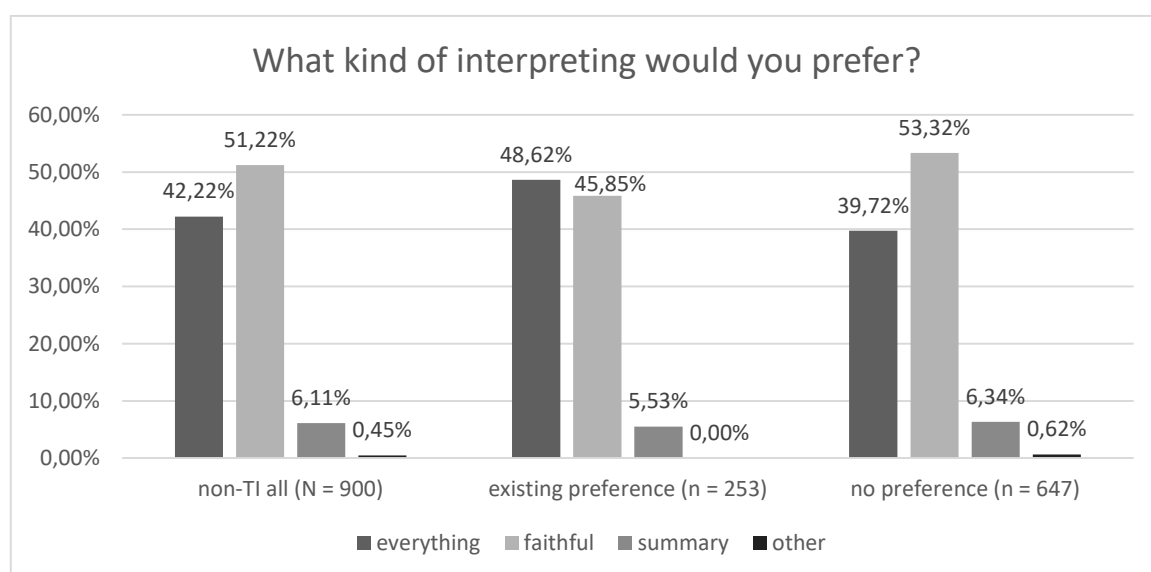

Figure 23: Interpreting styles – non-TI students with (no) gender preference

Students who are more lenient in terms of the interpreter's gender are also more lenient about the amount of information they receive. According to our results, students with no gender preference are 8% more likely to favour a more active role of the interpreter (either a free rendition or a summary of the ST).

12. Non-TI students & field of study:

	everything	free	summary	other
art/aesthetics (n = 33)	18.18%	**66.67%**	**15.15%**	0%
civil engineering (n = 31)	**51.61%**	41.94%	6.45%	0%
finance/economy/management (n = 80)	42.5%	56.25%	1.25%	0%
foreign languages (n = 32)	25%	**71.88%**	3.12%	0%
forestry/agriculture/wildlife management (n = 43)	41.86%	48.84%	9.3%	0%
healthcare (n = 89)	42.7%	50.56%	5.62%	1.12%
international relations (n = 46)	**52.17%**	41.31%	6.52%	0%
information technologies (n = 40)	47.5%	42.5%	7.5%	2.5%
journalism/media (n = 23)	39.13%	**60.87%**	0%	0%
law (n = 52)	**55.77%**	40.38%	3.85%	0%
medicine (n = 69)	**55.07%**	42.03%	2.9%	0%
natural sciences (n = 27)	**55.56%**	44.44%	0%	0%
physical education (n = 33)	30.30%	57.58%	**12.12%**	0%
pharmacy (n = 34)	38.24%	52.94%	8.82%	0%
political science (n = 30)	**50%**	36.67%	10%	3.33%
public administration (n = 27)	44.45%	44.45%	**11.11%**	0%
Slovak language (n = 33)	27.27%	**66.67%**	6.06%	0%
social work (n = 42)	33.33%	**59.53%**	7.14%	0%
special pedagogy (n = 20)	**60%**	35%	5%	0%
technical engineering (n = 39)	43.59%	46.15%	**10.26%**	0%
tourism (n = 33)	30.30%	**66.67%**	3.03%	0%
transport/logistics/postal services (n = 24)	25%	**58.33%**	**12.5%**	4.17%
non-TI all (N = 900)	42.22%	51.22%	6.11%	0.45%

Table 20: Interpreting styles – non-TI study groups

The table clearly demonstrates that there are differences between individual study fields. The seven strictest groups (i.e. the groups which demand full rendition the most) are: **special pedagogy** (60%), **law** (55.77%) **natural sciences** (55.56%),

medicine (55.07%), **international relations** (52.17%), **civil engineering** (51.61%), and **political science** (50%).

On the other hand, the seven groups most in favour of a free rather than a complete rendition are: **foreign languages** (71.88%), **art/aesthetics** (66.67%), **Slovak language** (66.67%), **tourism** (66.67%), **journalism/media** (60.87%), **social work** (59.53%), and **transport/logistics/postal services** (58.33%).

We also found five groups in which more than 10% students preferred a summary of the ST. These are: **art/aesthetics** (15.15%), **transport/logistics/postal services** (12.5%), **physical education** (12.12%), **public administration** (11.11%), and **technical engineering** (10.26%).

Nevertheless, there does not seem to be a specific reason for the differences, at least not one which is readily identifiable. Either both natural and social sciences or both highly academic and skill-based fields are represented in all groups while others are not, which makes us think that the prediction of a specific group's preferences based solely on the field of study is complicated, if not impossible. Further research would have to be conducted in order to gain more accurate and informative data.

