Defining the Elusive –

Interpreting Quality Research and Its Pragmatic Relevance.

Eine dolmetschwissenschaftliche Arbeit in englischer Sprache.

Masterarbeit

im Masterstudiengang Konferenzdolmetschen.

Angefertigt an der Fachhochschule Köln,
Institut für Translation und Mehrsprachige Kommunikation (ITMK)

Betreuerin: Prof. Dr. Sylvia Kalina
Zweitprüferin: Dorothea Dahl

Vorgelegt von Anne-Kristin Krämer aus Bochum

Datum der Abgabe: 28.12.2006

Index

Table of illustrations

List of abbreviations

CI: conference interpreting

EVS: ear-voice span

IR: interpreting research

MI: media interpreting

QA: quality assurance

SCL: skin conductance level

SI: simultaneous interpreting

SL: source language

ST: source text

TC: Training Committee (AIIC)

TT: target text

Introduction

It was in my first semester of the MA course in Conference Interpreting at the Cologne University of Applied Sciences that I was assigned the task of giving a presentation about two of Barbara Moser-Mercer's articles: "Quality in interpreting: some methodological issues" (Moser-Mercer 1996) and "Searching to define expertise in interpreting" (Moser-Mercer et al. 2000). They both centred on the notion of quality in interpreting, and I was struck by two aspects.

The 1996 article dealt with basic methodological issues (e.g. the advantages of structured versus unstructured interviews) and would have been unlikely to be published in any of the traditional research disciplines, the natural sciences for example. In her paper, Moser-Mercer elaborates on advantages and disadvantages of certain research methods. When reading it, I realised that the discipline of interpreting research (IR) is still a very young one which seems to be struggling with methodological problems that other disciplines have long overcome, or that other disciplines are not even faced with.

The other aspect that came to my mind was that whilst it is important to define optimum quality – and Moser-Mercer's definition as "complete and accurate rendition of the original" (Moser-Mercer 1996: 44) is frequently quoted in the literature – the pragmatic implications need to be looked at, too. I was curious to learn more about aspects that interpreting researchers had delved into, and eager to compare these empirical findings with my own experience and views.

The thesis deals with the issue of quality in interpreting. As its scope is limited, I will concentrate on the most frequent – and most widely researched – form of interpreting, which is simultaneous conference interpreting. So both community interpreting – as the other form of interpreting to conference interpreting – and consecutive interpreting as a subcategory of conference interpreting will not be covered.

In this thesis, I will give a comprehensive overview on research about quality in SI, the pragmatic relevance of such research, and methodological issues that the research community is faced with.

1 Interpreting quality research: a theoretical framework

1.1 Defining the elusive

> We live in an era obsessed with quality control.
> *(Kahane 2000)*

As Kurz points out, Total Quality Management, benchmarking, best of class, best practice sharing, and process optimisation are some of the key words that are omnipresent in today's business world (Kurz 2001: 394 f.), and their omnipresence serves to show that the quality of products and services is subject to intensive scrutiny (e.g. Mack 2002: 110; Chiaro & Nocella 2004: 278). From my own experience as a staff interpreter for the German subsidiary of a US-American company I could add many more of these buzzwords: KPI, cost control, Balanced Scorecard – and the list goes on. If products and services are put to the efficiency and quality test, why should the interpreting profession remain untouched by this development? There is ever-increasing cost pressure. As English is becoming more and more the *lingua franca* of our times, high quality standards gain even higher importance, with interpretation being a luxury item (Shlesinger et al. 1997: 131). These factors as well as higher demand and tougher competition give rise to the question of how interpreting quality can be measured.

Quality assurance has become a matter of the profession's reputation as well as a basis for assuring good working conditions and adequate pay since those who pay for interpreting services have a justified interest in getting good quality (Kalina 2005: 769). Both professional interpreters and clients should therefore be interested in QA (Pöchhacker 1994: 246). But what is ideal quality?

> With experience, you learn to tell the difference between quite
> good, very good and excellent interpreters.
> *(Gebhard 1999 as quoted by Kalina 2004: 3)*

The European Organization for Quality Control defines quality as “the totality of features and characteristics of a product or service that bear on its ability to satisfy a given need” (Wenger 1981: 24). As to the concept of quality in interpreting, there are various references to its elusiveness: It is hard to grasp and difficult to define (e.g. Shlesinger et al 1997: 122). Ackermann compares the notion of interpreting quality with the Loch Ness Monster – because of its "elusiveness and the fact that once seen both are immediately recognizable" (Ackermann et al. 1997: 262). The only point on which there seems to be general agreement is that quality may mean very different things to different persons (Kalina 2005: 771). Or, even if there is a certain consensus on quality recognising its importance, this consensus lacks substance (Kahane 2000). Also, there seems to be a "gap between ideal (academic) quality and situated (real-world) quality" (Straniero Sergio 2003: 135). Literature on SI tends to express quality in terms of an ideal world and falls prey to the wishful thinking syndrome (Straniero Sergio 2003: 171).

Interestingly, as Barbara Moser-Mercer rightly points out, the notion of interpreting has not been defined whereas there exists a commonly accepted definition for the (conference) interpreter as a person who "is a qualified specialist in bilingual or multilingual communication. He/she[1] makes this communication possible between delegates of different linguistic communities … by comprehending the concepts of the speaker's message and conveying them orally in another language. … Conference interpreters are moreover bound to respect the code of professional ethics, including the strictest professional secrecy" (AIIC 1994 as quoted by Moser-Mercer 1996: 44[2]).

An early quality definition from the late 1970s sees good simultaneous interpreting as "the art of re-expressing in one language a message delivered in another language at the same time as it is being delivered; the re-expression should be clear, unambiguous and immediately comprehensible, that is to say perfectly idiomatic, so that the listener does not have to mentally re-interpret what reaches him through the earphones" (Namy 1978: 26). Another definition, frequently quoted, is that by Barbara Moser-Mercer: "Optimum quality in professional interpreting implies that an interpreter provides a complete and accurate rendition of the original that does not distort the original message and tries to capture any and all extralinguistic information that the speaker might have provided subject to the constraints imposed by certain external conditions" (Moser-Mercer 1996: 44).

---

[1] Apart from literal quotations, I have refrained from using *he/she* or *his/her* expressions – for the sake of readability.

[2] Moser-Mercer does not mention the source text in her list of references which is why I could not include it in my bibliography.

Quality in interpreting means that more than just the words are rendered in the target language. According to Seleskovitch and her *théorie du sens*, or deverbalisation principle, “interpretation focuses on the ideas expressed in live utterances rather than on the language itself” (Seleskovitch 1976: 93). The interpreter has to “[re-express] the original thought in terms that are independent from the original wording” (Seleskovitch 1976: 113); he should grasp the sense, and render the ideas behind the words (Seleskovitch 1978: 341).

Some researchers see compression and abstraction as quality parameters. According to Gile, abstracting and compressing the source message causes a risk of information loss but also the possibility of converting a linguistically inadequate source speech into a high-quality target speech, “une transformation qualitative du discours susceptible de renforcer son impact” (Gile 1983: 239). Alexieva claims that the optimal way to ensure a successful communication act is to “produce a TL text of a greater semantic load and of a shorter length, i.e. to render the information volume of the SL text with the optimum degree of compression” (Alexieva 1983: 234). According to Levý, all interpreters should strive to deliver the maximum effect with a minimum of effort (Levý 1969: 37 as quoted by Viezzi 1993: 396).

One could follow Reiss’s theory of translation and state that an interpreter’s performance is good if it serves its purpose, adequacy being the decisive criterion applied (Reiss 1983: 208), an idea which is taken up by Bühler who claims that an ideal interpreter is “one who supplies an ‘ideal interpretation’ in a given situation for a given purpose” (Bühler 1986: 233). Interpreting is a service to those who do not understand a message delivered in the original language. Therefore, quality must not be seen in the narrow and abstract sense as something only inherent in the message delivered (Moser-Mercer 1996: 45). Pöchhacker thus suggests taking a product and service-oriented approach (Pöchhacker 2001: 413).

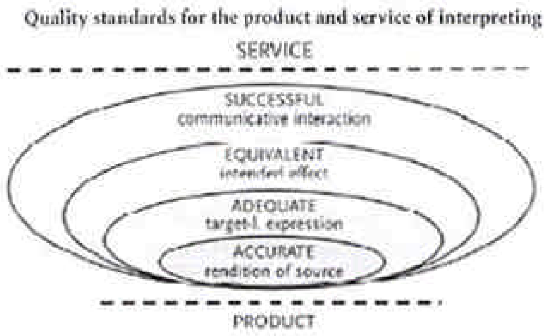

Fig. 1: Quality standards for the product and service of interpreting (Pöchhacker 2001: 413)

Déjean le Féal focuses on the effect that the interpretation creates in the users' minds: "What our listeners receive through their earphones should produce the same effect on them as the original speech does on the speaker's audience. Its language and oratory quality should be at least on the same level as that of the original speech, if not better, given that we are professional communicators while many speakers are not" (Déjean le Féal 1990: 155). Vuorikoski follows this line of thought and postulates that studying interpreting quality means to examine to what extent the substance of the ST argumentation, including the speech act of the original speech, has been conveyed in such a way as to allow listeners to create an impression which is equal to the impression they would have created had they been listening to the original – the effect of the TT on those listening to interpretation should be identical to the effect on those listening to the original (Vuorikoski 2004: 71). One has to bear in mind, however, that the effect might be different for other reasons of interpreting, i.e. on cultural grounds (Kalina 2005: 774).

Shlesinger raises the issue of perspectives from which quality can be assessed and asks: "Quality for whom? Who should judge quality" (Shlesinger et al. 1997: 126)? It seems that one has to look at different perspectives, e.g. that of the organiser of the conference, of the target-text recipient, of the interpreter himself, of the interpreter's colleagues, of trainers/teachers, of the interpreting researcher (e.g. Moser-Mercer 1996: 46; Kutz 1997: 243; Kutz 2000: 192 f.; Pöchhacker 2001: 412). Regarding the market perspective, Shlesinger asks the provocative question: "Do our clients know what's good for them? What do they expect, and what will make them happy with the service and product we provide" (Shlesinger et al. 1997: 126)? She also refers to the research perspective, asking what kind of research is needed in order to gain a better appreciation of the strengths and weaknesses of an interpreter's performance. A question from the didactic perspective might be: What aspects of the interpreter's output should be stressed by those of us who teach it? How can the skills which add up to quality be taught more effectively (ibid.)?

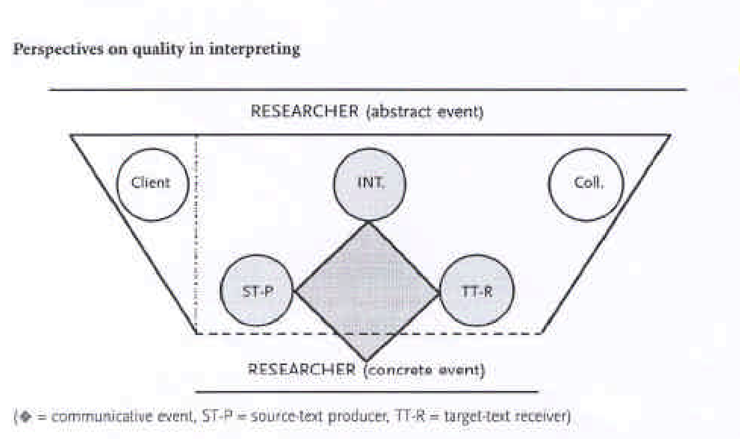

Fig. 2: Perspectives on quality in interpreting (Pöchhacker 2001: 412)

According to Pöchhacker, interpreting research can investigate the recordable product or the overall process of communicative interaction: one distinguishes between product-orientation and interaction-orientation. These two perspectives are crucial to the key issues of quality standards and assessment criteria (Pöchhacker 2001: 412).

Kopczynski underlines that the notion of quality can be approached from two angles, namely the linguistic and the pragmatic one. Linguistically, quality in interpretation is defined by applying rigid standards and criteria such as equivalence, congruence and correspondence. Quality from a pragmatic perspective is not an absolute value but dependent on certain situational variables, thus requiring different solutions and priorities in different situations (Kopczynski 2004: 87 f.).

So if interpreting quality is a (multi-)dependent variable, not only different perspectives have to be looked at but also different factors that can affect an interpreter's performance since quality is a relative concept (e.g. Kopczynski 1994: 88; Riccardi 2003: 256; Kalina 2005: 778). Kalina postulates that no two interpretations will ever occur under the same conditions as to direction of interpretation, producers of source texts, recipients of target texts, situation, etc. (Kalina 1998: 129 f.). There are many aspects that come into play: physical and mental factors, subject knowledge, preparation, and team skills – as Déjean le Féal points out: "Simultaneous interpretation is, by definition, teamwork"[3] (Déjean le Féal 1990: 154).

Kalina outlines the importance of processes taking place before and after the actual interpretation, processes that partly fall within the interpreter's scope of responsibility (Kalina 2005: 768 f.). She sets up a model that takes all perspectives in a conference communication situation into consideration: not only that of interpreters and listeners, but also that of speakers and organisers/clients/employers. Hers is a holistic approach that encompasses all stages of the interpreting assignment, going beyond the limits of the interpreting product and process as such. She distinguishes four different dimensions: pre-process (prerequisites), peri-process (conditions), in-process (requirements) and post-process (efforts) (Kalina 2005: 776 ff.).

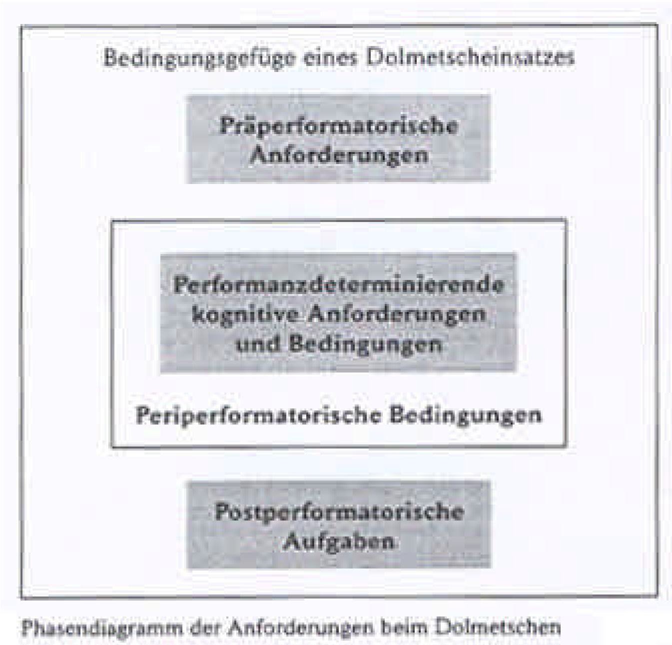

Fig. 3: Pre-process, peri-process, post-process and in-process factors (Kalina 2004: 7)

Pre-process prerequisites include all factors that are important prior to the assignment, such as the interpreter's skills, contract and task specifications, and preparation work. Peri-process conditions refer to the overall setting, i.e. number of participants, technical matters, team members, length of the confer-

[3] Interestingly enough, team skills were not rated as highly important by interpreters themselves in the Bühler study (Bühler 1986: 233). However, they ranked third in the survey carried out by Chiaro & Nocella almost 20 years later (Chiaro & Nocella 2004: 289).

ence etc., while in-process requirements, coming into play at the same point in the timeline, refer to the requirements that the interpreter is confronted with during the actual interpreting process. Post-process efforts, unlike pre-, peri- and in-process factors, fall within the interpreter's exclusive responsibility: working on one's terminology, documentation, quality control, specialisation, further training etc. (Kalina 2002: 126).

So it is obvious that research into interpreting quality must not be limited to the product. Interpreting quality has to be investigated from several perspectives, and it needs to be seen as a dependent variable, with certain factors having a clear impact on an interpreter's performance. Both aspects will be covered in the next chapters. But first, let us take a brief look at a theoretical concept: norms in interpreting.

1.2 Norms in interpreting

1.2.1 Theoretical concepts

> When clients' or users' preferences are not known, the interpreter's norm by default seems to give the highest priority to restitution of informational content, followed by clear and pleasant delivery.
> *(Gile 2006: 19)*

Interpreting, as a behavioural activity, is undoubtedly a norm-governed activity: interpreters need norms to help them select appropriate solutions to the problems they are confronted with (Schjoldager 1996: 302). Yet the work of translators and interpreters is often governed by unspoken negotiations of trust between specialist and client rather than externally determined standards (Chiaro & Nocella 2004: 278 f.). While there are specific standards that apply to translations (ISO 9002, DIN 2345, EN 15038), no such standards exist for the interpreting product.

There are two ways to define norms. One can either adopt a descriptive approach: defining how a text "has been" translated and not how it "could be", "could have been" or "should have been" translated. In that case a norm is a standard model or pattern regarded as typical (Straniero Sergio 2003: 135). Or, one can take the definition of norms as a prescriptive element. For example, a norm can govern what the interpreter ought or ought not to do when the task becomes difficult or impossible (Schjoldager [1995]2002: 303).

The concept of norms as a "category for descriptive analysis of translation phenomena" was first put forward by Gideon Toury (Toury 1980: 57). Toury, a translation scholar, defines norms as "values or ideas shared by a certain community as to what is right and wrong, adequate and inadequate". He distinguishes initial norms (aimed at adequacy, adherence to source-system norms or acceptability, adherence to target-system norms) and operational norms that come into play during the actual translation process and the respective on-the-spot decision-making process. Another distinction he makes is that of textual norms that are established through a comparison of source and target texts, and extra-textual norms which are found in explicit, normative statements in the literature about translation (Toury 1980: 53 ff.).

Regarding the identification of textual norms, the interpreting research community is faced with the difficulty of obtaining a sufficient number of interpreting performances (Schjoldager [1995]2002: 303). Extra-textual norms are equally difficult to identify since there is only scarce literature about the norms governing interpretation. Miriam Shlesinger in 1989 and Brian Harris in 1990 are two rare examples of interpreter-researchers who explored the issue of norms in interpreting.

Harris lists various norms governing the interpreting process, one of them being that the interpreter speaks in the first person as if he was the orator – not in the third person, another that interpreters take turns after 20 or 30 minutes, meaning that a speech which is longer than 20-30 minutes will be interpreted by a succession of several interpreters. Harris coins the term of the "honest spokesperson" norm. Interpreters – speaking on behalf of others – have to re-express the speaker's ideas and the manner of expressing them as accurately as possible and without any significant omissions. They must not bring in their own ideas and expressions. Yet Harris claims that many norms are not being recognised as such. As commonly happens with norms that are very familiar to us, it takes an exception to make us aware of the rule. The "honest spokesperson rule", for example, is so unquestioned that it mostly goes without saying, except in the oaths which court interpreters have to swear, and in interpreting schools where this principle is dunned into interpreting students (Harris 1990: 115 ff.).

Shlesinger points out that definitions of quality are likely to reflect norms that have been internalised by members of the profession – although it is questionable if all members have the same concepts of quality. An exception are centres such as Brussels, New York, Geneva, Paris, or interpreting schools where groups of interpreters or trainers are apt to share the same set of norms (Shlesinger 1989: 112; Shlesinger et al. 1997: 124). Especially in interpreting training it is crucial that commonly agreed norms exist. Chesterman claims that the existence of norms has pedagogical relevance, giving trainees the opportunity to learn norms and then integrate them into their knowledge-base (Chesterman 1993: 18 f.). However, if students have to cope with different sets of norms established by teachers, this can disrupt their learning curve (Gile 2001a: 387).

Gile suggests approaching the issue of norms from the extra-textual side. He believes that research on interpreting norms should not rely on large speech corpora ("textual approach") but should be done by asking interpreters about norms, by reading didactic, descriptive and narrative texts about interpreting, by asking interpreters and non-interpreters to assess target texts and by analysing user responses (Gile 1998b: 101).

This latter aspect is referred to by Chesterman. Although he defines norms in translation, his claims might be applicable to interpreting output as well. Chesterman defines three types of norms: a) the accountability norm, i.e. loyalty to the original writer, commissioner, readership, integrity; b) the communication norm, i.e. optimisation of communication between original writer and readership; c) the relation norm, appropriate relation between target and source text. But for Chesterman, there is also a hierarchy of norms, i.e. regular norms are governed by higher-order norms: expectancy norms (Chesterman 1993: 8 ff.). He concludes that the theory of norms must not be an end in itself but serve a purpose. Ultimately, target texts must be assessed based on expectancy norms. Expectancy norms need to be brought into a systematic hierarchical structure, covering both language-specific and language-independent norms. (Chesterman 1993: 18 f.). Chesterman thus places the user at the top of the hierarchy.

1.2.2 Variability and changeability of norms

As mentioned above, norms can differ across markets, between interpreting communities and centres, between or even within interpreting schools. One example illustrating the relativity of norms is the area of media interpreting.

