The impact of fluency on the subjective assessment of interpreting quality

SYLVI RENNERT

University of Vienna

Abstract

Prosodic features such as fluency are key components of natural speech and, thus, also of simultaneous interpreting. Disfluencies, such as hesitations, vowel lengthening and repairs, are particularly significant in the output of simultaneous interpreters, which presents a pattern of pausing and disfluencies that differs from other forms of spontaneous speech. This paper provides an overview of aspects of fluency and a brief introduction to previous research in the area of fluency and user perceptions, and describes a study conducted by the author at the University of Vienna. The results of this experiment indicate that there may be a link between perceived fluency of an interpretation and users' assessment of the interpreter's accuracy. There also appears to be a link between self-assessed comprehension and assessment of the interpreter's performance.

1. Introduction

Fluency has been studied as one among many aspects of quality in interpreting since the 1980s. However, only recently has research in the field of interpreting studies begun to focus on fluency as an individual factor in order to investigate what constitutes fluency and whether or not it has an impact on intelligibility and user perception.

This paper focuses on fluency in the subjective assessment of interpretations. While it can be argued that listeners hearing an interpreted speech usually cannot really evaluate the quality of the interpretation (cf. Bühler 1986: 233), practising interpreters nevertheless come across listeners every so often who do judge the quality of the interpretation and appear to see no reason why they should not be capable of doing so. This, coupled with the frequent experience that listeners often lay the blame for unclear formulations or errors squarely at the feet of the interpreter and rarely consider that the problem might, in fact, originate in the source speech (cf. Kurz 1989: 144, Ng 1992: 37), demonstrates the need for interpreters to come to terms with the fact that the audience will judge their interpretation, and that factors that users perceive as errors or problems can cast a bad light on their performance. Therefore, this paper takes a closer look at the impact of the quality factor *fluency* on audience evaluation.

The first sections focus on the definition of fluency and on previous research. The second part describes a study conducted by the author, which, among other things, aimed to investigate whether fluency has an impact on the way users evaluate an interpretation with regard to comprehensibility and the interpreter's performance.

2. Fluency

Fluency as a concept is somewhat elusive, as there is no generally recognised definition (cf. Aguado Padilla 2002: 13, Guillot 1999: 3). I will therefore describe some of the many different perspectives on fluency and the individual factors that may contribute to fluency or a lack thereof in subsection 2.1. Subsection 2.2. contains my working definition of fluency and its constituent parameters as used in my doctoral thesis and this paper.

2.1. Fluency in research

While some researchers consider fluency a part of language "proficiency that indicates the degree to which speech is articulated smoothly and continuously without any 'unnatural' breakdowns in flow" (Ejzenberg 2000: 287), others see it as the interaction of temporal variables of speech (such as pause length and length of uninterrupted speech runs) with other, less objective factors such as "clarity of voice", "enunciation" and "ease/confidence in speech" (cf. Freed 2000: 261).

Even among those researchers who focus on objective, measurable parameters there is no general consensus on the exact definition of fluency. However, a large number of researchers in linguistics,

interpreting studies and second language learning mention factors such as pauses, speech rate and hesitations. While these definitions differ in the number of parameters they use and their exact classification, some core features occur in numerous definitions (cf. Aguado Padilla 2002). These include speech rate, pauses, hesitations, lengthened syllables, repetitions, self-corrections and false starts (cf. Ahrens 2004, Chambers 1997: 538, Hedge 1993: 275, Kurz and Pöchhacker 1995: 354, Laver 1994: 537, Lennon 2000: 25, Mead 2005: 45, Tissi 2000: 112).

The factors listed above are key components of natural speech and, thus, also of simultaneous interpreting (cf. Ahrens 2004: 76f), which is a special form of spontaneous speech that is characterised by online planning but dependent on input from an external source. This dependence may lead to a pattern of pauses, breathing and other disfluency factors that distinguishes interpreting from other forms of spontaneous speech (cf. Shlesinger 1994). The following paragraphs describe the individual variables as they relate to interpreting in general. Where there are several definitions, I have described the ones used in my study in 2.2.