13. Non-TI students of various academic fields – summary of criterion importance:

Sense consistency received the highest average score from all non-TI students (3.528). It was rated the most important criterion by 10 groups and placed second in the ratings of six groups. **Civil engineering** students rated it as the **ninth** most important criterion, which is the lowest rank it received. They were also the only group which gave this criterion an average score of less than 3 points (2.935). **Healthcare** students also differed significantly from the whole sample (3.281; #7). On the contrary, students of **foreign languages** perceived this criterion very strictly, giving it an average score of 3.844, followed by **law** students and their score of 3.769.

Logical cohesion was the second highest-scoring criterion, with a mean score of 3.487. Seven groups perceived it as the most important criterion, five groups put it in the second place, and four groups in the fourth place. **Forestry/agriculture/wildlife management** students gave it the lowest mean score (3.165, #4), followed by students of **physical education** (3.182; #7). Both of these scores were significantly different from the average score. The highest mean score, 3.769, came from the group of **law** students, followed by students of tourism (3.727) and **medicine** (3.696). However,

while the scores of law and medicine students are statistically different, the same cannot be said for students of tourism, due to the small number of members in their group.

Fluency of delivery received an average score of 3.434, which placed it in the third place. However, its most frequently assigned rank was #4 (eight groups). Seven groups saw it as the third most important criterion, and five groups put it in the fifth place. The highest rank for this criterion was #1 (civil engineers, 3.516), while the lowest was #6 (three groups). Students of **medicine** gave fluency of delivery the highest average score of all groups (3.638), which made them the only group with a significantly different opinion on this criterion. The lowest average score was 3.148 (public administration), although we do not consider it to be statistically significant.

Correct terminology very closely followed fluency of delivery, with a mean score of 3.43. It was most commonly seen as the third or fourth most important criterion (by seven groups in both cases). Two groups perceived it as the most important criterion – students of **political science** (3.7, the highest score) and students of forestry (although only the former group's score was significantly different). On the contrary, **art/aesthetics** (#6) students gave it the lowest score of 3.061 (which was also a significant difference).

Clear articulation placed fifth with an average score of 3.397 and it was also ranked as the fifth most important criterion by eight groups. Four groups placed it at #2 and four groups at #4. The highest rank for clear articulation was #1 (**special pedagogy** and **transport/logistics/postal services** students) while the lowest was #6 (three groups). There were no significant differences in the perception of this criterion's importance. It received the highest score from special pedagogy students (3.65), and the lowest (3.148) from students of public administration.

Completeness of delivery received an average score of 3.327, which put it at the sixth place. This is consistent with its ranking from individual groups – seven groups placed it at #6, followed by #3 and #7 (both five groups). The highest rank for this criterion was #2 (special pedagogy students) and the lowest rank was #8 (forestry and PE students). Students of **law** differed significantly from the whole sample and gave this criterion its highest score, 3.596. On the other hand, two groups gave completeness of delivery significantly lower scores – **forestry** and **public administration** students (the average score in both cases was exactly three points).

Confident voice was the lowest-ranking criterion with an average score of more than three points – 3.226. This put it in the seventh place, just like exactly half off the

22 groups. However, it ranked as high as #2 in the case of two groups – political science and PE students. **Political science** students also gave it the highest mean score – 3.633, closely followed by students of **special pedagogy** (3.6), both of these groups significantly stricter than the average. The score went below three points in two cases – **public administration** (2.815) and **art/aesthetics** (2.939) students – both significantly more lenient.

No booth noises were given an average score of 2.962. Despite placing eight according this score, its most common rank of this criterion was in fact #10 (six groups), followed by #9 (five groups), and only then #8 (four groups). Students were not united in their perception of this criterion, with as many as eight groups having significantly different opinions on it. Stricter scores came from students of **tourism** (3.394, #7), **physical education** (3.273, #6), **civil engineering** (3.226, #4), **social work** (3.214, #9/10), and **Slovak** (3.212, #9). More lenient scores were given to this criterion by respondents studying **public administration** (2.481, #11), **informational technologies** (2.575, #12), and **natural sciences** (2.63, #10).

Synchronicity with the speaker placed ninth with a mean score of 2.946. Most groups put it at #8 (eight), followed by #9 and #10 (six groups in both cases). Students of **international relations** and **social work** were significantly stricter in their perception of this criterion's importance (3.304 and 3.214, respectively). On the other hand, three groups did not perceive this criterion as important as the average respondent – students of **art/aesthetics** (2.333), **public administration** (2.519), and **foreign languages** (2.563) gave it significantly lower scores.

Correct grammar received an average score of 2.922 and most commonly placed as the ninth or the eleventh most important criterion (6 groups in both cases). Five groups put it at #10. Here, we can observe a very interesting situation – two groups ranked this criterion as significantly more important and both of them study languages at university. Students of **Slovak** gave correct grammar an average score of 3.242 (#8), while students of **foreign languages** a somewhat lower score of 3.219 (#6). The only other groups which gave this criterion a score of 3.1 and more were students of **journalism** (3.13, #8) and **special pedagogy** (3.1, #9), i.e. future professionals in fields where language is also very important. There was only one significantly lower score – 2.641 from **technical engineering** students (#11).

Pleasant voice placed eleventh according to its average score of 2.731, although it was most commonly ranked as #13 (seven groups) and #12 (five groups). Two groups

ranked it as #8 – IT and **social work** students – while the latter group also gave it a significantly higher score of 3.31. On the contrary, three study groups were significantly more lenient in their perception of this criterion's importance – students of **political science** (2.4, #13), **law** (2.462, #12), and **medicine** (2.507, #13).