In 1990, Harris assumed that TV interpreting was going to develop some different norms from the established ones of conference interpreting. For instance, a norm in MI could be that the interpreter on the programme should have a similar voice to that of the person interviewed (Harris 1990: 115). This view is supported by a recent analysis by Straniero Sergio that deals with norms in the interpreting of Formula One Press conferences as a live TV broadcast and revealed, i.a., that the third-person norm is sometimes ignored in MI (Straniero Sergio 2003: 169). Straniero Sergio also postulates that the norm in MI is the adoption of emergency strategies. He cites the example of Ross Perot's answers to interviewers' questions. In Straniero Sergio's sample, answers were almost invariably interpreted using emergency strategies, interpreters reconstructing the sense of the message with a large number of omissions. However, the interpreting community and broadcasters were favourably impressed by these performances. It thus appears that the norm in MI is to put form over content. It is not so much important what is being said but rather that communication does not break down. According to Straniero Sergio, quality standards should be adjusted to concrete SI situations and focus on the *achievable* quality of the interpretation (Straniero Sergio 2003: 171 f.).

An example of how norms can change over time is the "Into-A-Language Norm", i.e. the norm that interpreters should only work into their mother tongue – a controversially discussed topic. Harris points out that in some places and under certain circumstances this norm is being ignored, e.g. whenever there is a shortage of native-speaker interpreters into foreign languages (as was the case in the former Soviet Union), or in Canada where many speakers are bilingual and switch from one language to another in the middle of a sentence (Harris 1990: 117). Garzone also mentions the Soviet Union as an exception to this norm. In the Soviet Union it was widely believed (for obvious ideological reasons) that the highest quality of performance could be achieved if the interpreter worked from his mother tongue into the B language – intertextual consistency with the ST was deemed more important than the linguistic effectiveness of the TT (Garzone 2002: 112).

This view clashes somewhat with Déjean le Féal's perspective, namely that of advising caution when it comes to excessive use of active B-language interpreting. According to le Féal, this practice puts the integrity of the language at risk: If listeners whose native language is not English were to receive basic English through their earphones less often, they would be less convinced that they 'understand' English and would insist on getting interpretation into their own language. She also makes the point of even distribution of job opportunities for members of the interpreting community (Déjean le Féal 1990: 158). To me, this approach – albeit understandable – seems highly questionable since interpreters are in the market to provide service to their customers, not to protect their *chasse gardée*.

In 1990, when Harris wrote his paper, the norm in EU institutions (Harris calls it a "dogma") was that professional conference interpreters should only work into their A language (Harris 1990: 116). Nowadays, the picture has changed. Before the 2005 enlargement, the European Parliament launched plans to introduce the exclusive use of the bi-active interpretation system for the languages of new member states. As outlined in the Podestà Report (named after the Vice-President of the EP, Guido Podestà), this system was to be gradually extended to the languages of the 'old' member states and was to be employed across the board before the end of the next parliamentary term, in 2009 (European Parliament 2001: 8). This plan for greater use of the retour system as well as the intention to step up on the use of relay interpreting met with strong resistance on the part of EU interpreters who claimed that "retour interpretation cannot be seen as a sustainable working method for the European Parliament in the long run" (Gebhard 2001). The protest of EU interpreters and AIIC culminated in the "Resolution on the Management of Multilingualism at the European Parliament after Enlargement" which was issued by the Joint General Assembly of Staff Interpreters and Conference Interpreters Auxiliaries (A.I.C.) in October 2001 (AIIC 2001). Yet it appears that despite these efforts, the retour system is being more widely used and is becoming a formalised practice. This is reflected in the SCIC admission test which also includes tests in a retour language (SCIC 2004). It appears that the "Into-A-

Language Norm" has given way to a more flexible idea of language direction in interpreting, requiring an adjustment of the norm-based idea of quality.

Let me add as a sideline remark that the interpreting reality seems to look different outside EU institutions. Chiaro and Nocella found in their study, in which 167 interpreters from all over the world participated, that the majority of respondents do not interpret into their mother tongue (Chiaro & Nocella 2004: 286). In Germany, for instance, it is common practice for freelance interpreters to work from their A into their B language.

2 Quality as a dependent variable

> Quality was no longer seen as an absolute value but rather as a dependent variable of the specific environment and user type.
> *(Gile 2001a: 380)*

As mentioned in the first chapter, there is general agreement among interpreting researchers that quality is dependent on the interpreter's working conditions, which, to a large extent, determine the quality of a given SI. Mack points out that there is not necessarily a direct correlation between the interpreter's skills and qualifications and the actual performance in a specific situation – interpreting has to be regarded as a dependent variable (Mack 2002: 117; Kalina 2002: 124).

The list of framework conditions affecting quality is long: physical environment, complexity of subject matter, change in subject matter, adversarial nature of meetings, discourse characteristics, delivery, preparation of interpreter, availability of speech manuscripts, team size, length of turn, load during working day, number of meetings, speakers speaking simultaneously, interpreter's emotional response, competence and availability of technician – and the list goes on (Moser-Mercer 1996: 44 f.). The quality of the original is a crucial factor. Shlesinger and Pöchhacker have underlined the importance of delivery characteristics such as segmentation and speed which can affect interpreting quality (Shlesinger 1994; Pöchhacker 1994b: 234). In her dissertation based on authentic material, Vuorikoski confirmed this statement by showing that certain factors do indeed have an impact on interpreting quality: e.g. monotony of speech, hesitations, and non-native accents (Vuorikoski 2004: 25, 51).

Since quality depends on so many external factors, interpreters cannot be held totally responsible for achieving high quality, and quality in interpreting will always have to be evaluated against the background of the respective working conditions (Moser-Mercer 1996: 45; Straniero Sergio 2003: 171).

The question is: To what extent and under what circumstances is it possible for the interpreter to achieve 'ideal quality'? Researchers have tried to assess the impact of certain parameters. Some answers to the above question can be provided when looking at their findings. In much of the empirical research on SI, the general approach has been to manipulate some aspect of the input information and to observe the effects of the manipulation on the interpreter's output – this line of research can be called 'quality-as-a-dependent-variable' research. Some research initiatives, although not as frequent as experimental research projects, were survey-based.

2.1 Survey research

Not much survey research is available on the question as to what conditions are apt to enable interpreters to provide an ideal interpretation.

Cenková carried out a survey among 34 conference interpreters in 1998, aimed at identifying the elements of speech that they regard as having a major (positive or negative) impact on their performance. Fast reading of a text, fluency/intelligibility of delivery, and the speaker being a good speaker received the highest scores for importance. One can assume that fast reading would have a negative impact, the other two elements a positive impact on the performance although this was not specified. Interpreters also attached high importance to the speaker being a native speaker and speaking without a text (Cenková 1998: 165 f.).

The AIIC Workload study conducted in 2001 included a mail survey among interpreters; the questions particularly referred to the effects of stress. Between 40% and 60% of participants in the mail survey reported a drop in performance levels when subject to work-related stress, particularly as regards accuracy and overall quality of work. 66% reported high or very high levels or work-related stress, 40% saw it as positive and 36% as harmful. 70% considered that they dealt with stress successfully and 6% unsuccessfully. The two stressors most frequently mentioned were 'fast speaker' (78% of respondents) and 'speaker reading from text' (71% of respondents). These two factors were also rated highest in terms of their perceived level of stressfulness. Further factors mentioned as contributing to stress were: 'textual complexity' (50%), 'subject of meeting' (48%), 'difficult accent' (31%) and 'booth discomfort' (24%). Almost half of the subjects reported a decline in their performance during a stressful turn as compared to a 'normal' turn. A decline in performance was more frequently reported by subjects in mobile booths. 73% of respondents also believed that videoconferencing has a negative impact on performance (Mackintosh 2002).

2.2 Experimental research

Most empirical research on quality as a dependent variable has been experimental, although there are also a few examples of observational research (Straniero Sergio 2003; Vuorikoski 2004). Researchers have explored what happens with interpreting quality if certain framework conditions are manipulated. In the following, I will give an overview of the different factors that affect quality as a dependent variable, starting with a brief summary of early research steps in the "Experimental Psychology Period" (Gile 2000: 300).

2.2.1 Research pioneers – the Experimental Psychology movement

The first steps in interpreting quality research were taken by psychologists and centred on comparative linguistic issues, with experimental recordings and transcription of results which were then compared with their originals. Although these experiments did not specifically address the issue of quality as such, the question as to how well interpreters perform under certain circumstances is closely linked to quality assessment.

Treisman and Barik focused on a quantitative analysis of interpreting output and 'errors', using different error categories such as omissions, additions and substitutions (Treisman 1965; Barik 1971, 1972, 1973). The communication situation was not included in these analyses, nor were parameters on a macro-textual level, i.e. cohesion and coherence, taken into account. Barik defined an error as a "substitution that is a considerable variance with the original" (Barik 1971: 21). Omissions were simply rated as errors even though they might have been the product of a conscious decision taken by the interpreter to provide high-quality output – an approach which was strongly criticised by interpreter-researchers who argued that omissions could be an element of interpreting strategies (e.g. Tommola & Lindholm 1995: 130; Kalina 1998: 119; Chiaro & Nocella 2004: 281), i.e. an 'omission tactique' (Gile 1995a: 132). In these early days of interpreting research, researchers were wondering how interpreters could manage the miraculous simultaneity of listening, thinking and speaking. Barik, no interpreter himself, assumed that interpreters used speakers' pauses to produce their output (Barik 1972: 6). He also noted that an adequate ear-voice span was critical since lagging too far behind the speaker could lead to difficulties or omissions, staying too close to the speaker can lead to errors and false starts (Barik 1975: 290).

Another aspect investigated in this early research period was the information density of the source speech. In her study about the effects of redundancy, Treisman found that performance decreased proportionally more for the passages with high informational content. One has to bear in mind, however, that the study was conducted with bilingual subjects who were not interpreters. Treisman also discovered that the ear-voice span of interpreters is much longer than those of shadowers[4] (Treisman 1965: 369 ff.). This was confirmed by Oléron and Nanpon as well as Gerver (Oléron & Nanpon [1965]2002; Gerver [1969]2002).

Oléron and Nanpon specifically addressed the issue of *décalage* when comparing source text and interpreters' output (Oléron and Nanpon [1965]2002). They compared their data on time lag in interpreting to those of a previous study (Oléron & Nanpon 1964) and found that the EVS in the interpretation of words was 50% longer than in the shadowing of words. They also noted that this time lag became much longer in interpreting. The increase ranged from 1.2 seconds to 3 seconds (which, in their view, was a time delay that corresponded to satisfactory performance), or even up to 10 seconds (Oléron & Nanpon [1965]2002: 48f.).

Gerver investigated how the source language presentation rate affected performance during simultaneous interpretation and shadowing. He found that an increased presentation rate had a significant effect on the ear-voice-span, and that interpreters tended to make more errors than shadowers at faster presentation rates. Interpreters also made more pauses and spoke in shorter bursts (Gerver [1969]2002: 65). He – anticipating Gile's effort model (Gile ([1997]2002) – assumed that the interpreter's attention is shared between several processes (receiving the input message, translating, monitoring of feedback); a balance that the interpreter can maintain under normal circumstances but which is disrupted when the total capacity is exceeded, e.g. because of faster presentation rates (Gerver [1969]: 66). He also investigated the impact of noise on interpreting quality and confirmed its adverse effects (Gerver 1974).

[4] Shadowing: a word-for-word repetition in the same language of a message presented through headphones (definition in Lambert 1991: 587).

Aside from the above-mentioned error definition that Barik was criticised for, there were many methodological flaws in this early research period that provoked protest among 'true professionals', e.g. the fact that source texts were not always speeches, settings were not always conferences, interpreters were not always professionals, and research parameters were not always conclusive for interpreting research (Kalina 2005: 770). The main point of criticism was that transcripts were used as a basis for evaluation (Kalina 2002: 122). However, as can be seen in papers edited in Pöchhacker and Shlesinger, these pioneering researchers were aware of the complexity of the task (Pöchhacker & Shlesinger 2002). In some respects, their approaches were more realistic and less ambitious than those of their successors from within the interpreting community. Oléron and Nanpon, for instance, referred to the "complex nature of the variables and the cumbersome nature of the analytical procedures applied" and ended their paper expressing their hope that further studies would follow, "investigations in which psychologists, linguists and professionals could ... usefully cooperate" (Oléron & Nanpon [1965]2002: 50).

2.2.2 View of the speaker

Interpreters claim that a clear view of the speaker is a prerequisite to high-quality interpretation. An early study to test this hypothesis was conducted by Anderson in 1979 (Anderson [1979]1994). The study examined interpreters' output in two different visual conditions – with a video showing the speaker ('video-on' setting), and without video ('video-off' setting). It emerged that scores for the 'video-on' setting were slightly higher but not statistically significant (Anderson [1979]1994: 104 ff.). However, one has to bear the methodological issues in mind: With only 12 participating interpreters, the sample was very small. Anderson herself pointed out that high subject variability might have led to a distortion of results, and that stricter control over level of skills might be required since the only criterion applied in this study was the 5-year-experience threshold (Anderson [1979]1994: 108 f.).

In 1995, an experiment by Tommola and Lindholm investigated the interaction between visual and acoustic quality. Gerver had shown that noise affects the interpreter's performance (Gerver 1974). Tommola and Lindholm now set out to explore whether the deterioration of acoustic quality had a particularly adverse effect if visual information was not present. Interpreters worked under four different conditions: high acoustic and high visual quality, low acoustic and high visual quality, high acoustic and low visual quality, low acoustic and low visual quality. Low acoustic quality meant that white discourse was mixed in with the source discourse. In the 'low visual quality' condition, the video screen was blanked out. Although – as could be expected from Gerver's findings – a significant effect emerged for acoustic quality, neither the effect of visuality nor the interaction between visual and acoustic quality reached significance (Tommola & Lindholm 1995: 127 ff.). These findings are problematic insofar as they challenge interpreters' claim that a clear view of the speaker is an integral part of good working conditions – *for quality's sake*. On the other hand, Tommola and Lindholm admitted that their method of propositional count may not have been totally valid since omissions were counted as errors – an approach similar to Barik's which had been strongly criticised (Barik 1971, 1972, 1973). Another methodological problem was the small sample group: only eight individuals took part in the experiment (Tommola & Lindholm 1995: 130).

2.2.3 Remote interpreting

Closely linked with the speaker-view issue is the increasing trend of remote interpreting, i.e. interpreting in a video-conferencing set-up or through a cabled arrangement close to the meeting facilities – a setting where interpreters are not in the same room as delegates, and only have a limited view of the speaking environment via monitor – if they have monitors at all.

Riccardi, Marinuzzi and Zecchin (1998) established the hypothesis that remote interpreting with no direct view of the speaker would impose a higher strain on conference interpreters, compared to 'regular' conference interpreting. A study carried out by the UN in 2001 corroborated this hypothesis. Firstly, interpreters participating in the experiment underlined that higher psychological and physiological effort is necessary if they are to keep up the same level of performance during remote interpreting. This was proved to be true in a second experiment, carried out the same year (Report of the Secretary General 2001a, 2001b as quoted by Kurz 2003: 55 f.).

The first controlled experiment to evaluate human factors and technical arrangements in remote interpreting was carried out by Moser-Mercer in collaboration with the International Telecommunications Union. Interpreters' output was sampled three times during each half-day, once in the beginning, once in the middle and once towards the end of a turn both in a normal live setting and in a remote setting. The comparison of the effects of fatigue on performance for the same interpreters working either live or in a remote condition revealed significant differences. The same interpreters were less tired, hence worked at a higher level of quality, in live conference conditions as opposed to remote conditions. One can conclude that remote interpretation increases an interpreter's mental workload and leads to fatigue and decline in performance faster than live interpretation. The experiment also demonstrated that for the same group of interpreters, working live in a conference room was psychologically less stressful (according to interpreters' self reports), less tiring as evaluated via performance indicators, and conducive to better overall performance. The remote interpreting situation appears to represent not only a novel environment for interpreters they need to cope with by using more effortful problem-solving strategies but also seems to cause more than the usual physiological and psychological strain "in that the coordination of image and sound, the piecing together of a reality far away and the concomitant feeling of lack of control all draw on mental resources already overcommitted in this highly complex skill" (Moser-Mercer 2003).

2.2.4 Preparation

In her previously mentioned study conducted in 1979, Anderson did not only examine the impact of visual conditions but also the question as to whether interpreters perform better when prior information about either the context or total content of the message is made available to them. The parameters 'intelligibility' and 'informativeness' were measured, and the resulting data showed no significant effects for prior information on either performance (Anderson [1979]1994: 104 ff.). The study thus suggested that interpreters did not perform better when preparation material was available. Yet if the same study was to be repeated today, it would most likely bear totally different results. The fact that interpreters did not feel that preparation material had a positive influence on their performance may have been due to the fact that they saw themselves as *généralistes* who could cope with any difficult and even specialised subject, even unprepared (Kalina 2004: 4). Today, such an attitude is unthinkable. Conferences are highly specialised, and good preparation is of paramount importance since interpreters are in the position of ‘outsiders’ who know less – generally much less in technical and scientific settings – about the subject (Gile 1991: 194).

In 1984, a study by Gile demonstrated that the generally held belief that preparation documents play a pivotal role in the QA process is justified. The study revealed that an error rate of around 50% in the interpretation of names was not infrequent, showing that preparation documents are indispensable (Gile 1984: 79). Vuorikoski’s findings also support this claim (Vuorikoski 2004).

2.2.5 Interpreting with text

A special form of preparation is the availability of a speech manuscript. Lamberger-Felber investigated four hypotheses with regard to simultaneous interpretation with manuscript, namely that interpreters would make fewer mistakes regarding names and figures, fewer omissions and errors in general, fewer meaning errors, and work with a longer ear-voice span because they concentrate on the written manuscript. All four hypotheses were confirmed. Interpreters with text made fewer errors and omissions, both generally speaking and with regard to numbers and names. The number of serious errors in meaning was also lower. However, the conclusion that the availability of a speech manuscript ultimately leads to higher interpretation quality is only partially valid: Interpreters showed a tendency to interpret with an excessively long ear-voice span which led to long omissions. This means that while a speech manuscript helps interpreters produce a more accurate output it might also have a negative impact on quality, albeit only to a limited extent (Lamberger-Felber: 1998; Lamberger-Felber 2003: 151 f.).

2.2.6 Background knowledge

Mackintosh found in her experiment that figures which did not offer any possibility of using background knowledge were almost invariably reproduced incorrectly in SI. If they were correct, it was at the expense of the surrounding text (Mackintosh 1983). This seems to underline the importance of background knowledge for interpreting quality – as is confirmed by Vuorikoski in her study who found that shortcomings in the TT are at least partly due to the fact that interpreters do not share enough background knowledge with the speakers (Vuorikoski 2004: 207). Déjean le Féal requests that minimum requirements with regard to background knowledge that interpreters have to possess be defined (Déjean le Féal 1990: 160).

These findings and views contrast with Bowen's result from a study among student interpreters which revealed that information acquired in secondary education is not necessarily easily and efficiently retrievable for use in SI (Bowen 1990: 552). The reason for this divergence may be that students fail to establish a link between their background-knowledge acquired in school, and the specific interpreting situation – a pattern which may be an element of the novice-typical micro-processing approach (cf. 4.1).

2.2.7 Length of turns

Moser-Mercer, Künzli and Korac investigated the effect of prolonged turns (longer than 30 minutes) on the quality of interpreters' output and measured physiological and psychological stress that the interpreters experienced. They found a significant increase of meaning errors after 30 minutes, which represented a strong decline in quality. They also showed that interpreters' own judgement of quality is extremely unreliable after increased time on task. Interpreters chose to stay on task even though they could have ended the turn (Moser-Mercer et al. 1998: 55).

If interpreters cannot necessarily be trusted as to how much time they should spend on the task, it is unlikely that conference organisers can (Moser-Mercer et al. 1998: 61).

2.2.8 Delivery rate

One aspect regarding the speaker's role and responsibility that has been taken up in research is the effect of presentation rate on interpreters' output.

Sonia Pio explored the question of how source text delivery rate affected the interpreter's output. The findings confirmed those of Gerver insofar as there appears to be a link between ST delivery rate and interpreting quality. This was reflected particularly by the increase of filled pauses (as an indicator of interpreters' hesitations and doubts), unfilled pauses (as an indicator that –

when delivery rate increases – interpreters take longer to fully elaborate the incoming message but that this may lead to more omissions), and corrections. However, Pio found that it is neither possible to confirm nor to reject Gerver's theory according to which SL delivery rate and EVS increase are directly linked (Pio 2003: 96 ff.).

Shlesinger tested the hypothesis that a higher presentation rate can have a positive effect on performance because of the reduced strain on interpreters' working memory. The hypothesis proved to be true: Interpreters were able to reproduce more modifiers when the source text had a higher presentation rate. Higher delivery speed seems to have a positive impact on the interpreter's working memory since a lower presentation rate entails the risk of greater trace decay. However, Shlesinger conceded that her findings – which contrast with those of other studies – need to be tested in more natural settings (Shlesinger 2003: 43 ff.).

