

Assessing spoken-language educational interpreting: Measuring up and measuring right

Lenelle Foster

Language Centre, Stellenbosch University, South Africa
E-mail: lfoster@sun.ac.za

Adriaan Cupido

Language Centre, Stellenbosch University, South Africa
E-mail: adriaancupido@sun.ac.za

Abstract

This article, primarily, presents a critical evaluation of the development and refinement of the assessment instrument used to assess formally the spoken-language educational interpreters at Stellenbosch University (SU). Research on interpreting quality has tended to produce varying perspectives on what quality might entail (cf. Pöchhacker 1994, 2001; Kurz 2001; Kalina 2002; Pradas Marcías 2006; Grbić 2008; Moser-Mercer 2008; Alonso Bacigalupe 2013). Consequently, there is no ready-made, universally accepted or applicable mechanism for assessing quality. The need for both an effective assessment instrument and regular assessments at SU is driven by two factors: Firstly, a link exists between the quality of the service provided and the extent to which that service remains sustainable. Plainly put, if the educational interpreting service wishes to remain viable, the quality of the interpreting product needs to be more than merely acceptable. Secondly, and more important, educational interpreters play an integral role in students' learning experience at SU by relaying the content of lectures. Interpreting quality could potentially have serious ramifications for students, and therefore quality assessment is imperative. Two assessment formats are used within the interpreting service, each with a different focus. The development and refinement of the assessment instrument for formal assessments discussed in this article have been ongoing since 2011. The main aim has been to devise an instrument that could be used to assess spoken-language interpreting in the university classroom. Complicating factors have included the various ways in which communication occurs in the classroom and the different sociocultural backgrounds and levels of linguistic proficiency of users. The secondary focus is on the nascent system of peer assessment. This system and the various incarnations of the peer assessment instrument are discussed. Linkages (and the lack thereof) between the two systems are briefly described.

Keywords: assessment; assessment instrument; educational interpreting; peer assessment; quality

1. Introduction

Spoken-language educational interpreting at Stellenbosch University (SU) received institutional sanction late in 2011 and, as of 2012, simultaneous educational interpreting formed part of the University's attempts at managing learning and teaching in a multilingual environment. Working into either Afrikaans or English, the educational interpreter acts as a communication facilitator between lecturer and student during formal lectures in cases where the student is either not proficient enough in the language of instruction or would prefer to receive tuition in the other language. The interpreting provided in the lecture venue should be of such a standard that it helps to ensure effective communication between lecturer and student.

In order both to safeguard and improve the quality of the interpreting service provided to students, a system of regular assessments has been instituted. Regular assessment is not done because interpreting quality and performance have been conflated – with factors such as “individual qualifications and skills, [...] professional ethics and the conditions under which [interpreters ...] work” being ignored in favour of the “discourse produced in real time” (Pöchhacker 2013:33) – but because measuring interpreting quality by means of assessment provides an opportunity to identify strengths and weaknesses and devise strategies for improvement on the part of the interpreter.

Two types of assessment are undertaken as a matter of course¹ by interpreters working at SU. Formal assessments, using an assessment grid developed at the University (described in section 2), are done by a team of senior interpreters with a record of exceptional interpreting over a period of at least three years. As detailed in section 2.1, this assessment grid is regularly revised to ensure that it can deal with changing demands resulting from shifts in the language policy at SU and in the higher education landscape in South Africa. The formal assessment process is described in section 3. The system of coordinated peer-to-peer feedback is discussed in section 4, as it, too, provides the interpreters with information on where they could improve their performance, although it has a greater emphasis on tracking progress on specific objectives set by the interpreters themselves.

