Quality and norms in conference interpreting

E. Macarena Pradas Macías and Cornelia Zwischenberger

Introduction

Quality and norms have always been a matter of concern to the conference interpreting research community, and the question “Quality for whom?” (Shlesinger 1997: 127) is constantly present.

Despite the fact that “quality has been approached from different perspectives, there is not yet a universal consensus on what it is” (Collados Aís & García Becerra 2015: 368). The complexity of the concept was addressed by Pöchhacker (2001) in his description of the multiple perspectives from which quality may be approached. He proposed a model for defining quality involving two analytical perspectives. The first aims to determine the relation between participants in an abstract interpreting context using their attitudes, needs and points of view (‘norms’), while the second focuses on the product of interpreting (within a specific event).

Keeping in mind the importance of this dual perspective for aspects such as quality standards and criteria for evaluation, this chapter surveys the studies that have contributed to forming a consensus on a range of key issues: determining what quality is, the criteria by which it is defined, how technology has influenced the methods of investigating quality, and which norms, proposed by whom, are intended to ensure that quality is achieved.

Clearly, the cognitive dimension is highly relevant both in the effective generation of an interpretation and in the interaction between participants in an interpreting event. As such, cognition also has a central role in determining and affecting quality. However, this chapter concentrates on the product-related rather than the process-related aspects of interpreting quality (for cognitive aspects, see Hodzik & Williams, Chapter 26 and Riccardi, Chapter 27, in this volume).

The chapter begins with studies of expectations and evaluation on the part of users and experts. It focuses on empirical studies in quality research that emphasise evaluating an interpreting product as objectively as possible via a more or less stable set of criteria. Up to the end of the twentieth century, such studies were concerned with research methodology and sought to establish evaluation criteria and determine how and by whom interpretations should be evaluated. In a later section we focus on developments during the twenty-first century, emphasising how new technologies have facilitated new methods in observational and
experimental studies and the possibility of studying larger sample populations. The beginning of the twenty-first century was also when conference interpreting research started approaching quality from a sociological perspective, thereby focusing on the relations between participants at an interpreting event. This is achieved by analysing the norms that shape the expectations of such participants.

Another issue discussed in this chapter is common to research in both centuries: the difficulty of balancing the influence of linguistic and extra-linguistic aspects in the expectations and evaluations of quality. Moreover, if we agree that a vision of quality should take the interaction in an interpreting event into account, we also need to consider the factors that condition the production of the interpretation. The various factors influencing processing and performance will, in the final analysis, not only affect the quality of the end product/speech, but also have a crucial impact on the receivers’ perceptions (see, for example, Barranco-Droege (2015), discussed below). This is because receivers’ perceptions depend not only on the characteristics of the interpretation itself, but also on sociological factors inherent in human communication. Indeed, the sociological component may also provide an explanation for the variability that studies identify in expectations and evaluation, ensuing from the operation of norms.

Empirical approach to interpreting quality: expectations versus evaluation

Interpreting research based on user and expert questionnaires has provided the most methodologically consistent approach to analysing quality (Gile 2000: 305). In the 1980s, Gile (1989: 25) and other researchers recognised the importance of carrying out empirical studies that reflect the needs and expectations of users in order to answer practical questions about quality, such as for whom? and with what criteria? Broadly speaking, these studies may be classified into two groups: those focusing on expected/perceived quality, and a smaller group concerned with actual quality, where the source text and target text are compared, and the interpreter’s cognitive processes may be considered.

The very first empirical study of expectations was conducted by Bühler (1986). She addressed one of the practical questions in quality research, namely, who should evaluate the interpretation? After discussing the pros and cons of evaluation being performed by interpreting trainers, agencies, clients or users, she presented two groups with a questionnaire. The first group consisted of interpreters who were members of AIIC, and the other group of members of the Committee for the Admission and Language Classification of Applicants (CACL), set up by AIIC. She was particularly interested in this second group because she supposed that as evaluators, they would take the expectations and needs of potential users into account.