Many different ways of defining and measuring *pauses* have been used in research. Some researchers (cf. Stuckenberg and O'Connell 1988) argue for measuring pauses as interruptions of the speech signal, which has the advantage that it can be easily detected and automatically measured with the aid of computers. However, this way of measuring does not necessarily correspond to our subjective experience: measured pauses are often not perceived as pauses, and pauses are perceived where there is no true interruption of the speech signal (cf. Ahrens 2004: 102, Laver 1994: 536, Pompino-Marschall 1995: 174, Stuckenberg and O'Connell 1988). Therefore, some researchers differentiate between pauses at the psychoacoustic level, where a certain percentage of listeners perceive an interruption of the speech signal, and instrumentally detectable breaks (cf. Butcher 1981: 61). Pauses can be distinguished by position, as either *syntactic* – e.g. between sentences or phrases – or *non-syntactic* – in positions that do not usually call for a pause (cf. Ahrens 2004, Chambers 1997, Goldman-Eisler 1968). Researchers have defined the length of pauses in different ways, and it is important to note that the position has a strong impact on the minimum length necessary for a pause to be noticed.

Audible breathing includes any intake or exhalation of breath that is perceived as out of the ordinary. While breathing in natural speech mostly occurs in syntactic pauses and is usually barely audible (cf. Ahrens 2004: 186f., Butcher 1981: 112, Chambers 1997: 539), interpreters cannot always adjust the position of syntactic pauses to their breathing requirements owing to their dependence on the source speaker (cf. Ahrens 2004:187).

Hesitations are sounds such as "umm" and "ahh". Vowels and consonants can be *lengthened* at the beginning, end or in the middle of a word. These phenomena are signs of spontaneous speech planning processes and are,

therefore, quite common in the various forms of spontaneous speech (cf. Arnold and Tanenhaus, forthcoming, Henderson *et al.* 1966, Hieke 1981).

False starts are created by interrupting a sentence and beginning a new one without completing or correcting the previous sentence, while *repairs* include corrections of errors in pronunciation, grammar, structure, content or style (cf. Pöchhacker 1994: 135f., Tissi 2000: 114). *Repetitions* of words can be used as a method of stalling and speech planning; however, they can also bridge two speech segments that have been separated by pauses or hesitations (cf. Hieke 1981: 152ff).

The *tempo of speech* can be measured in various ways: the speech rate is measured in syllables per minute of total speech time, including pauses, whereas the articulation rate corresponds to the number of syllables per minute of all vocalised speech, i.e. words and hesitations, but without pauses (cf. Ahrens 2004: 101, Goldman-Eisler 1958: 61, Laver 1994: 539, Möhle 1984: 27). The rates can also be measured in words, but this makes it hard to compare the speed of utterances in languages with differing word lengths (cf. Pöchhacker 1994:131). As the perceived local speech rate can apparently differ radically from the calculated average, measuring the speed in audible syllables is considered more accurate for perception studies than a transcription-based syllable count (cf. Pfitzinger 2001).

2.2. Fluency: a working definition

Based on the numerous definitions of fluency and its individual constituents as described in 2.1, I have defined fluency as a prosodic feature of speech that can be viewed as a function of a number of temporal variables. It is the complex interaction of pauses, audible breathing, hesitations, vowel and consonant lengthening, false starts, repairs, repetitions and speech rate that creates the impression of fluency or a lack thereof.

With regard to *pauses*, as my research centres around the subjective impression of the target audience, only perceived interruptions of the acoustic signal were counted as pauses for the purposes of the study. As the minimum length where pauses were perceived as such varied from 0.05 to approx. 2 seconds in my study, no cut-off point for pauses is given in this paper.

Audible breathing is defined as in 2.1, and Ahrens' (2004: 187) remark on the interpreter's dependence on the source speaker may be an explanation for the frequent sharp and clearly audible intakes of breath found in my material.

As discussed above, there are many ways of measuring the tempo of speech. For my research material, a manual syllable count was used to calculate the *global articulation rate* (syllables or sounds per minute without pauses calculated over the entire interpretation) and the *net articulation rate* (average speed between two pauses).

Sylvi Rennert

Hesitations, vowel and consonant lengthening, false starts, repairs and repetitions were all found in the experimental material as well. For definitions see 2.1.

3.   User perceptions: previous research

User expectations of various criteria have been studied since Bühler's seminal 1986 survey among conference interpreters, and researchers such as Kurz (1989, 1993), Vuorikoski (1993, 1998), Kopczyński (1994) and Moser (1995) have shown that users tend to consider fluency important, though usually not a top priority. However, it was not until 1998 that user expectations were contrasted with user evaluations of interpretations. With an experimental study on intonation, Collados Aís (1998) initiated a line of research at the University of Granada that compares users' expectations with their ratings of actual interpretations and explores the impact of various factors on user evaluation.