2. Assessment instrument

As Kasandrinou (2010:195) and Wallmach (2013:394) argue, evaluation is a tool for collecting information. This information should not only signify that “procedures are being [...] performed effectively and that the desired levels of quality” are being achieved (Kasandrinou 2010:194) or inspire “confidence in the process and the final product” (Kasandrinou 2010:194). It should be used to “measure the gap between standards and actual practice, and work out ways to close the gap” (Wallmach 2013:394). The extent of this gap can only be determined, however, if the instrument used for measurement purposes is measuring the right gap. A change in institutional culture (or interpreting conditions) will have an impact on the way interpreting is conducted and on the way clients perceive their needs. This may well lead to interpreters using different strategies to meet these needs and possibly even to changes in what interpreters view as ‘successful’ interpreting. Kalina (2005:28) has emphasised the “ephemeral”, “irregular” nature of interpreting and noted how dependent it is “on external factors”. If this holds true for

¹ External assessments were instituted at SU early in 2017; they are not discussed in this article.

interpreting itself, it should also be true of the mechanisms used to evaluate the quality of that interpreting.²

The assessment instrument used at SU is not static. It is reviewed at least once per semester to consider changes, based on the practical experiences of the assessment team. Some of the changes are cosmetic, for example ensuring that text is wrapped in spreadsheet cells. As is explained in section 2.1, other changes are more significant and reflect changing ideas of what should be assessed and how it should be assessed. The ‘why’ is fixed – to identify what interpreters should focus on to improve their performance in the classroom.

2.1. Development

Rather than design its own assessment instrument *ex nihilo*, the interpreting service relied on the examination grid used by the South African Translators’ Institute (SATI) for the accreditation of conference interpreters, since it was a tried-and-tested assessment instrument used by a reputable professional body. The SATI examination grid was modified by a staff member of the SU Language Centre for use in one of the pilot projects conducted at SU in 2011 because of the difference in purpose (and application) of the assessment tool. The SATI grid was designed for and is used in determining whether a candidate passes the accreditation examination – a context in which a strong ‘yea’ or ‘nay’ is required and (detailed) feedback is unnecessary. Clearly a different approach was needed when assessing performance with a view to improving that performance.³ The modified SATI grid (initially used in 2011) had three main sections corresponding with those used by Ibrahim-González (2013:229): content, form and interpreting competency, with each section being marked out of 10.⁴

Subsequently, in March 2014, three major changes were made to the modified SATI grid of 2011. The first was the inclusion of a section on an aspect that was, at the time, described as the interpreter’s educational role, reflecting to what extent the interpreter conveyed questions from students (and responses from the lecturer), how well the interpreter conveyed the lecturer’s teaching style and character, and what the interpreter’s behaviour was toward lecturer and students. Secondly, subsections were added to each of the four sections. The modified SATI grid had contained keywords and -phrases, indicating what aspects should be considered when awarding a mark in a particular section. Interpreting competency would, for example, include listening skills, intonation, paraphrasing, lag and pronunciation. These keywords and -phrases made up the various subsections of the new in-house grid and each contributed 10 marks to the assessment total. Instead of the four sections each counting 10 marks (total score of 40), the various subsections would hence each be marked out of 10, with the total out of 340. (Table 1 summarises the differences between the 2011 and 2014 grids.) The third and, arguably, most significant change was the insertion of a separate section on context.

² An example: Changes in the language policy at SU have led to a shift in usage patterns, with an increase in the amount of interpreting from English to Afrikaans. Because of the differences between the two languages, a separate instrument was in development at the time of writing, with a focus on language issues pertinent to Afrikaans interpretation.

³ Bearing in mind Moser-Mercer’s (2008:146) appeal that quality “be broken down into more tangible components”.

⁴ For a more detailed discussion, see Foster (2014:93-94).

Table 1: Differences between modified SATI grid and first iteration of in-house grid

	Modified SATI grid (2011)	In-house grid (2014)
Number of sections	3	4
Section on interpreter's educational role	No	Yes
Record of contextual factors	No	Yes
Subsections	No	Yes (10 marks each)
Total score	40	340

As a number of interpreting theorists (notably Garzone 2002:117, Kalina 2002:124; 2005:34 and Pöchhacker 2004:156) have indicated, interpreting quality is influenced very strongly by the circumstances in which the interpreting takes place. The quality of the interpreting product cannot be evaluated solely on the basis of the output provided by the interpreter. The assessment process should also take other factors into account. In the new in-house grid, these factors were described as the lecturer's delivery speed and fluency, how clearly the lecturer spoke, the noise level in the venue, and how audible student questions and comments were.⁵ These five contextual factors could be graded from one (not particularly fast, noisy, audible etc.) to five (very fast, noisy, audible etc.), with an additional space provided for comments.