No filler words & hesitation noises were given an average score of 2.699 and their ranking varied from #8 (special pedagogy) to #13 (nine groups). Five groups ranked it at #11. This criterion also came with considerably different values given to it by the respondents. Three study groups were significantly stricter than the whole sample – **special pedagogy** students gave it the highest score of 3.1, followed by students of **tourism** (3.03) and **Slovak** (2.97). There were also three lenient groups, namely students of **public administration** (2.074), **law** (2.385), and **technical engineering** (2.436).

Lively intonation closely followed its predecessor with 2.669 points, which put it on the 13th place, although it most commonly ranked as #12 (12 groups). Foreign language students ranked it highest (#8), even though theirs was not the highest average score (2.844 vs. 2.905 from students of social work). Only one significant difference was found in the scores, and that was the low score of **public administration** students (2.259).

Native accent was rated as the least important criterion of all, with an average score of 2.123 points. All groups unanimously ranked it at #14. This criterion received significantly higher scores from three groups of students – those studying **social work** (2.523), **foreign languages** (2.438), and **physical education** (2.424). Three groups' scores were below two points – 1.692 from students of **law**, 1.938 from finance/economy/management students, and 1.986 from students of medicine. However, only the first score counts as significantly lower than the average.

14. Non-TI students – other criteria:

35 respondents added their own criteria to the list. These could be split into several categories:

- at least a basic knowledge of the interpreted field (5 respondents);
- pleasant appearance (6 respondents);
- nice/kind personality (6 respondents);
- unbiased output, not expressing own opinion (4 respondents);

- appropriate usage of gestures (3 respondents);
- good personality match between the speaker and the interpreter (2 respondents);
- not being nervous (2 respondents);
- comprehensible output (2 respondents);
- good sound quality (2 respondents);
- political orientation (1 respondent);
- appropriate speed of interpreting (1 respondent);
- no speech impediments (1 respondent).

From the responses, it is clear that some respondents had consecutive interpreting in mind when filling in Questionnaire A (or its last part at the very least). It is interesting that three out of four respondents who demanded unbiased output were students of political science or journalism/media (the fourth one studies social work). Furthermore, a student of political science was also concerned about the interpreter's political orientation (R786).

One respondent (R173, international relations, PhD) also left a comment explaining why he sees lively intonation as an unimportant criterion – according to him, it is *"much easier to take notes or do something else"* while listening to interpreting with a monotonous intonation. We think this is an interesting opinion.

Appendix D. Questionnaires.

The questionnaires found on the following pages are in this order:

   1. the original questionnaire for users (listeners)

   2. the original questionnaire for interpreters

   3. the new questionnaire for non-TI students

   4. the new questionnaire for TI students

The following accompanying information was stated on a separate piece of paper attached to the questionnaires:

1. the original questionnaire for interpreters:

Vážení tlmočníci,

tento dotazník slúži ako výskum k mojej diplomovej práci, ktorá je zameraná na kvalitu tlmočenia a očakávania užívateľov od výkonov tlmočníkov. Prosím Vás o jeho vyplnenie, ktoré by Vám nemalo zabrať viac ako dve minúty. Výsledky dotazníka sú anonymné a budú použité len na účely diplomovej práce.

Ak máte pocit, že vami žiadaná odpoveď nie je v ponuke, prosím, neváhajte a dopíšte ju pod otázku.

Ďakujem.

Andrea Tokárová (Filozofická fakulta UMB)

2. the new questionnaire for non-TI students:

Vážení respondenti,

tento anonymný dotazník slúži na účely mojej diplomovej práce, v ktorej sa snažím zistiť, čo slovenskí študenti (a budúci profesionáli vo svojom odbore) považujú za dôležité pri tlmočení. Aj keď ste sa s tlmočením možno ešte nestretli, snažte sa, prosím, predstaviť si, čo by ste od neho očakávali.

Dotazník Vám nezaberie viac ako päť minút. Za jeho vyplnenie Vám vopred ďakujem.

Pokiaľ ste študentom prekladateľstva a tlmočníctva, tento dotazník, prosím, NEVYPĹŇAJTE. Ďakujem.

Andrea Tokárová (FF UMB)

(andrea.tokarova4@gmail.com)

<u>3. the new questionnaire for TI students</u>

Vážení respondenti,

*tento anonymný dotazník slúži na účely mojej diplomovej práce, v ktorej sa snažím zistiť, čo slovenskí študenti prekladateľstva a tlmočníctva považujú za dôležité pri **simultánnom (konferenčnom)** tlmočení.*

Dotazník Vám nezaberie viac ako päť minút. Za jeho vyplnenie Vám vopred ďakujem.

Ak máte akékoľvek poznámky k dotazníku, napíšte ich, prosím, na jeho druhú stranu. V prípade otázok, prosím, uveďte aj svoju e-mailovú adresu, aby som Vám na ne mohla odpovedať.

Andrea Tokárová (FF UMB)

(andrea.tokarova4@gmail.com)

Prieskum spokojnosti s tlmočníckymi službami

Pohlavie: □ muž   □ žena

Vek: □ 0-29 □ 30-45 □ 46-60 □ viac ako 60

Využili ste už niekedy tlmočnícke služby?   □ Nie.   □ Áno, niekoľkokrát.   □ Áno, veľakrát.

Rozumiete jazyku rečníka? □ Nie.   □ Len trochu.   □ Áno, dobre.

Tlmočenie dnes využívate: □ Vôbec.   □ Niekedy.   □ Počas celého prejavu zahraničného rečníka.

Aké by malo byť pohlavie tlmočníka?

□ Preferujem tlmočníčky.   □ Preferujem tlmočníkov.   □ Rovnaké ako pohlavie rečníka.

□ Na pohlaví mi nezáleží.   □ Iné: _____

Aké tlmočenie uprednostňujete?

□ Tlmočník pretlmočí všetko tak, ako to rečník povie, nič nepridáva ani nevynecháva.