Collados Aís (1998) provided her test subjects, consisting of a user group (legal experts) and a practitioner group (interpreters), with a question-naire in which they were asked to rate the importance of ten criteria for interpretations. They were then asked to listen to and evaluate one of three interpretations. The videos were recordings of the same original speech with the interpretation recorded as a voice-over; one interpretation was accurate but presented in a monotonous way, the second version contained content errors but was presented with a lively intonation, and the third one contained no errors and had a lively intonation. The correct but monotonous version received the lowest rating for overall quality from both groups; the interpreters rated the two lively versions equally, while the user group gave the lively and correct version the highest rating and the lively but incorrect version received slightly lower marks but was nevertheless considered better than the correct but monotonous one.

The monotonous version was also evaluated worst on a number of aspects that had not been changed: quality of voice, logical cohesion, sense consistency, terminology, style and professionalism. While the number of subjects in this study was too small to justify generalisations, this was the first piece of experimental research that indicated that user perceptions of interpreter performance might be influenced by prosodic factors rather than accuracy.

In 2003, this line of research was extended to fluency with Pradas Macías' doctoral dissertation on silent pauses as a parameter of fluency. Similar to the methodology employed by Collados Aís (1998), Pradas Macías first elicited user expectations, asking subjects to rate the influence of ten quality-related features. Fluency was rated fifth, a ranking similar to that given in previous user expectation studies. Subjects were divided into three groups and asked to watch and listen to one of three videos with

a voice-over interpretation of the same speech used by Collados Aís (1998). The videos were recorded by a professional interpreter using a prepared translation to ensure that the content and choice of words did not vary between recordings. However, she introduced pauses of different lengths at previously determined positions. 13 pauses were added to video 1, whereas video 2 contained these same 13 and seven additional pauses. The third video was the control version without any additional pauses. The subjects were then asked to judge the interpretation based on 14 parameters: overall quality, impression of professionalism, impression of reliability, quality of original speech, accent, voice, logical cohesion, correct rendition of sense, completeness, terminology, style, diction, intonation, and fluency (cf. Pradas Macías 2003).

While the differences in the ratings were not very pronounced and not statistically significant, a number of interesting trends could be observed. The control video was rated better than the two versions with additional pauses not only with regard to fluency but also for factors such as professionalism, logical cohesion, completeness and diction. However, in some cases the control version was rated slightly lower, e.g. for style. The differences between videos 1 and 2 are less pronounced. In some cases video 1, the version with fewer additional pauses, was evaluated more favourably, such as in the case of accent, voice and logical cohesion, while video 2 with the largest number of pauses scored higher for other parameters, including overall quality, impression of reliability and, interestingly, fluency.

The number of subjects and the very small differences in ratings – which may also be attributable to the fact that all three versions are based on an "ideal" interpretation – again make it hard to extrapolate from this research, but the general trend of the results suggests that the manipulation of a prosodic factor may have influenced user perception of other aspects as well.

Following these studies, a more detailed investigation into all the individual parameters was conducted at the University of Granada (cf. Collados Aís *et al.* 2007), which was an important contribution to the systematic investigation of quality criteria of simultaneous interpretations. The study consisted of three phases. First, a user expectation survey on 11 quality-related features was carried out. Next, twelve versions of an interpretation were recorded as voice-over, 11 of which were manipulated for one of the quality-related features. The last video was not manipulated and served as control video. The subjects were divided into 12 groups, with each group watching and rating one of the versions. As a last step, the subjects were given a questionnaire in which they had to spontaneously define the parameter that had been manipulated for their group and indicate how much that parameter had bothered them in the interpretation they had just heard.

As in other studies, fluency ranked in the midfield as the fifth of 11 criteria (cf. Pradas Macías 2007: 60). The version that had been manipulated for fluency contained not only pauses but also repairs and false starts. It received low ratings compared to the control video not only for fluency but also for most other parameters, including those for correct rendition of sense and completeness (cf. Pradas Macías 2007: 64). No other manipulated factor elicited such consistently lower scores for all parameters compared to the control version, and fluency was one of the factors with the strongest negative impact on the evaluation of professionalism, reliability and overall quality (cf. Collados *et al.* 2007: 218).