2.2.9 Speaker's accent

Straniero Sergio examined the performance of media interpreters during Formula One Press Conferences. He collected a corpus of 80 performances by 11 interpreters between 1997 and 2002, with two interpreters accounting for 90% of the whole corpus. One can assume, therefore, that the majority of interpreting performances were delivered by interpreters who were experienced in and familiar with the subject and the situation. It emerged from his analysis that out of 512 drivers' answers, 249 were not interpreted correctly. Regarding the interpretation of questions, the performance was significantly better: 403 correct vs. 95 incorrect. One factor seems to have been the strong non-native accent of the 11 non-English speaking drivers (139 answers were not interpreted correctly), and the strong Scottish and Irish accents of Coulthard and Irvine (110 incorrect answers) (Straniero Sergio 2003: 139 f.).

2.2.10 Monaural vs. binaural input

Lambert investigated if there is a difference in performance when the message is received through the right ear, the left ear or both ears, and found that (right-handed) individuals performed far better when incoming information was presented to their right hemisphere – i.e. the left ear. She concluded that

– if one can assume that decoding and encoding activities, namely listening and speech production cannot be performed through the same hemisphere – interpreters use their left hemisphere (right ear) for what they consider to be the more critical of the two tasks: monitoring of output (Lambert 1989: 324 f.; Lambert 1993).

2.2.11 Language combination

The language combination can also play a role in how difficult it is to achieve good quality. Every interpreter working from German into another language has suffered the perennial problem of verbs at the end of a very long sentence: The interpreter has no choice but to use anticipation strategies. But as Fusco has shown it is not necessarily easier to interpret between cognate languages, e.g. Spanish and Italian. The interpreter might be tempted to go for the “nearest possible” solution, thus falling prey to the risk of producing sentences that are awkward because the similar-sounding solution might not be an idiomatic one, or of mistranslating paronyms, ‘false friends’ (Fusco 1990: 93 ff.).

2.2.12 Interpreters and stress

Job stress is caused by the interaction between the individual and the conditions of work. According to the National Institute for Occupational Safety and Health (NIOSH) in the US, working conditions play a primary role in causing job stress. However, individual and other situational factors also play a role. What one person perceives as stressful may not be stressful for someone else (NIOSH 1999: 7 ff.). Leaving individual factors out of consideration, there are many conditions that are stressful for most people, e.g. heavy workload, infrequent breaks, long working hours. In addition to the fact that interpreting is a task which requires an extremely high level of concentration over prolonged periods of time, the need to cope with different subjects (often highly divergent ones), different speakers and accents, and the possibility of failure are regarded as factors that contribute to interpreters’ stress (Kurz 2003: 51 f.). In interpreting literature, there are several references to the personality an interpreter should allegedly have: he should be able to work under stress and stay cool under pressure. Gravier expects “nerves of steel” from interpreters (Gravier 1978: 4 as quoted by Kurz 2003: 53), whilst Coughlin defines ideal interpreters as “programmed to be winners under adverse circumstances” (Coughlin 1988: 359 as quoted by Kurz 2003: 53).

There is wide consensus that stress is intrinsic to interpreting although only very few studies have been conducted to corroborate this hypothesis. Some of the earliest studies in this field examined the work environment, i.e. temperature, air quality and humidity in the booth and the resulting physical stress (Kurz 1981; Kurz 1983; Kurz and Kolmer 1984; all quoted by Kurz 2003: 54).

In 1997, Kurz used the STAI method (State and Trait Anxiety Inventory) to examine conference interpreters’ anxiety levels. The findings of her study suggested that interpreters have better situation-dependent control of their feelings and can use their anxiety productively by labelling it in a positive way – they are “consistent performers” (Kurz 1997: 55).

In their 2001 study Jímenez and Pinazo found that the level of anxiety does not have an effect on the interpreting performance – a fact which the two researchers put down to the interpreters’ coping tactics, or to “positive anxiety” which is needed to face stressful situations (Jímenez & Pinazo 2001: 114).

The largest study that dealt with stress as a performance-affecting factor was the Workload Study commissioned by AIIC in 2001, which investigated four sets of parameters: psychological, physiological/physical and performance aspects, as well as the correlation between them. The assumption was that the influence of several stressors should have a negative impact on interpreting quality (Mackintosh 2002). The data obtained from the study placed simultaneous interpretation in the category of high stress professions, and although no correlation was found between the physiological indices and performance levels, the data pointed to psychological and physiological costs. Performance data was sampled three times in the course of the working day: at the beginning and end of three interpretation turns occurring at the start, in the middle and at the end of the day. Interpreter performance was highly rated (4.12-4.4 on a scale from 1-5, 5 being the highest mark) with a slight decline towards the end of the day. No connection was found between CO_2 levels and performance quality although low humidity correlated weakly with a lower performance evaluation, as did working in mobile booths.

Although the correlations between measures of stress (objective and subjective) and performance are weak, this is in keeping with findings in the literature which indicate that highly competent and motivated workers maintain high levels of performance in the presence of a variety of stressors. Other studies show that there are physiological costs associated with maintaining these high levels in such conditions, namely exhaustion and post-work stress. It can be argued that the fact that the correlations between performance and stress are weak is evidence of coping strategies having been developed by the interpreters (Mackintosh 2002).

Stress can also be a resulting factor from prolonged turns in interpreting as shown in the study by Moser-Mercer, Künzli and Korac. While physiological stress indicators (cortisol and immunoglobulin A levels) showed a remarkable trend to adapt to the task, the psychological stress levels that interpreters have to cope with (as reflected in questionnaire results) should not be underrated. For example, interpreters showed much more passive and escapist behaviour and displayed patterns – differing from the norm – that are often encountered in pathological states such as high depressiveness and high anxiety. Longer turns – with the resulting psychological stress – can thus represent a threat to the mental health of interpreters (Moser-Mercer et al. 1998: 56 ff.).

To summarise the above findings: It seems that interpreters have a remarkable tendency to cope with stress. But one should be cautious in playing down its negative impact. Even though short-term quality effects are not in all cases identifiable, the long-term effects on interpreters' health (and thus on their professional ability to deliver high quality) must not be underestimated. Quoting the National Institute for Occupational Safety and Health: "[The] early signs of job stress are usually easy to recognize. But the effects of job stress on chronic diseases are more difficult to see because chronic diseases take a long time to develop" (NIOSH 1999: 10).

2.3 Pragmatic conclusions

What conclusions can be drawn from the research findings outlined above? They undoubtedly give ample support to the intuitive feeling of interpreters that certain conditions affect their performance. They also justify the demands that interpreters make when negotiating an assignment and the conditions under which it takes place.

The findings make it clear that interpreting quality is subject to appropriate standards in interpreters' working conditions (Kalina 2002: 124). Certain external factors that the interpreting performance hinges on can be made measurable. Thus interpreters and stakeholders can identify any room for improvement. The motto should be: Talk to your interpreters! Interpreters should be asked by clients or organisers to act as advisors in the planning stages of a conference (Pöchhacker 1994a: 118; Mack 2002: 117). And speakers should take into account that their speeches are being or are to be interpreted and make sure that they speak in a way which allows interpreters to render their message so that listeners comprehend it – i.e. "prioritize the needs of the audience" (Vuorikoski 2004: 249 ff.). The speaker should design the text so that the interpreter, rather than the speaker's equally knowledgeable peers, can understand it (Dillinger 1990: 185). A good overview of aspects that speakers should bear in mind is provided in the "10 Commandments for Speakers at International Conferences". Speakers are advised to speak freely (not read a manuscript), speak clearly, avoid rapid enumerations of names and figures, explain abbreviations, refrain from the use of puns, provide documentation to the interpreters etc. (AIIC & Universitas 1991: 17f.).

But even if we assume that the speaker is cooperative, the profession is faced with new challenges, for example the number of languages from and into which an interpreter is expected to work. As Kalina points out, quality problems arise when interpreters are expected to work from three or more languages: The higher the number of languages, the less interpreters can be expected to possess in-depth knowledge of all linguistic nuances and cultural aspects (Kalina 2005: 778). And there are more challenges to master, e.g. the current trend of greater distances between booths and speakers' platforms, remote interpreting, videoconferencing, conference call interpreting, Internet interpreting, non-native speakers, less preparation time and increasing cost pressure as an underlying reason for these trends (Kalina 2005: 778; Kutz 2005: 33). Many of the studies described above have proved that these new developments do indeed constitute a threat to the interpreting quality that can be achieved.

As early as twenty years ago, Gile predicted that specialist terminology would grow significantly, exponentially (Gile 1986: 36). And he was right. Conferences get more and more specialised, and more often than not preparation material is provided at short notice. Gile could not foresee that PCs and notebooks would become an integral part and indispensable tool in every interpreter's working method, not only for preparation but even in the booth. The Internet and the ensuing possibility to access a wealth of information through it make interpreters' lives much easier. But the challenge remains: Interpreters have to prepare thoroughly and in great detail for their assignments. No interpreter can be expected to memorise the 100,000s of potential terminological expressions associated with a given subject – this is why it is crucial that documentation be made available to the interpreter beforehand. If it is not, the risk of information losses is significant (Gile 1985: 200).

Kutz distinguishes different aspects of preparation. These do not only include preparation of the subject matter, linguistic preparation, and preparation for interpretation (such as sight translation, if possible using the speech manuscript) but also organisational preparation (i.e. the interpreter should only accept the assignment under certain – acceptable – conditions, as outlined in the AIIC Code of Professional Ethics, and see that these conditions are met, for example regarding preparation documents) (AIIC 2006). He also mentions psychological preparation and self-motivation (Kutz 2000: 10 f.). This is a good example to show that, in addition to purely linguistic skills and task-specific skills, further factors come into play. If one defines quality as the quality of the product that is possible under certain framework conditions, and if one defines the framework conditions as something that the interpreter can partly influence, then the interpreter's responsibility for the product does indeed start at a very early point: when accepting (or declining) the assignment. This means that interpreters have to live up to their responsibility by negotiating acceptable working conditions. Or, if negotiation is not successful, refuse to take an assignment.

Kutz's view is similar to that of Pöchhacker who stipulates that hypertext production begins when an interpreter accepts an assignment. The interpreter must then specify technical details as to how the conference is organised, whether preparation material will be made available, the briefing of conference organisers and actual preparation work. The more professional an interpreter's "hypertext production" at this stage, i.e. before the conference takes place, the better the conditions for providing high-quality interpretation. Not only preparation work but also follow-up work falls within the hypertext category, as well as activities during the conference, such as informal contact with participants, speakers, organisers before or after turns (Pöchhacker 1994a: 118).

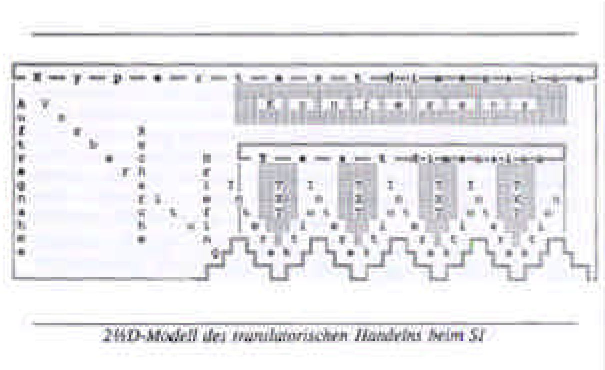

Fig. 4: 2½D-model of translational activities in SI (Pöchhacker 1994a: 120)

I firmly believe that the training of future interpreters must go beyond the limits of teaching interpreting skills and processes in the strict sense. Young interpreters must learn how to negotiate working conditions that give them the possibility to provide high-quality interpretation. Negotiating training, quasi-automatically scheduled for junior managers working for large companies, must become an integral part of every interpreting school's curriculum. Interpreters are managers, too. They manage their own assignments, terminology, and relationship with customers. Negotiating and (self-)management skills are an integral part of professionalism, and a basic set of crucial skills can be acquired at university to prepare students efficiently for the challenges they need to master.

If financial constraints do not allow for formalised training supported by a specialised trainer or consultant, practical exercises can replace such costly training. In the renowned interpreting schools most teachers are active interpreters who have acquired ample experience working as professionals on the market. They should let their students in on their expertise regarding the negotiation of contracts and working conditions, and this should happen not in an anecdotal way during normal interpreting classes but in separate courses so that student interpreters are taught systematically what they might otherwise have to learn the hard way at the beginning of their professional careers.

I also share Kutz's view that the technical management of an interpreting assignment is a skill which ought to be practised in class to train accuracy and speed of this management process. He recommends that it be practised once per semester in different constellations although he concedes that this skill can not be tested in an exam situation (Kutz 2005: 20).

Student interpreters should also be made aware of the fact that taking assignments from agencies can be a risky business as it means that the number of factors beyond the interpreter's control will invariably increase (preparation material, location of the booth etc.). If interpreters have no direct contact with their customer, they will risk receiving no information as to the situation, context, speakers and their intentions, as well as the audience (Kalina 2004: 6). Young interpreters must be aware of the need to assume an adequate degree of responsibility for the conditions under which they work. This is one of the key prerequisites for delivering excellent quality.

3. Interpreting quality from different perspectives

3.1 Quality from the interpreter's perspective

3.1.1 Survey research

3.1.1.1 General surveys

The first survey on the topic of interpreting quality was not carried out among users but among interpreters (Bühler 1986). Although the survey was much criticised, it is frequently quoted in the literature about interpreting quality and can be considered groundbreaking. Just as Barik's and Gerver's work met with criticism but triggered a new research wave, Bühler had many researchers follow in her footsteps.

The Bühler survey (1986) was conducted among 41 AIIC members plus six members of the AIIC Admissions Committee (CACL). Although the questionnaire was explicitly addressed to (and filled out by) professional interpreters, Bühler inferred that the results also represented "user expectations and needs in its evaluation standards" (Bühler 1986: 231). However, this assumption is questionable because of the feedback that interpreters gave regarding the importance of the users' perspective. The survey revealed that only 16% of interpreters considered a positive reaction of the user as highly important, 47% considered it as important, and a surprisingly high percentage of respondents (31%) claimed that it was less important or even irrelevant. When asked what linguistic and extra-linguistic criteria are important when evaluating the performance of a candidate for membership, interpreters' gave the following ranking: 1) sense consistency with the original message, 2) logical cohesion of utterance, 3) completeness of interpretation, 4) correct grammatical usage, 5) use of correct terminology, 6) use of appropriate style (Bühler 1986: 231ff.). However, this ranking of criteria has to be treated with caution since interpreters valued most of the 16 linguistic and extra-linguistic criteria as highly important; they seemingly found it difficult to assign an order of importance.

The survey carried out by Altman among EC and AIIC interpreters in 1990 yielded unexpected results insofar as it revealed that the majority of professionals did not consider the clear view of the speaker as a highly important factor for quality, only 41% rating it as enormously or considerably important, 44% as somewhat important, and 15% as hardly or very little important (Altman 1990: 32). Three quarters of respondents felt that they were at least occasionally blamed unjustifiably for communication breakdowns, i.e. used as scapegoats (Altman 1990: 23). Even though the ‘speaker view requirement’ was not confirmed in this study, another statement was once again underlined: Relay interpreting leads to a deterioration of quality. Interpreters virtually unanimously stated that they had a strong distaste for relay interpreting and shared the universal view that it affects quality negatively (Altman 1990: 23).

Feldweg carried out a survey among conference interpreters, not on the basis of questionnaires but using structured interviews (Feldweg 1996: 297 ff.). Inter alia, the survey explored the personality traits and capabilities that a good interpreter must possess. The same phenomenon that Bühler had witnessed in her survey occurred again: Interpreters considered almost all criteria as very important or important (except for ‘acting skills’ and ‘self-assuredness’) so that virtually no ranking emerged from the results (Feldweg 1996: 372).

In Bühler’s case, only 47 interpreters had returned their questionnaire. When Chiaro and Nocella conducted a similar survey 18 years later, they used a more innovative method and conducted their survey via the Internet. As a result, 286 respondents sent back their online questionnaires. Kurz’s samples had been very small and uneven as well as based on returns of a questionnaire administered at different points in time and different settings. In this new survey, respondents came from all over the world. Chiaro and Nocella adopted a multidisciplinary approach, with the support of one statistician and one linguist. They also strove to avoid obtaining similarly meaningless rankings such as those that the studies by Bühler and Feldweg had produced, with interpreters rating all criteria as (almost) equally important. They concluded that something had been wrong in the research design of these earlier studies and adopted a different approach (Chiaro & Nocella 2004: 283). Interpreters were asked to rank the criteria under investigation. The three most important factors were consistency with original, completeness of information, logical cohesion. The second most important set of criteria: fluency of delivery, correct grammatical usage, correct terminology. Appropriate style, pleasant voice, native accent were rated as relatively unimportant. Of the non-linguistic criteria, concentration and preparation of documents, i.e. two criteria that fall within the interpreter's scope of responsibility, were perceived as the two most important criteria (Chiaro & Nocella 2004: 287 ff.).

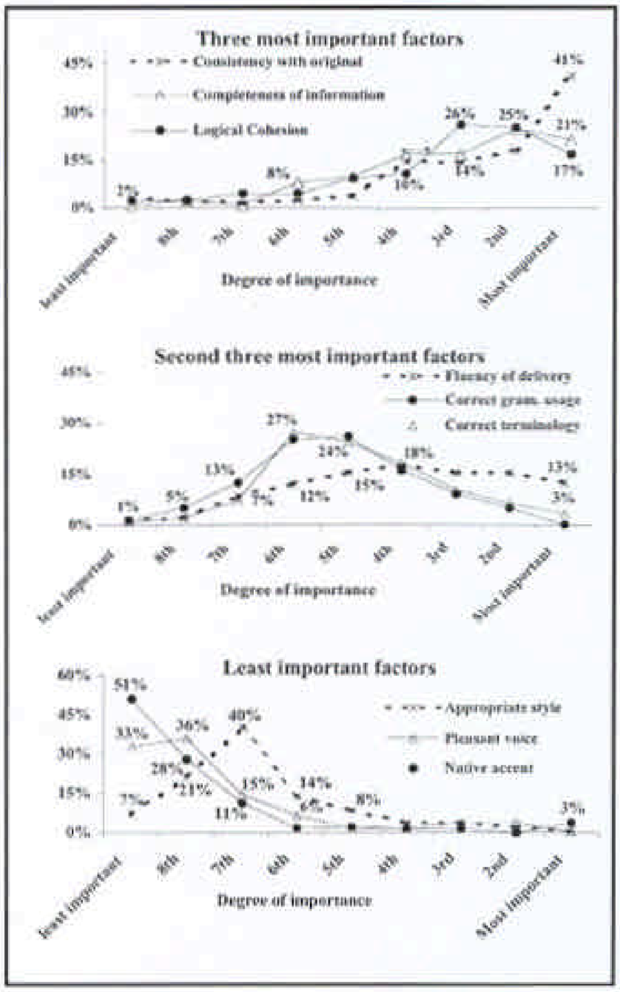

Fig. 5: Distribution of the degree of importance given to each linguistic criterion (Chiaro & Nocella 2004: 287)

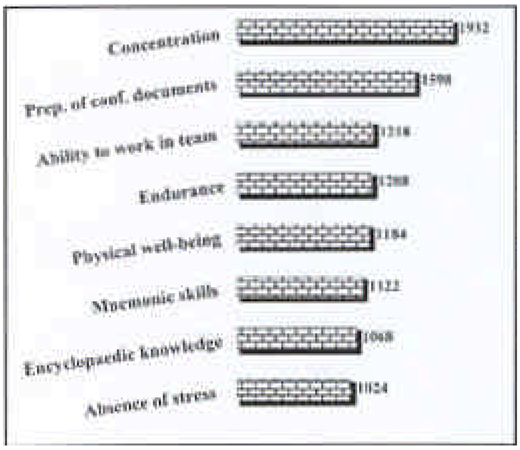

Fig. 6: Sum of the scores given to extra-linguistic criteria (Chiaro & Nocella 2004: 288)

3.1.1.2 Relay interpreting as seen by interpreters

The survey by Altman, as mentioned above, confirmed interpreters' strong distaste for relay interpreting and the view that it affects quality negatively (Altman 1990: 23). As Kahane points out: Pivots are also listeners. One question to be answered is whether the ideal quality for the 'genuine' listener is equally ideal for the 'pivot' listener who relies on the interpretation to produce another version. Should it be more literal, or recreated in the target language (Kahane 2000)?