□ Tlmočník tlmočí „verne", no môže pridať (vysvetliť) alebo vynechať informácie (napr. opakujúce sa).

□ Tlmočník len sumarizuje to, čo rečník povedal.

□ Iné:

Aké dôležité sú pre Vás nasledujúce kritériá pre poskytnutie kvalitného tlmočenia? (zakrúžkujte)

1 = nepodstatné; 4 = veľmi dôležité

1. plynulosť	1 – 2 – 3 – 4
2. rodný prízvuk	1 – 2 – 3 – 4
3. logická súdržnosť	1 – 2 – 3 – 4
4. správna terminológia	1 – 2 – 3 – 4
5. úplnosť prejavu	1 – 2 – 3 – 4
6. správna gramatika	1 – 2 – 3 – 4
7. významová zhoda s pôvodným prejavom	1 – 2 – 3 – 4
8. príjemný hlas	1 – 2 – 3 – 4
9. živá intonácia	1 – 2 – 3 – 4
10. nepoužívanie výplnkových slov („takže", „vlastne") a hezitačných zvukov („hmmm")	1 – 2 – 3 – 4
11. nerušenie tlmočenia zvukmi z kabíny (kašeľ, šuchotanie papiermi)	1 – 2 – 3 – 4
12. synchrónnosť s rečníkom (medzi tlmočníkom a rečníkom je krátky časový odstup)	1 – 2 – 3 – 4
13. čistá artikulácia	1 – 2 – 3 – 4
14. istota v hlase	1 – 2 – 3 – 4

Iné dôležité kritériá (vypíšte):

Ako hodnotíte dnešné tlmočenie? (zakrúžkujte; 1 = veľmi zlé; 5 = veľmi dobré)

1. celkový dojem	☹ 1 – 2 – 3 – 4 – 5 ☺
2. prednes (plynulosť, istota v hlase, …)	☹ 1 – 2 – 3 – 4 – 5 ☺
3. jazyk (gramatika, terminológia)	☹ 1 – 2 – 3 – 4 – 5 ☺
4. obsah (významová zhoda, logická súdržnosť)	☹ 1 – 2 – 3 – 4 – 5 ☺

Veľmi pekne Vám **ďakujem** za Váš čas. Ak máte akékoľvek poznámky k dotazníku alebo
k tlmočeniu, napíšte ich, prosím, na druhú stranu alebo priamo k otázkam.
Kontakt: andrea.tokarova4@gmail.com

Prieskum názorov o kvalite tlmočenia

Pohlavie: □ muž   □ žena

Vek: □ 0-29 □ 30-45 □ 46-60 □ viac ako 60

Ako dlho už aktívne tlmočíte (uveďte v rokoch)?   □ menej ako 5 □ 5-10 □ 10-20 □ viac ako 20

Aké sú vaše pracovné jazyky? Prosím, vypíšte ich a zaraďte do príslušnej kategórie A, B, C (A = rodný jazyk; B = jazyk, z ktorého a do ktorého tlmočíte; C = jazyk, z ktorého tlmočíte).

Špecializujete sa ako tlmočník na určitú oblasť?

Je podľa vás pohlavie tlmočníka dôležité? Ak áno, stručne popíšte, ako:

Ako by mal podľa vás tlmočník tlmočiť?

□ Tlmočník pretlmočí všetko tak, ako to rečník povie, nič nepridáva ani nevynecháva.

□ Tlmočník tlmočí „verne", no môže pridať (vysvetliť) alebo vynechať informácie (napr. opakujúce sa).

□ Tlmočník len sumarizuje to, čo rečník povedal.

□ Iné:

Aké dôležité sú pre Vás nasledujúce kritériá pre poskytnutie kvalitného tlmočenia? (zakrúžkujte)

1 = nepodstatné; 4 = veľmi dôležité

1. plynulosť	1 – 2 – 3 – 4
2. rodný prízvuk	1 – 2 – 3 – 4
3. logická súdržnosť	1 – 2 – 3 – 4
4. správna terminológia	1 – 2 – 3 – 4
5. úplnosť prejavu	1 – 2 – 3 – 4
6. správna gramatika	1 – 2 – 3 – 4
7. významová zhoda s pôvodným prejavom	1 – 2 – 3 – 4
8. príjemný hlas	1 – 2 – 3 – 4
9. živá intonácia	1 – 2 – 3 – 4
10. nepoužívanie výplnkových slov („takže", „vlastne") a hezitačných zvukov („hmmm")	1 – 2 – 3 – 4
11. nerušenie tlmočenia zvukmi z kabíny (kašeľ, šuchotanie papiermi)	1 – 2 – 3 – 4
12. synchrónnosť s rečníkom (medzi tlmočníkom a rečníkom je krátky časový odstup)	1 – 2 – 3 – 4
13. čistá artikulácia	1 – 2 – 3 – 4
14. istota v hlase	1 – 2 – 3 – 4

Iné dôležité kritériá (vypíšte):

Mení sa pre vás dôležitosť daných kritérií s typom podujatia alebo prejavu? Ak áno, stručne popíšte, ako:

Prosím, stručne zhodnoťte vaše dnešné tlmočenie a pracovné podmienky:

Ak máte akékoľvek poznámky k dotazníku, napíšte ich, prosím, sem alebo priamo k otázkam:

Veľmi pekne Vám ďakujem za Váš čas. V prípade akýchkoľvek otázok ma môžete kontaktovať

na e-mailovej adrese andrea.tokarova4@gmail.com.

Čo očakávam od kvalitného tlmočenia?