In a study by García Becerra (2007) at the University of Granada, fluency was named as one of the top three out of 14 factors by users when asked which factors were most important in forming first impressions of an interpreter. The study also found that first impressions appeared to influence the evaluation of subsequent performances by the same interpreters. In the same study, subjects were asked to choose the best of four interpretations and state their reasons for this choice. The main reasons given by the subjects were fluency, voice and intonation (cf. García Becerra 2007: 314).

These studies all indicate that users are influenced by a number of factors in their subjective quality judgements and that fluency may be among the top factors. The following chapter describes an experiment I conducted in order to measure whether such an influence does, in fact, exist for fluency and what it means for user perceptions of the quality of the interpretation.

4. Fluency and user perceptions: experiment

Within the framework of the research project entitled *Quality in Simultaneous Interpreting* (QuaSI) at the University of Vienna. I conducted an experiment in March 2010 to investigate the impact of fluency on quality. The main focus of the project, which is part of my doctoral thesis, is on objective, measurable quality criteria, such as equivalence of effect (assessed in terms of comprehension). Nevertheless, subjective user evaluation was also included in the research design, which builds on the work done by colleagues in Granada. Some results from this second part of the study will be presented here. As mentioned in the introduction, subjective assessments by the audience are part of the everyday reality of interpreters, and, therefore, have to be taken into account even though they are not necessarily a reliable measure of interpreting quality from the point of view of interpreting researchers and practitioners.

To accommodate both the comprehension and subjective evaluation aspects, a new design was developed that is based in part on previous

studies in interpreting research, but was tailored to the research questions.

4.1. Subjects and material

With comprehension as one of the key dependent variables, all participants had to have similar background knowledge in order to control for confounding factors. This was achieved by using a speech on a specialised topic and a group of subject-matter experts where similar background knowledge could be assumed. The source text used was a speech on an innovative marketing topic held in English by an expert in the field, and the subjects were business students specialising in marketing (N=47). The subjects were parallelised for language background in order to ensure an equal distribution of native and non-native speakers of German, and randomly assigned to one of two groups.

The experimental material was produced in advance using the audio software PRAAT. A professional conference interpreter was asked to interpret the source text several times from English to German, yielding interpretations with differing degrees of fluency. The most fluent of these was selected as a basis for manipulation, and all versions were mined for disfluencies – hesitations, lengthened vowels and consonants, audible breathing – and analysed to detect typical pausing patterns of this interpreter. One copy of the selected base version was then manipulated for increased fluency by removing hesitations, false starts and audible breathing that had a negative impact on fluency, and shortening or removing pauses in non-syntactic positions, while at the same time adding pauses and calm audible breathing to syntactic positions. A second copy was turned into the non-fluent version by adding pauses, hesitations and audible breathing to non-syntactic positions, lengthening existing pauses in non-syntactic positions and adding false starts, lengthened sounds and repairs.

The material was pretested to verify that it sounded natural and did not appear exaggerated in either version, making sure that the fluent version was not so smooth as to be unrealistic and that the non-fluent version still sounded like a professional interpretation one might hear at a conference.

The advantage of using audio manipulation to produce the material is that it allows changing only one parameter, in this case fluency, while leaving all other parameters intact, thus ensuring that the versions are completely comparable in all other respects, such as intonation, voice or choice of words.

Sylvi Rennert

4.2. Experimental design

A conference setting was simulated for the experiment in order to approach realistic conditions. Two interpreting booths were set up in the auditorium along with the necessary technical equipment, operated by a professional audio technician. Infrared receivers were distributed to the audience.

   The experiment was held during regular university classes without informing the students in advance. At the start of the class, the subjects were told they would hear a guest lecture on marketing and post-modernism, presented as a video with an interpretation which they would hear via the headphones, and that they would be asked to answer questions about it afterwards. It was stressed that the anonymous questionnaires would only be used to evaluate the quality of the interpretation, not their own skills or knowledge.

   Each group was assigned a separate channel on the infrared receiver. While the impression conveyed was that these channels corresponded to the two interpreting booths, they were in fact connected directly to audio equipment playing back the previously manipulated versions of the interpretation. Although they could not be heard, the interpreters in the booths shadowed one of the versions each in order to give the impression of a live interpretation to any subjects turning to watch them.