The issue of relay interpreting was further explored in a survey by Cenková (Cenková 1998). Interpreters rated quality of relayed interpreting as good if the relay interpreter speaks calmly, if he uses correct terminology and delivers his utterance in coherent sentences *with redundancies already eliminated* (Cenková 1998: 167). What they expect is concision, i.e. – according to Sunnari's definition "the sampling based on a global comprehension of the message and the consequent deletion of redundant local-level material (Sunnari 2003: 245). Interpreters taking a relay from their colleagues thus expect them to do some pre-processing of the speaker's message, i.e. like them to go beyond the 'ghost role' (Kopczynski 1994: 90), and assume an active role in communication (Cenková 1998: 167).

The question of relay interpreting was also touched upon by Collados Aís who compared quality expectations of interpreters in two roles: in their function as sender and recipient, i.e. as pivot-giving and pivot-taking interpreter. Although there was no difference in ranking of the different parameters (sense consistency came in first, followed by cohesion; least importance was attributed to monotonous delivery) – pivot-giving interpreters were more demanding and gave higher absolute importance scores to all parameters. Pivot-taking interpreters were less demanding as to the quality expected. This seems to indicate that interpreters apply high standards to their own performance on the one hand (as pivot) but are less demanding regarding their colleagues' output, probably because they are aware of the complexity of the task (Collados Aís 1998: 157 ff., 244).

3.1.1.3 Specialisation – yes or no?

> [Interpreters are] Jacks of all trades, and masters of none.
> *(Watson 1982: 140 as quoted by Schweda-Nicolson 1986: 68)*

Many translators specialise while in the interpreting community the situation is somewhat different. The survey that Schweda-Nicolson carried out among UN staff interpreters revealed that although many interpreters are in favour of specialisation (39.3%), the overwhelming majority see themselves as generalists (80.3%), less than one in five claiming to be a specialist (17.9%). One survey participant made the following statement: “It could be said that I specialize in the ‘general’, as opposed to the ‘specialized’” (Schweda-Nicolson 1986: 74 f.). It is interesting to note that this low degree of specialisation among interpreters was found among *staff* interpreters who, according to Watson, are the only group of interpreters who can undertake specialisation (Watson 1982: 140 as quoted by Schweda-Nicolson 1986: 74).

3.1.2 Experimental research – quality *assessment* by interpreters

> Interpreters’ evaluation of their own output is usually not very reliable.
> *(Déjean le Féal 1990: 154)*

So far, there have only been few experiments in which interpreters assessed their colleagues’ performance. Gile points out that many interpreters dislike being listened to by colleagues, because of their “situation de concurrence économique qui accentue leur vulnerabilité”. Some interpreters even like working alone in the booth – preferring the absence of a colleague who might help in cases of emergency to being judged by that same colleague (Gile 1983: 236). Yet no open evaluation of other team members takes place (Déjean le Féal 1990: 154). And maybe this is just as well: Pöchhacker characterises the evaluation of an interpreter’s output by a colleague as subjective, regardless of expert knowledge and shared situationality (Pöchhacker 1994a: 124).

Altman notes that many interpreters are reluctant to correct their colleagues’ mistakes, undoubtedly because they do not want to strain the working relationship with their colleagues. On the one hand, one might argue: The interpreting performance is a goal to be achieved by both boothmates. So why be shy about correcting your boothmates' mistakes? On the other hand, strained relations with colleagues can have a detrimental effect on one's own performance – the most commonly listed factor in addition to the factors mentioned in her above-mentioned survey (Altman 1990: 29).

Since many interpreters are reluctant to have their output scrutinised, the interpreting research community is faced with methodological problems, and many researchers have commented on this phenomenon (Kalina 1994: 225; Gile 1995b: 161; Gile 1998a: 87; Kahane 2000; Gile 2001b; Gile 2003: 112; Chiaro & Nocella 2004: 280; Gile 2006: 20). Also, knowing that one is being monitored and/or recorded can affect an interpreter's performance. Unlike translators who know that their product is subject to later inspection, interpreters believe their output to be evanescent (Shlesinger 1989: 114). Let me conclude that it is not surprising that only few experiments have been carried out so far with interpreters judging their interpreter-colleagues' performance.

Rare examples of such experiments are two laboratory-setting studies in the late 1990s. In his 1999 study about the variability of in the perception of fidelity in SI, Gile found that both professional interpreters and non-interpreters were more lenient in auditory quality assessment (listening to the interpreter's output) as opposed to the visual mode (reading the transcript of the interpreter's output), detecting fewer errors and omissions in the auditory mode, and giving a higher overall fidelity rating (Gile 1999: 64). However, he also discovered substantial variability, both in numbers of errors/omissions found and fidelity ratings. The implication is clear: When asking evaluators to assess fidelity without clear instructions as to which fidelity criteria should be applied, one will be likely to produce variable results (Gile 1999: 68).

Collados Aís, in her study about the impact of monotonous intonation on overall performance assessment, had both users and interpreters judge interpreting output. It emerged that the latter were far more sensitive to delivery parameters than users, rating an interpreting performance with non-monotonous delivery significantly better, a performance with monotonous delivery far worse than users. Even though one might expect interpreters to be able to make sound judgements about interpreting quality, this was not the case. Collados Aís found a clear separation between actual quality and the perceived quality or success of a simultaneous interpretation in the assessment of an interpretation – for users and interpreters alike. It appears that interpreters are just as sensitive to the influence to extra-linguistic criteria as non-professionals (Collados Aís [1998]2002: 336).

Another experiment was carried out by Ackermann, Lenk and Redmond (Ackermann et al. 1997). Even though the experiment took place in a consecutive interpreting setting, and assessors were students, not fully-fledged professionals, the results were quite revealing and are probably, in the light of Collados Aís' findings, transferable to other settings. In this study, students were asked to assess the consecutive interpretation of fellow students. In no case were all errors detected, instead the average detection rate was, at only 30%, surprisingly low. One explanation may be that fellow students turned a 'deaf ear' to any errors their peers made. Another finding was that the overall performance assessment was not the composite result of individual observations. Even in cases where many errors were detected, the overall rating was much better than to be assumed. It seems that once judgement on overall impression is reached, the number of errors made are being mentally adjusted accordingly (Ackermann et al. 1997: 264 ff.).

So Moser-Mercer may be right in saying that interpreters "have an intuitive feel for good (or bad quality) – and yet intuition may fail an interpreter more often than he likes" (Moser-Mercer 1996: 45 f.).

3.2 Quality from the user's perspective

> Interpretation should always be judged from the perspective of the listener and never as an end in itself. The chain of communication does not end in the booth!
> *(Seleskovitch 1986: 236)*

Interpreting is intended for those relying on interpretation because they do not understand the source language. Obviously, listeners' needs and expectations play a pivotal role and are central to the entire discussion of interpreting quality. For practitioners, who view interpreting as a communication service, the concept of quality is seen as attached to that service as provided to users (Gile 2003: 109).

Even though references to users' needs can be found in early articles about interpreting quality – one example being Gold who stressed that the interpreter's language should be adjusted to the audience's expectations (Gold

1973: 155) – the literature about interpreting quality only took scholars' and teachers' perspectives into account for a long time, neglecting actual market needs and users' expectations (Gile 2001a: 379). Stenzl was one of the first authors to point out that the research community had failed to consider user needs and expectations (Stenzl 1983: 31 as quoted by Kurz 2001: 396).

The first study on user expectations was published in 1989 by Ingrid Kurz, who later established a formula for quality centred on user expectations: "Quality = Actual Service – Expected Service" (Kurz 2001: 405). Kurz carried out a survey among end users to determine whether Bühler had been right in her assumption that the results discussed in her paper also reflected the expectations of users, or whether "very often a good interpreter is two quite different people, being one thing to a conference participant and another to a colleague" (Cartellieri 1983: 213). In turn, Kurz's study also struck a chord in other researchers who followed in her footsteps, many of them conducting survey-based research to further explore users' *expectations*, others engaging in experimental or fieldwork research to find out more about users' actual *evaluations* of interpreting performances. The majority of researchers focused on users as listeners. Some research approaches also investigated the question how listeners' expectations may differ from those of other users of the service, namely the client and the speaker.

3.2.1 The user's (= listener's) perspective

> An interpreter should always have the most demanding listener in mind and aim at fulfilling all the criteria of good quality interpreting.
> *(Vuorikoski 1995: 173)*

> Our ultimate goal must be to satisfy our audience.
> *(Déjean le Féal 1990: 155)*

3.2.1.1 Survey-based studies – users' *expectations*

In her above-mentioned survey (Kurz 1989; Kurz 1993), Kurz found a strong correlation between AIIC evaluation standards and user expectations with regard to those criteria that are essential for communication, i.e. use of correct terminology, logical cohesion of utterance, and sense consistency with the original message. Other criteria were attributed far higher importance by AIIC members as compared to users, namely native accent, pleasant voice, correct usage of grammar, fluency of delivery and completeness of interpretation (Kurz 1993: 17). These results suggest that users do not consider 'superficial criteria' as significant. While the relative importance was fairly similar, interpreters tended to attach higher importance to almost all criteria in absolute terms, thus applying stricter standards to interpreting performance (Kurz 1993: 15).

In her 1993 paper Kurz built on her survey's results and examined whether different user groups have different expectations. The three groups were medical doctors (at an international conference on general medicine), engineers (at an international conference on quality control), and CE representatives (at a Council of Europe meeting). It emerged that there was fairly high agreement by all groups on the importance of some of the criteria but not all of them. Differences in assessment were found both not only between interpreters and delegates but also among various groups of delegates themselves. Some examples: Engineers rated grammar as less important than medical doctors and – especially – CE delegates. CE delegates were also more demanding than engineers and doctors regarding terminology and completeness but attached less significance to logical cohesion than medical doctors (Kurz 1993: 20). She carried out further surveys among TV representatives and TV interpreters (Kurz 1996a; Kurz 1996b). Elsagir provided an overview of the results in her paper (Elsagir 2000)[5].

[5] Expectations in media interpreting will be covered in greater detail under section 3.5.

	Kriterium	AIIC-Dolmetscher		Mediziner		Techniker		Europarat		Fernseh-mitarbeiter		Fernsehdolmetscher	
		n=47		n=47		n=29		n=48		n=19		n=10	
1	native accent	2,900	(8)	2,300	(8)	2,200	(7)	2,080	(8)	2,840	(6)	3,100	(6)
2	pleasant voice	3,085	(7)	2,800	(6)	2,400	(6)	2,396	(7)	3,470	(3)	3,500	(3)
3	fluency of delivery	3,468	(4)	2,900	(5)	2,966	(4)	3,208	(5)	3,320	(4)	3,500	(3)
4	logical cohesion	3,800	(2)	3,600	(1)	3,100	(3)	3,300	(4)	3,680	(2)	3,800	(2)
5	sense consistency	3,957	(1)	3,600	(1)	3,655	(1)	3,600	(2)	3,840	(1)	3,900	(1)
6	completeness of interpretation	3,420	(5)	3,000	(4)	2,900	(5)	3,458	(3)	2,530	(8)	2,500	(8)
7	correct grammatical usage	3,380	(6)	2,400	(7)	2,030	(8)	2,688	(6)	2,790	(7)	3,000	(7)
8	use of correct terminology	3,489	(3)	3,400	(3)	3,138	(2)	3,729	(1)	3,320	(4)	3,200	(5)
	Durchschnitt	3,440		3,000		2,800		3,060		3,220		3,290	

Überblick über die Untersuchungsergebnisse zur Qualität von Dolmetschleistungen von Bühler (1986) and Kurz (1996a, 1996b) im Hinblick auf die Bedeutung der oben aufgeführten Kriterien, 4 = sehr wichtig, 3 = wichtig, 2 = weniger wichtig, 1 = unwichtig, n = Anzahl der Befragten, mit Angabe der Durchschnittswerte und der Rangfolge in Klammern

Fig. 7: Comparison of results obtained in the surveys of Bühler (1986) and Kurz (1996a, 1996b) (Elsagir 2000: 109)

Meak conducted a survey among participants in a medical conference. She found that these were relatively lenient regarding inadequate (non-scientific / too informal) terminology as they were used to patients using wrong terms in surgery (Meak 1990: 9). Marrone's survey confirmed the preference of 'content over form'. He found that users' expectations focused on substance, fidelity and completeness rather than pleasant voice and delivery (Marrone 1993: 38). Regarding the interpreter's role, listeners favoured the 'cultural mediator' over the 'scrupulous translator' role although this majority (at 47 vs. 40) was not an overwhelming one (Marrone 1993: 38).

The largest survey to date has been carried out by Moser (Moser 1995). His questionnaire was by far the most elaborate one. One reason may be that it was the only study funded by AIIC. Faithfulness to the original was the expectation that was most frequently mentioned spontaneously (Moser 1995: 8). In his survey, Moser distinguished between different conference types (large technical, small technical, large general, small general) as well as users' conference-going experience and found that, although expectations were somewhat different in various groups, the importance attached to the principal criteria remained constant. The main differences were that users in large conferences were slightly more in favour of the interpreter focusing on essentials as opposed to completeness (Moser 1995: 23). The smaller and more technical the conference, the greater the desire for completeness and literalness

(Moser 1995: 24). However, the average of responses showed a clear preference of faithfulness to meaning over literal translation – even in technical meetings (Moser 1995: 17). This raises the question of whether Moser-Mercer's definition of quality as a "complete and accurate rendition of the original" (Moser-Mercer 1996: 44) corresponds to actual users' needs and expectations, as pointed out by Lamberger-Felber (Lamberger-Felber 1998).

Variability of preferences could also be found among users of different age groups and different genders. Users under the age of 30 were less focused on 'essentials over completeness' than older users, the latter also attaching higher importance to terminological accuracy than their younger colleagues. Women attending technical meetings ranked terminological accuracy higher than their male counterparts, the same goes for synchronicity. They also proved more demanding on rhetorical skills and were more strongly disturbed by 'ums' and 'aahs' (Moser 1995: 16 ff.).

In an open-ended question in Moser's survey, synchronicity was the formal (as opposed to content-related) factor that was mentioned most often as a source of irritation; nearly 90% of the people who mentioned this aspect were disturbed by poor synchronicity (Moser 1995: 19). This preference had already been mentioned by Déjean le Féal in 1990: "Users prefer that the interpretation be truly simultaneous, i.e. without interruptions or delays in the onset of interpretation – they assume that silence is equivalent to a loss in information" (Déjean le Féal 1990: 155). A long ear-voice span, even though strategically efficient in producing high-quality output, may thus be not what users want.

Another interesting finding of the study was the audience's sympathy with and consideration for the interpreters. More than half of the respondents thought that a simultaneous interpreter should not work for more than 30 minutes without a break (Moser 1995: 13).

3.2.1.2 Performance *evaluation* by users as opposed to their *expectations*

3.2.1.2.1 Performance evaluation in natural settings

In 1990, Gile carried out a case study at a medical conference which included a survey among delegates. One purpose of the survey was to prove his hypothesis that users' *expectations* and the actual *evaluation* of a performance by users are not necessarily identical. He discovered that English-speaking listeners proved less demanding than French, rating all interpreting criteria unanimously as "excellent", while the French listeners did not (Gile 1990: 67). So it appears that there can be variability among users even in the same conference setting and the same discipline. Gile gives two potential reasons for the difference in evaluation: Either Anglophones are generally less critical than French users, or the English-speaking participants in that particular conference were less demanding, being the minority of participants in a non English-speaking environment, therefore all the more grateful that interpretation was provided (Gile 1990: 67 f.). Another of Gile's findings was that – as in the study by Kurz – users attached less importance to voice quality features than interpreters. He raised the question whether this was a specific characteristic of the user group (medical conference participants), who being "scientifiques et techniciens seraient moins sensibles à la qualité de la voix, du rhythme et de l'intonation ... que d'autres publics" (Gile 1990: 68).

Ng carried out a survey among Japanese native speakers, using both questionnaires and interviews, and asked them to evaluate five interpreters' performances. Although the purpose of the survey was to explore a Japanese-specific phenomenon, namely the use of speech levels, another aspect moved to the foreground: sex differences. Females placed significantly higher importance on correct grammatical structures and speech levels (Ng 1992: 38). This finding was corroborated by Peter Moser in his survey as quite a number of gender-related preferences emerged from the data (Moser 1995: 16 ff.; cf. 3.2.1.1).

Vuorikoski explored users' perception of SI as a 'fieldworker' during five seminars where interpretation was primarily provided for the Finnish audience. She used questionnaires that she had participants fill out directly after the seminar, and conducted phone interviews a few days later. Many users in

Vuorikoski's study among seminar participants in Finland (being Finnish, most of them had learnt English at school or elsewhere) used SI much in the same way as subtitles on TV. When they realised that the interpreter was lagging behind, they tried to listen to the original. When they had not understood a complex or technically difficult part of the discourse, they switched back to interpretation (Vuorikoski 1993: 324). Users clearly expressed that they preferred the interpreter to be well informed about the subject, as opposed to correct terminology only. Another finding of the study: Not only are there different expectations between different audiences, as Kurz suggested, but also significant variations among the members of one single audience (Vuorikoski 1995: 173). The question is to what extent it is possible to assume that a certain audience has certain expectations of SI – or "if [it is]at all possible to find any unity in the diversity of audience responses" (Vuorikoski 1998: 187 f.). In order to determine which factors influence the overall perception of quality it would be advisable to establish a thorough classification, a typology, encompassing different settings and forms of interpreting, such as media or relay interpreting (Kahane 2000).

The same multimethod research model employed by Vuorikoski, including phone interviews and questionnaire surveys, was also put to use by Mack and Cataruzza. A questionnaire asked users to rate certain criteria (informed, easy to follow, accurate, pleasant rhythm, fluent, terminology) that they expected in high-quality interpretation, but also assess the actual interpreting performance along these criteria. There were significant differences between inexperienced users of SI, and experienced ones. The group using SI for the first time tended to give lower marks but also to expect less whilst the other two groups (those having experienced SI 1-5 times and more often than 5 times) assigned higher marks but also expected more. It seems that the frequent use of SI sharpens users' awareness of quality (Mack & Cataruzza 1995: 45).

3.2.1.2.2 Performance evaluation in experimental settings – quality as an input variable

> Users are not good judges of quality, simply because they are not
> in a position to perform this task.
> *(Collados Aís [1998]2002: 336)*

After several surveys had produced similar results regarding the importance – or rather unimportance – of pleasant voice and intonation attributed by users when asked what they *expected* in good interpretation, Collados Aís set out in 1998 to investigate whether abstract user expectations match their actual performance evaluation in practice, and to what extent monotonous intonation in SI has a negative effect on a) judgements of overall quality and b) on other quality criteria. The survey that was part of her study confirmed that the participating users shared the same expectations as those in previous studies: The accurate rendition of the sense of a message was the most important expectation of users whilst accent, pleasant voice and intonation were rated as least important (Collados Aís 1998: 239 f.).

Collados Aís' findings were groundbreaking insofar as they challenged the validity of user expectations. She found that users, although claiming to put content over form, attributed primordial importance to form-related criteria in actual performance evaluation. She also discovered a clear separation between actual quality and the perceived quality or success of a simultaneous interpretation (Collados Aís [1998]2002: 336). Three videos were played to the subjects taking part in the experiment, one with monotonous intonation, and fully consistent with the sense of the original discourse, one with lively intonation, but not fully consistent with the sense of the original discourse, and one with lively intonation, and fully consistent with the sense of the original discourse. Version 2 received the highest rating in the overall assessment of the interpretation, and although a significant difference could only be established between versions 1 and 3, subjects clearly favoured version 2 over version 1 as well. Video 1 received lower scores for all quality criteria, which indicates that monotonous intonation has a negative effect on the evaluation of other quality criteria. Video 2, the most animated version, received the highest ratings for all of the individual criteria, followed by video 3. There was a the close link between the parameters 'monotonous delivery' and 'pleasant voice': the less monotonous the interpreter's output, the better the scores for pleasant voice, both among users and interpreters (Collados Aís 1998: 244 f.). One can thus conclude that the less monotonous the intonation the more positive the users' assessment of the interpretation will be (Collados Aís [1998]2002: 328 ff.).

The study revealed that monotonous intonation affects the overall rating of an interpreter's performance and the assessment of fidelity and professionalism.

This phenomenon applies to interpreters and users alike (Collados Aís 1998: 157 ff.). This means that neither users nor interpreters are judges apt to give a reliable performance assessment since significant content errors were not detected by users or interpreters, and – as a logical consequence – had no effect whatsoever on the overall quality scores (Collados Aís 1998: 239 ff.).

