
TRANSKULTURALITÄT - TRANSLATION -
TRANSFER



**Interpreting Quality:
A Look Around and Ahead**

Cornelia Zwischenberger/ Martina Behr (eds.)

F Frank & Timme

Verlag für wissenschaftliche Literatur

Cornelia Zwischenberger/Martina Behr (eds.)
Interpreting Quality: A Look Around and Ahead

Transkulturalität – Translation – Transfer, Band 19

Herausgegeben von

Dörte Andres / Martina Behr / Larisa Schippel / Cornelia Zwischenberger

Cornelia Zwischenberger / Martina Behr (eds.)

Interpreting Quality: A Look Around and Ahead

FFrank & Timme
Verlag für wissenschaftliche Literatur

Umschlagabbildung: © Aurélie Auzas

Gedruckt mit Unterstützung der Österreichischen Forschungsgemeinschaft

ÖFG // ÖSTERREICHISCHE
FORSCHUNGSGEMEINSCHAFT

ISBN 978-3-7329-0191-3

ISSN 2196-2405

© Frank & Timme GmbH Verlag für wissenschaftliche Literatur
Berlin 2015. Alle Rechte vorbehalten.

Das Werk einschließlich aller Teile ist urheberrechtlich geschützt.
Jede Verwertung außerhalb der engen Grenzen des Urheberrechts-
gesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar.
Das gilt insbesondere für Vervielfältigungen, Übersetzungen,
Mikroverfilmungen und die Einspeicherung und Verarbeitung in
elektronischen Systemen.

Herstellung durch Frank & Timme GmbH,
Wittelsbacherstraße 27a, 10707 Berlin.
Printed in Germany.

Gedruckt auf säurefreiem, alterungsbeständigem Papier.

www.frank-timme.de

Table of Contents

CORNELIA ZWISCHENBERGER & MARTINA BEHR A Look Around and Ahead: Manifestations and Interpretations of Quality in Interpreting.....	7
SYLVIA KALINA Measure for Measure – Comparing Speeches with Their Interpreted Versions	15
EMILIA IGLESIAS FERNÁNDEZ Making Sense of Interpreting Difficulty Through Corpus-Based Observation. Correlations between Speaker’s Speech Rate, Mode of Presentation, Delivery Profile and Experts’ Judgements of Difficulty	35
E. MACARENA PRADAS MACÍAS User Awareness of Sui Generis Pause Patterns in Simultaneous Interpreting	67
RAFAEL BARRANCO-DROEGE Probing the Perception of Time-Manipulated Speech. What Can Content Analysis Tell Us?	89
OLALLA GARCÍA BECERRA Order Effect, Impression Formation and Their Impact on the Evaluation of Interpreting Quality.....	123
MARTINA BEHR Quality and Feelings – How our Feelings Come into Play When We Evaluate Interpreting Performances.....	147
FABRIZIO GALLAI Quality in Legal Interpreting and Pragmatics: Are They Compatible?.....	167
AÍDA MARTÍNEZ-GÓMEZ Interpreting Quality in Prison Settings: a Proposal for an Evaluation Tool	205

CORNELIA ZWISCHENBERGER Bridging Quality and Role in Conference Interpreting. Norms as Mediating Constructs	231
FRANZ PÖCHHACKER & CORNELIA ZWISCHENBERGER Quality and Role in Conference Interpreting. Views from the East and South of Europe.....	269
JIM HLAVAC & MARC ORLANDO 'Capturing' and 'Prescribing' Desirable Attributes of Community Interpreting: Design and Implementation of Testing Systems in Four Anglophone Countries.....	297
About the Authors.....	325

A Look Around and Ahead: Manifestations and Interpretations of Quality in Interpreting

Introduction by the Editors Cornelia Zwischenberger & Martina Behr

The topic of quality undoubtedly holds a central position in the research landscape of Interpreting Studies (IS). This may partly be explained by the need for an unprotected (academic) profession to focus particularly on quality in order to promote its (further) professionalization. Thus, it may be no wonder that despite the young age of IS, the topic of quality has been around now for about three decades already (cf. BÜHLER 1986), both in an explicit and implicit manner, and has gone through various formations.

A look around the research landscape shows that quality established itself as a very explicit research topic in conference interpreting with a whole bulk of (mainly quantitative) empirical research, while in community interpreting it may be said to be more of an implicit nature. There, it usually manifests itself as a “by-product” of investigations into the agency and functions of community interpreters, their role definitions, as well as the requirements regarding professional interpreting as expressed by the various testing schemes for accreditation and certification. A closer look around the research landscape in the realm of quality also suggests that the quest for clear universal criteria or parameters that determine quality which has virtually solely dominated quality research in conference interpreting (e.g. BÜHLER 1986; KURZ 2001; CHIARO & NOCELLA 2004) is no longer happening. Thanks to the advances in research into quality in this domain (cf. RICCARDI 2002; KALINA 2005; GRBIĆ 2008), we no longer assume that quality in interpreting is something that actually can be grasped and pinned down so easily. Current research on quality increasingly defines quality no longer as something objectively measurable and therefore tangible but as strongly bound to the situation and/or context, and the general circumstances under which an evaluation takes place. Quality is thus one part of an interactive situation of communication which is generated and influ-

enced by the various social agents in the field of interpreting, their respective roles (cf. ZWISCHENBERGER 2013), and their interrelatedness (cf. BEHR 2013). Therefore, a new approach is needed: a sociological one.