4.3. The questionnaire

The questionnaire consisted of two parts: the first contained a set of 11 comprehension questions to investigate the first research issue, audience comprehension in relation to different degrees of fluency. The second part was dedicated to the subjective assessment of the interpretation and the gathering of background variables such as age, gender, familiarity with the topic, etc.

   One of the evaluation questions was designed to assess the subjective impression of fluency in order to determine whether the two versions manipulated to be either very fluent or less fluent were, in fact, experienced as such by the listeners. This question – "Please rate the fluency of the interpretation (pauses, hesitations)" – also provided context for two further items, worded as "How well did you understand the content of the interpretation?" and "How accurately did the interpreter render the content?". The rating was given on a 7-point scale with 1 being the best and 7 the worst result. The scale was chosen to permit nuanced answers and to have a mid-point (4) that allowed subjects to choose a neutral answer.

   Presenting the results of the full study would go beyond the scope of this paper, and as the findings of the first part, comprehension, will be

reported in detail elsewhere, I will focus here on the aspect of audience assessment.

While the objective of these questions was similar to that of the studies conducted in Granada (cf. Collados Aís 1998, Pradas Macías 2003, Collados Aís *et al.* 2007), they were posed in a different way, as the studies by Pradas Macías (2003) and Collados Aís *et al.* (2007), in particular, had shown that the definitions of the individual items varied widely from subject to subject (cf. Collados Aís *et al.* 2007: 219). To avoid misunderstandings and different interpretations of a generalised question on interpreting quality, which might have led to unreliable results, a narrower question was chosen. The question "How accurately did the interpreter render the content?" can be seen as covering one aspect of quality, the impression of fidelity, without the risk of a confounding influence of other possible readings of a question referring directly to "quality of the interpretation", such as voice, accent, or style.

As has been pointed out, the question about accurate rendition of content is one that listeners cannot answer unless they have heard both versions and understood them equally well, which was not the case in the study described here, as the original sound was not audible to the subjects. Nevertheless, only 5 out of 47 subjects added a comment to the effect that they were not able to judge this as they had not heard the original speech, and two of these still answered the question. All other subjects responded to this question without any further comments.

The purpose of the question on understanding was twofold: it allowed the comparison of the actual comprehension scores from the first part of the questionnaire with the impression the subjects themselves had of their comprehension, and at the same time it provided information on possible links between fluency (perceived or actual), perceived interpreting quality and perceived comprehension.

4.4. Results

The results of the subjective fluency assessment confirmed that the non-fluent version was indeed perceived as less fluent than the fluent one by listeners who did not have the possibility of comparing both versions, thus validating the manipulation method. Figure 1 compares the median fluency ratings of the two groups, given on a seven-point scale, where 1 was "very fluent" and 7 "not fluent at all". As the fluency rating is not on an interval scale, median values are used instead of means.

A U test showed the difference between the two groups to be significant (Mann-Whitney U=172.500, p=.021 < 0.05 two-tailed).

Correlating perceived fluency with the results of the question "How accurately did the interpreter render the content?", showed a low correlation between the two (r=.343): there was a slight tendency for lower
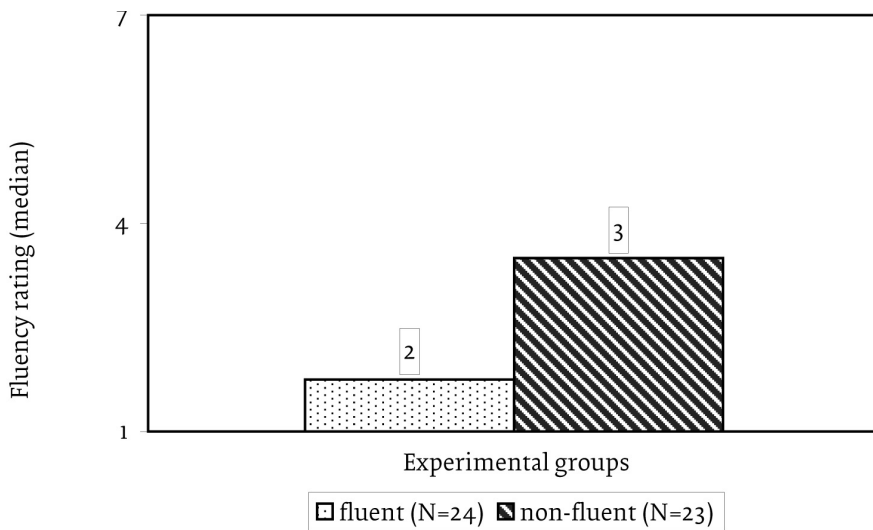
Figure 1: Fluency ratings

fluency ratings to correspond to a worse evaluation of the interpreting performance. The difference in scores for interpreter performance by fluency rating (see below) is significant at the 90 percent level in a two-tailed U test (Mann-Whitney U=166, p=.073 < 0.1 two-tailed).