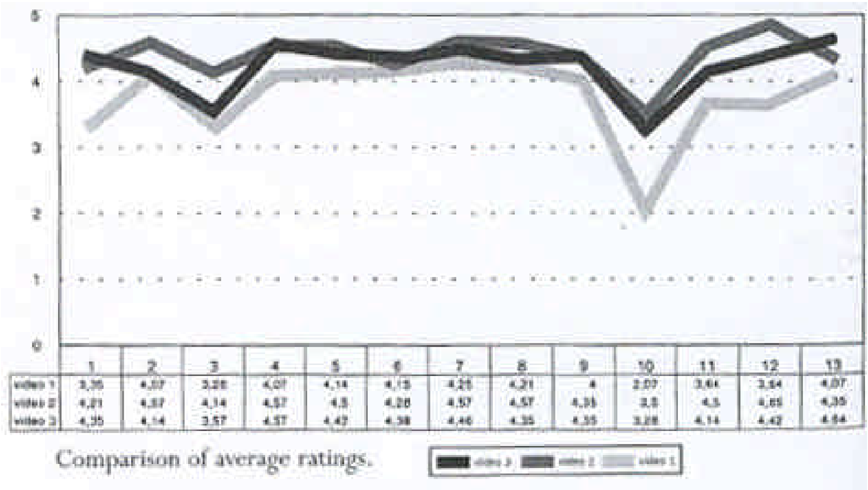

Fig. 8: Comparison of average ratings for the three videos (Collados Aís [1998]2002: 334)

Garzone discovered a similar pattern in her 2002 survey among medical doctors and other professionals. Analogous to Collados Aís' approach, she first asked respondents to rate quality criteria in theory, and then checked these expectations against actual performance evaluation scores. The results mirrored those of the earlier study by Collados Aís. Even though all respondents had rated 'pleasantness of voice' and 'fluency of delivery' as less important, these formal characteristics had a marked impact on the assessment of other aspects (Garzone 2002: 23 ff.). Garzone concluded that responses to questionnaires often reflect abstract ideas concerning quality criteria. Therefore, they are not reliable if used to formulate a concept that can be effectively applied in practice. Surveys also allow respondents to reflect on expectation criteria in an isolated manner. The situation is altogether different when users are asked to assess an individual performance as different elements will overlap and interfere with one another. As a consequence, a performance which is poor in prosody and fluency is perceived as less correct, even if it is not. Garzone leaves the question unanswered as to whether this is due to the objective difficulty in following a TT which is not well presented, or to a psychological effect that makes the text appear less reliable to the user if its form is not appealing (Garzone 2003: 28).

Pradas Macías built on Collados Aís' work and used the same speeches to explore the role of silent pauses in fluency – the former being a subparameter of the latter. The two videos into which additional pauses had been inserted (video 1 and 2) did not only receive lower scores for fluency but also for other parameters, such as correct rendition of sense (Pradas Macías 2006: 36 f.). Another finding referred to the evaluation of intonation that received the lowest scores of all criteria in all three videos. Pradas Macías explains this result with the fact that the simulated video was recorded as voice-over, and the German original speech – with a somewhat monotonous intonation – was audible in the background (Pradas Macías 2006: 37). This could confirm what Ng pointed out in his paper, namely that users have a tendency to 'blame the interpreter', and that his subjects never attributed the lack of clarity of the message to the original speaker but rather held the interpreter responsible for it (Ng 1992: 37). And it would support interpreters' self-perception as 'scapegoats' which they expressed in Altman's survey (Altman 1990: 23; cf. 3.1.1.1). Surprisingly, video 2 (manipulated) received some of the higher ratings, e.g. on overall quality. However, the study may have been methodologically flawed as video 2 was viewed and evaluated by the group having comparatively more experience with SI than the other two groups, i.e. this result could be in line with Peter Moser's finding that more experienced users of SI tend to give higher scores (Pradas Macías 2006: 38).

3.2.1.3 The effect of interpretation on users

Shlesinger conducted a study on the effects of intonation on the perception of interpretation. The study confirmed what could be inferred from the literature on the effects of intonation on communication in general: that intonation has an effect on the perception of interpreting output. Shlesinger tested the hypothesis that good rhythm and intonational patterns are "conducive to information retrieval" (Shlesinger 1994: 227). In the study, one group of listeners was confronted with an interpreted version of a text, the other group with the read-aloud version of that same text. Salient features of interpretational tonality included pauses in final positions, stress incompatible with semantic contrast, mismatch between pitch movement and discourse pattern, and lengthening of segments. Both groups had to answer questions designed to test comprehension and recall afterwards. There was a 20% difference of correct answers between group 1 and 2, suggesting that information processing and retrieval was significantly facilitated by ‘natural’ (as opposed to ‘interpretational’) intonation (Shlesinger 1994: 228 ff.).

3.3 The speaker’s perspective

> Not only the listener but the speaker, too, must be served.
> *(Shlesinger et al. 1997: 127)*

This claim is justified. As a mediator between source and target language, the interpreter is required to render the speaker’s message, one norm being that of the “honest spokesperson” (Harris 1990: 118, cf. 1.2.1). The question is: To what extent should the interpreter leave the ‘honest spokesperson’ role or and ‘intrude’ by omitting, summarising, adding information to the text? Interpreters are experts in several languages and cultures and thus have a tendency to intrude. They will usually try to please the target audience: “All interpreters are likely to embellish, enrich and generally raise the tone to project a positive image, which will also increase the possibility of more work” (Katan & Straniero Sergio 2001: 224).

But the question remains: Should interpreters really intrude, or refrain from intrusion (Kopczynski 1994: 90)? Kopczynski found the answer in his survey among 19 speakers and 38 listeners. While both groups preferred the ghost role to the role of intruder, listeners were more strongly in favour of the ghost role than speakers – 58% (speakers) versus 84% (receptors) – although there was a slightly more liberal attitude in the ‘humanities’ group while diplomats were more restrictive. Interestingly, most respondents would allow corrections of the speaker – seemingly unaware of the fact that this contradicted the (allegedly preferred) ghost role of the interpreter (Kopczynski 1994: 96 f.). Jones claims: better intervene too rarely than too often (Jones 1998: 122).

Schjoldager holds a different view. The speaker is chiefly interested in the interpreter’s loyalty to him whereas the listener is primarily focused on the interpreter’s skills as a speaker and on his succeeding in faithfully rendering the speaker’s message (Schjoldager 1996: 189). The issue of loyalty is raised by Gile: “The Translator[6] is working for the Sender but also for the Receiver and

[6] Gile uses the term ‘Translator’ with a capital T to denote both translators and interpreters.

the Client, the interests of whom may not coincide" (Gile 1991: 192). Communication is successful, from the sender's point of view, if he manages to bring the message across to the listener. According to the widely accepted norm that the translator acts as 'alter ego' or 'honest spokesperson', he should act in the interest of the speaker, even against the aims of the listener or client. Although the client is not directly part of the sender-to-receiver communication process, and strictly speaking his interests should not be taken into account, the interpreter might be tempted to do so, for practical and psychological reasons. It is a fact that the interpreter's livelihood depends on the client, and it is only understandably human that the interpreter will try to satisfy his needs. It is the same reason why many conference interpreters who are recruited by colleagues more often than by clients attach the highest importance to their reputation in the profession, as opposed to users' feedback (Gile 1991: 192 f.).

3.4 The employer's perspective

> Do our clients know what's good for them?
> *(Shlesinger et al. 1997: 126)*

Moser-Mercer points out that the same methods used to evaluate quality from the user perspective can also be used for the employer or client perspective. Accuracy of content and technical terminology, rhetorical skills but also team discipline, adaptability to different situations, flexibility in scheduling, price, availability and loyalty to a particular employer are among the most important criteria. Obviously, different employer categories must be considered: agencies, professional conference organisers (PCOs), large language services of international organisations (Moser-Mercer 1996: 50 f.).

Seleskovitch claims that "employers are not competent judges since most of them never have occasion to actually use interpreters' services" (Seleskovitch 1986: 236). However, as Pöchhacker points out, the observation whether a discussion went smoothly may very well give the employer a certain idea of interpreting quality, possibly complemented by users' feedback that the employer will try to obtain (Pöchhacker 1994a: 124). One has to bear in mind, however, that people choosing interpreters are often not measured against the quality of the interpretation provided but rather on the basis of their budget compliance (Feldweg 1996: 151).

There is one example of an employer systematically measuring quality of interpretation: the DG Interpretation, the European Commission's interpreting service and conference organiser. To quote from its Annual Activity Report: "In terms of quality assurance and evaluation, 2005 was a landmark year" (European Commission 2006: 4). The electronic reporting system for interpreter evaluation (SERIF) which had been launched in 2004 was refined and further developed. A total of over 1,800 reports were filed covering almost 800 different interpreters. Also, quality ratings incorporating the above data were produced for over 2,000 individual interpreters and taken into account in offers of recruitment" (ibid.). "The assessment system added cost efficiency and management considerations to the list of quality-related concerns. Here, the trends already recorded were confirmed. Customer feedback was monitored systematically and satisfaction rates were running at 97%, similarly high to the 98% of 2004" (European Commission 2006: 20). As Kahane points out, it is not surprising that the DG Interpretation should perform systematic analyses of interpreting performances, being the world's largest client of interpreting services (Kahane 2000).

3.5 Excursus: User expectations and performance assessment in media interpreting

Media interpreting (MI) is the best-known form of interpreting, or at least the form of interpreting that most users have experienced. Elsagir compared user expectations regarding non-content-related features from Bühler's, Kurz's and Moser's surveys with her own findings obtained in a survey among TV interpreters, TV professionals and TV viewers. She found that TV viewers were more demanding than conference delegates since three out of four factors were rated as more disturbing by viewers than conference delegates (speaking with a low voice, monotonous intonation, use of 'ums' and 'aahs'). The only factor regarding which TV users were less demanding than delegates was the frequent occurrence of long pauses. The comparison between TV professionals and TV viewers revealed that they had very similar quality expectations which, however, contrasted significantly with those of the other groups (Bühler 1986; Moser 1995; Kurz 1996a; Kurz 1996b; all quoted by Elsagir 2000: 117 ff.).

Kriterium	AIIC-Dolmetscher n=47	Mediziner n=47	Techniker n=29	Europarat n=48	Fernseh-mitarbeiter n=19	Fernsehdolmetscher n=10	Fernsehzuschauer n=57
1 native accent	2,900 (7)	2,300 (7)	2,209 (6)	2,090 (7)	2,840 (5)	3,100 (5)	2,579 (7)
2 pleasant voice	3,085 (6)	2,600 (5)	2,409 (5)	2,396 (6)	3,470 (2)	3,500 (2)	3,123 (4)
3 fluency of delivery	3,458 (3)	2,900 (4)	2,966 (3)	3,208 (4)	3,320 (3)	3,500 (2)	3,158 (3)
4 logical cohesion	3,800 (1)	3,600 (1)	3,103 (2)	3,300 (3)	3,680 (1)	3,800 (1)	3,404 (2)
5 completeness of interpretation	3,428 (4)	3,000 (3)	2,900 (4)	3,458 (2)	2,530 (7)	2,500 (7)	3,509 (1)
6 correct grammatical usage	3,360 (5)	2,400 (6)	2,533 (7)	2,688 (5)	2,790 (6)	3,000 (6)	3,070 (6)
7 use of correct terminology	3,489 (2)	3,400 (2)	3,138 (1)	3,729 (1)	3,320 (3)	3,200 (4)	3,123 (4)
Durchschnitt	3,364	2,886	2,676	2,980	3,138	3,229	3,138

Fig. 9: Comparison of results obtained in Elsagir's survey with those from the surveys of Bühler (1986) and Kurz (1996a, 1996b) (Elsagir 2000: 120)

Fluency might have obtained higher ratings from TV representatives because of the overwhelming time pressure in TV broadcasting, every minute 'on air' being expensive, a fact that both interpreters and TV employees seem to be acutely aware of (Kurz & Pöchhacker 1995: 357).

In the case of Formula One press conferences, users are fans who are keen on seeing their idols and sharing their emotions and not so much focused on pure informational content. SI therefore has a mainly phatic function, i.e. it performs a social task. The focus is not so much on what is being said but rather on the fact that communication does not break down. As Straniero Sergio puts it: "In MI the interpreters are judged not for interpreting a speech correctly but convincingly well. It is the form and not the content that both broadcasters and viewers respond to" (Straniero Sergio 2003: 169 ff.).

In film interpreting, synchronicity is even more important to interpreting quality as shown in a survey on users' expectations for simultaneous film interpreting by Guardini. While completeness and rendition of general content ranked first and second, the two criteria that came in third and fourth were fluency and synchronicity (Guardini 1995 as quoted by Russo 2005: 13).

It thus appears that TV and film viewers are the only groups who openly admit to the high importance of form when asked about their expectations whereas other groups are more 'ideologically biased' and place content over form – even though the reality in actual performance evaluations might look completely different.

A possible explanation lies in the fact that conference interpreting can be regarded as an 'intra- or inter-professional discourse', i.e. it takes place in a setting where the discourse to be interpreted is being produced by professionals. Unlike in CI, the MI text can be described as a 'professional-lay' discourse, i.e. a discourse which is directed at a mass audience which is not embarrassed to admit that form matters (Linell 1998 as quoted by Straniero Sergio 2003: 167). As a great number of media texts may be defined as 'limited-quality' texts, it is not at all surprising that listeners want this limited quality of the original to be improved to a better-quality interpretation (Straniero Sergio 2003: 171).

3.6 Pragmatic implications

In interpreting, as a communication activity, interpreters cannot and must not neglect any parameter that touches upon the concept of quality or successful interpretation (Collados Aís 1998: 251). In other words: Even criteria that are seemingly unimportant deserve the interpreter's attention. And experience has shown that unimportant criteria from a theoretical standpoint may well become very important in real life when it comes to assessing an actual performance. One conclusion to be drawn is that intonation and voice need to become an integral part of any interpreting training course, and that interpreters should, as a continuous QA effort, work regularly on their voice quality.

There have been claims to further proceed along the user research lines to identify the expectation profiles of different groups of recipients so that interpreters can devise strategies to meet these expectations, as for example at the Saarbrücken Symposium in 2000 where several speakers expressed this opinion (Kurz 2001: 396 f.). The findings of surveys seem to confirm this need, as do other references in the literature. According to Viezzi, the Chinese delegation to the UN favours a literal rendering – i.e. their expectations do not match those of interpreters themselves (Shlesinger et al. 1997: 127). A similar situation is described by Kahane, with Northern Europeans favouring a more literal rendering of the original and Mediterraneans more prone to an interpretation that adjusts the target speech to the users' cognitive skills, in his specific case prisoners who otherwise would not have been able to understand the judges' lofty terminology (Kahane 2000).

One aspect of surveys that might prove useful in strengthening the profession's viability is to ask participants whether they think interpretation is useful. Not being aware of financial constraints, they will (hopefully) say: *Yes, it is!*

Yes, we need it! (Vuorikoski 2003: 321). A similar finding to Vuorikoski's seminar feedback emerged in Mack's and Cataruzza's study; 91.4% of participants stated that SI should be offered in multilingual meetings to those who want to use it (Mack & Cataruzza 1995: 43). Such statements give interpreters some leverage in the interminable cost-benefit discussion. Listeners are often the customers of an interpreter's client. By obtaining knowledge about what listeners – the client's customers – want, interpreters can strengthen their negotiating position. Only one methodological obstacle remains, as Feldweg points out: There are many situations where one can expect users to be reluctant to give honest feedback, e.g. participants from a developing country attending a seminar in Germany who may not openly criticise the organisers of the seminar for fear of not being invited again (Feldweg 1996: 151).

Yet the findings from surveys and studies have been very valuable. The discrepancy between results of surveys and actual performance assessments has shown that users are likely to base their judgment on form rather than content. And this does not come as a surprise when accepting the view that users are not capable of judging quality (Collados Aís 1998: 250; Collados Aís [1998]2002: 336; Chiaro & Nocella 2004: 281). It offers an explanation for the "surprisingly favourable assessment of quality in conferences in which the interpreters themselves feel they have done a poor job" that Gile describes (Gile 1991: 198).

Respondents in Moser's 1995 survey also gave interesting feedback regarding the ear-voice span. It is generally acknowledged that an appropriate *décalage* is crucial; initially it was assumed that a longer ear-voice span was tantamount to better interpreting quality (e.g. Barik 1975). This statement has to be seen from a new perspective. In our times, as powerpoint presentations are widely used, listeners want to be able to follow such presentations closely, i.e. they want the interpretation to match the timing of the original (Moser 1995: 8). Thus the only way to use the ear-voice span as a quality parameter would be to measure an interpreter's flexibility regarding his *décalage*, and monitor to what extent he succeeds in achieving a quasi-zero EVS when required, e.g. when a speaker uses charts and points to certain graphics or diagrams, and switching back to a longer EVS when truly simultaneous interpretation is no longer necessary.

Quality assessment of the live – and evanescent – product that interpreting represents is difficult and extremely unreliable. This is also and especially true for interpreters monitoring their own output and detecting a deterioration in performance (Déjean le Féal 1990: 154; Moser-Mercer et al. 1998: 55). On the other hand, users are even less suited for quality assurance as they do not even know the target language. As Gile points out, fidelity is a quality parameter that can only be assessed through constant comparison with the source text (Gile 1995b 152). It is this very reference to the source-language speech that makes fidelity assessment in SI so difficult as noted by many authors (e.g. Cartellieri 1983). And how do users try to perform the appraisal process? A) through sampling, i.e. by listening to selected segments of the original and checking whether these are found in the target text, and b) through accidental error detection when a segment heard in the target text does not seem plausible to the listener, which leads the listener to the assumption that an error has been made (Gile 1995b: 152 f.). Users thus have no objective elements to check the intertextual coherence between TT and have to rely on indirect indicators.

Therefore, as Garzone concludes, quality assurance rests almost exclusively on interpreters' shoulders – even though, as mentioned above, they might not be reliable judges of quality themselves (Garzone 2003: 29). One solution to the dilemma may be self-recording one's own output as an efficient means of quality assurance (Kalina 1994: 229; Schjoldager 1996: 188). It is recommended to students but should be a part of every professional interpreter's routine. Self-recording – although to a certain extent subjective – will also allow an interpreter to be more objective insofar as he will compare the colleagues' performance with his own (recorded) output rather than with his (perceived) ideal performance (Déjean le Féal 1990: 156 f.).

We have also learned that interpreters might be rendering themselves a disservice when – "although confident about their ability to improve upon the quality of a speech, interpreters hesitate to do so, for fear of overstepping the limits of their role in the communication process" (Altman 1990: 29). Collados Aís, drawing the consequences from her study on the impact of monotonous intonation on performance evaluation, concludes that "users not only desire but demand ... a certain degree of intrusion or active involvement on the part of the interpreter". To her mind, the interpreter should assume the conscious role of professional communicator and go beyond the ‘ghost role’ (Collados Aís [1998]2002: 336).

On the other hand, interpreters should not give in to the temptation of resting on their laurels, even if they know that their rhetorical skills are excellent and they can assume that any errors they make will go unnoticed. In the long run, they will not. Every interpreter who has ethical standards knows that his job is to “[comprehend] the concepts of the speaker’s message and conveying them orally in another language” (AIIC 1994 as quoted by Moser-Mercer 1996: 44). Even though the user might prefer ‘attractive packaging’, the interpreter must bring the substance of the message across. He must attach the highest importance to the faithful rendition of the original.

4 Interpreting quality in the training environment

4.1 Are interpreters made or born? The expert-novice paradigm

> Interpreters are born, not made.
> *(van Hoof 1973 in 'The Times', published on 26-09-73, quoted by Weischedel 1977: 101, as quoted by Kalina 1998: 233)*

There is general agreement that this statement does not hold true (Kalina 1998: 233). The quality of interpreting depends on certain skills and strategies that need to be acquired over time, usually as part of a university training programme.

One issue inherent in the question of what makes a good interpreter, and closely related to the teaching activity, is the question as to how the interpreting output of experts and novices differs, both in quality and in processing terms (Sunnari 2003: 235). Teaching is aimed at turning novices into experts – students into professionals – who are able to provide high-quality interpretation. Before one can talk about what it necessary to achieve this goal, one first has to look into what sets experts and novices apart since "the models that are characteristic of skilled interpreting do not apply to novices" (Dillinger 1990: 155).

Studies on expertise go back to DeGroot and Chase & Simon who investigated the different behaviour patterns of expert and novice chess-players (DeGroot 1966; Chase & Simon 1973; both quoted by Moser-Mercer et al. 2000: 109). They concluded from their experiments that experienced chess players perceive relations between chess pieces as 'chunks' and are consequently able to memorise a larger number of pieces and valid board configurations. It appears that experts and novices differ in how their knowledge base is organised, and information is being processed.

This principle seems to apply to interpreters as well as chess-players. Expert interpreters display a better organisation of their factual knowledge and make more associative connections. Experts' semantic interpretation is almost always tied to the context whereas that of novices is often unrelated to the context. As regards interpreting strategies, experts tend to proceed from known to unknown information whereas beginners often focus on the unknown and then get stuck. Experts' approaches are more global; novices tend to favour low-level (microcontextual) plans (Moser-Mercer et al. 2000: 108 f.). On their way to expertise, novices move from micro to macro level (Moser-Mercer 1996: 53). This view is shared by Kurz who states that professional conference interpreters have, through their training and experience, acquired sufficient expertise – defined by Kurz as "a combination of knowledge and better strategies" – which is reflected in the ability to process larger segments, and to adopt the right strategy quickly, or even automatically (Kurz 2003: 60).