Bühler (1986) divided her questionnaire into three parts. The first aimed to determine how much importance respondents attributed to a positive opinion from delegates when considering a candidate for admission. The second part presented a list of 16 linguistic criteria (e.g. sense consistency with original message, logical cohesion of utterance, completeness of interpretation, correct grammatical usage) and extra-linguistic criteria (e.g. native accent, fluency of delivery and ability to work in a team), and asked respondents to rate the importance of each on a four-point scale from “highly important” to “irrelevant”. In the third section, respondents were invited to add further criteria of their own and comment on their importance.

Respondents rated linguistic criteria more highly than extra-linguistic ones. The most highly rated criterion was sense consistency with original message, followed by logical cohesion and the extra-linguistic criterion reliability. In practice, however, when it came to judging the quality of a specific interpretation, Bühler (1986) observed that users do not normally have
access to the original speech and are therefore likely to judge by “superficial criteria”, such as a pleasant voice. Bühler (1986: 233) formulated the following claim: “The criteria … reflect the requirements of the user as well as fellow interpreter in a (hopefully) well-balanced mixture”.

As Kahane (2000: 2) argued, the value of Bühler’s (1986) study lies above all in its establishment of a series of linguistic and extra-linguistic criteria for assessing an interpretation. This has been adopted with minor modifications by a wide range of scholars, allowing the results of their studies to be compared.

In the following paragraphs we will describe studies which adopted Bühler’s questionnaire-based methodology and her criteria. As mentioned above, the basic aim of all of these studies was to establish the a priori opinions of users with regard to interpreting quality and to the role played by each quality criterion.

Inspired by Bühler’s hypothesis (1986) but reducing the number of criteria from sixteen to eight, Kurz (1989) aimed to compare the expectations of users at a medical conference, and to determine the differences between their expectations and those of the AIIC and CACL professionals. Although her respondents were not interpreting experts, they echoed Bühler’s (1986) results in attributing more importance to linguistic factors than to extra-linguistic ones, and even reflected the same ranking when deciding the relative importance of each criterion.

Intrigued by these results, Kurz (1993) undertook a further study with exactly the same questionnaire in order to compare the expectations of three different user groups with the interpreters in Bühler’s (1986) study. In this case, the users were participants in a medical conference, a conference on quality control, and a Council of Europe meeting. The study revealed some very interesting results. While the interpreters and users agreed on the ranking of the different criteria, the interpreters were shown to be more demanding than the users. Thus, Kurz (1993) ultimately disproved Bühler’s (1986) claim. This led Kurz (1993: 20) to conclude that “the target language receivers or listeners must be seen as an essential element in the process”, since they have expectations arising from the specific interpreting situation. Thus, while applying the same priority ranking, they allocated more or less importance to each criterion depending on whether the situation was oriented more towards content, as in a medical conference, or more towards presentation, as in a political meeting. Different user groups thus have different expectations.

At about the same time, other authors were using the questionnaire method to shed new light on quality research. One example is Marrone (1993), who graded questions progressively from the most general aspects to the most specific in an attempt to determine the role of subjectivity. For her part, Vuorikoski (1993) used various research tools simultaneously, in addition to the questionnaire, in order to protect her results against the potential deficiencies of any single method. This multi-method approach was replicated in Italy by Mack & Cattaruzza (1995), who found that first-time users of interpreting were less demanding in their expectations than frequent users, but at the same time awarded a lower score to the interpretation. Kopczynski (1994) also made an interesting contribution with his analysis of the expectations of users from the dual perspective of those who merely attended a conference and those who participated as speakers.

It may be claimed that all this research paved the way for the key study by Moser (1996). His objective was to ascertain if user expectations remain consistent across different types of conference. Moser’s research (1996) was innovative at the time for the extent and processing of its data—the questionnaire was circulated at a total of 84 conferences and the results subjected to rigorous statistical analysis. The findings showed that, in general, respondents attached maximum importance to fidelity, but not always to the same degree. For instance, terminological accuracy was clearly considered to be more important at technical conferences,
and significantly less so at general conferences. A further finding related to clarity of expression. Moser (1996) compared the expectations of users with different levels of experience of conferences: newcomers, those with little experience, and old hands. Remarkably, the most demanding users with regard to clarity of expression were those with little experience, while women always clearly considered it to be more important than men did (see also Defrancq, Collard, Magnifico & Iglesias Fernández, Chapter 30, in this volume).