While the interactionist perspective on interpreting with its strong focus on the various interpreter roles and the ensuing questions of power distribution has been an integral part of research into community interpreting from its very inception, it is only now beginning to gain ground in conference interpreting and quality research in this domain. The interactionist approach challenges the concept of a stable sense that is conveyed from one side to the other. Instead, it places emphasis on communication, thereby underlining that there must be at least two parties interacting with each other and constructing a sense in a common effort. In the case of interpreting, this necessarily involves a speaker, an interpreter, and a listener who all – in the hermeneutic sense – interpret what is being said (cf. BEHR 2013; ZWISCHENBERGER 2013).

Professional associations such as the International Association of Conference Interpreters (AIIC), however, remain engaged in promoting the idea that the interpreter faithfully and completely renders the sense of the message as intended by the speaker, thereby still adhering to the myths of objectivity and transparency in determining quality. But thanks to research in community interpreting and its influence on conference interpreting, the ideas of neutrality and transparency are challenged (cf. PRUNČ 2012³: 350ff.).

For current quality research, the content of the source speech itself is still an issue but is integrated within the framework of more holistic approaches where the central focus is on the (quality of the) various components that generate the respective communicative situation in which an interpretation assignment takes place. Therefore, the source text itself and its characteristics become a subject of research and a relevant precondition of every interpreting performance whose quality it also ultimately determines (cf. KALINA 2011).

The interpreter's primary task to convey content and to do this in the appropriate form is assured by professionally trained interpreters in conference interpreting. Academic training is engaged in conveying an ethics-based habitus which allows the students to develop a profession's ethical "self-concept" (ANDRES 2011) that guarantees adherence to the respective quality standards that are possible under the given conditions.

While such an approach already calls on the budding interpreter to adopt social responsibility, and therefore a sociological approach is needed, the aim to assure understanding on the side of the listeners calls for this even more:

investigation into how the content conveyed by the interpreter is ultimately received, cognitively processed, and appropriated on the other side by the recipient also calls for a sociological, or rather socio-cognitive, perspective to be adopted by IS scholars in the future.

On the one hand, a holistic perspective leads to the development of macrostructural approaches in quality research on the horizontal line and efforts to be as broad as possible by taking all sorts of factors and aspects of quality on board. On the other hand, however, we also witness the formation of approaches on a microstructural level going into the vertical depth with the focus on an in-depth analysis of very specific parameters, namely formal pragmatic ones (e. g. COLLADOS AIS et al. 2011).

In line with this, all of the volume's contributions, regardless of the interpreting domain they represent, are united by presenting the topic of quality as a multifaceted, complex, and highly relative construct. In this respect, this volume features quality "not as a self-contained topic but as a complex, overarching theme in which all aspects of the interpreter's product and performance – textuality, source-target correspondence, communicative effect, and role performance – play an integral part" (PÖCHHACKER 2004: 153). The majority of the contributions refers to conference interpreting; this seems to confirm the continuing more implicit existence of quality in community interpreting showing at the same time the importance of scientific events such as the Critical Link 5 conference in 2007 and its proceedings "Quality in interpreting – a shared responsibility" (HALE et al. 2009) in order to improve research in this area.

A closer look at the content of this volume shows that the articles are indicative of a trend in the research on quality of interpreting. The fact that the complexity of the (conference) interpreting process already starts with the quality of the original speech is shown by **Sylvia Kalina** in her article *Measure for Measure – Comparing Speeches with their Interpreted Versions*. It is based on the initial assumption that these days interpreting, like everything else, is regarded as a service that is supposed to be as inexpensive as possible. For a calculation of costs, this requires greater clarity with regard to what actually constitutes a 'good' interpreting performance. The product of interpretation is, according to Kalina, "a function of the original speech and its characteristics." In her article, Kalina provides profiles of the source and the target texts in which individual characteristics can be evaluated by first analysing the quality of the source text and subsequently assessing the quality of the interpretation

dependent on the source text, i.e. as the relative quality achievable under the respective conditions. Another essential notion is that “a source text presented by a speaker (and any interpreted version of it) is, in its entirety, more than the sum total of measurable components.” Kalina’s contribution in this volume is an extended and updated version in English of her paper published in German (cf. KALINA 2011).

Emilia Iglesias Fernández goes in a similar direction by examining the original speech in order to be able to make assertions about its level of difficulty. In her article *Making Sense of Interpreting Difficulty Through Corpus-Based Observation. Correlations between Speaker’s Speech Rate, Mode of Presentation, Delivery Profile and Experts’ Judgements of Difficulty*, she juxtaposes the characteristics of the original speech and the evaluation of the interpretation in order to determine more precisely which aspects are crucial for the level of difficulty perceived by the listener, since “the assumption that high source speech rate as counted by the number of words or syllables per minute/second is not a reliable predictor of perceived difficulty.” On the contrary, the latter seems to strongly depend on the connection between source text speech rate, the speaker’s degree of pre-planning and expressiveness, and the presentation mode.

Other studies of the ECIS research group in Granada, Spain, take a similar approach, examining additional variables in order to be able to make assertions about the equivalent effect. **Macarena Pradas Macías’s** *User Awareness of Sui Generis Pause Patterns in Simultaneous Interpreting* investigates the role of pause patterns and the influence they have on the evaluation of an interpreting performance. It seems that different pause patterns are due to certain lexical categories in the original, and that a manipulation of these patterns can be detected by the listeners. This highlights the importance of listeners’ expectations, assessing “whether this manipulation also affects their evaluation and their perception of standard quality criteria measured against their expectations.”

One possibility to improve research on quality in conference interpreting regarding the role of the listener is proposed by **Rafael Barranco-Droege** who examines the speed of a delivered speech as an independent variable: *Probing the Perception of Time-Manipulated Speech. What Can Content Analysis Tell Us?* Focusing on the perception of the listener in particular, the article questions the reliability of previous studies in this context and contributes to a critical debate about the reliability of the methods in the research on quality, partly

due to “doubts on the reliability of evaluation by untrained users when answering ordinal-scale questionnaires.” Examining a specific variable and applying a proposed *user education* allows not only to carry out research on quality, but also to categorise and evaluate it.