Evaluation of the interpreting performance and subjective comprehension, however, showed a moderate correlation (r=.533), meaning that lower subjective comprehension correlated with a worse impression of the interpreter's performance.

While there is only a very weak correlation between perceived fluency and subjective comprehension (r=.164), a comparison of the median ratings for subjective comprehension shows a difference both between the experimental groups (see Figure 2) and by fluency rating (see Figure 3).

As perceptions of fluency can vary among individuals – and as judgements seem to be influenced by subjectively experienced fluency rather than by any external definition or measurement – the respondents were grouped by perceived fluency for a further part of the analysis. Subjects were assigned to one of two groups depending on their fluency rating, regardless of the experimental group they were in. Fluency was given a rating of 1 or 2 by 26 subjects (17 from the "fluent" and 9 from the "non-fluent" experimental group), and ratings between 3 and 5 by 21 subjects (7 from the "fluent" and 14 from the "non-fluent" experimental group).

It was interesting to see that the result of the comparison of the experimental groups (Figure 2) was borne out by the subjective ratings (Figure 3).
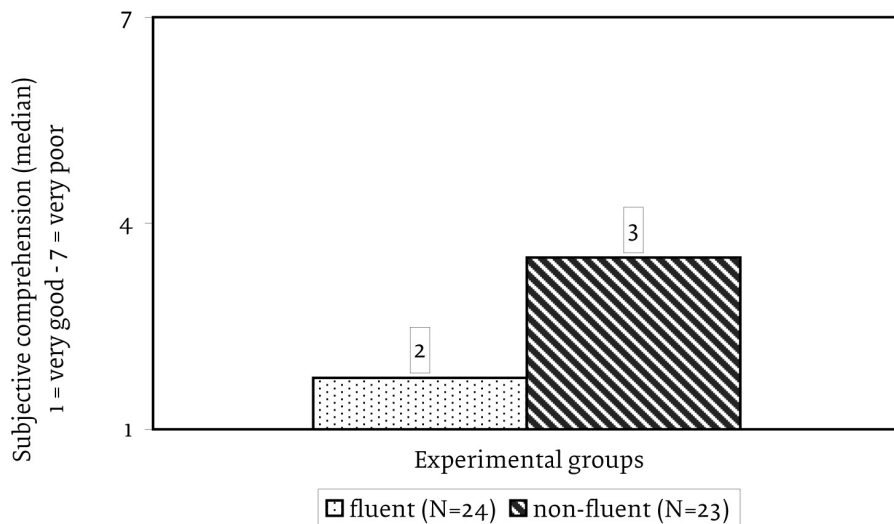
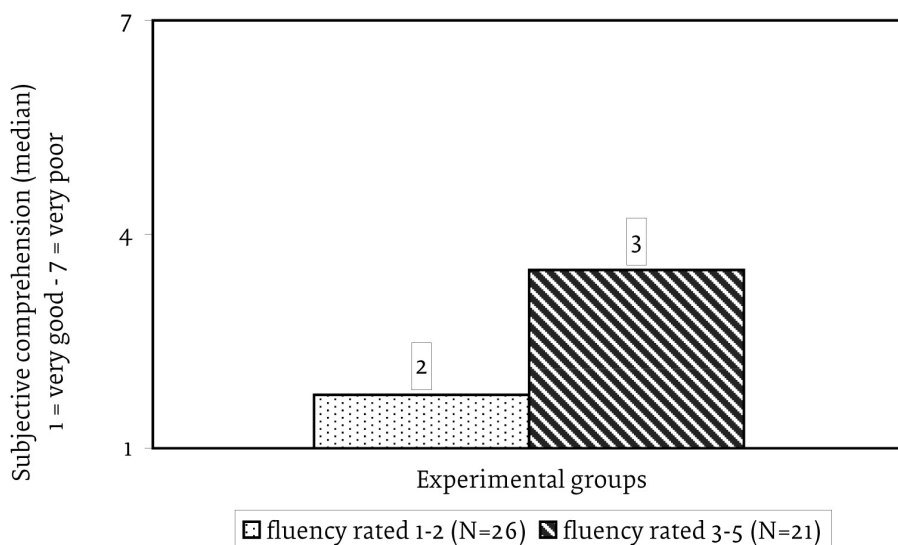Figure 2: Subjective comprehension by experimental groups