As mentioned, the research described up to this point focused on a priori expectations about quality. In this context, Gile (1990) opened up a new line of research by focusing on a posteriori evaluation of interpreting quality, whereby users were asked to evaluate specific interpretations after the event. Although Vuorikoski (1993) and Mack & Cattaruzza (1995) included a posteriori evaluation in their studies, it was not linked in any detail to the source speech, with the result that the authors could not determine which elements of the interpretation had been evaluated positively or negatively by the respondents. In contrast, a study by Pöchhacker (1995) not only incorporated a comparison of source and target speech, but also a comparative analysis of linguistic and extra-linguistic elements. In terms of the relation between source and target text, Kalina (2015: 31) argued that quality assessment should always involve a linguistic comparison of the two. However, she advocated comparing source and target text delivery within a framework that considers all the elements intervening in the interpretation, in order “to make the relationship between external factors, processing conditions and source and target discourse more transparent”.

**External conditioning factors**

One obviously crucial ‘external’ factor is the source speech, which conditions processing at the most basic level. Lamberger-Felber (1998) found that when interpreters receive the transcript of the source speech well in advance, their performance is noticeably better.

A further processing conditioning factor is the speaking rate. In this respect, Barranco-Droege (2015) studied the effect of the speaker’s pauses on the interpreter’s cognitive load. He found that longer pauses at syntactical boundaries may lead to a more complete rendering of the speech, which is compatible with Gile’s Tightrope Hypothesis (2009). For their part, Baigorri-Jalón & Travieso-Rodríguez (2017: 53–59) studied the effect of speed and other external variables that have changed in UN interpreters’ working conditions in recent years. The main results of their study (2017: 65) suggest that extremely high speed of speech delivery is the most salient feature affecting interpreting performance. As high speed tends to jeopardise the interpreter’s cognition, it has a significant impact on the quality of output.

However, when considering user evaluation, we should also pay attention to the filters through which each user receives and consequently evaluates the interpretation, including his or her level of comprehension. For instance, Reithofer (2013) compared user comprehension of a speech by a speaker whose mother tongue was not English with that of its interpretation into German by a professional interpreter. The results showed that the interpretation led to a better cognitive end-result in the audience than the source speech in non-native English (see Albl-Mikasa, Chapter 39, in this volume). This suggests that users’ evaluations about complete or correct transmission are open to question unless their level of comprehension is taken into account.

The need to give proper consideration to contextual elements led Collados Ais (1998) to carry out the first experimental study of quality in interpreting. Her aim was to determine how consistently users evaluate a given interpretation by comparison with their stated expectations about quality criteria. Two groups of subjects, interpreting experts and university lecturers in
law, were asked to assess one out of three interpretations of a speech on law given at a real conference. Two of the interpretations were experimental versions and the third functioned as a control. The experimental versions were manipulated to focus on a particular criterion—intonation and fidelity (the latter comprising both accuracy and completeness).

Collados Aís (1998) found that although users consider extra-linguistic criteria (such as intonation) less important a priori, the use of a poor intonation by interpreters does indeed impact their assessment of overall quality and of linguistic aspects (such as fidelity). This research approach, followed up by associates in her research group ECIS (Evaluación de la Calidad en Interpretación Simultánea) (Collados Aís et al. 2007), was acknowledged by Pöchhacker (2013: 33) as one of “the major milestones in the development of empirical research on quality in interpreting” and dubbed the “Granada paradigm”. Indeed, in partnership with ECIS, Franz Pöchhacker conducted a research project entitled QuaSI (Quality in Simultaneous Interpreting), adopting this same empirical research line.

In sum, whether the method is observational or experimental, any attempt to define quality in interpreting needs to take all the factors that intervene in any given interpretation into account. As Behr (2013: 96) maintained, the study of quality evaluation at events should start from a common research ground and take into account what she calls the “evaluation framework”, in which the setting, disposition of participants and criteria consensus are interlinked.