The intention was twofold. Firstly, highlighting the context in this way should help to ensure fair assessments because the assessors are forced to acknowledge the context and its impact on interpreting performance, rather than awarding marks based on their opinion of what the ideal interpreted text should be. Secondly, by indicating where a particular interpreter was performing badly in a particular context, the assessment could be used to guide training and interpreter management, thereby ensuring a better match between lecturer and interpreter, and consequently improving the quality of the service.

Another round of significant changes to the assessment grid occurred in 2015, when the weighting of the various subsections was altered and the subcategories themselves were revised. The revisions included the removal or reconstitution of some subsections. *Voice quality* and *Intonation*, for example, were grouped together and *Interpreter has prepared* was removed, as preparedness (or the lack thereof) would already be indicated by the interpreter's performance in a number of other subsections.

Changes to the weighting of the various subsections has proved to be a source of debate in the interpreting service, since the weighting reflects the dominant discourse on what constitutes 'good' educational interpreting.⁶ The greatest weight – 25 marks – was given to *Message accuracy and cohesion* (indicating the extent of cohesive, error-free interpretation) and *Equivalent meaning conveyed fully* (indicating the degree of loss of information). The smallest

⁵ In one of the pilot studies conducted in 2011, both assessors and interpreters had to complete an additional form, grading various contextual aspects, including the five listed here. These contextual aspects were, however, ancillary and not part of the formal assessment.

⁶ There is also the additional dilemma – highlighted by Pradas Macías (2006:37) – “that individual parameters do not necessarily add up to a general effect”. A listener, especially an experienced user of interpreting (Pradas Macías 2006:38), may rate overall interpreting quality as being “of high quality” “in spite of specific shortcomings” (Pradas Macías 2006:37).

scores – five marks – were given to aspects such as *Behaviour toward users, Intonation and voice quality, Equipment management, Suitable register* and *Accent*. Intermediate weights of 10 or 15 were assigned to subsections such as *Complete and coherent sentences* (10), *Subject terminology* (10), *Conveying class experience, humour, idioms* (15) and *Pronunciation and general clarity* (15). Further alterations have been made on a regular basis and the format of the assessment instrument at the time of writing is discussed below.

2.2. Present format (2017)

Considering the importance of context in determining interpreting quality, the first section of the 2017 assessment form covers several contextual factors discussed in the preceding subsection: some information on the lecturer’s delivery (rapidity, fluency and audibility), some information on the audibility of questions and remarks by students, and an indication of noise levels in the classroom. This latter point not only includes noise levels as generated by students, but also sounds from outside the classroom. The assessor also has to provide information on where the assessment occurred – module and lecturer – and what the interpreting conditions were during the assessment. When being assessed, each interpreter has to complete a context sheet, noting how often they interpret the module and whether there is preparatory material available. They may also wish to bring other matters to the assessor’s attention. At the end of the assessment, the assessor collects the context sheet and includes the information in their report, under the *Interpreting conditions* heading. None of this contextual information should influence the marks awarded – context is not an excuse. Rather, it is an explanation of the interpreter’s behaviour and may even provide an opportunity to highlight the interpreter’s “adaptation”, “flexibility” and “problem-solving ability” (Alonso Bacigalupe 2013:28). A noisy environment may lead to information loss because the interpreter cannot hear what is being said, for example. The contextual information reflects the noisy environment and, as a result, the interpreter may be asked to practise interpreting in a noisy environment and to develop coping strategies, rather than being made to do general exercises aimed at preventing information loss. Figure 1 shows the first section of the 2017 assessment form and its focus on context.

Educational interpreting: Assessment								
Interpreter:				Lecturer:			Date:	
Assessor:				Module:				

Class context	Min					Max		Remarks:
	1	2	3	4	5	N/A		
Lecturer: How fast?								
How fluent?								
How audible?								
Students: How audible?								
Classroom: How noisy?								