Olalla García Becerra similarly points out gaps in the research on quality of conference interpreting by analysing *Order Effect, Impression Formation and Their Impact on the Evaluation of Interpreting Quality*. Her article clearly shows the many factors that play an important role in the evaluation of conference interpreting performances that are still not being taken into account sufficiently. The question of the relevance of order effects can be used as an opportunity to think about prevailing evaluation practices, where several performances are evaluated consecutively without reflection on the possible influence this order may have on the evaluation outcome. García Becerra concludes that “the position of an interpreting performance in the sequence influences the subjects’ quality perception and the process of impression formation regarding the interpreters.” The issue of impression formation thereby incorporates a socio-cognitive element into this essentially holistic view-point.

Martina Behr takes this socio-cognitive approach one step further. Her article *Quality and Feelings – How our Feelings Come into Play When We Evaluate Interpreting Performances* focuses on the equivalent effect and the active role of the listener. She shows that the listeners’ or examiners’ understanding and evaluation processes can be influenced by their feelings, and tries to answer the question: “Can the correlation between the listeners’ feelings and their evaluation of interpreting performances be measured?” Analysing the issue from a cognitive-psychological perspective, Behr determines that expertise as well as intersubjectivity play a decisive role for more objectivity in grading interpreting performances.

The question of the equivalent effect for the active listener is also the focus of this volume’s first article in the area of community interpreting when **Fabrizio Gallai** asks: *Quality in Legal Interpreting and Pragmatics: Are They Compatible?* With reference to Relevance Theory, he examines to what extent discourse markers add to a better equivalent effect by reducing the interpreter’s visibility and thus giving the listener greater responsibility in the understanding process. Successful communication – in this case in legal settings – is achieved by “the ‘mutual cognitive environment’ of speaker and hearer, i.e. a shared set of assumptions which constitute the ‘context’ of communication.”

The importance of the expectations of all parties involved in the triad is emphasised by **Aída Martínez-Gómez** in her article *Interpreting Quality in Prison Settings: a Proposal for an Evaluation Tool*. Her tri-partite evaluation model illustrates the role of the different expectations of prisoners, officers, and interpreters, and especially the interpersonal factors when unqualified interpretation leads to communication break-downs, as may happen for interpreting in prison settings. Quality assurance in these settings is a highly challenging goal, because “bringing non-professionals into the quality equation adds an extra layer of complexity to the evaluation of interpreting quality.”

Role expectations also play a central role in **Cornelia Zwischenberger’s** empirical study and the theoretical framework it is based on. In her article *Bridging Quality and Role in Conference Interpreting. Norms as Mediating Constructs* she describes quality as a social construction that builds on the foundation of normative role expectations of the interpreter. “This social constructionist view is enriched by a further perspective on quality which integrates the concept of ‘social role’. This concept contributes to explaining the variability of quality judgements.” Zwischenberger presents the results of two large-scale web-based surveys amongst the members of the International Association of Conference Interpreters (AIIC) as well as the German Association of Conference Interpreters (VKD) on the roles of conference interpreters and thereby the expectations of the quality of interpreting. She looks at how active or passive conference interpreters judge their own roles to be in the construction of quality.

The article *Quality and Role in Conference Interpreting. Views from the East and South of Europe* by **Franz Pöchhacker & Cornelia Zwischenberger** builds on this preliminary work and presents the results of three replication studies of the global AIIC study. These web-based surveys were carried out with conference interpreters in the Czech Republic, Italy, and Poland. The aim of these replication studies was to ascertain differences as well as consistencies within the responses of the conference interpreter communities within the Czech Republic, Italy, and Poland, and to compare these results to the answers provided by the members of AIIC and VKD. Interestingly, the study revealed a number of differences in role perceptions and thus eventually in the definitions of quality which can be explained by dissimilar professional socialisations and therefore habitus.

The volume closes with a meta-approach on quality evaluations by **Jim Hlavac & Marc Orlando**. In their article ‘*Capturing’ and ‘Prescribing’ Desirable*

Attributes of Community Interpreting: Design and Implementation of Testing Systems in Four Anglophone Countries, quality management itself undergoes quality control and is put to the test. The article shows the particularities of community interpreting compared to conference interpreting, which are also reflected in the quality management in this area: “Observations about the social construction of quality become evident in the cross-national description of certification procedures where credentialing authorities with congruent aims employ varying approaches and measures in the ‘pursuit for quality.’” A central aspect of this article is a very precise examination of the validity and utility of the four different testing systems.

While a sociological turn has been taken in Translation Studies and research on community interpreting (cf. ANGELELLI 2012), this was not entirely the case for conference interpreting until now. However, conference interpreting and quality research in interpreting are now about to take this turn, as shown by this volume. Whether it is the socio-cognitive perspective in Behr, Gallai and García Becerra, the sociological approach in Martínez-Gómez and Zwischenberger, or the situative social variability of the entire source context in Iglesias Fernández and Kalina – all of these contributions allude to the social or sociological dimension of quality research. The articles compiled in this volume are certainly not representative of all current trends, but they are at least indicative of future developments: the sociological turn in quality research will increasingly shift our attention to the interpreters as social agents and to all “the social factors permeating the communicative and social act of [...] interpreting” (ANGELELLI 2012: 125). This will significantly extend the field of variables that come into play in quality research, and will furthermore confirm that there is and cannot be the one and only definition of quality. Because of this increasing complexity of research into interpreting quality, it may safely be assumed that only inter- and transdisciplinary approaches will prove to be successful in the future. Thus, progress in Interpreting Studies, and in particular in the field of quality in interpreting, will also largely depend on future developments of the disciplines we are adopting and appropriating concepts and methods from. This will require keeping pace with any advances in both close and more distant neighbouring disciplines.