We would add to this that the better we define all the elements involved in the evaluation of quality, the closer we will be to achieving a stable list of evaluation criteria.

**IT tools: expanding methods and samples in empirical quality research**

Research into expectations and evaluation has continued into the twenty-first century, but with one major difference: the availability of technology permitting innovative design of studies of interpreting quality (Pöchhacker 2010). The benefits of IT include the use of highly sophisticated speech and sound analysis programmes with ample possibilities for experimental manipulation and providing access to much bigger samples.

There has also been a dramatic increase in the size of corpora (see Bendazzoli, Chapter 32, in this volume). Regarding sample size, a study by Chiaro & Nocella (2004) was the first web-based survey on interpreting quality. The authors sent an electronic questionnaire to 1,000 conference interpreters and yielded a total of 286 responses. The questionnaire adopted nine of Bühler’s (1986) criteria and requested that survey participants rank the importance of the nine criteria listed. The results were in line with Bühler (1986) and many of the other studies we have described: the two content-related criteria of sense consistency with original and logical cohesion were ranked highest, while the two delivery-related criteria of pleasant voice and native accent were considered least important.

IT tools have likewise facilitated a closer look at the criteria definitions that have been broadly accepted by researchers in empirical studies since Bühler introduced them in 1986. This has involved a more precise measurement of the impact on quality of criteria such as fluency (Pradas Macías 2003; Rennert 2010), as well as aspects that determine the quality criteria themselves, sometimes known as sub-criteria, such as speed (Barranco-Droege 2015) and pauses (Pradas Macías 2015).

Rennert (2010) used the audio software PRAAT to explore the relationship between prosodic features specific to simultaneous interpreting and the listeners’ perception of fluency and accuracy, as well as their comprehension of the source speech. She asked 47 interpreters and 40 users to listen to a 20-minute lecture in German and its interpretation into English under two conditions: the actual interpretation and a version manipulated to enhance its fluency. The
analysis of the prosodic features of the two conditions showed differences regarding the temporal organisation of speech, disfluencies and pitch register. Based on participants’ answers to comprehension and rating questions, Rennert (2010) concluded that interpreting-specific prosodic features affect the perception of fluency, which in turn affects the perception of accuracy.

These results are in line with Pradas Macías (2015), who explored user awareness of *sui generis* pause patterns in simultaneous interpreting. Seeking to establish which lexical categories in a source text are more likely to prompt interpreters to insert pauses, she found that “pauses appeared more frequently and more consistently following the rendering of nouns”. She then used the sound-processing program Audacity in an experiment to determine whether deviations from this pattern could be detected by listeners and whether the deviations influenced their evaluation of interpreting quality. The study showed that “when the pauses following ST-noun phrases were removed from the TT, the participants described the pause pattern as unexpected” (Pradas Macías 2015: 67).

Also using Audacity, Barranco-Droege (2015) prepared study material based on an excerpt from Barack Obama’s inaugural address as President of the United States in 2009. The material consisted of the original audio recording (Clip 1) and two manipulated versions, one time-expanded (Clip 2) and the other pause-expanded (Clip 3). The aim was to determine how well the subjects were able to detect time and pause expansion, and whether these expansions altered their perception of any other features of the speech (Barranco-Droege 2015: 93); moreover, the author hoped to establish whether subjects’ perceptions change depending on which audio they listened to. But he also set up an instrumental objective, namely, to test whether content analysis can outperform ordinal-scale questionnaires in terms of validity and reliability when probing the perception of time-manipulated speech.

With regard to the non-instrumental aspects, “the results suggest that time expansion may have confounding effects on other variables and should be used with care, while pause expansion does not seem to present such a risk, although it may be detected more easily by the subjects” (Barranco-Droege 2015: 114). But the “most useful findings”, as the author himself suggests (Barranco-Droege 2015: 115), “are probably on the methodological level” since they “raise the question of whether survey-based studies on user perception need to build more strongly on shared and negotiated meaning to put results on a firmer footing”.