Content	0	0%	
Interpreting competency	0	0%	
Language use	0	0%	

Interpreting conditions

Figure 1: Contextualisation

The next section deals with content (see Figure 2). The arrangement of the subsections according to total marks is intended to guide assessors as to relative importance. This section is similar to the version discussed in the preceding subsection, the major change being the inclusion of *Conveying class experience* and *Conveying interaction between lecturer and*

students. Previously, these were part of the fourth section on the interpreter's educational role. This section was removed from the 2017 version, as it had been argued that the interpreter's role cannot be neatly distilled and separated from other aspects related to interpreting performance. Rather, the interpreter's role in the educational context is fulfilled in a multitude of ways, all of them expressed through the content included or omitted, the way that content is conveyed (competency) and the language used to do so (cf. Kotzé 2014, 2016).

Content			
Message accuracy and cohesion		25	Remarks:
Equivalent meaning conveyed fully		25	
Ability to deverbilise message		15	
Detail (names, dates, figures, acronyms etc.)		15	
Conveying class experience (humour, idioms, lecturer's style/character)		15	
Conveying interaction between lecturer and students		10	
Tidying and repackaging message where necessary		10	

Figure 2: Content

The total marks awarded for each section corresponds with the rankings identified by Kurz (2001:406), with *Content* having a total of 115, *Interpreting competency* 85 and *Language* 55 marks.⁷ Interpreting competency (see Figure 3) covers a number of aspects which could either distract students – inadequate breath control or incorrect equipment management – or hinder their ability to follow the lecture and recall content at a later stage. If students using the service are unable to comprehend or recall lecture content,⁸ then the service has failed in its primary task of making lectures accessible to students who do not understand the language of instruction.

Interpreting competency			
Fluency of interpreting product (hesitations, silences etc.)		15	Remarks:
Pronunciation and general clarity		15	
Lag		15	
Use of coping techniques		10	
Error correction		5	
Intonation and voice quality		5	
Breath control and volume		5	
Confidence and credibility		5	
Managing equipment		5	
Teamwork / "booth behaviour"		5	

Figure 3: Interpreting competency

Placement as the third section (see Figure 4) does not mean that language use is not important, but rather that language errors are considered less important than omitting information or interpreting in a halting fashion. This position has its limitations, however. One may well ask to what extent content can be conveyed accurately if incorrect terminology is used, for instance.

⁷ Interpreters receive a mark for each of the three sections, not an average mark. Consequently a good command of the target language should not 'make up' for inaccurate interpretation when the interpreter is assessed.

⁸ Work by Shlesinger (cited in Pöschhacker 2001:419), Collados Aís (2002:335), Ahrens (2005:2), Bartłomiejczyk (2007:249) and Holub (2010:124-125) has indicated that intonation and delivery fluency could influence both listener comprehension and recall, and listener confidence in the accuracy of the interpretation.

Language use			
Complete and coherent sentences		10	Remarks:
Vocabulary	Subject terminology (knowledge & fluency of use)	10	
	General vocabulary	5	
	Academic (conjunctions, operators)	5	
Suitable register		5	
Grammar	Syntax	5	
	Tense	5	
	Concord	5	
	Prepositions	5	

Figure 4: Language use

To help the assessors account for the various aspects they need to balance while undertaking an assessment, an area for remarks on overall impression has also been added. This enables the assessor to comment on the interpreting product as a whole, as there may be some aspects that could detract from the overall user experience without being adequately reflected in the marks awarded.

3. Assessment process

3.1. Current process

Prior to the second semester of 2015, assessments had been conducted by two senior interpreters and/or the head of the interpreting service on a rotational basis so that every interpreter would be assessed twice a year. However, rapid growth – both in the number of interpreted lectures and the number of interpreters on duty – by the interpreting service meant that it was well-nigh impossible logistically to conduct regular formal assessments. To evaluate the work of a team of 27 interpreters, the pool of assessors was expanded to five senior interpreters. To facilitate systematic reporting on the quality of interpreting in as many interpreted modules as possible, each interpreter would be assessed once per term by one of the five assessors.

Since 2015, the size of the interpreting team has stabilised at 25 interpreters. The assessment team consists of nine senior interpreters with an average of five years' experience in interpreting.⁹ Each interpreter is assessed twice per term by two different assessors according to a list drawn up by the assessment coordinator. The lists are compiled in such a way that – barring serious logistical mishaps – each interpreter will be assessed by each of the assessors during the course of the academic year.¹⁰ Figure 5 provides an extract from the assessment list for the second term of 2017. It does not contain any proposed dates or subjects, as these are determined by the assessors themselves.