Riccardi identifies differences between experienced interpreters and novices as different levels of creativity. Novices tend to keep the same word order in both languages. The more the target-text structure resembles the source-language structure, the less it is creative. Novices should build up a greater association capacity to produce a more creative interpreting performance (Riccardi 1998: 176).

4.1.1 Empirical studies

4.1.1.1 Strategic processing

Kalina tested the hypothesis that professionals' strategic processing differs from that of beginners by comparing students' and interpreters' output. She used an authentic corpus (from a three language, three-booth simultaneous conference, consisting of 20 tapes of 90 minutes each) to compare professionals' performance with the interpreting output of students in a mock conference where the same source text recordings were used (Kalina 1994: 225 ff.). Transcripts revealed that strategies of professionals were at a higher level. Experts had a lower correction rate for minor errors and a higher correction rate for severe errors. Increased cohesion and connectivity in professionals' output showed that they were more user-oriented. They also tended to have a different attitude towards their own deficiencies: Problems that had not been solved successfully did not have a far-reaching effect on other parts of the text, unlike in the case of students' performances. Moreover, experienced interpreters made less use of the sentence-splitting strategy; they even tended to connect more utterances. It thus appears that the total capacity of professionals is higher than that of students (Kalina 1994: 229 ff.).

Sunnari's 1995 analysis of expert performances revealed that expert interpreters construct the macrostructure of what has been heard by applying certain macrostrategies to the micropropositions such as selection, deletion and generalisation (Sunnari 1995 as quoted by Sunnari 2003: 239). Observations in her study suggested that experts resort to strategies that lead to more effective content communication than novices' strategies. Experts were able to reproduce in their output a higher number of ST propositions and established more accurate links between these propositions. They also succeeded better in filtering out redundant information, processed larger chunks of information and not only showed better accuracy but also greater economy of performance (Sunnari 2003: 244).

4.1.1.2 The shadowing experiment

An interesting experiment was carried out by Moser-Mercer as part of the Geneva Project on Aptitude Testing. Five students and five professionals shadowed a text in French and English, with the instruction to repeat the incoming message as quickly and accurately as possible. Although it might have been expected that interpreters – who are experienced in listening and speaking simultaneously – would perform better than students, the contrary was the case. Both in terms of delay and errors, professionals did not shadow the text more efficiently. Professionals' errors, however, were not errors distorting the sense of the message; professionals made greater use of substitutions but did not alter the meaning. The delays could well reflect the difficulty interpreters had in suppressing automated processing strategies and in adapting to changed processing requirements (Moser-Mercer et al. 2000: 113 ff.).

4.1.1.3 Students and stress

A study by Kurz investigated what sets novices and experts apart in the way they perceive, deal with and are affected by stress. In her pilot study designed to measure stress levels among students and professionals, pulse rate and skin conductance level (SCL) of two interpreters interpreting at a – highly technical and demanding – medical conference and three students in an interpreting class were measured. Although the SCL method failed to discriminate between experts and novices, the average pulse rate varied significantly

(97.6 students; professionals 74.4), clearly indicating higher stress levels among students. Kurz concludes that conference interpreters have learned to overcome their ‘stage fright’ with experience and are more stress-resilient than beginners. Student interpreters still grapple with problems of anxiety and stress (Kurz 2003: 62 ff.). This relatively high stress-resilience of professional interpreters was also confirmed in studies on prolonged terms in interpreting and the AIIC Workload Study (Moser-Mercer et al. 1998; Mackintosh 2002; cf. 2.1.12).

4.2 Teaching quality

The implications of the expert-novice paradigm are clear. If novices’ and experts’ performances differ particularly in the use of automated, strategic macro-processing skills, teaching must focus on the process, not the product. Expertise cannot be acquired over night – it is a developmental process. This has to be recognised by teachers, or they risk “comparing apples and oranges” (Moser-Mercer 2000: 339 as quoted by Kurz 2003: 64).

In-training quality assessment differs significantly from quality assessment in the field. The trainer’s ultimate objective may be the same as that of users and selectors, but ideally he engages in the process, not only in the product (Gile 2001a: 379; Gile 2003: 110). Teachers must enable students to acquire strategic processing skills at their own pace, instead of encouraging attention to detail over meaning (Moser-Mercer 1996: 53).

Viaggio, a proponent of the deverbalisation principle, underlines the importance of abstraction and compression as a crucial skill that novices have to acquire. Students should refrain from what he terms the “saying-it-all” approach (“conveying the whole of the same sense with as many of the stylistic and semantic nuances as can be possibly reproduced on the spot without abusing one’s target language”) but compress the message whenever possible (Viaggio 1992: 51). They should not open their mouth until they are absolutely sure that they have understood the sense of the speaker’s message and have also planned their own output. In Viaggio’s words: “Death by silence is better and more dignified than death by inanity” (Viaggio 1992: 49). Viezzi follows this line of thought and calls the version to be aimed at “synthetic interpreting” (Viezzi 1993: 396).

4.2.1 Quality assessment

4.2.1.1 Entrance tests (aptitude tests) and final examinations

In many countries, recruitment tests for interpreters consist of 'on the spot assessment' – no recordings or transcripts are made, and there is no uniform set of criteria. The situation is similar when it comes to university entrance tests. Although the objectiveness of the assessment is somewhat helped by assessing candidates in panels to avoid individual judgements, assessors find it difficult to put their judgements in the wording of legally incontestable, formal reporting language (Kalina 2005: 769 f.). It is therefore crucial that a clear set of criteria be established to make entrance test procedures both more transparent and more reliable. I will not go into greater detail on this topic as this would exceed the scope of this thesis. Suffice to say that a number of parameters are considered as important for the success of a trainee interpreter: high proficiency in the active and passive language and cultures, the ability to grasp rapidly the meaning of a source text and convey the essential meaning adequately, the ability to project information with confidence, good voice, wide general knowledge, the ability to memorise new information rapidly, and team skills (Lambert 1991: 586). The AIIC Training Committee mentions linguistic competence, general knowledge, maturity, basic code-switching ability, resourcefulness, and the ability to cope with stress as some of the criteria that can shed some light on the level of preparedness of a candidate wishing to enter an academic training programme (AIIC Training Committee 2006).

The selectors' perspective may be somewhat different. While selectors examine the product in a similar way as users, they will have some flexibility since it is not only the product 'here and now' that counts but also the potential, and they will allow for a certain adaptation period (Gile 2003: 110). To avoid subjective judgements, aptitude tests should be carried out by a panel of professional conference interpreters (who are also teachers). They should offer feedback to the candidates regarding their ability to perform the designated tasks and their readiness to begin training (AIIC Training Committee 2006).

Regarding final examinations, the jury principle is equally important. However, the TC suggests that this jury should not only consist of instructors but should include external interpreters who are also active conference interpreters (ibid.).

4.2.1.2 In-class assessment

4.2.1.2.1 Assessment by teachers

One can conclude that students perceive the artificial monitoring in class – as opposed to interpreting for real-life audiences – as stressful. This is true both for performing in front of the group and in front of teachers, who will usually take notes when listening to a student's performance (Ackermann et al. 1997: 266).

In the teaching process, assessment is essential to guide students on their way from novices to experts. Therefore, some kind of systematic approach is required. To develop individual student profiles, Moser-Mercer suggests that students use a diary on their performance in class, and this information – together with the teacher's perception of a student's performance – can then be fed into a matrix to get a comprehensive picture of the learning status (Moser-Mercer 1996: 53). Curriculum and learning curve must obviously be taken into account. Overcorrection can have a demotivating effect on students, undercorrection can raise false expectations (Moser-Mercer 1996: 53).

Dillinger postulates a complex scoring system where each filler (concept) produced by the interpreter is evaluated in terms of how closely it corresponds to the filler concept in the source language text, i.e. if it represents a paraphrase, a semantic change, or a verbatim reconstruction of the source text (Dillinger 1990: 185). Altman establishes a hierarchy of errors: excessive concentration on a preceding item due to processing problems, attempt to improve TL style, difficulty in finding the correct contextual equivalent for a given lexical item, drawing erroneously upon one's store of background knowledge, compression of two information items into one, thereby producing a third, incorrect item, shortcomings in mastery of the foreign language, leading to misunderstandings and therefore misinterpretations of the original speech. The further down this list an error type occurs, the more likely it seems to have a detrimental effect on the communication (Altman 1994: 34). Schjoldager suggests using a feedback sheet as a diagnostic tool that is used for assessment of performance from both the speaker's and the listener's perspective. Interestingly, she deliberately excludes the client from the array of perspectives as to her mind the client's point of view is a difficult parameter in the teaching situation. The feedback form focuses on four criteria: a) the listener can understand what the interpreter is saying, b) the interpreter's language is adequate, c) the interpreter's rendition is coherent and plausible, d) the interpreter is a loyal communicator of the speaker's message. An additional benefit of a feedback form is that the very use of it makes it clear to the class what the assessment criteria are – there is higher transparency of what is being expected, and the students' awareness of what makes an interpreting performance a 'good' one is raised (Schjoldager 1996: 188 ff.).

These are just a few examples of systems for the assessment of a student's performance. There are many more. However, what counts is not only that the right criteria are being applied but also that the assessment is being done effectively. It is crucial that evaluation is being performed by a native speaker of the language, especially for interpreters working into their B language (Déjean le Féal 1990: 157). This principle is also reflected in the paper published by the AIIC Training Committee which states that "the faculty should include native speakers for all the courses offered" (AIIC Training Committee 2006). Paneth described the common practice of an interpreting teacher reading out a manuscript, listening to several interpreting students, and writing comments on their assessment on the written manuscript (Paneth 1957: 54 as quoted by Kalina 1998: 235). As Kalina points out, there are countries and interpreting schools where this practice still exists. We have seen earlier that even professional interpreters find it difficult to reliably assess an interpreter's performance and to detect all errors. The use of new media is indispensable, as is the recording of students' output. Having students perform their fellow-students' evaluation is also an option – albeit not a very reliable one.

4.2.1.2.2 Student-student assessment

Another type of assessment in class is the quality assessment by fellow students (Kutz 2005: 15). It plays an important role in the learning environment – not only to use the time of 'inactive' students more efficiently, but also to benefit from the group's collective mind to make a comprehensive, multi-dimensional performance assessment plus reduce the students' stress (Gile 2001a: 389). Schjoldager also stresses the need for constructive criticism that is interactive, i.e. in which the class is involved (Schjoldager 1996: 188).

However, students tend to be unsystematic and inconsistent in judging their peers' interpreting output as shown in a study on student-student assessment in liaison interpreting. The assessment was made first live, on the spot, then on the basis of the tape recording, thus allowing a revision of the initial assessment. Nevertheless, responses diverged dramatically. In no case were all errors discovered. Detection rates amounted to a mere average of 30%. Very interesting was the comparison of the responses to section A (general impression), B (content errors) and C (individual merits / shortcomings). For example, one interpreter-student received the highest scores for professionalism and excellent scores for general impression although terminology and fidelity scores were not particularly good. The teacher therefore can – and must – play a crucial role in the process, using his unique combination of knowledge and skills – linguistic, didactic, theoretical, methodological, professional – to enhance students' performance (Ackermann et al. 1997: 264 ff.).

4.2.1.3 In-class assessment from the students' perspective

Another part of the above-mentioned study by Ackermann dealt with the stress students are likely to experience in class. In the questionnaire, students mentioned 'artificiality of the classroom situation' (50%), 'having to perform in front of the group' (59%) and 'noting down of errors by instructors' (77%). Even though this shows clearly that the artificial classroom situation and monitoring by fellow students and instructors is perceived as stressful, it is highly interesting to review students' suggestions on how to relieve this stress: Only 9% of students asked for the format of the class to be relaxed, and private sessions were an option for an equally low percentage (9%). No student demanded that instructors stop taking notes during sessions. These answers are a clear sign that – although students are aware of and affected by the artificiality of the classroom monitoring situation – they are willing to put up with this 'necessary evil' for the sake of their learning progress. They are "satisfied to grin and bear it" (Ackermann et al. 1997: 266).

4.2.2 Pragmatic aspects

Teachers can be faced with a dilemma. In order to foster quality performance, they will feel the need to address all content errors made by the student although fully aware of the fact that many of these errors would not have been noticed by users in a real-life situation. Even in a workshop hosted by the Department of Translatology in Leipzig with instructors, researchers and practitioners, it emerged that – when given the choice – workshop participants favoured packaging over content, giving higher scores to an interpreting performance that contained several content errors which were skilfully glossed over by the interpreter, on the grounds that this was “what clients look for” (Kutz et al. 1995 as quoted by Ackermann et al. 1997: 263). However, Ackermann underlines the importance of not being led astray by statistics on user expectations. In a way, teachers have to master a tightrope walk. Users’ expectations have to be borne in mind at all times. Yet the teacher’s unique critical perspective goes beyond what users expect (and detect) in interpreting quality. It might be interesting to learn that a certain percentage of users do not find a foreign accent irritating. But does this dispense the teacher from trying to help students work on their accent? It certainly does not (Ackermann et al. 1997: 266). Viaggio sounds a little like a father figure when he stipulates that “our students will acquire the right to bend the rules only once they have fully mastered them” (Viaggio 1992: 50). But he may be right.

So far, no systematic quality research has been carried out to answer the question as to how effective particular training methods are. And even if it had been, a potential distorting effect would need to be considered. Experimenters are often teachers themselves, so they will test methods they believe are good. One can expect these experimenters to be biased (Gile 2003: 116).

According to Kalina, empirical research is crucial to the development of efficient teaching methods: “Adequate teaching of SI must be based on a theory of the processes involved, and for developing such a theory, empirical study is necessary” (Kalina 1994: 231). Professionals and researchers alike have to make their contribution to develop adequate training standards and methods. The research community needs to define professional standards, and teaching needs to make students fully aware of what quality criteria and expectations they have to fulfil, and enable them to deliver high quality (Kalina 1994: 225 ff.).

Bühler states that active conference interpreters can contribute to the training success by working out a typology of input texts they encounter in their professional lives which should then form the basis of interpreter training at university. In this claim, Bühler supports Kopczynski’s typology that he outlined in 1982: a) unprepared oral monologue or dialogue, b) semi-prepared oral monologue with notes, c) reading of a written monologue that was intended for the spoken medium, d) written text intended for the written medium (Kopczynski 1982: 256 as quoted by Bühler 1989: 134). Bühler is certainly right in saying that such a typology established by those who are active in the field should form the basis of all interpreter training. But I believe her claim also makes another thing very evident: Interpreter trainers should be interpreters themselves. Professionals know the market, users' expectations, and the actual working life. If teachers are not experienced interpreters, this entails various risks, e.g. that of focusing too much on linguistic errors instead of the delivery of content, ignorance of interpreting strategies, and ignorance of the market and its norms (Gile 2001a: 390).

This view is also shared by the AIIC Training Committee: "Courses should be designed and interpretation classes taught by practising conference interpreters. Professional interpreters who serve as teaching faculty provide the essential interface between the classroom and the profession" (AIIC Training Committee 2006).

5 Methodological issues

> So alt das Dolmetschen als menschliche Tätigkeit ist, so jung ist das Simultandolmetschen, noch jünger seine Erforschung.
> *(Salevsky 1985: 191)*

> How do you conduct hundreds of experiments on a population of less than 5,000 individuals scattered over many countries, with different language combinations, different training and qualification backgrounds, of different ages, many of whom will refuse to take part in any study in which their output will be scrutinised?
> *(Gile 2001b)*

One cannot talk about interpreting quality research without touching upon the methodological issues that the research community is faced with. Gile is one of the researchers who have heavily (self-)criticised the profession and its research approaches. An early statement regarding quality assessment dates back to 1983: “les jugements… sont intuitifs, personnels, et ne se réfèrent pas à des critères objectifs” (Gile 1983: 236).

5.1 Observational vs. experimental research

The advantage of observational research is that it allows the investigation of phenomena as they occur naturally, with no distortion induced by the study. But this is difficult in IR since due to the evanescent nature of both input and output, on-site observation is often required (Gile 1998a: 88).

One basic scientific rule is that there has to be a large sample if one wants to perform analyses that yield statistically valid results. The data sample must be large enough to isolate individual from across-the-board phenomena, idiosyncratic from general tendencies (Shlesinger 1989: 113). Furthermore, many researchers have claimed that research has to be carried out in a natural setting – an authentic corpus is required (e.g. Gile 1983: 243). But it has proved to be difficult to assemble a representative corpus. Meetings are restricted, if not confidential. An ever greater problem, much alluded to in the literature about interpreting quality, is that professional interpreters are reluctant to have their work closely monitored and compared to that of their colleagues (Kalina 1994: 225; Gile 1998a: 87; Kalina 1998: 131 f.; Kahane 2000; Gile 2003: 112; Chiaro & Nocella 2004: 280). Not informing interpreters that observation is taking place is ethically problematic, informing interpreters might lead to a change in behaviour – or even interpreters’ refusal to allow such observation (Gile 1998a: 77). Another challenge is the question of how to involve conference participants. These are usually more interested in what speakers have to say and do not care about interpretation research, which makes it difficult to submit them to an experiment (Gile 1995b: 161).

As Straniero Sergio points out, experimental studies can only evaluate an interpreter's performance in a 'sheltered' environment, and studies are often based on self-prophecising hypotheses (Straniero Sergio 2003: 171). In experimental settings, moreover, the researcher will only gain insights on one set of features or dimension of quality, not assess quality as such (Pöchhacker 2001: 419). In contrast, Gile demands that experimental studies and observational studies be regarded as mutually reinforcing, not as mutually exclusive (Gile 1998a: 88).

But even if researchers resort to the second-best option of laboratory-controlled experiments, it is not easy to find practitioners who are willing to take part, probably for the same reason for which corpus-data are difficult to obtain. Interpreters simply do not like their performance being scrutinised.

For the reasons outlined above, it becomes obvious that large samples are often difficult to obtain. Data thus tend to be insufficient for drawing any general conclusions; the 1979 study by Anderson study is a good example (Anderson [1979]1994; cf. 2.2.2).

One possible compromise solution is researching interpreting quality in the training environment. This approach yields several advantages: there are no acceptability problems since assessment is part of the training process, thus accepted – or even welcomed – by students, and a larger number of subjects will be available. Assessment variability will be lower, the assessment will be made by one or two experimenters who can agree on the norms and criteria that will be applied. At the same time, sensitivity is higher because teachers will exclusively focus on quality assessment. A disadvantage might be the fact that neither the skills nor the strategies of students are likely to match those of professionals (Lamberger-Felber: 1998; Gile 2003: 115 ff.; Chiaro & Nocella 2004: 280).

A further option is to aggregate data from smaller samples and make inferences – a method that is particularly useful when findings on smaller samples all point in the same direction (Gile 2003: 115 ff.).

5.2 What should be assessed, and how?

Another shortcoming in interpreting research is that so far it has primarily dealt with quantifiable textual features such as errors and omissions but has neglected complex psycho-communicative relationships and effects. Most publications on the topic have focused on the interpreter's output. Various error scales have been developed (e.g. Barik 1971; Mackintosh 1983; Falbo 1998; Kopczynski 1983a as quoted by Altman 1994: 26), but even if the more recent ones provide for a weighting of errors (e.g. Bowen 1990) and do not regard changes and omissions as errors when these form part of a strategic approach, the evaluation of quality using error scales is not sufficient.

Pöchhacker asks: Should the concept of quality not extend far beyond the mere product of the interpreting activity? In his view, case studies would be the most efficient – and a holistic – tool to evaluate interpreting quality, not just as a textual but also as a multidimensional socio-psychological phenomenon (Pöchhacker 2001: 420). So far, only one study in the area of community interpreting can be called a case study in all respects: the work of Cecilia Wadensjö who recorded and analysed a large corpus of authentic discourse (Wadensjö 1998 as quoted by Pöchhacker 2001: 420). No such case study exists for the area of conference interpreting. Marrone used a questionnaire but did not analyse the corpus (Marrone 1993). Gile used a questionnaire and reported on his impressions of textual output quality but did not engage in systematic corpus analysis either (Gile 1990). Pöchhacker used corpus-based data analysis, participant observer notes and documentary analysis but did not gain sufficient access to conference participants with his user assessment survey (Pöchhacker 1994).

Quality control of interpretation is even more difficult than for translation given the transient nature, or evanescence, of the product (e.g. Lederer 1985: 28; Garzone 2002: 107; Chiaro & Nocella 2004: 279). This evanescence can be remedied by putting an interpretation down in writing: as a transcript. However, this approach is regarded as problematic insofar as it "truncates and distorts the semiotically complex textual product" (Pöchhacker 2001: 420). In his 1999 study on variations in fidelity assessment, Gile found that transcripts are not assessed like auditorily presented material, even by interpreters, and advised much caution when assessing quality on the basis of transcripts (Gile 1999: 54, 67 f.). To quote Mack: "Transcription is always an artificial product"

(Mack 2002: 112). A better option is the use of recordings, if possible supported by the interpreters' own feedback on the recorded interpretation so as to obtain additional information on the strategies used (Moser-Mercer 1996: 52).