References

- ANDRES, Dörte (2011): "Dolmetschwissenschaft zu Beginn des 21. Jahrhunderts: Ein integrative konzipiertes Dolmetschprozessmodell", *Studia Universitatis Babeş-Bolyai. Philologia*, No. 1/2011, 81-103.
- ANGELELLI, Claudia (2012): "Introduction. The sociological turn in translation and interpreting studies", *Translation and Interpreting Studies* 7 (2), 125-128.
- BEHR, Martina (2013): *Evaluation und Stimmung. Ein neuer Blick auf Qualität im (Simultan-)Dolmetschen*. Berlin: Frank & Timme.
- BÜHLER, Hildegund (1986): "Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters", *Multilingua* 5 (4), 231-235.
- CHIARO, Delia & NOCELLA, Giuseppe (2004): "Interpreters' perception of linguistic and non-linguistic factors affecting quality: A survey through the World Wide Web", *Meta* 49 (2), 278-293.
- COLLADOS Afís, Ángela; IGLESIAS FERNÁNDEZ, Emilia; PRADAS MACÍAS, E. Macarena & STÉVAUX, Elisabeth (eds.) (2011): *Qualitätsparameter beim Simultandolmetschen Interdisziplinäre Perspektiven*. Tübingen: Gunter Narr.
- GRBIĆ, Nadja (2008): "Constructing interpreting quality", *Interpreting* 10 (2), 232-257.
- HALE, S.; OZOLINS, U. & STERN, L. (eds.) (2009): *The Critical Link 5. Quality in Interpreting – a shared responsibility*. Amsterdam/Philadelphia: John Benjamins.
- KALINA, Sylvia (2005): "Quality assurance for Interpreting processes", *Meta* 50 (2), 769-784.
- KALINA, Sylvia (2011): "Maß für Maß. Eine vergleichende Profilanalyse von Diskursen beim Dolmetschen", *trans-kom* 4 (2), 161-175.
- KURZ, Ingrid (2001): "Conference Interpreting: Quality in the Ears of the User", *Meta* 46 (2), 394-409.
- PÖCHHACKER, Franz (2004): *Introducing Interpreting Studies*. London/New York: Routledge.
- PRUNČ, Erich (2012³): *Entwicklungslinien der Translationswissenschaft*. Berlin: Frank & Timme.
- RICCARDI, Alessandra (2002): "Evaluation in interpreting. Macrocriteria and microcriteria". In: HUNG, E. (ed.): *Teaching Translation and interpreting, building bridges*. Amsterdam/Philadelphia: John Benjamins, 115-125.
- ZWISCHENBERGER, Cornelia (2013): *Qualität und Rollenbilder beim simultanen Konferenzdolmetschen*. Berlin: Frank & Timme.

Measure for Measure – Comparing Speeches with Their Interpreted Versions

The question of what constitutes quality in interpreting has been answered in different ways by linguists, psychologists and interpreters themselves. Measurements were initially based on transcripts and found insufficient, but obtaining authentic audio and video data that can be analysed is difficult. Even surveys among users appear to yield inconclusive results. If interpreting is regarded as a service whose quality needs to be assessed, it is necessary to choose a process perspective which includes all phases of an interpreting assignment. For the purposes of quality assurance, a checklist that includes the different phases helps document how an assignment has been carried out. In addition to this practical instrument it is necessary to develop methods of measuring those components of performance that can be measured and make their dependence on external factors transparent.

This article attempts to develop profiles of speeches made in conferences and their versions in other languages as produced by conference interpreters. Within this framework, an interpreting product can be assessed as a function of the original speech and its characteristics. Thus, a degree of measurability should be achieved, and the results of measurements should reflect not only source and target texts and their presentation but also the conditions and factors that influence their perceived quality. The profiles were tested with professional interpreters and may serve as a component of quality assurance in the field of conference interpreting, and hopefully in the future for all types of interpreting.

Quality Assurance for Interpreters

In all types of interpreting across all settings, quality is a central concern, as is quality control and assurance. Too many people who have a command of two languages think that they can offer interpreting services, and too many bodies in search of such services are not aware of the skills needed to provide reliable

interpreting. This is true not only for community settings but also for conference-like assignments where interpreting services must, above all, be cheap and, in the second place, of good quality. This is one of the reasons why professional interpreters will have to prove that their work fulfils ambitious quality criteria. The question is what these criteria are or should be? For audiences and conference organisers there is too little transparency as far as those criteria are concerned, and very often, after a meeting, a questionnaire has to be filled in by participants with one question on interpreting where either “good” or “poor” has to be ticked, without any regard to the circumstances in which the interpreters in question have been working. Even interpreting studies is not really in a position to provide a clear answer that is applicable across the board, as the following statement illustrates:

Those who would evaluate quality in interpreting ‘across the board’ are faced with the fact that interpreting is not a single invariant phenomenon but a (more or less professionalized) activity which takes different forms in different contexts. Therefore, the concept of quality cannot be pinned down to some linguistic substrate but must be viewed also at the level of its communicative effect and impact on the interaction within particular situational and institutional constraints (PÖCHHACKER 2001: 421).

The difficulties in measuring and controlling or even assuring interpreting quality are obvious. Interpreting is a multifaceted service, it is temporary in character, it is rendered under conditions that are often adverse (KOHN & KALINA 1996), and it is provided for the purpose of being used once and on the spot; unlike a translation, it is not intended to be consulted again later. Owing to the situation in which interpreting takes place, its quality is of a temporary and evanescent nature. Accordingly, the rules governing quality assurance in translation cannot be transferred to the field of interpreting.