Figure 3: Subjective comprehension by fluency rating

5. Discussion

The results presented above suggest that there is a link between perceived fluency and perception of the interpreter's accuracy, confirming previous studies that suggested that lower fluency may impact negatively on the perceived quality of an interpretation. As the material used in this study differed only in terms of fluency and the two versions were identical with regard to voice, intonation, wording and information content, it appears

that fluency cannot be ignored as a factor that influences audience perception. This is also corroborated by the difference in perceived comprehension that was visible for both subjective fluency and the two versions of the experimental material.

As there is also a correlation between subjective comprehension and perception of the interpreter's performance, it appears that users may indeed show a tendency to blame the interpreter for problems with the interpretation.

6.   Concluding remarks

It appears from the results of this study that fluency is more than just a matter of style and may, in fact, impact users' and clients' opinion of the quality of an interpretation in terms of performance and intelligibility.

It must be borne in mind that these results are valid only for the business students who participated in this study and cannot be generalised to other groups of students or subject-matter experts. However, as these results are in line with previous small-scale studies, more large-scale studies in this area would be welcome to test and corroborate these findings.

References

Aguado Padilla K. (2002) *Imitation als Erwerbsstrategie. Interaktive und kognitive Dimensionen des Fremdsprachenerwerbs,* Habilitationsschrift, Universität Bielefeld, http://karin.aguado.de/publ/downloads/ImitationAlsErwerbsstragie.pdf (last accessed 12 November 2010).
Ahrens B. (2004) *Prosodie beim Simultandolmetschen*, Frankfurt/Main, Peter Lang.
Arnold J.E. and Tanenhaus M.K. (forthcoming) "Disfluency effects in comprehension: how new information can become accessible", in E. Gibson and N. Perlmutter (eds.) *The Processing and Acquisition of Reference*, MIT Press, http://www.unc.edu/-jarnold/papers/arnoldtan2007.pdf (last accessed 27 July 2010).
Bühler H. (1986) "Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters", *Multilingua* 5 (4), 231-235.
Butcher A. (1981) *Aspects of the Speech Pause. Phonetic Correlates and Communicative Functions* (Institut für Phonetik, Universität Kiel, Arbeitsberichte 15), Kiel, Institut für Phonetik.
Chambers F. (1997) "What do we mean by fluency?", *System* 25 (4), 535-544, http://docs.ksu.edu.sa/KSU_AFCs/yousif/What%20do%20we%20mean%20by%20fluency.pdf (last accessed 07.01.2009)

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

Collados Aís A., Pradas Macías M., Stévaux E. and García Becerra O. (eds.) (2007) *La evaluación de la calidad en interpretación simultánea: Parámetros de incidencia*, Granada, Comares.

Ejzenberg R. (2000) "The juggling act of oral fluency: a psycho-socio-linguistic metaphor", in H. Riggenbach (ed.) *Perspectives on Fluency*, Ann Arbor, University of Michigan Press, 287-313.

Freed B.F. (2000) "Is fluency, like beauty, in the eyes (and ears) of the beholder", in H. Riggenbach (ed.) *Perspectives on Fluency*, Ann Arbor, University of Michigan Press, 243-265.

García Becerra O. (2007) "La incidencia de las primeras impresiones en la evaluación de la calidad de la interpretación. Estudio piloto", in M.M. Fernández Sánchez and R. Muñoz Martín (eds.) *Aproximaciones cognitivas al estudio de la traducción y la interpretación*, Granada, Comares, 302-326.

Goldman-Eisler F. (1958) "Speech analysis and mental processes", *Language and Speech* 1 (1), 59-75.

Goldman-Eisler F. (1968) *Psycholinguistics. Experiments in Spontaneous Speech*, London/New York, Academic Press.

Guillot M.-N. (1999) *Fluency and its Teaching*, Clevedon/Philadelphia/Toronto/Sydney, Multilingual Matters.