⁹ Assessor A = 5,5 years; B = 4,5 years; C = 6 years; D = 3,5 years; E = 3,5 years; F = 5 years; G = 4,5 years; H = 12 years; I = 4 years.

Virtually all the assessors (the exception being Assessor H) gained their experience as interpreters in an educational interpreting environment, but all of them have had exposure to working as conference interpreters.

¹⁰ In unusual cases – such as marked differences between assessment reports, where complaints have been received about a particular interpreter or where an interpreter repeatedly fares badly in assessments – a team consisting of the head of the interpreting service, the assessment coordinator and the most experienced assessor evaluates the performance of the particular interpreter. This is done during a single lecture to ensure fairness.

Interpreter	Assessor 1	Assessor 2
Interpreter A	Assessor A	Assessor D
Interpreter B	Assessor A	Assessor G
Interpreter C	Assessor C	Assessor F
Interpreter D	Assessor D	Assessor C
Interpreter E	Assessor A	Assessor I
Interpreter F	Assessor F	Assessor A
Interpreter G	Assessor G	Assessor D
Interpreter H	Assessor H	Assessor C
Interpreter I	Assessor I	Assessor H

Figure 5: Extract from the assessment list for the second term of 2017

Once the assessment list has been sent to the assessors, each assessor determines their own assessment schedule. Assessments are conducted during normal lectures and are unannounced. On arriving in the classroom, the assessor is expected to give the relevant interpreter a context sheet, which the assessor will later include in the *Interpreting conditions* box on their assessment report.

Performing the assessments during a lecture foregrounds the contextual elements which should be considered when evaluating interpreting quality. However, if, as Kopczynski (1994:88) argues, “context ‘complicates’ the problems of quality” because it “introduces situational variables that might call for different priorities in different situations”, then immersion in the context can produce additional complications for an assessor. The assessor needs to compare the interpreted lecture to the original as it is being delivered, but also needs to note any “situational variables” and evaluate whether the “different priorities” caused by these variables have been accounted for in a satisfactory manner. Discounting the contextual elements would, however, be unacceptable, as the way in which the interpreter balances the linguistic and pragmatic demands of the particular lecture (cf. Kopczynski 1994:87-88) is integral to determining how successful the interpreter was in conveying the source text in the target language.

While conducting the evaluation, the assessor should include as many remarks as possible, explaining why certain marks were given or providing examples of successes or failures, since the intention is not only to provide a snapshot of interpreter performance at a particular time, but also to indicate how and where the interpreter can improve.

Once assessments have been concluded, the assessors send the completed assessment forms to the assessment coordinator, who in turn sends the forms to the relevant interpreters once all the assessments for the term have been concluded – ideally by the end of the term. One of the major flaws in the current system is the lag between the assessment date and the date on which interpreters receive feedback. At the end of the semester, the information from the various assessments is condensed into a report, summarising each interpreter’s performance. These reports are discussed by the interpreter, the assessment coordinator and the head of the interpreting service at the end of each semester, and areas for improvement and possible strategies for doing so are considered.

3.2. Current shortcomings

There are some areas where the current assessment process is lacking and, while their impact both on the smooth running and the fairness and effectiveness of the assessment system has been posited, their impact has not been measured. Some of these deficiencies can be considered either minor or limited in scope, such as the fact that some assessors leave their assessments until the end of term and run into logistical problems in terms of available lecture slots in which to do their assessments. Other deficiencies are more serious: to date, only one calibration exercise has been conducted; a number of the assessors currently on the assessment panel have therefore not been part of a calibration exercise. This means that there could potentially be differences in how strictly assessors award marks and how they interpret the assessment form. Although assessment briefings are held at the start of each semester, these may well be insufficient to ensure consistency.