Pohlavie: □ muž   □ žena   **Vek:** _____

Uveďte, prosím, Váš študijný odbor a program. V prípade, že študujete vo viac než jednom študijnom programe, vyberte si, prosím, jeden: _____

V ktorom roku štúdia sa nachádzate? Ak ste študentom v medziročníku, uveďte, prosím, ktorý ročník dokončujete: □ 1. (Bc.)   □ 2. (Bc.)   □ 3. (Bc.)   □ 4. (Mgr./Ing.)   □ 5. (Mgr./Ing.)   □ 6.   □ PhD.

Zúčastnili ste sa už niekedy konferencie, kde ste využili tlmočnícke služby?

□ Nie.   □ Áno, niekoľkokrát.   □ Áno, veľakrát.

Predstavte si, že sa zúčastňujete medzinárodnej konferencie **vo Vašom odbore** a nerozumiete jazyku rečníka. K dispozícii sú tlmočnícke služby (simultánne tlmočenie cez slúchadlá). Odpovedzte, prosím, na nasledujúce otázky, týkajúce sa tohto hypotetického tlmočenia.

Aké by malo byť pohlavie tlmočníka?

□ Preferujem tlmočníčky.   □ Preferujem tlmočníkov.   □ Rovnaké ako pohlavie rečníka.

□ Na pohlaví mi nezáleží.   □ Iné: _____

Aké tlmočenie by ste uprednostnili?

□ Tlmočník pretlmočí všetko tak, ako to rečník povie, nič nepridáva ani nevynecháva.

□ Tlmočník tlmočí „verne", no môže pridať (vysvetliť) alebo vynechať informácie (napr. opakujúce sa).

□ Tlmočník len sumarizuje to, čo rečník povedal.

□ Iné:

Aké dôležité by pre Vás boli nasledujúce kritériá pre poskytnutie kvalitného tlmočenia? (zakrúžkujte)
<u>1 = nepodstatné; 4 = veľmi dôležité</u>

1. plynulosť	1 – 2 – 3 – 4
2. rodný prízvuk	1 – 2 – 3 – 4
3. logická súdržnosť	1 – 2 – 3 – 4
4. správna terminológia	1 – 2 – 3 – 4
5. úplnosť prejavu	1 – 2 – 3 – 4
6. správna gramatika	1 – 2 – 3 – 4
7. významová zhoda s pôvodným prejavom	1 – 2 – 3 – 4
8. príjemný hlas	1 – 2 – 3 – 4
9. živá intonácia	1 – 2 – 3 – 4
10. nepoužívanie výplnkových slov („takže", „vlastne") a hezitačných zvukov („hmmm")	1 – 2 – 3 – 4
11. nerušenie tlmočenia zvukmi z kabíny (kašeľ, šuchotanie papiermi)	1 – 2 – 3 – 4
12. synchrónnosť s rečníkom (medzi tlmočníkom a rečníkom je krátky časový odstup)	1 – 2 – 3 – 4
13. čistá artikulácia	1 – 2 – 3 – 4
14. istota v hlase	1 – 2 – 3 – 4

Boli by pre Vás dôležité aj nejaké iné kritériá? Ak áno, prosím, vypíšte ich:

Veľmi pekne Vám **ďakujem** za Váš čas. Ak máte akékoľvek poznámky k dotazníku, napíšte ich, prosím, na druhú stranu alebo priamo k otázkam (v prípade otázok nechajte aj svoj kontakt).

Čo pre mňa znamená kvalitné tlmočenie?

Pohlavie: □ muž   □ žena   **Vek:** ____

Uveďte, prosím, na ktorej univerzite študujete tlmočníctvo a prekladateľstvo: _____

V ktorom roku štúdia sa nachádzate? Ak ste študentom v medziročníku, uveďte, prosím, ktorý ročník

dokončujete: □ 1. (Bc.)   □ 2. (Bc.)   □ 3. (Bc.)   □ 4. (Mgr.)   □ 5. (Mgr.)   □ PhD.

Ktoré cudzie jazyky študujete? _____

Aké sú vaše skúsenosti so <u>simultánnym</u> tlmočením?

□ Zatiaľ žiadne nemám.   □ Mám skúsenosti len z hodín tlmočenia.

□ Simultánne som tlmočil v rámci povinnej praxe (pod záštitou univerzity).

□ Simultánne som už niekoľkokrát tlmočil aj vo vlastnej réžii (nie v rámci praxe).

□ Často chodím (simultánne) tlmočiť mimo školy.

□ Iné: _____

Chceli by ste sa v budúcnosti venovať tlmočeniu? □ Určite nie. □ Skôr nie. □ Možno áno. □ Určite áno.

Predstavte si, že simultánne tlmočíte **odbornú medzinárodnú konferenciu**. Odpovedzte, prosím, na nasledujúce otázky, týkajúce sa tohto hypotetického tlmočenia.

Ako podľa Vás tlmočí dobrý tlmočník?

□ Tlmočník pretlmočí všetko tak, ako to rečník povie, nič nepridáva ani nevynecháva.

□ Tlmočník tlmočí „verne", no môže pridať (vysvetliť) alebo vynechať informácie (napr. opakujúce sa).

□ Tlmočník len sumarizuje to, čo rečník povedal.