To sum up, we can see how increasingly sophisticated computer programs have enabled researchers to measure and manipulate different aspects for investigation.

**Sociological approaches to research into quality in conference interpreting (and their long exclusion)**

The beginning of the twenty-first century also saw the first investigations into conference interpreting from a sociological perspective. This second part of the contribution will look at conference interpreting quality from a sociological angle. Quality can fundamentally be defined as a social construct generated by a particular time and its people, the context and culture in which it is evaluated. From this perspective, it is therefore not something that can ever be pinned down and definitively measured (Grbić 2008; Zwischenberger 2013). This is where the sociological concept of norms and their changing nature come in, linked as they are to quality via the roles the interpreter has to play according to the expectations held by others and the interpreters themselves. Norms may ultimately account for the variability found not only in the results of empirical studies into quality but also the theoretical frameworks and methods underlying these studies. Quality may then be considered a normative social construct (Zwischenberger 2013).
This section will start by trying to explain the reasons for the neglect of the social dimensions of interpreting quality. As early as the 1970s, at a time when (empirical) research on interpreting was still in its infancy, the sociologist Bruce Anderson (1976) pointed out the need for conference interpreting research to go social:

"translation also occurs in social situations—situations amenable to sociological analysis. In any such setting the role played by the interpreter is likely to exert considerable influence on the evolution of group structure and on the outcome of the interaction."

(1976: 209; emphasis in the original)

This call, however, went completely unheeded at the time. Back then, the interpretive theory developed by Seleskovitch (1968) put the transfer of sense from the source speech to the target speech centre stage, while cognitive research tried to shed light on the internal workings of the simultaneous interpreter’s ‘black box’ and/or mind. The latter shift started with work coming from outside interpreting studies by, for example, Goldman-Eisler (1967) or Gerver ([1969] 2002, 1975). This research was later taken on by interpreting studies scholars such as Moser (1978) and Gile (1985) and became the dominant research focus for conference interpreting for the years and decades to come.

Pöchhacker (1994) was the first to integrate the social context of simultaneous conference interpreting into his investigations, and ultimately also into the quality of the output produced by interpreters. He conceived a conference as a “hypertext”—an overarching text consisting of a number of individual texts.

However, the clear breakthrough for studying conference interpreting from a sociological perspective and with a focus on role-related issues was relatively recent (Angelelli 2004; Bartłomiejczyk 2017; Beaton 2007; Diriker 2004; Monacelli 2009). These studies, while not dedicated directly to the issue of quality, at least touched upon it.

**Explaining variability in the expectations and evaluations of quality**

The shift towards sociological studies in quality research, explicitly bringing together social roles and quality, was advanced relatively recently by Zwischenberger (2013, 2015a), against the backdrop of two large-scale web-based questionnaires on quality and role conducted among the members of AIIC and the German Association of Conference interpreters (V KD). Zwischenberger (2013, 2015a) was also the first to undertake a major conceptual analysis of the ‘social role’ in interpreting and to dissect the concept by bringing to light the various concepts associated with and constituting it. Despite the various role-theoretical perspectives taken by scholars, a consensus has been reached regarding the definition of the social role. Social roles may be conceived of as a bundle of normative behavioural expectations that are directed at the occupants of social positions (Joas 1991: 146; Peuckert 2006: 242). It is either single-role others such as the boothmate or reference groups (Turner 1956) like the listeners directing their expectations towards the occupants of social positions that mould and co-construct the various roles the conference interpreter is expected to play from their social position. Thus, ‘the role of the interpreter’ usually used in the singular and unspecified, does not exist since, for example, the social position of a simultaneous conference interpreter at a specific conference requires her or him to play as many roles as there are role others and reference groups. Furthermore, interpreters hold certain expectations themselves. Single-role others such as the booth colleague, the teamleader or the conference organiser, as well as reference groups such as the listeners at the conference, all bring their various expectations on quality. This explains
the variability in quality evaluations, both between interpreters and users and within these 
groups, found in the various questionnaires and experimental studies described above.