One outcome of this lack of calibration is the potential for confusion on the part of the assessors with regard to their role. Kotzé (2014:128) argues that the role of the interpreter – particularly in educational interpreting – is ill-defined. Following on from work done by Angelelli (2015), Kotzé contends that perceptions, both among interpreters and users, influence the role the interpreter will play or is expected to play during any communicative interaction (Kotzé 2016:784) and refers to the need for “sanctioned and agreed-upon conventions” (Kotzé 2014:129). Particularly in her arguments relating to role, she echoes arguments made by Hale (cited in Svongoro and Kadenge 2015:50) about the ease with which interpreter performance can be influenced very strongly by externalities if the interpreters do not have clarity about their role. The same is likely to hold true for those who assess interpreting performance, except that they require not only clarity about the role the interpreter should play, but also about their own role and the function of the assessment.

Another major defect is the delay between the first assessment and the progress discussion, which can be as much as five months. Even with the shortest possible delay – three months, under unusual conditions – a significant amount of time elapses and a substantial number of lectures are interpreted before formal feedback can be given and discussed. This poses a grave risk to the quality of the work done by the interpreting service. To some extent, this risk is mitigated by the system for more immediate peer-to-peer feedback, which is discussed in the next section.

4. Peer assessment

Since interpreting is a performance-based activity, it becomes paramount that interpreters constantly strive for improvement in their interpreting product to avoid complacency, especially in the educational context.¹¹ Peer assessment has been used as a means to stimulate and encourage self-directed and collaborative learning. Self-directed or self-regulated learning implies that the interpreters share responsibility for their own improvement and interpreting performance, while collaborating with their fellow interpreters, who conduct the peer assessments, since improvement is a joint effort to achieve stated goals (Van Zundert, Sluijsmans and Van Merriënboer 2010:270).

¹¹ This approach to quality assurance lies somewhere between the democratisation of quality (Grbić 2008:246) and “[q]uality as mission” (Grbić 2008:250).

The peer assessment system at SU follows a formative approach that is based on the seven principles that Nicol and Macfarlane-Dick (2006:205) identified as constituting good feedback practice. The primary objective of the peer assessment system at SU is to identify areas of improvement and to create opportunities to “close the gap between current and desired performance” (Nicol and Macfarlane-Dick 2006:205), gathering information on areas where interpreters struggle and use those to develop training opportunities.

4.1. Development

In the educational context interpreters at SU often have less than 10 minutes to move from venue to venue between classes and as a result immediate peer feedback on interpreting performance is neglected. A systematised approach to peer feedback was instituted in 2015. Using a simple system of three emoticons to indicate performance, interpreters could rate their colleagues’ performance during a particular lecture. This could either be handed over immediately or, should the interpreter prefer to remain anonymous, the feedback slip could be deposited in a feedback box. Every second week, the slips would be collected and their content summarised electronically before being sent to the relevant interpreter.

Although the three emoticons were simple to use, many interpreters indicated that they would prefer more detailed feedback. To this end, the system was revised in 2016. After initial problems in ensuring that interpreters regularly used the revised peer assessment tool (cf. Foster 2016), the system seems to have stabilised. Although it is based on the assessment tool, the simplified format used for peer assessment has meant that fewer alterations are needed, compared to the formal assessment grid.

4.2. Current format

The new format – see Figure 6 – focuses on three main areas (scored from one to five): technical aspects, message and personal objectives. The intention is to have a peer assessment tool that is as easy and quick to complete as possible, since the main focus of the passive interpreter is to support their active colleague. The electronic version is therefore set up to provide drop-down menus and interpreters can pick from the list of modules and interpreters, rather than having to type these out. Contextual information is limited to interpreters – using a drop-down menu – deciding whether the lecture, as a whole, was easy, average or difficult. The inclusion of fewer subsections (see Figure 6), compared with the in-house assessment grid, is also intended to optimise the usability of the peer assessment tool, while gaining as much insight into interpreter performance as possible.

Although interpreters are free to ask colleagues for feedback at any time, every second week of the semester is an official peer assessment week. An interpreter will track the performance of each colleague with whom they are on duty in that week and upload their feedback onto a spreadsheet template via Google Drive by the end of the week. The information from the various peer assessments is then collated for each interpreter and a summary is sent to the interpreter on the following Monday. Interpreters have the option of submitting feedback anonymously, but many see no need to do so. The summary indicates worst and best performance for the preceding week and how interpreters performed in pursuit of their stated fortnightly objectives.