□ Iné:

Aké dôležité by pre Vás boli nasledujúce kritériá pre poskytnutie kvalitného tlmočenia? (zakrúžkujte)

<u>1 = nepodstatné; 4 = veľmi dôležité</u>

1. plynulosť	1 – 2 – 3 – 4
2. rodný prízvuk	1 – 2 – 3 – 4
3. logická súdržnosť	1 – 2 – 3 – 4
4. správna terminológia	1 – 2 – 3 – 4
5. úplnosť prejavu	1 – 2 – 3 – 4
6. správna gramatika	1 – 2 – 3 – 4
7. významová zhoda s pôvodným prejavom	1 – 2 – 3 – 4
8. príjemný hlas	1 – 2 – 3 – 4
9. živá intonácia	1 – 2 – 3 – 4
10. nepoužívanie výplnkových slov („takže", „vlastne") a hezitačných zvukov („hmmm")	1 – 2 – 3 – 4
11. nerušenie tlmočenia zvukmi z kabíny (kašeľ, šuchotanie papiermi)	1 – 2 – 3 – 4
12. synchrónnosť s rečníkom (medzi tlmočníkom a rečníkom je krátky časový odstup)	1 – 2 – 3 – 4
13. čistá artikulácia	1 – 2 – 3 – 4
14. istota v hlase	1 – 2 – 3 – 4

Boli by pre Vás dôležité aj nejaké iné kritériá? Ak áno, prosím, vypíšte ich:

Myslíte, že dôležitosť jednotlivých kritérií by sa menila v závislosti od témy konferencie?
Ak áno, stručne popíšte, ako:

Appendix E. The Evaluation Form.

Vážení respondenti,

prosím, vypočujte si krátke (5 min.) tlmočenie o placebo efekte a priraďte tlmočníčke body za nasledujúce kritériá. Za každé kritérium môžete dať minimálne jeden bod (ak podľa Vás nebolo splnené) a maximálne päť (ak bolo veľmi dobre splnené). Príslušné body napíšte číslicou do pravého stĺpca. Ďakujem.

KRITÉRIUM	POČET BODOV (min. 1, max. 5)
plynulosť	
rodný prízvuk	
logická súdržnosť	
správna terminológia	
úplnosť prejavu	
správna gramatika	
významová zhoda s pôvodným prejavom	
príjemný hlas	
živá intonácia	
nepoužívanie výplnkových slov (napr. „takže", „vlastne") a hezitačných zvukov („hmmm")	
nerušenie tlmočenia zvukmi z kabíny (kašeľ, šuchotanie papiermi)	
synchrónnosť s rečníkom (medzi tlmočníkom a rečníkom je krátky časový odstup)	
čistá artikulácia	
istota v hlase	

Celkový dojem z tlmočenia (uveďte číslom na stupnici od 1 do 10, kde 1 = veľmi zlý dojem a 10 = výborný dojem):

Priestor na komentáre (nepovinné):

Appendix F. Glossary for the recording & ST and TT transcripts.

Glossary given to students before interpreting:

complementary therapies – doplnkové liečebné metódy
double-blind trial – dvojito slepá skúška
efficacy – účinnosť
conventional treatment – štandardná liečebná metóda
branding – značka, branding
viral infections – vírusové ochorenia
sham treatment – predstieraná/fingovaná liečba
The Royal College of GP's – profesný organ britských všeobecných lekárov

Key:

- … = short pause (< 3s)
- …… = long pause (3s <)
- @ = short hesitation noise
- significant omissions (marked in the ST)
- significant additions (marked in the TT)
- significant content errors (marked in both the ST and the TT)
- other errors (e.g. incorrect pronunciation or grammar, marked in the TT)

ST	TT
Ladies and gentlemen, probably like me, as a child you fell over and grazed your knee and your mother or father picked you up and gave you a kiss and said: "There, there." And that is the ultimate placebo effect, because it makes you feel better, even though there's nothing in the kiss that can actually cure you or heal you and it's the expectation of getting better that actually gets rid of the symptoms.	Dámy a páni…... Možno tak, ako ja, vo svojom detstve... ste niekedy spadli a… mama alebo oco vás zobral na ruky a pobozkal vás. Aj toto je taký placebo efekt. V bozku nie je nič, čo by vám mohlo pomôcť zotaviť sa, ... ale je to iba to očakávanie, vy očakávate, že sa váš stav zlepší...

Many people are very dismissive of the placebo effect and in the context of complimentary therapies, for example aromatherapy or homeopathy, they will say things like: "There's no scientific evidence that it works, there haven't been any double-blind trials, there's no proof of efficacy." And then they'll go on perhaps to add something like: "So, if it works, it's just because of the placebo effect." Well, if people feel better and if the treatment is relatively cheap and it doesn't cause any harm – because we all know that medication can cause side effects, which are sometimes harmful – if all these conditions are fulfilled, then, surely, that's a good thing! Possibly in these cases, a placebo would even be better than a conventional treatment.

And there've been a lot of studies on the placebo effect, it's well researched, it's known that the placebo effect can work even if you know that you are receiving a placebo, and that the size, the colour, and the branding of a pill, for example, that you are given, can influence its effectiveness. And I think that shows just how powerful the mind is, how powerful suggestion can be in affecting our health.

Now, a new study has been published, by the universities of Oxford and Southampton, with

M-mnoho ľudí placebo efektu neverí…… @ Neverí aromaterapii alebo homeopatii á hovoria o tom, že žiadne… žiadne fakty nepotvrdzujú, že tieto metódy skutočne fungujú... A teda ak to predsa len funguje, tak musí tam byť nejaký placebo efekt…… Ak sa ľudia cítia lepšie á metóda je rl-relatívne lacná, pretože... vieme, že @ rôzne metódy môžu spôsobiť rôzne vedľajšie účinky, a ak táto metóda tieto vedľajšie účinky nemá, môžeme povedať, že @ placebo by mohlo byť aj lepšie ako taká štandardná liečebná metóda...

Už bolo spravených mnoho štúdií o placebo efekte a je dokázané, že aj keď vy viete o tom, že dostávate nejaké placebo, vaša- váš zdravotný stav sa môže zlepšiť... Napríklad aj farba pilulky, ktorú dostávate, môže ovplyvniť, či funguje, alebo nie...... Je to spojené s tým, že vaša myseľ pracuje @ s farbami a... teda toto placebo vám môže pomôcť aj bez toho, aby malo nejakú lieč- nejakú liečebnú látku.......