Hedge T. (1993) "Key concepts in ELT (Fluency)", *ELT Journal* 47 (3), 275-276.

Henderson A., Goldman-Eisler F. and Skarbek A. (1966) "Sequential temporal patterns in spontaneous speech", *Language and Speech* 9 (4), 207-216.

Hieke A.E. (1981) "A content-processing view of hesitation phenomena", *Language and Speech* 24 (2), 147-160.

Kopczyński A. (1994) "Quality in conference interpreting: some pragmatic problems", in M. Snell Hornby, F. Pöchhacker and K. Kaindl (eds.) *Translation Studies: An Interdiscipline*, Amsterdam/Philadelphia, John Benjamins, 189-198.

Kurz I. (1989) "Conference interpreting – user expectations", in D. Hammond (ed.) *Coming of Age. Proceedings of the 30th Annual Conference of the American Translators Association*, Medford, NJ, Learned Information, 143-148.

Kurz I. (1993) "Conference interpretation: expectations of different user groups", *The Interpreters' Newsletter* 5, 13-21.

Kurz I. and Pöchhacker F. (1995) "Quality in TV interpreting", *Translatio – Nouvelles de la FIT – FIT Newsletter* 15 (3/4), 350-358.

Laver J. (1994) *Principles of Phonetics*, Cambridge, Cambridge University Press.

Lennon P. (2000) "The lexical element in spoken second language fluency", in H. Riggenbach (ed.) *Perspectives on Fluency*, Ann Arbor, University of Michigan Press, 25-42.

Mead P. (2005) "Methodological issues in the study of interpreters' fluency", *The Interpreters' Newsletter* 13, 39-63.

Möhle D. (1984) "A comparison of the second language speech production in different native speakers", in H.W. Dechert, D. Möhle and M. Raupach (eds.) *Second Language Productions*, Tübingen, Gunter Narr, 26-49.

Moser P. (1995) *Survey on Expectations of Users of Conference Interpretation. Final report*, http://aiic.net/ViewPage.cfm/article525 (last accessed 27 July 2010).

Ng B.C. (1992) "End-users' subjective reaction to the performance of student interpreters", *The Interpreters' Newsletter,* Special Issue 1, 35-41.

Pfitzinger H.R. (2001) "Phonetische Analyse der Sprechgeschwindigkeit", in *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München (FIPKM)* 38 (2001), Munich, Universität München, 117-264, http://www.ipds.uni-kiel.de/hpt/pub/f38＿hp＿1.pdf (last accessed 27 July 2010).

Pöchhacker F. (1994) *Simultandolmetschen als komplexes Handeln* Tübingen, Gunter Narr.

Pompino-Marschall B. (1995) *Einführung in die Phonetik*, Berlin/New York, de Gruyter.

Pradas Macías E.M. (2003) *Repercusión del intraparámetro pausas silenciosas en la fluidez: Influencia en las expectativas y en la evaluación de la calidad en interpretación simultánea*, Dissertation, Universidad de Granada.

Pradas Macías E.M. (2007) "La incidencia del parámetro fluidez", in Collados Aís *et al.* (eds.), *op. cit,.* 53-70.

Shlesinger M. (1994) "Intonation in the production and perception of simultaneous interpretation", in S. Lambert and B. Moser-Mercer (eds.), *Bridging the Gap. Empirical Research in Simultaneous Interpretation*, Amsterdam/Philadelphia, John Benjamins, 225-236.

Stuckenberg A. and O'Connell D.C. (1988) "The long and short of it: reports of pause occurrence and duration in speech", *Journal of Psycho-linguistic Research* 17 (1), 19-28.

Tissi B. (2000) "Silent pauses and disfluencies in simultaneous inter-pretation: a descriptive analysis", *The Interpreters' Newsletter* 10, 103-127.

Vuorikoski A.-R. (1993) "Simultaneous interpretation – user experience and expectations", in C. Picken (ed.) *Translation – The Vital Link. XIII FIT World Congress, Proceedings*, Brighton, Institute of Translation and Interpreting, 317-327.

Vuorikoski A.-R. (1998) "User responses to simultaneous interpreting", in L. Bowker, M. Cronin, D. Kenny and J. Pearson (eds.) *Unity in Diversity? Current Trends in Translation Studies*, Manchester, St. Jerome Publishing, 185-194.