When confronted with requests for quality control, conference interpreters have traditionally invoked this volatile nature of what they do. In the digital age, however, this point of view can no longer be defended, as interpreting products are regularly recorded and offered for downstream or download via the Internet or on CD-ROMs for participants and/or the public at large after a conference. They are therefore available and accessible for product inspection. Nevertheless Pöchhacker (1994b: 233) and, several years later, the same author (2010: 1) deplores the limited number of studies based on authentic products.

Those studies that have focussed on interpreting products in the interim have been mostly concerned with one single phenomenon or some interrelated phenomena such as omissions and additions or delivery characteristics. Moreover, they have rarely performed source text analysis as thoroughly as they have analysed target texts.

I shall therefore endeavour to discuss ways of analysing interpreting quality in a framework that includes both source text parameters and the conditions in which target texts are produced. With this approach, I hope to gain acceptance among practising interpreters and overcome the scepticism they have expressed vis-à-vis previous attempts at measuring the quality of their work on the basis of parameters such as errors, omissions, hesitations, awkward syntax and others.

Past Approaches to Measuring Interpreting Products

Early research into interpreting quality focused on conference interpreting, usually the simultaneous mode, as other interpreting services were not as much in demand as today and there was little awareness of their importance. Initially, the analysis took the form of a comparison of transcripts of an original and the interpreted version. As it was difficult to obtain authentic material, especially when the aim was to compare various interpreted versions of one original, researchers chose to proceed experimentally, sometimes using subjects who had not even undergone any training in simultaneous interpreting. Barik (1971) was criticised for his focus on errors and the categories he established (omissions, additions, substitutions, errors and other changes that can be detected in the transcript). As he had neglected delivery parameters, he was unable to identify whether, in the interpreter's delivery, a word of the original omitted was perhaps made up for by, say, prosodic resources. Nevertheless, his study was the first to establish categories of differences between a source and a target text and to propose types of deviation.

Another attempt at identifying quality criteria was made by Bühler (1986) who surveyed professional conference interpreters and established a list of priorities that they regarded as crucial for quality.¹ Interestingly, reliability,

.....
1 In Bühler's questionnaire, the following criteria were included: native accent, pleasant voice, fluency of delivery, logical cohesion of utterance, sense consistency with original message, com-

teamwork and thorough preparation, i.e. criteria other than pure product data, were rated higher than some of the product parameters. After a number of interpreting scholars had conducted their own surveys of specific groups of participants, an AIIC study by Moser (1995) extended the scope of the survey to all types of conference participants who were asked to state their priorities with regard to interpreting quality. As Pöchhacker (2004: 27f) comments, participants' judgements are mostly somewhere between gratitude and mistrust. Results are heterogeneous or even self-contradictory. All these attempts provided insight into user expectations and their conceptions of interpreting quality, but they could not solve the problem of measurability. Their objective was rather to gain an overall picture of what quality is, or should be made up of, than to assess individual interpreting products.

Focus on Target Text

Even before but mainly after, Barik's study, several scholars devoted their efforts to the analysis of authentic material, from Oléron and Nanpon (1965) whose interest was in ear-voice span, and later Lederer (1981), who described different ways in which interpreting problems were solved and used her insights to propose a theory of simultaneous interpreting, to Pöchhacker (1994a), who developed a method for establishing text delivery profiles, and Vuorikoski (2004), who analysed interpreter performance in different languages at the European Parliament. They initially worked with audio material which needed to be transcribed and in some cases were later able to use video recordings of originals and interpreted versions. Due to the particular circumstances prevailing in each of the events analysed, there could not be any common yardstick or description by which to measure or judge a) which criteria any good-quality interpretation must fulfill, b) who is in a position to evaluate it and c) what the optimum product used as a yardstick for interpreting quality would look like. What made quality studies with authentic material from freelance interpreters even more difficult was that professional conference inter-

.....
pleteness of interpretation, correct grammatical usage, use of correct terminology, use of appropriate style, thorough preparation of conference documents, endurance, poise, pleasant appearance, reliability, ability to work in a team, positive feedback from delegates, other criteria (BÜHLER 1986: 234).

preters were more than reluctant to have their performance recorded and analysed.

Moser-Mercer (1996) proposes that for an interpreting product to be of ideal quality, it should be complete, accurate, free from any distortions of the original, and should also account for extralinguistic factors as far as the external conditions permitted. The importance of external factors and the impact that they have on interpretation are emphasised by Kopczynski (1994), according to whom quality depends on contextual and situational variables; these factors have to be considered and integrated into any evaluation of interpreting quality and the expectations on which that evaluation is based.

According to Pöchhacker's approach, based on the General Theory of Translation and Interpreting (REISS & VERMEER 1984), the translator's (and also the interpreter's) "discourse must first and foremost conform to the standard of intratextual coherence, and only in the second place must there be intertextual coherence, i.e. some relation of fidelity to the original" (1992: 213f). Pöchhacker (1994b: 241) is referring here to what Bühler (1986) termed "logical cohesion", meaning that the interpreted version is, above all, intralingually coherent, i.e. logical in itself. Nevertheless, he emphasises the intertextual dimension by pointing to the need to verify an interpreting product for "logical consistency" with the source text. This means giving priority to the functioning of the target text in the situation and sociocultural context of the target language over its dependence on the source text. It should be noted that the term "logical consistency" appears to be more appropriate than "logical cohesion", which, after all, can only be obtained if such logic (or coherence) can be identified in the source text – for the user of the original and the interpreter alike. Viezzi's quality model (2007) centres on four goals: equivalence, accuracy, appropriateness and usability; quality is defined as the degree to which these goals are achieved. The last two goals refer to the target text; it remains open how appropriate and usable the source text has been.