...... @ ... @ V Británii napríklad zo sedemsto osemdesia...tich...troch

interesting figures. It shows that 97% of 783 doctors, GPs, admitted recommending a sugar pill to a patient, or a treatment with no established efficacy. For example, a food supplement or probiotics, you know those little drinks that look like yoghurt, things like Yakult, and one in ten GPs said that they had given a sugar pill or even an injection of salty water as a placebo at some time in their career. Indeed, one in a hundred said they did so more than once a week!

And the reason they offer these placebos is typically to reassure a patient or because the patient requests treatment. And a prime example of this, actually, is people who go to the doctor with a cold and they say: "Doctor, I want you to give me antibiotics." And we all know, or we should know, that antibiotics don't work against viral infections, there's no point taking antibiotics if you have a cold, but still people ask for them. And sometimes doctors give them.

And the Royal College of GPs says that this type of sham treatment may be inappropriate or it could cause drug resistance, but apart from that category, in general, the Royal college of GPs says that there is a place for placebos in medicine. And as I was saying at the

lekárov 97% odporúča @ užívanie probiotík, alebó aj nejakého placeba...... a m-mnoho lekárov hovorí o tom, že niekedy dalo ľuďom injekciu s cukrovou alebo so slanou vodou svojim pacientom... A... okolo sto doktorov toto robilo dokonca raz týždenne...

Robia to preto, aby uistili pacienta, že mu niečo podali; dokonca niektorí ľudia si vyžadujú, aby im doktor niečo podal. Mnoho ľudí má napríklad len nádchu a príde k doktorovi a povedia: „Pán doktor, prosím vás, dajte mi antibiotiká," a doktor im vysvetľuje, že na antibakteriálne infekcie antibiotiká nezaberajú...... No a práve preto im niekedy dajú takúto predstieranú liečbu...

Niektorí lekári sa ale domnievajú, žé... to nie je správne... Ale mnoho lekárov sa domnieva, že placebo má svoje miesto v medicíne... Sú lacné, predsa zažívame ekonomickú krízu, a môžu byť

beginning, placebos can sound like a very good solution. They're cheap, after all this is an economic crisis, they can be very effective, they can make people feel a lot better. But this is also the era in the UK of patient choice and of empowerment. And being prescribed a placebo by a doctor without your knowledge goes against this trend. Some people think it can harm the doctor-patient relationship, it can harm that relationship of trust that you have with your doctor. And, personally, I would feel uneasy at the thought of a doctor giving me a placebo without my knowledge. As if he knew better than I did what decisions to take. As I see it, the doctor's there to give advice, but not necessarily to take all your decisions on your behalf. But perhaps that's just because I'm a control freak. Thank you.	veľmi efektívne, môžu spôsobiť, že sa cítite oveľa lepšie...... Ale v Spojenom kráľovstve teraz prežívame éru, kedy si doktori môžu... kedy- pardon, kedy si pacienti môžu vybrať, akú liečbu chcú. A preto, do-doktor vám musí povedať, že vám podáva placebo... Keď vám to p-, keď vám to nepovie, predsa už nemusíte svojmu doktorovi tak dôverovať...... @ Predsa ja si myslím, že doktori by vám mali podávať rady a nemali by za vás robiť všetky rozhodnutia, preto si myslím, že... mali by ste byť informovaní.

Table 21: ST and TT transcripts

Appendix G. Recording assessment.

Assessor #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Int.
fluency of delivery	4	5	3	2	4	5	4	3	4	4	4	3	3	3	3	3	4	4	3	2	2.5
native accent	5	5	4	4	5	5	5	5	5	4	5	4	4	5	5	4	5	5	3	5	5
logical cohesion	4	5	5	4	4	4	4	4	4	4	5	3	3	3	5	4	5	5	4	2	4
correct terminology	2	5	4	4	5	5	5	3	3	5	4	3	3	4	5	3	5	5	4	2	4
completeness of delivery	3	5	3	3	4	5	4	3	4	4	4	4	3	3	3	3	5	4	3	1	3
correct grammar	4	5	4	3	4	5	4	3	4	4	5	4	4	5	5	4	5	5	4	3	4
sense consistency	4	5	4	4	4	n/a	n/a	4	4	5	5	3	3	3	4	4	5	5	4	3	4
pleasant voice	5	5	4	3	4	5	5	4	5	4	4	5	3	4	2	5	4	5	5	4	n/a
lively intonation	5	5	3	4	3	5	4	3	4	4	4	3	3	2	2	4	5	4	4	5	5
no filler words & hesitation noises	5	4	4	4	5	4	4	3	5	3	5	4	4	3	5	3	4	4	4	5	4
no booth noises	5	4	5	5	5	5	5	4	5	5	5	5	5	5	4	4	5	4	5	4	5
synchronicity with the speaker	3	3	2	2	4	4	4	4	4	3	4	3	4	3	3	3	4	5	3	1	3
clear articulation	5	4	5	4	5	5	5	4	5	5	5	5	4	5	5	3	5	5	5	4	5
confident voice	5	5	4	4	4	5	4	3	5	3	3	2	4	3	2	5	5	5	4	2	4
overall impression	7	9	8	6	8	9	9	8	9	8	8	7	6	6	7	8	9	9	8	3	6

Table 22: Scores given by the 20 assessors and the interpreter

Verbal evaluations and other comments (quoted as they were originally written, including grammatical mistakes):

Assessor #1:

„Tlmočníčka má príjemný prejav, jej hlas sa dobre počúva, a aj keď si necháva dlhšiu dekaláž vie na seba jednotlivé vety logicky napojiť. Vety dokončuje, nepoužíva takmer žiadne výplnkové slová. Až na terminologické nedostatky, ktoré som si istá, že by v reálnej situácii boli vyriešené riadnou prípravou na tlmočenie, pôsobí nahrávka profesionálne."