5.3 Variability

In Gile's experiment with students judging the quality of a consecutive interpretation, the following findings emerged: Participants were not very reliable error detectors, and there did not seem to be a clear correlation between the number of errors reported and the assessors' fidelity ratings (Gile 1995b: 160; Gile 2006: 19). There is general agreement that low inter-rater reliability can be a problem. Gile even states that considerable variability is an intrinsic part of quality performance and assessment (Gile 2003: 114; cf. 3.1.2.2).

To ensure that the assessment of quality also takes place at a high-quality level, it is important to use conference interpreters as jurors since only they can assess information loss correctly knowing that during simultaneous interpreting a variety of compression strategies are applied by the interpreter which do not necessarily lead to meaning/information loss, but might produce an 'overt' deviation from the original for listeners/readers/jurors unfamiliar with interpreting (Moser-Mercer 2003). In addition to involving several judges who are familiar with interpreting strategies an analysis of interjudge reliability should be conducted (Moser-Mercer 1996: 51 f.).

However, there is not only inter-judge variability but also inter-interpreter variability to cope with. Kalina points out that no two interpretations are ever the same (Kalina 1998: 129 f.). Yet little research has been carried out to investigate variability among interpreters and its potential implications for research methodology.

Lamberger-Felber reports on high variability among conference interpreter performances. She had intended to test the influence of working conditions on interpreter performance but found that – although she had invested considerable efforts to ensure that the recruited group of interpreters was homogeneous – performance variability proved to be a major methodological issue. She claims that variability must be taken into consideration for the set-up of empirical SI studies. Even though some might regard the sample of – in this instance – twelve interpreters as large, it may well be too small to obtain statistically viable data if variability among performances attains a high level (Lamberger-Felber 2003: 155).

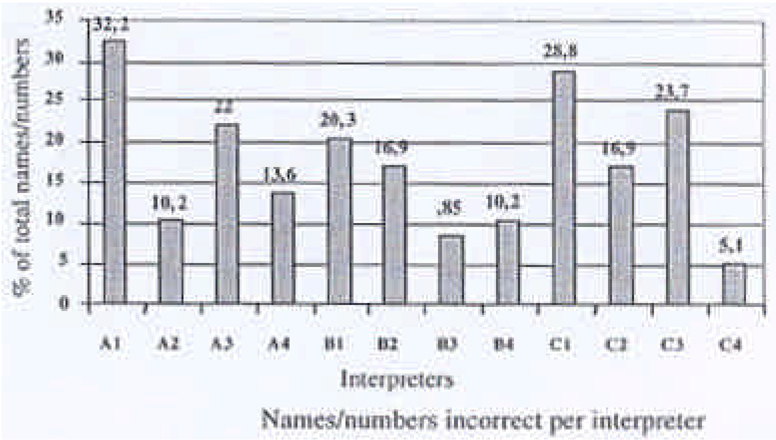

Fig. 10: Names/numbers not interpreted correctly – overview by interpreter (Lamberger-Felber 2003: 151)

5.4 Researchers

Gile puts it bluntly. The problem in IR is that almost all scholars have little or no formal training in empirical research, including statistics, and prominent personalities have not undergone a strict methodological screening process – unlike in established disciplines where young researchers go through rigorous training. Even though interpreting researchers may enjoy an excellent reputation in their field, many of them fall prey to the PhG syndrome. The abbreviation PhG stands for Gurus in Philosophy – a status that researchers may well deserve for their creativity, writing, personality and productivity but not in terms of their methodological skills and knowledge. This is dangerous as PhGs have a lot of influence over aspiring researchers and may misguide them because they trust them and do not know better (Gile 2005). Academic titles and publications generate prestige, and researchers may well reach positions higher than those that could be reached in a more competitive, selective environment. In other disciplines, researchers are trained over several years – many of the interpreter researchers have earned doctoral degrees – “but how many of us have gone through formal research training” (Gile 2001b)? In IR, because of the relatively small number of researchers, the mechanisms that promote quality and filter out poor texts are less effective than they should be.

It is not only the lack of methodological training of researchers that constitutes a problem but also the very scarcity of researchers, regardless of their expertise. Only a few dozen researchers are active over several years and produce more than one study (Gile 2000: 307; Gile 2001b; Gile 2006: 15). The reason is simple: Most authors in IR are also active interpreters. In most countries, interpreting is more lucrative than academic research. The conflict of interest is obvious (Gile 2000: 311).

5.5 Interdisciplinarity

One viable approach to overcome some of the above-mentioned problems is to join forces with other disciplines. Kahane wants to "broaden the field by moving from purely linguistic issues to pragmatic, communication issues" and recommends including elements of information theory, the communication situation, or even group dynamics (Kahane 2000).

Methodological difficulties clearly point to the need for interdisciplinary approaches and for the joint definition of criteria that are decisive for quality in order to facilitate comparability. Researchers from different research centres, countries and disciplines need to work together, collaborating in projects that have clearly defined and coordinated research objectives, and using the same methodology (Gile 1983: 243; Collados Aís 1998: 250; Kahane 2000). Methodology and questionnaires should be harmonised (Marrone 1993: 39; Collados Aís 1998: 55, 250).

What interdisciplinary approaches have been taken so far? In 1977, a symposium organised by Gerver and Sinaiko brought together researchers from psychology, linguistics and sociology on the one hand, and practitioners of interpreting on the other, hoping to initiate successful cooperation. However, the initiative failed, largely due to the interpreters' unwillingness to have their actual performance examined and scrutinised by researchers who were not interpreters and therefore – as they felt – could not grasp the true essence of interpreting. This conference may have been the symbolic starting point of the Practitioners' Period – interpreters taking over research – ignoring their predecessors' methods and findings. The researchers of the Practitioners' Period strongly opposed experimental disciplines and quantitative experimental methods. This nipped any interdisciplinary ventures in the bud (Gile 2006).

The number of non-interpreter researchers has rather declined than risen over the years. *Interpreting* as an interdisciplinary journal published by Massaro and Moser-Mercer following the Trieste Symposium, or the Turku Conference, were hopeful signs but not necessarily sufficient to engender a fruitful cooperation between different disciplines (Gambier et al: 1997). Some further interdisciplinary initiatives with neurophysiology were taken by Ingrid Kurz in Vienna, Jorma Tommola in Finland and Barbara Moser-Mercer in Switzerland. A good – or rather concerning – example of the lack of interdisciplinarity is Gile's Effort Model. It is basically a conceptual model, built intuitively but untested experimentally. In 1997, Massaro and Shlesinger pointed out that little had been done to test it systematically (Massaro & Shlesinger et al. 1997: 43). The same observation still holds true almost 10 years later.

One can only hope that continuing attempts to foster cooperation between the IR community and other disciplines will be fruitful.

5.6 The future of interpreting research – challenges and opportunities

I have outlined the methodological issues this discipline is faced with and that, as a consequence, many research findings lack empirical validity. This is only one problem. Gile raises another question: What if faced with the dilemma of research not proving that interpreters' claims are justified? "Do we accept the data and use them to change our claims, do we disregard them, or choose to only quote those findings that corroborate our positions – thus running the risk of employers or non-AIIC interpreters use the evidence – or lack of evidence – against us" (Gile 2001b)?

Many findings from empirical research indicate that interpreters' output contains many errors and omissions, regardless of their competence (Gile 1985: 204). Are interpreters willing to acknowledge that fact, although it might affect the very image of conference interpreters as high-quality providers (Gile 2001b)? The remarkably high number of errors in Formula One press conferences as reported by Straniero Sergio demonstrates clearly that actual interpreting performances do not correspond to the ideal of quality (Straniero Sergio 2003: 139 f.), as do Pöchhacker's or Vuorikoski's findings (Pöchhacker 1994: 246; Vuorikoski 2004). Vuorikoski even states that various types of non-correspondence between the originals and the interpreter's versions are a normal element of SI (Vuorikoski 2004: 23). Research strives to be objective. Professional associations do not, and their respective interests may clash (Gile 2001b).

On the other hand, non-ideal quality is an argument that interpreters can leverage to make clients aware of an important aspect: That interpreting performance and quality will always depend on a multitude of factors. Working environment and working conditions determine what level of quality is achievable. Some studies on interpreting as a dependent variable have shown that interpreters have a right to insist on certain conditions, not only for their personal well-being but also in the client's interest, i.e. for quality's sake. It is crucial, therefore, and in the interest of all participants in the communication process, that client and interpreter cooperate to ensure the best possible conditions for achieving the highest quality standards (e.g. Gile 1984: 84; Pöchhacker 1994: 246).

Kalina points out that – even though valuable findings have been obtained in the 'quality-as-a-dependent-variable' research line – more efforts will be necessary to gain further insight into the weighting of specific factors. The interpreting research community has not provided a clear answer to the question *which factors* can have *what degree of influence* on an interpreter's performance in an authentic setting (Kalina 2004: 4; my italics). She also suggests the pragmatic approach of using a QA data sheet for interpreting assignments. Such a data sheet serves a dual purpose as it can be used as a quality control tool for interpreters but can also help researchers obtain more data on real-life scenarios, i.e. paint a clearer and more refined picture of interpreters' professional activity (Kalina 2005: 779 ff.).

As a final remark: One can assume that the current trend of increasing cost pressure will continue. Research into interpreting can also benefit interpreters insofar as it is associated with academia and can thus can help raise the social status of interpreters as well as support their claims for adequate remuneration (Gile 2001b).

Bibliography

Ackermann, Dorothea; Lenk, Hannelore; Redmond, Mary: "Between Three Stools – Performance Assessment in Interpreter Training". In: Fleischmann et al. (1997), 262-267

AIIC (2001): *Resolution on the management of multilingualism at the European Parliament after enlargement.* Retrieved from http://www.aiic.net/ViewPage.cfm/page531, accessed on 20-11-06

AIIC (2006): *Code of Professional Ethics*. Retrieved from http://www.aiic.net/ViewPage.cfm?article_id=24&plg=1&slg=1, accessed on 4-12-06

AIIC Training Committee (2006): *Conference Interpreting Training Programmes: Best Practice*. Retrieved from http://www.aiic.net/ViewPage.cfm/page60.htm, accessed on 16-12-06

AIIC/Universitas (1991): "10 Gebote für Redner bei internationalen Tagungen". In: Fremdsprachen 4, 17-18

Alexieva, Bistra (1983): "Compression as a means of realisation of the communicative act in simultaneous interpreting“. In: Fremdsprachen 4, 233-238

Alexieva, Bistra (1990): "Creativity in simultaneous interpretation". In: Babel 36-1, 1-6

Altman, Janet (1990): "What helps effective communication? Some interpreters' views": In: The Interpreters' Newsletter 3, 23-32

Altman, Janet (1994): "Error Analysis in the Teaching of Simultaneous Interpreting: A Pilot Study". In: Lambert/Moser-Mercer (1994), 25-38

Anderson, Linda ([1979]1994): "Simultaneous interpretation: Contextual and translation aspects". In: Lambert/Moser-Mercer (1994), 101-120

Arntz, Reiner; Thome, Gisela (eds.) (1990): *Übersetzungswissenschaft – Ergebnisse und Perspektiven. Festschrift für Wolfram Wilss*. Tübingen: Narr

AUSIT (ed.) (1996): *New Horizons. Proceedings of the XIV World Congress of FIT*. Melbourne: AUSIT

Barik, Henry C. (1971): “A description of various types of omissions, additions, errors of translation encountered in simultaneous interpretation”. In: Meta 16-4: 199-210

Barik, Henri C. (1972): “Interpreters talk a lot among other things”. In: Babel 3, 10

Barik, Henri C. (1973): “Simultaneous interpretation: Temporal and quantitative data”. In: Language and Speech 16, 237-270

Barik, Henri C. (1975): “Simultaneous Interpretation: Qualitative and Linguistic Data”. In: Language and Speech 18, 272-297

Best, Joanna; Kalina, Sylvia (eds.) (2000): *Übersetzen und Dolmetschen. Eine Orientierungshilfe*. Tübingen: Francke Verlag

Beylard-Ozeroff, Ann; Králová**,** Jana**;** Moser-Mercer**,** Barbara (eds.) (1998): *Translators' strategies and creativity*. Amsterdam/Philadelphia: John Benjamins

Bowen, David; Bowen, Margareta (eds.) (1990): *Interpreting – Yesterday, Today, and Tomorrow, ATA Monograph Series IV*. Binghamton: State University of New York

Bowen, Margareta (1990): “The Weighting of Errors in Simultaneous Interpreting“. In: Arntz (1990), 545-554

Bowker, Lynne; Cronin, Michael; Kenny, Dorothy; Pearson, Jennifer (eds.) (1998): *Unity and Diversity? Current Trends in Translation Studies*. Manchester: St. Jerome Publishing,

Brislin, Richard W. (ed.) (1976): *Translation. Applications and Research*. New York: Gardner Press

Brown, Keith (ed.) (2006): *Encyclopedia of Language and Linguistics*. Oxford: Elsevier

Bühler, Hildegund (ed.) (1985): *Translators and their position in Society. X. FIT World Congress, Proceedings*. Vienna: Braumüller

Bühler, Hildegund (1986): "Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters". In: Multilingua 5-4, 231-235

Bühler, Hildegund (1989): "Discourse Analysis and the Spoken Text – A Critical Analysis of the Performance of Advanced Interpretation Students". In: Gran/Dodds (1989), 131-137

Cartellieri, Claus (1977): "Über einen experimentellen Ansatz zur Analyse der Qualitätsbestimmung beim Simultandolmetschen". In: Vermittelte Kommunikation, Sprachmittlung, Translation (Übersetzungswissenschaftliche Beiträge 1). Leipzig: Verlag Enzyklopädie, 78-85

Cartellieri, Claus (1983): "The inescapable dilemma. Quality and/or quantity in interpreting". In: Babel 29-4, 209-213

Cenková, Ivana (1998): "Quality of interpreting – A binding or liberating factor?". In: Beylard-Ozeroff et al. (1995), 163-170

Chernov, Ghelly V. (1987): Osnovy sinkhronovo perevoda, Vysshaja Shkola, Moscow 256

Chase, William G.; Simon, Herbert A. (1973): "Perception in chess". In: Cognitive Psychology 4, 55-81

Chesterman, Andrew (1993): "From 'Ought' to 'Is'". In: Target 5-1, 1-20

Chiaro, Delia; Nocella, Guiseppe (2004): "Interpreters' Perception of Linguistic and Non-Linguistic Factors Affecting Quality: A Survey through the World Wide Web". In: Meta 29-2, 278-293

Collados Aís, Angela (1998): *La evaluación de la calidad en interpretación simultánea. La importancia de la comunicación no verbal.* Granada: Editorial Comares

Collados Aís, Angela ([1998]2002): “Quality assessment in simultaneous interpreting: the importance of nonverbal communication“. In: Pöchhacker/Shlesinger (2002), 327-336

Collados Aís, Angela; Fernández Sanchez, Manuela; Gile, Daniel (eds.) (2003): *La evaluación de la calidad en interpretación: investigación*. Granada: Editorial Comares

Coughlin, Josette M. (1988): “Inside or Between Languages, Oral Communication Equals Interpretation”. In: Hammond (1988), 355-362

Darò, Valeria; Lambert, Sylvie; Fabbro, Franco (1996): “Conscious monitoring of attention during simultaneous interpretation”. In: Interpreting 1-1, 101-124

DeGroot, Adrianus Dingeman (1966): “Perception and memory versus thought: Some old ideas and recent findings”. In: Kleinmuntz (1966), 19-50

Déjean le Féal, Karla (1990): “Some thoughts on the evaluation of simultaneous interpretation”. In: Bowen (1990),154-160

Dillinger, Mike (1990): “Comprehension during Interpreting: What do Interpreters know that Bilinguals don’t?”. In: Lambert/Moser-Mercer (1994), 155-190

Dollerup, Cay; Lindegaard, Anne (eds.) (1994): *Teaching Translation and Interpreting 2. Insights, Aims, Visions*. Selected Papers from the Second ‘Language International’ Conference, Elsinore, Denmark, 4-6 June 1993. Amsterdam/Philadelphia: John Benjamins

Dollerup, Cay; Loddegaard, Anne (eds.) (1992): *Teaching Translation and Interpreting. Training, Talent and Experience.* Selected Papers from the First ‘Language International’ Conference, Elsinore, Denmark, 31 May-2 June 1991. Amsterdam/Philadelphia: John Benjamins

Dollerup, Cay; Appel, Vibeke (eds.) (1996): *Teaching Translation and Interpreting 3*. Amsterdam/Philadelphia: John Benjamins

Elsagir, Ina M. (2000): “Anforderungen an Dolmetschleistungen aus Zuschauersicht”. In: Kalina et al. (2000), 107-123

Englund Dimitrova, Birgitta; Hyltenstam, Kenneth (eds). (2000): *Language Processing and Simultaneous Interpreting. Interdisciplinary Perspectives*. Amsterdam/Philadelphia: John Benjamins,

Enkvist, Nils Erik (1982) (ed.): *Impromptu Speech. A Symposium*. Åbo : Åbo Akademi

European Commission, DG Interpretation (2006): *Annual Activity Report 2005.* Retrieved from http://www.ec.europa.eu/atwork/synthesis/doc/scic_aar.pdf, accessed on 12-17-06

European Parliament (2001): *Preparing for the Parliament of the Enlarged European Union*. Retrieved from http://www.europarl.europa.eu/interp/online/lf99_one/v04_no9/podesta_en.pdf, accessed on 23-11-06

Falbo, Caterina (1998): "Analyse des erreurs en interprétation simultanée". In: The Interpreters' Newsletter 8, 107-120

Falbo, Caterina; Russo, Mariachiara; Straniero Sergio, Francesco (eds.) (1999): *Interpretazione Simultanea e Consecutiva – Problemi Teorici e Metodologie Didattiche*. Milano: Ulrico Hoepli Editore

Feldweg, Erich (1996): *Der Konferenzdolmetscher im internationalen Kommunikationsprozeß*. Heidelberg: Julius Groos

Fleischmann, Eberhard; Kutz, Wladimir; Schmitt, Peter A. (eds.) (1997): *Translationsdidaktik: Grundfragen der Übersetzungswissenschaft.* Tübingen: Narr

Fusco, Maria Antonietta (1990): "Quality in conference interpreting between cognate languages: A preliminary approach to the Spanish-Italian case". In: The Interpreters' Newsletter 3, 93-97

Gambier, Yves; Gile, Daniel; Taylor, Christopher (eds.) (1997): *Conference interpreting: current trends in research*. Amsterdam/Philadelphia: John Benjamins

Garzone, Giuliana (2002): “Quality and norms in interpretation”. In: Garzone/Viezzi (2002), 107-119

Garzone, Giuliana (2003): “Reliability of quality criteria evaluation in survey research”. In: Collados Aís et al. (2003), 23-30

Garzone, Giuliana; Viezzi, Maurizio (eds.) (2002): *Interpreting in the 21st Century. Proceedings of the 1st Conference on Interpreting Studies, Forlì, University of Bologna (9-11 November 2000).* Amsterdam/Philadelphia: John Benjamins

Gebhard, Silke (1999): “Learning to tell the difference“. In: AIIC Bulletin 27-2, 7

Gebhard, Silke (2001): “Building Europe – or Back to Babel?”. Retrieved from http://www.aiic.net/ViewPage.cfm?page_id=52, accessed on 23-11-06

Gerver, David ([1969]2002): “The effects of source language presentation rate on the performance of simultaneous conference interpreters“. In: Pöchhacker/Shlesinger (2002), 53-66

Gerver, David (1974): “The effects of noise on the performance of simultaneous interpreters: Accuracy of performance“. In: Acta Psychologica 38, 159-167

Gerver, David; Sinaiko, H. Wallace (eds.) (1978): *Language Interpretation and Communication*. New York: Plenum Press

Gile, Daniel (1983): “Aspects méthodologiques de l’évaluation de la qualité du travail en interprétation simultanée”. In: Meta 28-3, 236-243

Gile, Daniel (1984): “Les noms propres en interprétation simultanée”. In: Multilingua 3-2, 79-85

Gile, Daniel (1985): “Les termes techniques en interprétation simultanée”. In: Meta 30-3, 199-210

Gile, Daniel (1986): “Le travail terminologique en interprétation de conférence”. In: Multilingua 5-1, 31-36

Gile, Daniel (1990): “L’évaluation de la qualité de l’intérpretation par les délégués: une étude de cas”. In: The Interpreters’ Newsletter 3, 66-71

Gile, Daniel (1991): “A Communication-Oriented Analysis of Quality in Nonliterary Translation and Interpretation”. In: Larson (1991), 188-200

Gile, Daniel (1995a): *Regards sur la recherche en interprétation de conférence*. Lille: Presses Universitaires de Lille

Gile, Daniel (1995b): “Fidelity assessment in consecutive interpretation: an experiment”. In: Target 7-1, 151-164

Gile, Daniel ([1997]2002): “Conference interpreting as a cognitive management problem”. In: Pöchhacker/Shlesinger (2002), 163-176

Gile, Daniel (1998a): “Observational Studies and Experimental Studies in the Investigation of Conference Interpreting”. In: Target 10-1, 69-93

Gile, Daniel (1998b): “Norms in Research on Conference Interpreting: A Response to Theo Hermans and Gideon Toury”. In: Multilingual Matters 5-1/2, 99-106

Gile, Daniel (1999): “Variability in the Perception of Fidelity in Simultaneous Interpretation”. In: Hermes 22, 51-79

Gile, Daniel (2000): “The history of research into conference interpreting. A scientometric approach”. In: Target 12-2, 297-321

Gile, Daniel (2001a): “L’évaluation de la qualité de l’interprétation en cours de formation”. In: Meta 46-2, 379-393

Gile, Daniel (2001b): “Interpreting Research. What you never wanted to ask but may like to know“. Retrieved from http://www.aiic.net/ViewPage.cfm/article229.htm, accessed on 24-11-06

Gile, Daniel (2003): “Quality Assessment in Conference Interpreting: Methodological Issues”. In: Collados Aís et al. (2003), 110-123

Gile, Daniel (2005): Editorial of the CIRIN Bulletin No. 31, December 2005. Retrieved from http://cirinandgile.com/Bulletin%2031rtf.htm, accessed on 28-11-06

Gile, Daniel (2006): “Conference Interpreting”. In: Brown (2006), 9-23

Gold, David L. (1973): “On Quality in Interpretation“. In: Babel 19-4, 154-155

Goldman-Eisler, Frieda (1972): “Segmentation of input in simultaneous translation”. In: Journal of Psycholinguistic research 1, 127-140

Gran, Laura; Dodds, John (eds.) (1989): *The Theoretical and Practical Aspects of Teaching Interpretation. Proceedings of the Trieste Symposium, 1986.* Udine: Campanotto

Gravier, Maurice (1978): Preface to Seleskovitch (1978), 4

Guardini, Pascal (1995): *La traduzione simultanea del film: produzione e percezione. Unpublished Thesis*. Trieste: University of Trieste

Hammond, Deanna L. (ed.) (1988): *Languages at Crossroads. Proceedings of the 29th Annual Conference of the American Translators Association*. Medford, NJ: Learned Information Inc.