Even when analysing textual data alone, the question that arises is whether it is at all possible to assess an instance of interpretation without taking full account of the source text serving as its basis. Given the significance of the relationship between source and target text, I would suggest that we need a profile for the source text and its delivery that includes all the factors that may have an impact on interpreting quality and which should be as fine-honed as the corresponding target text profile.

Pöchhacker (1994b) has suggested a number of parameters for the establishment of a (source) text delivery profile that encompasses addressees, spontaneity vs. planned speech, media used and oral presentation variables. To some extent, Bühler (1990) had paved the way for this by describing characteristics of delivery in interpreting contexts (loose structure, un-/filled pauses, hesitations, false starts, repetitions), which she relates to discourse parameters such as situationality, textuality including intertextuality and intentionality.

Much later, Zwischenberger (2013) (see also PÖCHHACKER 2012) conducted a comprehensive online survey, inviting conference interpreters to listen to an audio recording of an interpretation and to answer questions on parameters based on Bühler's criteria. The informants had no opportunity to listen to the original, and the result is an assessment of quality focused on the product and not taking into account the characteristics of the source text. This method is appropriate if the aim is to find out what priorities interpreters themselves have. Pöchhacker (2012) discusses whether the findings of these studies can be generalised on a global scale with many different languages and cultures. My approach goes in the opposite direction, trying to analyse individual performance. In this, if the goal is to investigate the quality of the entire process, all phases and situational and contextual factors involved will have to be considered and accounted for. This is indispensable for quality assurance, where the aim is to overcome weaknesses and improve the entire process.

The Process Model as a Basis for Evaluating Interpreting Quality

If we wish to consider the entirety of factors that (may) have an immediate or indirect impact on the product, we need to include data on communication situation, conference equipment standard, participant composition, roles and characteristics of speakers, preparation of subject and terminology and many other factors (see KALINA 2002) related to different process dimensions (pre-, peri-, in- and post-process, see also KALINA 2005).

The assessment of quality cannot do without an assessment of all these factors. Kalina (2002) analyses two scientific conferences with starkly contrasting communication profiles and conditions, and the success of the interpreters who worked there. Comparison shows how strongly interpreting quality is influenced by external factors, i.e. factors beyond the control of the interpreters. The result was that the quality achievable under the circumstances (PÖCH-

HACKER 1994: 242) was extraordinary in one case and rather poor in the other. It is therefore indispensable for us to establish a method of description of external factors and conditions as well as a profile of each of the source texts to be interpreted, and to examine the interpreting product against this background.

What this rough analysis of the data relating to the two events also revealed was the significance of source text discourse and its delivery characteristics.

Inspired by this analysis and the ideas that resulted from it I set out to develop a purely descriptive process model of interpreting and its quality assessment. It has four dimensions and is intended, on the one hand, to facilitate in-depth analysis of interpreting quality and, on the other, to provide a basis for quality assurance (ideally) for all settings and modes of interpreting. The first practical instrument to be used by professional conference interpreters is a checklist published in Kalina (2005) covering all factors, phases and components and intended for regular use by the profession. Some of its components have been tested and the checklist is used by a number of interpreting colleagues.

The first dimension of the process model is the pre-process phase, where a number of studies are available on preparatory activities performed by interpreters. Gile (1995) describes the different phases of preparation and the main activities involved in each of them. With this in mind, I conducted an observational study of how exactly interpreters proceed in the advance preparation phase. Its results are discussed in Kalina (2009). Data was also collected and analysed for the peri-process and post-process phases, leaving the most difficult and delicate dimension, the in-process phase, to be addressed last. For this, it is necessary to find a solution making it possible to identify and measure the relationship between source text and target text and their different parameters.

One consideration that has to be kept in mind in this context is that, just as in any discourse, a source text presented by a speaker (and any interpreted version of it) is, in its entirety, more than the sum total of measurable components. Some factors defy measurement or assessment as such but may well have a bearing on quality. How would one measure the degree of empathy displayed by a speaker or, even more importantly, by an interpreter? Nevertheless, it seems useful to obtain at least a rough profile of the characteristics found in the two related pieces of discourse and in their histories if one wants to learn more about how interpreting quality is achieved.

Components of a Framework for In-Process Analysis

If we assume that there are standards for an ideal presentation (especially delivery standards) by a speaker at a multilingual conference with interpreters (see the multitude of tips for speakers), we should think that such standards also exist for an interpreting product. As far as they can be defined at all for target text production, the relevant standards are certainly not identical with standards for a source text and its delivery. An original presentation may be attractive or even spectacular in some way (owing e.g. to the speaker's accent or a very personal style of delivery), but this is by no means the case for the interpreted version. And though an original speaker may indulge in linguistic peculiarities (self-correction or use of unusual metaphors), the interpreter is certainly not expected to do the same. The speaker may slow down or speed up whereas the interpreter is bound by the speed and rhythm of the original; and a speaker may decide to change the text of his/her speech at the last moment, whereas the interpreter is semantically dependent on what the speaker says (KOHN & KALINA 1996). The interpreter's language and delivery are generally under closer scrutiny than the speaker's, as the interpreter is rightly regarded as a language professional. However, analysing and assessing an instance of interpretation cannot be done by counting and comparing source and target text elements, as cultural and contextual factors may prompt an interpreter to opt for solutions by means of explicitation or omission and other modifications or adaptations.

One has to keep in mind that texts delivered in the oral mode with a certain degree of spontaneity and born of a real relationship with an audience that is present at the same time are usually linguistic performances intended exclusively for that particular audience and situation. This applies even more decisively to interpreting products. They are representations of intentions and utterances created in one culture and with audiences of that same culture in mind, but they have to function in a different culture where recipients will generally base their comprehension on different mental models (see SETTON 1998).