Assessor #2:

„Vzhľadom na to, že tlmočníctvo študujem, viem si dlhšiu dekaláž odôvodniť a neprekáža mi, prejav sa mi počúval dobre, významovo sedel a s tlmočením som bola spokojná. Ale keby sa na to pozerám z laického hľadiska, je veľmi pravdepodobné, že by mi dlhšia dekaláž trochu prekážala, keďže to vyvoláva dojem, že je z prejavu mnohé vynechané. Hlavne, pokiaľ rečník/čka rozprávajú takmer neustále."

Assessor #3:

„Zatiaľ čo sa tlmočníčka snažila zachytiť rečníka/rečníčku, prerušila tlmočenie na dlhšiu chvíľu, než je vo väčšine prípadov prípustné. Taktiež nebolo veľmi špecifikované zakončenie prejavu. V niektorých úsekoch prejavu tlmočníčky počuť rodný prízvuk, zmäkčovanie /ďi/, /ťi/, /ňi/, /ľi/. Terminologická chyba nastala (aspoň sa logicky domnievam) pri výraze „antibakteriálne infekcie". Gramatika je miestami príliš „poangličtená", no to je (podľa môjho názoru) často súčasťou nedokonalého charakteru tlmočenia. Tlmočníčka má hlas, ktorý nie je sám o sebe veľmi živý, čo sa týka intonácie, no nie je ani úplne fádny, má príjemnú farbu hlasu a pekne artikuluje. Niekedy znie trochu neisto, no pri chybe sa dokázala pohotovo a správne opraviť. Výborne sa kontroluje pri používaní hezitačných zvukov a doplnkových slov, ktoré som nezachytila takmer žiadne, resp. minimum. Taktiež nenastávajú žiadne rušivé momenty z prostredia kabíny."

Assessor #4:

„páčila sa mi istota v hlase tlmočníčky, na druhej strane ma dosť rušili hluché miesta, ktoré narúšali plynulosť prejavu a pôsobilo to, že tlmočníčka

153

nepretlmočila všetky vety, ktoré odzneli v origináli a nebolo počuť originálny prejav, takže som nevedela s istotou vyplniť dve políčka.“

Assessor #5:

„Tlmočenie sa veľmi dobre počúvalo, a aj napriek pár zaváhaniam bola podstata vystihnutá a zrozumiteľná.“

Assessor #8:

„Dojem z tlmočenia bol príjemný. Vadili mi niektoré vynechané pasáže a ticho v niektorých častiach. Hlas bol príjemný, ale niekedy chýbala istota v hlase.“

Assessor #9:

„Veľmi dobré tlmočenie!“

Assessor #12:

„Na základe nahrávky je komplikované ohodnotiť významovú zhodu s pôvodným prejavom.“

Assessor #13:

„slovné spojenie antibakteriálne infekcie (infekcie proti(anti-) baktériám?), správne je nebakteriálne (t.j. vírusové, protozoárne) infekcie

-možnú logiku spojenia antibakteriálne infekcie vidím v súvislosti s bakteriofágmi, čo sú vírusy infikujúce a usmrcujúce bakteriálnu bunku (teda pôsobia anti-bakteriálne), avšak s takýmto výrazom sa nestretneme ani v mikrobiológii, pôsobí to zavádzajúco, no pripúšťam, že možno v budúcnosti, keď bakteriálna rezistencia vyeskaluje do kritických rozmerov, bude terapia s využitím vírusov jednou z mála možností boja proti patogénnym druhom baktérií“

Assessor #14:

„Z prejavu bola zrozumiteľná hlavná myšlienka, avšak niektoré myšlienky prednášajúceho neboli dokončené, resp. sa stratili v preklade (napríklad farba tabletky a dopad na pacienta), zo záveru napríklad nebol jasný postoj prednášajúceho k efektu placeba a povinnosti lekára informovať o liečbe. Niektoré tvrdenia si vyložene odporovali alebo nedávali zmysel.“

Assessor #20:

„Som farmaceut a vyznám sa v problematike danej témy. Predpokladám že prednáška, ktorú tlmočníčka tlmočila bola určená širokej verejnosti alebo laikom a nie odborníkom. Ja ako poslucháč by som bola sklamaná, keby som mala počúvať toto tlmočenie. Radšej by som počúvala rečníka. Veta v tlmočení: „...na antibakteriálne infekcie antibiotiká nezaberajú...“ ČO? :D sú buď bakteriálne alebo nebakteriálne (vírusové) infekcie. Keby toto tlmočenie počúval odborník, bol by v celku zmätený.“

Interpreter:

„Tlmočilo sa mi celkom dobre, bola som oddýchnutá a prekvapivo, napriek tomu, že som tlmočenie tohto prejavu neočakávala, nebola som v strese. Možno to bolo tým, že už som bola rozcvičená a mali sme za sebou už niekoľko minút tlmočenia.

Nie som spokojná so svojou dekalážou, myslím, že bola moc dlhá, pretože som sa v problematike nevyznala a bála som sa anticipovať. Táto dlhá dekaláž spôsobila, že tlmočenie nie ju úplne plynulé. V niektorých pasážach je v hlase počuť nervozitu (hlavne na miestach, kde som si dekaláž nechala až príliš dlhú a snažila som sa dobehnúť rečníka).

Som spokojná s tým, že v nahrávke nie je veľa hezitačných zvukov (kedysi som s tým mala veľký problém).

Ohodnotila som sa na 6 z 10, pretože po obsahovej stránke som so sebou celkom spokojná, ale neplynulosť prejavu výrazne zhoršuje celkový dojem.“