Roles and their related expectations are not value-free but rather loaded with norms. This 
calls for the concept of norms to be integrated into the quality equation, since any evaluation is 
necessarily based on how well certain (norm-loaded) expectations are fulfilled.

**Norms as determining factors**

In neither translation studies nor sociology, where the norm concept is used as a core concept, 
is there a consensual and uniform definition of norms. Very often, norms in both sociology and 
translation studies are deduced from the values shared by a certain community:

> Norms have long been regarded as the translation of general values or ideas shared by a 
group—as to what is conventionally right and wrong, adequate and inadequate—into per-
formance instructions appropriate for and applicable to particular situations, specifying 
what is prescribed and forbidden, as well as what is tolerated and permitted in a certain 
behavioural dimension. (Toury 1999: 14; our emphasis)

The idea of studying social norms in conference interpreting was raised as early as the late 
1980s by Shlesinger (1989) but was deemed to be an insurmountable task due to the lack of 
readily available representative electronic corpora in those days. Influenced by Descriptive 
Translation Studies (Toury 1995), Shlesinger (1989) focused very much on using “textual 
sources”, i.e. the interpretations themselves, to find empirical proof for norms. Harris (1990), 
replying to Shlesinger (1989), challenged her conclusions that it was “presumptuous to specu-
late as to the existence of norms in interpretation” (Shlesinger 1989: 111). Harris (1990) listed 
a number of norms applying to Western conference interpreting practice, including in par-
ticular that the interpreter speaks in the first person, the norm of working into one’s A language, 
or the “true interpreter” norm that

> requires that people who speak on behalf of others, interpreters among them, re-express the 
original speakers’ ideas and the manner of expressing them as accurately as possible and 
without significant omissions, and not mix them up with their own ideas and expressions. 
(Harris 1990: 118)

Here, Harris (1990) hinted at norms which can be found in “extratextual sources” (Toury 
1995: 65), that is in any metatexts on interpreting such as those generated by professional asso-
ciations, interpreters themselves in their blogs, memoirs, or in questionnaires or interviews. 
Gile explicitly argued that research on norms should primarily rely upon extratextual sources 
for reasons of efficiency:

> I believe that research about norms does not necessarily have to rely on large speech corpora. In 
the field of interpreting, such research is probably more efficiently done by asking interpreters 
about norms, by reading didactic, descriptive and narrative texts about interpreting. 
(1999: 100)

According to the sociologist Eichner (1981: 43ff.), we may assume the actual existence of a 
social norm if it appears with a high frequency and in the imperative form in various metatexts
generated by a community. The sociologist Eichner (1981: 46–75) explicitly identified behavioural predications as a further indicator for the existence of a social norm. Judgements about whether a certain behaviour, product, or person is good, correct, appropriate, etc. are informed by normative expectations. Thus, any sort of quality feedback is normative in its very essence. Further indicators for the existence of norms named by this sociologist are sanctions and regularities in social behaviour (Eichner 1981: 15–38).

Since the late 1980s and early 1990s, a relatively small number of scholars have undertaken research into norms in conference interpreting. Schjoldager ([1995] 2002) focused on norms in the didactic setting by analysing the transcribed output of interpreting students. Diriker (2003, 2004) focused on the metadiscourses on conference interpreters and interpreting generated by interpreting scholars, trainers and practitioners, and the media. Marzocchi (2005) provided a theoretical discussion of norms and their implications for ethics in the field of conference interpreting. Duflou (2009), in addition to discussing methodological aspects of norm research into conference interpreting, also discussed the “first person norm” used in conference interpreting. Tiselius (2010) linked the issue of norms and habitus to the conference interpreters’ expertise. She analysed the metadiscourse generated by EU institutions and conducted two focus group interviews among a dozen interpreters working in the Swedish booth in EU institutions. Lenglet (2015) presented a survey conducted among scholars, trainers and students of conference interpreting on translational norms when face-threatening acts occur as a result of the source speaker and their speech.