What Napier stipulates for sign-language interpreting, i.e. that interpreters have to be able "[...] to determine what something means to their target audience and the best way to meaningfully interpret a message so it makes sense with respect to the audience's cultural norms and values" (NAPIER 2003: 102), is just as valid for spoken language and thus also for conference interpreting.

The question raised by Kopczynski “[...] should s/he be the ghost of the speaker or should s/he intrude, i.e. omit, summarize or add portions of the text?” (1994: 90), finds its answer in the fact that the interpreter is one of those who comprehend and thus interpret what s/he has understood. In this sense, the interpreter is part of the group of listeners (cf. also BEHR 2013).

By its very nature, interpreter-mediated communication and interaction differs from monolingual communication, as cultural differences exist at different levels. This is why an addition or omission is not necessarily to be rated as an error: an omission may even support the speaker’s intention and help him/her reach their goal. Conversely, an addition may be a necessary explicitation enabling the audience to comprehend what a speaker says against the background of his/her own culture and mental model. Moreover, the importance of prosodic means of communication must be considered when assessing interpreting quality and comprehension on the part of the audience, even when we know that many recipients are not even aware of prosodic effect (cf. e.g. MOSER 1995; AHRENS 2005).

Pre- and Peri-Process Factors to Be Considered

Two factors serve to illustrate the significance of pre- and peri-process conditions. One is the type of event which, as Pöchhacker’s approach (1994a: 52) suggests, can be identified by characteristics of source text types distinguished by the degree of formal structure or technicality, homogeneity of group culture, information density/intensity, visual material used and flow of information.

The second factor is the person delivering the source text. If possible, the interpreter will have sketched out a profile of a speaker in the pre-process phase (preparation) and will thus be familiar with some details about the speaker, the subject, his/her attitude or other relevant information. If a manuscript or presentation file is available for preparation, it will be used in pre-process activity, unless the document does not arrive before the event has started.

Measuring pre-process activity can take the form of indicating the number of hours invested in the preparation of each presentation/speech, plus general knowledge preparation plus the time invested in preparing for the assignment as such. It is highly probable that there will be a positive correlation between

hours invested in general, linguistic preparation and quality of interpreting, although repeated assignments of the same kind tend to need less preparation time. The relationship between preparation and quality is confirmed by Lamberger-Felber (1998). In the following, I have therefore chosen to address in-process factors and especially the relationship between source and target text presentation.

Measurable Components of Delivery Profiles

Before addressing the problem of measuring interpreting and its quality, it is necessary to point out, in line with the above, the necessity of developing a profile for a source text presentation in situation and context reflecting those factors that may influence interpreting.

If a purely linguistic analysis is aimed at, one can use audio recordings (available e.g. from the European Parliament, from numerous company General Assemblies etc.) and transcribe them. But as we assume that interpreting quality is not a purely linguistic entity, even source text recordings will not suffice to put together all components of the profiles in question. Factors that may have an impact on interpreting but are not reflected in a recording include the documents made available by the speaker prior to speaking, the interpreters' knowledge about a speaker's background, presuppositions and interest regarding the audience as well as such things as the significance a speaker attributes to his/her own delivery. These are aspects of the pre-process phase. But there are also nonverbal means of communication, audience response, and many other facets that an audio recording does not reveal. It is therefore essential to have at least video recordings of the speaker and, if possible, also of the setting as such.

Once a source text delivery profile is established, the same has to be done for the target text, and if the two profiles lend themselves to comparison at least in part, it may be possible to assess interpreting quality. As profiling the two presentations is very complicated and error-prone, both will have to be scrutinised by audience members, professional interpreters and researchers, especially those working in the fields of interpreting studies, as well as by specialists in psycholinguistics and the linguistics of spoken language. It goes without saying that both the profiling and the subsequent scrutiny stages continue to be the most difficult stage in attempting to develop a model of inter-

preting quality and its assurance. So far, publicly available speeches have been used and it was therefore felt that there was no need to inform speakers of the analysis performed with reference to their presentations. It is, however, necessary to obtain the support of professional interpreters the quality of whose performance is the object of the study. The reasons for scepticism on their part are: (1) Interpreting is error-prone, volatile and should therefore not be analysed from a perspective that aims at long-term validity. (2) They reject the idea of being assessed by non-interpreters who are not sufficiently acquainted with the constraints and peculiarities of their work. (3) They are extremely sceptical when it comes to research aimed at defining an ideal state as they know that their everyday working life, i.e. interpreting, is far from ideal. (4) An open assessment of an interpreting product as a whole has so far been regarded as absolutely taboo, even when the assessors are professional interpreters; this may even be regarded as a violation of professional ethics. (5) An assessment of quality always tends to raise the question of interpreter liability for the interpreting product.

Accordingly, it is a delicate undertaking to acquaint professional conference interpreters with any kind of measurement methodology that considers a large number of parameters. I have done so in several workshops, and the best way of breaking the ice was to put my own interpreting quality to the test. This was done with a semi-experimental piece of simultaneous interpreting from a video recording then unknown to me that I had done spontaneously at the request of my students (so there was no pre-process phase) and which they had recorded. In the workshops we managed to establish a first rough profile of both source and target presentation with the matrix that I had prepared. As the source text had some flaws of its own (e.g. poor sound) and no preparation had been possible, the interpreting product was bound to be sufficiently far from the ideal to demonstrate that the aim was not to count errors and assess quality on that basis alone.

The graphs in table 1 and table 2 (see below) give an idea of the factors we attempted to measure when assessing my interpreting product against the background of the source text and its presentation. I would suggest that such profiles should be established for a number of source texts with ratings by as many people as possible to check whether the result is sufficiently reliable to be used as an assessment of its quality.