Hammond, Deanna L. (ed.) (1989): *Coming of Age. Proceedings of the 30th Annual Conference of the American Translator Association*. Medford, NJ: Learned Information Inc.

Harris, Brian (1990): “Norms in Interpretation”. In: Target 2, 115-19

Jímenez Ivars, Amparo; Pinazo Calatayud, Daniel: “‘I failed because I got very nervous.’ Anxiety and performance in interpreter trainees: an empirical study”. In: The Interpreters' Newsletter 11, 105-118

Jones, Roderick (1998): *Conference interpreting explained*. Manchester: St. Jerome Publishing

Kadric, Mira; Kaindl, Klaus; Pöchhacker, Franz (eds.) (2000): *Translations-wissenschaft. Festschrift für Mary Snell-Hornby zum 60. Geburtstag*. Tübingen: Stauffenburg

Kahane, Eduardo (2000): “Thoughts on the quality of interpretation”. In: Communicate 4. Retrieved from http://www.aiic.net/ViewPage.cfm/page197.htm, accessed on 26-11-06

Kalina, Sylvia (1994): “Analyzing interpreters’ performance: methods and problems”. In: Dollerup/Lindegaard (1994), 225-232

Kalina, Sylvia (1998): *Strategische Prozesse beim Dolmetschen. Theoretische Grundlagen, empirische Fallstudien, didaktische Konsequenzen.* Tübingen: Gunter Narr

Kalina, Sylvia (2002a): “Quality in interpreting and its prerequisites – A framework for a comprehensive view”. In: Garzone/Viezzi (2002),121-130

Kalina, Sylvia (2002b): “Interpreters as Professionals”. In: Across Languages and Cultures 3-2, 169-187

Kalina, Sylvia (2004): “Zum Qualitätsbegriff beim Dolmetschen”. In: Lebende Sprachen 49-1, 2-8

Kalina, Sylvia (2005): “Quality assurance for interpreting processes“. In: Meta 50-2, 768-784

Kalina, Sylvia; Buhl, Silke; Gerzymisch-Arbogast, Heidrun (eds.) (2000): *Dolmetschen: Theorie · Praxis · Didaktik. Arbeitsberichte des Advanced Translation Research Center (ATRC) an der Universität des Saarlandes.* St. Ingbert: Röhrig

Katan, David; Straniero Sergio, Francesco (2001): “Look who's talking. The Ethics of Entertainment and Talkshow Interpreting”. In: The Translator 7-2, 213-237

Keiser, Walter (1978): “Selection and Training of Conference Interpreters”. In: Gerver/Sinaiko (1978), 11-24

Klaudy, Kinga; Kohn, János (eds.) (1997): *Transferre Necesse Est. Proceedings of the 2nd International Conference on Current Trends in the Studies of Translation and Interpreting*. Budapest: Scholastica

Kleinmuntz, Benjamin (ed.) (1966): *Problem Solving: Research, Method, and Theory*. New York: Wiley

Kopczynski, Andrej (1982): “Effects of some characteristics of impromptu speech on conference interpreting”. In: Enkvist (1982), 255-266

Kopczynski, Andrej (1983a): “Deviance in Conference interpreting”. In: Kopczynski (1983b), 399-404

Kopczynski, Andrej (ed.) (1983b): *The Mission of the Translator Today and Tomorrow*. *Proceedings of the IX World Congress of the International Federation of Translators.* Warsaw: Interpress

Kopczynski, Andrzej (1994): “Quality in conference interpreting: some pragmatic problems”. In: Lambert/Moser-Mercer (1994), 87-99

Krawutschke, Peter W. (ed.) (1995): *Connections. Proceedings of the 36th Annual Conference of the American Translators’ Association. November 8-12, 1995.* Medford, NJ: Information Today Inc.

Kurz, Ingrid (1981): “Temperatures in Interpreters’ Booths – A Hot Iron?”. In: AIIC Bulletin 9-4, 39-43

Kurz, Ingrid (1983a): “Temperatures Inside and Outside Booths – A Comparative Study”. In: AIIC Bulletin 11-3, 67-72

Kurz, Ingrid (1983b): “CO_2 and O_2 Levels in Booths at the End of a Conference Day – A Pilot Study”. In: AIIC Bulletin 11-3, 86-93

Kurz, Ingrid (1986): “Dolmetschen im Alten Rom”. In: Babel 32-4: 215-220

Kurz, Ingrid (1989): “Conference interpreting – user expectations”. In: Hammond (1989), 143-148

Kurz, Ingrid (1993): “Conference interpretation: expectations of different user groups”. In: The Interpreters’ Newsletter 5, 13-21

Kurz, Ingrid (1996a): *Simultandolmetschen als Gegenstand der interdisziplinären Forschung.* Vienna: WUV-Universitätsverlag

Kurz, Ingrid (1996b): "Special features of media interpreting as seen by interpreters and users". In AUSIT (1996), 957-965

Kurz, Ingrid (1997): "Interpreters: Stress and Situation-Dependent Control of Anxiety". In: Klaudy/Kohn (1997), 201-206

Kurz, Ingrid (2001): "Conference Interpreting: Quality in the Ears of the User". In: Meta 46-2, 394-409

Kurz, Ingrid (2003): "Physiological stress during simultaneous interpreting: a comparison of experts and novices". In: The Interpreters' Newsletter 12: 51-67

Kurz, Ingrid (2003): "Quality from the User Perspective". In: Collados Aís et al. (2003), 3-22

Kurz, Ingrid; Kolmer, Herbert (1984): "Humidity and Temperature Measurements in Booths". In: AIIC Bulletin 12-2, 15-16

Kurz, Ingrid; Pöchhacker, Franz (1995): "Quality in TV interpreting". In: Translatio. Nouvelles de la FIT – FIT Newsletter XIV (3/4), 350-358

Kutz, Wladimir (1997): "Gut für wen? Zur Bewertung von Konsekutivleistungen". In: Fleischmann et al. (1997), 243-254

Kutz, Wladimir (2000): "Kommunikative Dolmetschsituation – Dolmetschauftrag – Bewertungsmodus: Wie bewertet man das Konsekutivdolmetschen situationsgerecht?". In: Kalina et al. (2000), 191-211

Kutz, Wladimir (2000): "Training für den Ernstfall. Warum und wie sich die Vorbereitung auf den Dolmetscheinsatz lohnt". In: MDÜ 46-3, 8-13

Kutz, Wladimir (2005): "Zur Bewertung der Dolmetschqualität in der Ausbildung von Konferenzdolmetschern". In: Lebende Sprachen 1, 14-34

Kutz, Wladimir et al. (1995)[7]: "Protokoll zum Workshop "Evaluierung von Konsekutivdolmetschleistungen". Universität Leipzig, 24-25 November 1995 (Circular)

Lamberger-Felber, Heike (1998). *Der Einfluss kontextueller Faktoren auf das Simultandolmetschen. Eine Fallstudie am Beispiel gelesener Reden*. Abstract of her unpublished doctoral dissertation, Karl-Franzens-Universität Graz. Retrieved from http://www-gewi.uni-graz.at/uedo/research/dissabstr/ lamberger.html#en, accessed on 22-11-06

Lamberger-Felber, Heike (2003): "Performance variability among conference interpreters. Examples from a case study". In: Collados Aís et al. (2003), 147-168

Lambert, Sylvie (1989): "Simultaneous interpreters: one ear may be better than two". In: Lambert/Moser-Mercer (1994), 225-236

Lambert, Sylvie (1991): "Aptitude testing for simultaneous interpretation at the University of Ottawa". In: Meta 36-4, 586-594

Lambert, Sylvie; Moser-Mercer, Barbara (eds.) (1994): *Bridging the gap. Empirical research in simultaneous interpretation*. Amsterdam/Philadelphia: John Benjamins

Larson, Mildred L. (ed.) (1991): *Translation: Theory and Practice. Tension and interdependence*. Binghamton: State University of New York

Lederer, Marianne (1985): "L'interprétation, manifestation élémentaire de la traduction“. In: Meta 30-1 (special issue), 25-29

Levý, Jirí (1969): *Die literarische Übersetzung. Theorie einer Kunstgattung*. Frankfurt am Main: Athenäum

Linell, Per (1998): "Discourse across Boundaries: On Recontextualizations and the Blending of Voices in Professional Discourse". In: Text: 18-2, 143-157

---

[7] as quoted by Ackermann et al. who do not name the co-authors of the paper.

Mack, Gabriele (2002): "Die Beurteilung professioneller Dolmetschleistungen". In: Best/Kalina (2002), 110-119

Mack, Gabriele; Cattaruzza, Lorella (1995): "User Surveys in Simultaneous Interpretation: A Means of Learning about Quality and/or Raising some Reasonable Doubts". In: Tommola (1995), 37-49

Mackintosh, Jennifer (1983): *Relay Interpretation: An Exploratory Study.* Unpublished M.A. Thesis. London: University of London

Mackintosh, Jennifer: "The AIIC Workload Study – Executive Summary". Retrieved from http://uk-ireland.aiic.net/ViewPage.cfm?page_id=888, accessed on 01-12-06

Marrone, Stefano (1993): "Quality – a shared objective". In: The Interpreters' Newsletter 5, 35-41

Massaro, Dominic; Shlesinger, Miriam (1997): "Information Processing and a Computational Approach to the Study of Simultaneous Interpretation". In: Interpreting 2-1/2, 13-53

Meak, Lydia (1990): "Interprétation simultanée et congrès médical: attentes et commentaires". In: The Interpreters' Newsletter 3, 8-13

Moser, Peter (1995): Survey: Expectations of Users of Conference Interpretation. Final Report, commissioned by AIIC, Vienna, SRZ Stadt- und Regionalforschung GmbH. Retrieved from http://www.aiic.net/ViewPage.cfm/article525, accessed on 26-11-06

Moser-Mercer, Barbara (1996): "Quality in interpreting: some methodological issues". In: The Interpreters' Newsletter 7, 43-55

Moser-Mercer, Barbara (1997): "The expert-novice paradigm in interpreting research". In: Fleischmann et al. (1997), 255-261

Moser-Mercer, Barbara ([1997]2002): "Process models in simultaneous interpretation". In: Pöchhacker/Shlesinger (2002), 149-160

Moser-Mercer, Barbara (2000): "The rocky road to expertise in interpreting: Eliciting knowledge from learners". In: Kadric et al. (2000), 339-352

Moser-Mercer, Barbara (2003): "Remote interpreting: Assessment of human factors and performance parameters". Retrieved from http://www.aiic.net/ViewPage.cfm/articles879.htm, accessed on 15-12-06

Moser-Mercer, Barbara; Künzli, Alexander; Korac, Marina (1998): "Prolonged turns in interpreting: Effects on quality, physiological and psychological stress." In: Interpreting 3-5, 47-64

Moser-Mercer, Barbara; Frauenfelder, Uli H.; Casado, Beatriz; Künzli, Alexander (2000): "Searching to Define Expertise in Interpreting". In: Dimitrova/Hyltenstam (2000), 107-131

Namy, Claude (1978): "Reflections on the Training of Simultaneous Interpreters: A meta-linguistic approach". In: Gerver/Sinaiko (1978), 25-34

Ng, Bee Chin (1992): "End Users' Subjective Reaction to the Performance of Student Interpreters". In: The Interpreters' Newsletter, Special Issue 1, 35-41

NIOSH (1999): *Publication No. 99-101: Stress at Work.* Retrieved from http://www.cdc.gov/niosh/stresswk.html, accessed on 24-11-06

Oléron, Pierre; Nanpon, Hubert (1964): "Recherches sur la répétition orale de mots présentés auditivement". In: L'Année Psychologique 64, 397-410

Oléron, Pierre; Nanpon, Hubert ([1965]2002): "Research into simultaneous translation". In: Pöchhacker/Shlesinger (2002), 43-50

Paneth, Eva (1957): *An investigation into Conference Interpretation: an Exploratory Study*. Unpublished M.A. Thesis. London: University of London

Picken, Catriona (ed.) (1993): *Translation – the Vital Link. Proceedings of the XIIIth FIT World Congress. Vol. 1.* London: Institute of Translation and Interpreting

Pio, Sonia (2003): "The relation between ST delivery rate and quality in simultaneous interpretation". In: The Interpreters' Newsletter 12: 69-100

Pöchhacker, Franz (1994a): *Simultandolmetschen als komplexes Handeln.* Tübingen: Narr

Pöchhacker, Franz (1994b): "Quality assurance in simultaneous interpreting". In: Dollerup/Lindegaard (1994), 233-242

Pöchhacker, Franz (2001): "Quality Assessment in Conference and Community Interpreting". In: Meta 46-2, 410- 425

Pöchhacker, Franz (2002): "Researching interpreting quality: Models and methods". In: Garzone/Viezzi (2002), 95-106

Pöchhacker, Franz; Shlesinger, Miriam (eds.) (2002): *The Interpreting Studies Reader*. London/New York: Routledge

Pradas Macías, Macarena (2006): "Probing quality criteria in simultaneous interpreting: The role of silent pauses in fluency". In: Interpreting 8-1, 25-43

Reiss, Katharina (1983): "Quality in Translation oder: Wann ist eine Übersetzung gut?". In: Babel 29-4, 198-208

Report of the Secretary General (2001a): "Provision of interpretation services to other locations from permanent interpretation structures in New York, Geneva, Vienna and Nairobi", United Nations, General Assembly, 54th Session A/54/176, 21 July 1999. In: The Interpreters' Newsletter 11, 153-162

Report of the Secretary General (2001b): "Remote interpretation", United Nations, General Assembly, 56th Session A/56/188, 16 July 2001. In: The Interpreters' Newsletter 11, 163-180

Riccardi, Alessandra (1998): "Interpreting strategies and creativity." In: Beylard-Ozeroff et al. (1995), 171-179

Riccardi, Alessandra (2003): "The Relevance of Interpreting Strategies for Defining Quality in Simultaneous Interpretation". In: Collados Aís (2003), 257-265

Russo, Mariachiara (2005): "Simultaneous film interpreting and users' feedback". In: Interpreting 7-1, 1-26

Salevsky, Heidemarie (1985): "Probleme des Simultandolmetschens in Theorie und Praxis". In: Bühler (1985), 191-194

Schjoldager, Anne ([1995]2002): "An exploratory study of translational norms in simultaneous interpreting". In: Pöchhacker/Shlesinger (2002), 301-311

Schjoldager, Anne (1996): "Assessment of Simultaneous Interpreting". In: Dollerup/Vibeke (1996), 187-195

Schweda Nicholson, Nancy (1986): "A United Nations interpreter survey: The specialist/generalist controversy I". In: Multilingua 5-2, 67-80

SCIC (2004): "Giving you a voice in Europe". Retrieved from http://europa.eu.int/comm/scic/interpreting/tech_retour_en.htm, accessed on 28-11-06

Seleskovitch, Danica (1976): "Interpretation, a psychological approach to translation". In: Brislin (1976), 92-116

Seleskovitch, Danica (1978): "Language and Cognition". In: Gerver/Sinaiko (1978) 333-341

Seleskovitch, Danica (1986): "Who should assess an interpreter's performance? Comment on Hildegund Bühler". In: Multilingua 5, 236

Shlesinger, Miriam (1989): "Extending the Theory of Translation to Interpretation: Norm as a Case in Point". In: Target 1, 111-115

Shlesinger, Miriam (1994): "Intonation in the Production and Perception of Simultaneous Interpretation". In: Lambert/Moser-Mercer (1994), 225-236

Shlesinger, Miriam (2003): "Effects of presentation rate on working memory in simultaneous interpreting". In: The Interpreters' Newsletter 12, 37-49

Shlesinger, Miriam et al. (1997): "Quality in simultaneous interpreting". In: Gambier et al. (1997), 123-131

Stenzl, Catherine (1983): *Simultaneous Interpretation – groundwork towards a comprehensive model.* M.A. Thesis. London: University of London

Straniero Sergio, Francesco (2003): “Norms and quality in media interpreting: the case of formula one press conferences”. In: The Interpreters’ Newsletter 12, 135-174

Sunnari, Marianna (1995): “Processing strategies in simultaneous interpreting: saying it all vs. synthesis”. In: Tommola (1995), 109-120

Sunnari, Marianna (2003): “Expert and novice performance in simultaneous interpreting: implications for quality assessment”. In: Collados Aís et al. (2003), 235-247

Tommola, Jorma; Lindholm, Johan (1995): “Experimental research on interpreting: which dependent variable”. In: Tommola (1995), 121-133

Tommola, Jorma (ed.) (1995): *Topics in Interpreting Research.* Turku: University of Turku

Toury, Gideon (1980): *In Search of a Theory of Translation.* Tel Aviv: The Porter Institute for Poetics and Semiotics

Treisman, Anne M. (1965): “The effects of redundancy and familiarity on translating and repeating back a foreign and a native language”. In: British Journal of Psychology 65, 369-379

Viaggio, Sergio (1992): “Teaching beginners to shut up and listen”. In: The Interpreters’ Newsletter 4, 45-58

Viezzi, Maurizio (1993): “Considerations on Interpretation Quality Assessment”. In: Picken (1993), 389-397

Viezzi, Maurizio (1999): “Aspetti della qualità dell’interpretazione”. In: Falbo et al. (1999), 140-151

Vuorikoski, Anna-Riita (1993): “Simultaneous interpretation – User experience and expectation”. In: Picken (1993), 317-327

Vuorikoski, Anna-Riita (1995): “Simultaneous interpretation as experienced by the audience”. In: Krawutschke (1995), 165-174

Vuorikoski, Anna-Riita (1998): "User Responses to Simultaneous Interpreting". In: Bowker et al. (1998), 184-197

Vuorikoski, Anna-Riita (2004*): A Voice of its Citizens or a Modern Tower of Babel? The Quality of Interpreting as a Function of Political Rhetoric in the European Parliament*. Tampere: Tampere University Press. Retrieved from http://acta.uta.fi/pdf/951-44-5878-8.pdf, accessed on 06-12-2006

Wadensjö, Cecilia (1998): *Interpretation as Interaction*. London/New York: Longman

Watson, Hilde (1982): "Conference Interpreting". In: The Incorporated Linguist 21-4, 139-141

Wenger, Liselotte (ed.) (1981): *Glossary of Terms Used in the Management of Quality*. Bern: EOQC Secretariat

Weischedel, Gisela (1977): "The conference interpreter – a tentative psycholinguistic investigation". In: Lebende Sprachen 22, 101-102

Eidesstattliche Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbständig angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Köln, den 28.12.2006