**Excursus: interpreting norms as the result of standardisation**

The technical norms issued by standardisation bodies such as the various ISO norms or national norms such as DIN norms for Germany are not the focus of this contribution, but they will be briefly examined here. Even though interpreting quality is not dealt with per se in these norms, such norms are implemented in order to guarantee the best working conditions and in turn facilitate high-quality interpreting. These technical norms are also social norms in the wider sense in as much as they are the result of social negotiation, albeit by a clearly definable group of people that form part of a certain standardisation committee.

In recent years, there has been a great deal of standardisation in the field of conference interpreting in particular. Norms such as the ISO 20108: Simultaneous Interpretation—Audio Transmission Quality issued in 2017 (PCS 2020a), ISO 20109: Simultaneous Interpretation—Equipment Requirements (PCS 2020b), ISO 2603 for permanent interpreting booths (PCS 2020c) and ISO 4043 for transportable/mobile booths (PCS 2020d), all issued in 2016, all regulate the technical premises for successful simultaneous interpreting and are directed at providers of simultaneous interpreting equipment and conference organisers (PCS 2020c, 2020d). Unlike these standards, DIN 2347—Conference Interpreting, issued in 2017, is directed not at agencies but at individual conference interpreters, who can attain certification and work according to this standard. The standard stipulates how a conference interpreting assignment should be managed, ranging from its preparation to its postprocessing (ATICOM 2020). Standards are usually intended to ensure satisfactory minimum requirements for quality are fulfilled.

Apart from technical norms focusing specifically on interpreting, there is also the broad ISO 8402 on quality assurance and quality management in general. Quality according to this norm is the relation between the achieved and expected property of a unit. This norm was applied to interpreting by Mack (2002) and suggests that quality can indeed be measured and is objectively identifiable. This conception of quality following a standardised norm is in stark contrast to the notion of quality as a social construct (Zwischenberger 2013: 17f.).
Linking norms and quality

The first interpreting studies scholar to explicitly link social norms (understood here in the narrower sense) to quality was Garzone (2002). She traces the variability in expectations and specific evaluations of interpreting quality back to norms:

The concept of quality in interpretation can thus be defined as a construct embodying the norms which are deemed appropriate to guarantee the intrinsic and extrinsic properties considered ideal for an interpretation performance in a given social, cultural and historical situation.

(Garzone 2002: 110)

Garzone (2002) took up the tenets of social constructionism, taking Toury’s (1995) categorisation of norms and applying them to the field of conference interpreting.

Zwischenberger (2015a: 231) described norms as “mediating constructs” between role(s) and quality, since expectations are loaded with norms, as explained above. Zwischenberger (2013, 2015b), in her large-scale survey conducted among 811 conference interpreters, focused on “extratextual sources” of norms. She attributed norm-setting power to professional associations and to AIIC in particular. The association’s website and its bulletins containing metatexts on interpreting are a great extratextual source for extracting norms (Diriker 2004). These are specifically social norms issued by an institution. By explicitly and formally articulating social norms in their various metatexts, institutions “give order to social relations, reduce flexibility and variability in behaviour, and restrict the possibilities of a one-sided pursuit of self-interest or drives” (March & Olsen 2008: 4). Furthermore, the norms expressed by an institution always constitute the structured pursuit of aims or an end (McCormick 1998: 333). It is AIIC’s aim to ensure the profession’s prosperity and survival.

One norm which can be found frequently in AIIC’s metatexts on what makes good conference interpreting and/or a good conference interpreter is absolute loyalty regarding the speaker and their speech:

Professional conference interpreters speak in the first person, on behalf of the speaker, and, as such, their primary loyalty is always owed to the speaker and to the communicative intent that the speaker wishes to realise, whatever the speaker’s position or point of view.

(AIIC 2016; emphasis in the original)

Far from being a description of the task of a conference interpreter, this is highly prescriptive, defining the basis required to deliver high-quality interpreting and to be worthy of being considered a professional conference interpreter. The following example explicitly uses the imperative form in its description:

As conference interpreting is a professional communication service, quality in interpreting is a function of communication. It is your job to communicate the speaker’s intended messages as accurately, faithfully, and completely as possible. At the same time, make it your own speech, and be clear and lively in your delivery. A conference interpreter is a communication professional who needs to be a good public speaker, so make your interpretation fluent, expressive, and communicative.