Date		
Interpreter		
Module		
Conditions		
Remarks:		
Is the interpreter suited to the module?		
Is the interpreter easier to follow?		
Technical		
Microphone technique		
Lag		
Message		
Completeness		
Language use		
Vocabulary		
Role		
Personal Objectives		

Figure 6: Peer assessment tool

The ease and rapidity with which interpreters can obtain constructive feedback from a number of colleagues makes the peer assessment system suitable for tracking improvement in certain aspects of interpreting. Interpreters are encouraged to identify up to three personal objectives on which they intend to focus during the week. These may differ from week to week or may remain the same for a number of weeks. The aim is to promote self-awareness and a culture of self-improvement outside organised training sessions.

The current peer assessment system has provided an opportunity for much more rapid feedback than is provided by the formal assessment system, which it should supplement rather than supplant. It lacks the opportunity to account for various contextual factors, but this is considered acceptable as it is not intended to provide as nuanced an impression of the interpreter’s performance as the formal assessment instrument.

At least two major challenges remain. The first is ensuring that all interpreters complete and submit peer assessments for each module they interpret during a peer assessment week. The second is that there is, at present, no formalised way of correlating the information gathered during formal assessments and peer assessments. The two assessment systems run in parallel.

5. Conclusion

The aim of educational interpreting at SU is to assist students who are not sufficiently proficient in the language of instruction during formal lectures. As formal lectures are an integral part of learning, it is necessary to ensure that interpreting is of the highest standard possible. The two

systems used to measure interpreting quality are formal assessments by senior interpreters and peer feedback. Both of these systems use assessment tools developed at the University.

Although both assessment tools have been refined with reference to literature on interpreting assessment and experience-based praxis at SU, they are not without their shortcomings and neither are the assessment procedures. Some of these shortcomings are related to logistics – colleagues not submitting their peer assessments on time or delays in the feedback process after formal assessments, for instance – and may improve to a limited extent, given time and improved management. Others – such as the wording of particular criteria or the total marks assigned to subsections – are tied to particular perspectives on interpreting quality and seem likely to remain debated.

If the interpreting service at SU is to continue a comprehensive assessment system in order to monitor interpreting quality on a regular basis, calibration exercises by the assessors to ensure the consistent application of set quality standards will be necessary, as will rapid feedback to interpreters. Adequate dovetailing between formal assessments and peer assessments should also be ensured. A final requirement is that the results of all assessments – both formal and peer-based – should be incorporated into exercises to improve deficiencies in interpreting performance, thereby ensuring that students are helped and not hindered by educational interpreting.

References

- Ahrens, B. 2005. Analysing prosody in simultaneous interpreting: Difficulties and possible solutions. *The Interpreters' Newsletter* 13: 1-14.
- Alonso Bacigalupe, L. 2013. Interpretation quality: From cognitive constraints to market limitations. In R. Barranco-Droege, E.M. Pradas Macías and O. García Becerra (eds.) *Quality in interpreting: Widening the scope, Vol. 2*. Granada: Editorial Comades. pp. 9-34.
- Angelelli, C. 2015. Justice for all? Issues faced by linguistic minorities and border patrol agents during interpreted arraignment interviews. *Monographs in Translation and Interpreting* 7: 181-205.
- Bartłomiejczyk, M. 2007. Interpreting quality as perceived by trainee interpreters: Self-evaluation. *The Interpreter and Translator Trainer* 1(2): 247-267.
- Collados Aís, Á. 2002. Quality assessment in simultaneous interpreting: The importance of nonverbal communication. In F. Pöchhacker and M. Shlesinger (eds.) *The interpreting studies reader*. London and New York: Routledge. pp. 327-336.
- Foster, L. 2014. Quality-assessment expectations and quality-assessment reality in educational interpreting: An exploratory case study. *Stellenbosch Papers in Linguistics Plus* 43: 87-102.
- Foster, L. 2016. A matter of calibre: Developing a useful assessment gauge in an educational environment. Poster presentation at the Eighth European Society for Translation Studies Congress, 15-17 September, Aarhus, Denmark.
- Garzone, G. 2002. Quality and norms in interpreting. In G. Garzone and M. Viezzi (eds.) *Interpreting in the 21st century: Challenges and opportunities*. Amsterdam: John Benjamins. pp. 107-119.
- Grbić, N. 2008. Constructing interpreting quality. *Interpreting* 10(2): 232-257.