Source Text Presentation Profile

In table 1 the left-hand column (1) lists the characteristics of the original, broken down into addressee directionality, linguistic and delivery parameters and exemplifying (in italics) typical properties of each of these characteristics. Column (2) indicates the assessment grid and its direction (normally from left to right) to make assessment rating easier, and column (3) indicates the scale of possible values; negative values stand for factors that make interpreting difficult, positive values refer to the interpreter-friendliness of the presentation. Column (4) serves for entry of the actual value, and column (5) is a multiplier for the value of column (4); it should be an indication of the significance each characteristic has for successful interpreting. Column (6) calculates the result for each line (automatically if an Excel format is used) and displays the calculated total in the bottom line; a high value should suggest interpreter-friendly presentation by the speaker.

(1) PROFILE SOURCE TEXT (ST) <i>PARAMETER ST PRODUCER</i>	(2)	(3) ASSESSMENT SCALE from -4 → +4 -4 = LOW, POOR +4 = HIGH, GOOD	(4) VALUE	(5) MULTI- PLIER	(6) RESULT
<i>DISCOURSE TYPE</i>					
<i>TOTAL LENGTH IN MINUTES</i>	In minutes short → long	No assessment			
ADDRESSEE ORIENTEDNESS					
Clear outline, structural signals <i>Overview, announcement of new sections</i>	few → many	-4 → +4			
ST-audience orientedness <i>References to knowledge and expectations of ST listeners</i>	few → many	-4 → +4			
TT-audience orientedness <i>References to knowledge, cultural differences and expectations of TT listeners</i>	few → many	-4 → +4			
Comprehensibility for ST listeners <i>Coherence, information content, use of connectives</i>	weak → strong	-4 → +4			

LINGUISTIC FEATURES					
Register and style <i>inappropriate, colloquial, technical, appropriate</i>	poor → good	-4 → +4			
Syntax appropriateness for oral presentation <i>Ease of recognition of syntactical structures, sentence boundaries and length</i>	low → high	-4 → +4			
Amount of technical terms and concepts not communicated in advance	high → low	-4 → +4			
Amount of explanations of potentially unknown terms and concepts	low → high	-4 → +4			
Amount of complex figures, proper names	high → low	-4 → +4			
Amount of other names, culture-specific terms	high → low	-4 → +4			
Amount of puns, jokes or anecdotes, metaphors	high → low	-4 → +4			
DELIVERY CHARACTERISTICS					
source language delivery quality <i>Non-native speaker, dialect</i>	poor → good	-4 → +4			
Handling of media <i>Manuscript reading, media combination with comments, presentation, spontaneity</i>	poor → good	-4 → +4			
Amount of hesitations, (un/filled)	few → many	-4 → +4			
Articulation, segmenting, pauses	Poor, negligent → clear, pronounced	-4 → +4			
Prosody, meaning-based intonation	poor → good	-4 → +4			
Nonverbal means of communication <i>Extent of supportiveness</i>	few → many	-4 → +4			
Degree of spontaneity <i>Speed variation, unfinished sentences, insertions</i>	many → few	-4 → +4			

Overall delivery speed <i>Words (100, 120, 140, 160), syllables per minute (too fast/slow, appropriate)</i>	Inappropriate → appropriate	-4 → +4			
Enunciation or grammar errors, slips of the tongue	many → few	-4 → +4			
Amount of self-repair operations	many → few	-4 → +4			
Handling of microphone	poor → good	-4 → +4			
Degree of redundancy, amount of repetitions	few → many	-4 → +4			
Amount of unexpected propositions <i>Surprising utterances, self-contradictions</i>	many → few	-4 → +4	--		
Sum total					

Table 1: Profile of source text delivery

Target Text Presentation Profile

In column (1), the table for the target text (table 2) lists the typical characteristics which, according to the literature, an interpreting product is expected to have, with subcategories of consistency with original, language and delivery. Again, properties attributed to each of the characteristics are exemplified in italics to facilitate rating. Column (2) indicates the direction of assessment (usually from left to right, meaning from poor to good). Column (3) indicates the scale of possible assessment values. Column (4) is for entry of the assessment, column (5) for the weighted significance of the characteristic in question and column (6) presents the result for each line; the bottom line again presents the calculated overall result. A high value will generally suggest good interpreting quality. To reflect the one-way relationship between source and target text, the overall result for the original must be set in relation to the result for the interpreting product. If the original scores better than the interpreting product, interpreting quality would be regarded as poor.

(1) PROFILE TARGET TEXT – TT Parameters	(2)	(3) ASSESSMENT SCALE from -4 → +4 -4 = <i>low, poor</i> +4 = <i>high, good</i>	(4) VALUE	(5) MULTI- PLIER	(6) RESULT
CONSISTENCY WITH SOURCE TEXT					
Appropriateness for type of discourse	poor → good	-4 → +4			
Lack of completeness <i>Unjustified emissions, additions</i>	many → few	-4 → +4			
Accuracy <i>Details, nuances, replacements</i>	poor → good	-4 → +4			
Rendition of figures, proper names	poor → good	-4 → +4			
Adaptations to target culture <i>Amount of successful renditions</i>	low → high	-4 → +4			
Content of message	poor → good	-4 → +4			
Micropropositions rendered appropriately <i>Possible with different means (prosody etc.)</i>	poor → good	-4 → +4			
Macropropositions rendered <i>Comprehension by TT listeners achieved</i>	poor → good	-4 → +4			
LINGUISTIC FEATURES					
Style, register <i>Adequateness for target audience</i>	poor → good	-4 → +4			
Grammar <i>Syntax, morphology</i>	poor → good	-4 → +4			
Terminology <i>Correctness, appropriateness</i>	poor → good	-4 → +4			
Finished sentences	few → many	-4 → +4			
Elegance, idiomaticity	poor → good	-4 → +4			