(AIIC 2016; emphasis in the original)
Again, absolute loyalty to the speaker and their intended message is demanded and must be achieved in a complete, fluent and lively way. There is, however, also a contradiction when absolute loyalty and a mimetic representation are demanded, on the one hand, while, on the other hand, the interpreter is supposed to make the speech their own. What if the speaker is not at all clear, lively and fluent? Is the interpreter in the name of loyalty and accuracy supposed to not fully represent this aspect of the speaker? There are also other contradictions in these metatexts, which explains the often highly contradictory and vague nature of norms in the field of conference interpreting (Zwischenberger 2015a: 239ff.).

In her empirical study, Zwischenberger (2013) asked professionals to rate a set of quality criteria, in the spirit of a more traditional way of investigating quality, and asked respondents to directly evaluate the quality of a sample of an interpreting performance. Furthermore, the main part of the questionnaire consisted of various role-related items for which respondents were asked to state their degree of (dis-)agreement. The questionnaire results suggest that conference interpreters strongly cling onto the “supernorm of interpreters as conduits” (Zwischenberger 2015b), which prescribes that interpreters should channel the message as intended by the speaker in a complete, emotionally detached and fluent manner. Fluency may be regarded not just as a prosodic feature but is also to be understood in Venuti’s (1995) sense that the interpretation sounds like an original.

According to Zwischenberger (2013), this expectation that interpreters should act as conduits may be understood as a supernorm in two ways: it is a supernorm due to its normative strength, but also in its role as metanorm consisting of and covering multiple subnorms, such as sense consistency with the original, completeness, fluency of delivery, etc. In any event, it is a fiction upon which quality is defined as a normative construct, but one nevertheless upheld by the profession as a “useful ideal” (Gile 2017: 241) because it is linked to the vital issue of trust. This trust is based on the assumption that (conference) interpreting is a powerless and thus innocent act. All interpreters do is channel an original message in a complete and undistorted manner. The authority, and thus power, of the original speaker remain completely untouched.

Conclusion

Quality research in conference interpreting covers an amazing timespan of 35 years since Bühler’s (1986) pioneering work. The first studies among both interpreter experts and users used a seemingly stable set of criteria rated as a priori expectations and directed mainly at the interpreting product. Later studies, although fewer in number, used an a posteriori approach.

Furthermore, there has been a widening to include potential factors and/or conditions influencing quality, such as the source speech and whether it is made available to interpreters in advance or not, the speaker’s delivery rate, the use of English as a lingua franca, etc. All these external conditioning or situational factors certainly have a bearing on comprehension, cognitive processing and capacity management.

The twenty-first century saw huge technological and digital advances that also had an influence on the methodological possibilities by providing more sophisticated ways to conduct research into interpreting quality and to generate larger samples. Similarly, only at the beginning of the twenty-first century was there a beginning of sociological research into conference interpreting and interpreting quality, despite an early call. The reasons for this are historically motivated.

Explicitly sociological approaches to studying interpreting quality that focus on describing quality as a social construct and linking it directly to roles and norms are still rare. It is, however, precisely such sociological research into conference interpreting quality via the concepts
of social roles and norms that enables us to account for the variability in quality expectations, and ultimately that of evaluations of quality.

In general, it is of importance to increasingly widen quality research towards the long-neglected non-verbal, situational, and ideological factors that influence and co-construct quality expectations and ultimately also its evaluations and assessments. This widening of the scope of quality research may be necessary not only in order to be able to pin down the elusive nature of quality, but also to more fully understand the various factors co-constructing the social construct, or rather the various existing social constructs, of quality.

**Further reading**


Zwischenberger, Cornelia 2015. Simultaneous conference interpreting and a supernorm that governs it all. *Meta* 60 (1), 90–111.

**References**


Quality and norms


Zwischenberger, Cornelia 2015b. Simultaneous conference interpreting and a supernorm that governs it all. *Meta* 60 (1), 90–111.