- Holub, E. 2010. Does intonation matter? The impact of monotony on listener comprehension. *The Interpreters' Newsletter* 15: 117-126.
- Ibrahim-González, N. 2013. Learner autonomy via self-assessment in consecutive interpreting for novice learners in a non-interpreting environment. In R. Barranco-Droege, E.M. Pradas Macías and O. García Becerra (eds.) *Quality in interpreting: Widening the scope, Vol. 2*. Granada: Editorial Comades. pp. 223-238.
- Kalina, S. 2002. Quality in interpreting and its prerequisites: A framework for a comprehensive view. In G. Garzone and M. Viezzi (eds.) *Interpreting in the 21st century: Challenges and opportunities*. Amsterdam: John Benjamins. pp. 121-130.
- Kalina, S. 2005. Quality in the interpreting process: What can be measured and how? *Communication & cognition. Monographies* 38(1-2): 27-46.
- Kasandrinou, M. 2010. Evaluation as a means of quality assurance: What is it and how can it be done? In V. Pellat, K. Griffiths and S.C. Wu (eds.) *Teaching and testing interpreting and translating*. Bern: Peter Lang. pp. 193-206.
- Kopczynski, A. 1994. Quality in conference interpreting: Some pragmatic problems. In S. Lambert and B. Moser-Mercer (eds.) *Bridging the gap: Empirical research in simultaneous interpretation*. Amsterdam and Philadelphia: John Benjamins. pp. 87-99.
- Kotzé, H. 2014. Educational interpreting: A dynamic role model. *Stellenbosch Papers in Linguistics Plus* 43: 127-145.
- Kotzé, H. 2016. Die rolpersepsies van opvoedkundige tolke in Suid-Afrika. *Tydskrif vir Geesteswetenskappe* 56(3): 780-794.
- Kurz, I. 2001. Conference interpreting: Quality in the ear of the user. *Meta* 46(2): 394-404.
- Moser-Mercer, B. 1996. Quality in interpreting: Some methodological issues. *The Interpreters' Newsletter* 7: 43-55.
- Moser-Mercer, B. 2008. Construct-ing quality. In G. Hansen, A. Chesterman and H. Gerzymisch-Arbogast (eds.) *Efforts and models in interpreting and translation research*. Amsterdam and Philadelphia: John Benjamins. pp. 143-156.
- Nicol, D.J. and D. Macfarlane-Dick. 2006. Formative assessment and selfregulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education* 31(2): 199-218.
- Pöchhacker, F. 1994. Quality assurance in simultaneous interpreting. In C. Dollerup and A. Lindegaard (eds.) *Teaching translation and interpreting 2: Insights, aims and visions*. Amsterdam and Philadelphia: John Benjamins. pp. 233-242.
- Pöchhacker, F. 2001. Quality assessment in conference and community interpreting. *Meta* 46(2): 410-425.
- Pöchhacker, F. 2004. *Introducing interpreting studies*. London and New York: Routledge.
- Pöchhacker, F. 2013. Researching quality. A two-pronged approach. In R. Barranco-Droege, E.M. Pradas Macías and O. García Becerra (eds.) *Quality in interpreting: Widening the scope, Vol. 1*. Granada: Editorial Comades. pp. 33-56.

Pradas Marcías, M. 2006. Probing quality criteria in simultaneous interpreting: The role of silent pauses in fluency. *Interpreting* 8(1): 25-43.

Svongoro, P. and M. Kadenge. 2015. From language to society. An analysis of interpreting quality and the linguistic rights of the accused in selected Zimbabwean courtrooms. *South African Linguistic and Applied Language Studies* 33(1): 47-62.

Van Zundert, M., D. Sluijsmans and J. van Merriënboer. 2010. Effective peer assessment processes: Research findings and future directions. *Learning and Instruction* 20(4): 270-279.

Wallmach, K. 2013. Providing truly patient-centred care: Harnessing the pragmatic power of interpreters. *Stellenbosch Papers in Linguistics Plus* 42: 393-414.