CONTROL OF PAUSES BY TRAINEE INTERPRETERS
IN THEIR A AND B LANGUAGES

Peter Mead
SSLiMIT, University of Bologna

They spell it Vinci and pronounce it Vinchy; foreigners always spell better than they pronounce
Mark Twain

1. Introduction

This study examines pauses, as an index of fluency, in consecutive interpretations into Italian and English. The analysis focuses on target speeches by fifteen final year interpreting students, all Italian native speakers. A basis is thus provided for comparison of performance in the A and B languages, a subject which has prompted scholarly discussion but little systematic research (Dejean Le Féal 1998).

In written translation, Stewart’s (1999) counterblast to those who state that the translator should work only into his or her native language calls for greater recognition of “inverse” translation and the need to train translators accordingly. Stewart observes that, despite the familiarity of the debate regarding directionality in translation, translation scholars offer little systematic, reasoned discussion of the matter.

Perennial, sometimes unconstructive argument about the acceptability of translating into a foreign language is not the prerogative of those concerned with written translation. In conference interpreting, the “querelle du A et du B” (Seleskovitch & Lederer 1989: 134) was highlighted by Denissenko’s (1989) paper on interpreter training in Moscow and the ensuing discussion with a number of western scholars (Gran & Dodds 1989: 199-200). While the amenability of scholars like Denissenko to an active B language may be interpreted as a reflection of practical constraints specific to certain countries and language combinations (ibid.), it can equally be argued that the opposing view is based on traditional orthodoxy rather than firm evidence.

The demarcation line between opposing attitudes to an active B language does not depend only on divergences of opinion among different schools. There can also be differences between the practice in international organisations and on the private market – for example, the European Parliament generally favours three passive languages and active use of the mother tongue (http://www.europarl.eu.int/interp/public), whereas the mainstay of professional
practice for members of major associations operating on the private market is often “aller-retour” between A and B languages.

Given the increasing recognition by scholars in western Europe that the traditional opposition to an active B language hardly reflects the current needs of the market (Gran and Snelling 1998), the “querelle du A et du B” is more than ever a topic of potential interest for research. Against this background, the present study is part of an ongoing doctoral project at the Université Lumière-Lyon 2, France. The focus of the project is the interpreter’s control of pauses, as a factor in fluency, at different stages of training and professional experience. Ultimately, this research will provide a small – but empirically based – contribution to both the “querelle du A et du B” and discussion of interpreter training.

2. Fluency

Fluent speech or public speaking ability has been identified as a mark of aptitude for conference interpreting (Gile 1995: 172; http://www.aiic.net/en/tips: 10) and a criterion by which the interpreter’s output can be evaluated (Gile 1995: 162; Jones 1998: 40). Indeed, Altman (1994: 36) states that “fluency […] is the one single aspect of an interpretation which most palpably distinguishes a professional performance from that of a trainee”.

Surveys among interpreters and conference participants confirm the importance of fluency as a determinant of quality in interpreting (Bühler 1986; Kurz 1993). However, references in the literature do not go into detail about how fluent speech can actually be defined and developed. For example, Seleskovitch and Lederer simply comment on the importance of speaking ability and the desirability of controlling “effets de voix” (Seleskovitch and Lederer 1989: 229, 110). Similarly, Weber (1990: 47) states that “one of the main reflexes to develop […] is the ability to enunciate ideas almost automatically”, but does not break down such ability into its constituent parts.

Seleskovitch and Lederer (op. cit.: 137) consider that, while students’ awareness of fluency should be heightened, time constraints do not allow the interpreting trainer to do so. This need can be fulfilled by public speaking courses (Katz 1989: 218; Weber 1989), though interpreter trainers have expressed reservations about trainees receiving instruction in public speaking from non interpreters (AIIC 1979: 13).

The above sources do not offer detailed advice about how to incorporate development of speaking skills into the interpreting curriculum, or specify how trainee interpreters can profitably focus on problems of fluency. This lack of detailed comment on the development and evaluation of fluency arguably reflects limited attention to the issue in traditional approaches to language training.
2.1. Control of lapses in fluency

One perspective on fluency is to consider evidence of its absence, i.e. faults of production which an accomplished public speaker should not allow to exceed “acceptable limits for pauses, restarts, repetitions, redirections […]” (Goffman 1981: 172). Goffman considers these “linguistically detectable faults” (ibid.) as manifestations of the efforts of reasoning and formulation which accompany linguistic production. The skill of professional speakers such as the lecturer or radio announcer is to control output in such a way as to hide these efforts and any hesitations they may entail; no “production crisis” or “backstage considerations” (ibid.: 172) are allowed to betray moments of doubt or distraction.

Goffman’s lecturer thus maintains control of any hesitations which could surface as “linguistically detectable faults” or “influencies” (ibid.: 209). Admittedly, this perspective is related to Goffman’s interest in the concept of “face”, i.e. respect of the expectations associated with social status or roles. Pauses and other influences are thus seen as inconsistent with expectations of how a lecturer or radio announcer should speak. Nevertheless, Goffman’s discussion provides an interesting theoretical basis for evaluation of fluency. Given that interpreters can to all intents and purposes be considered professional speakers, or “professionnels de l’oralité” (http://www.aiic.net/en/tips/voix/trottier.htm), the definition of fluency by default (i.e. absence of influences) can also prove relevant to evaluation of interpreting (Mead 1996).

It is probably more common to describe influences as “disfluencies”, a term explained as follows: “The most common types of disfluency are hesitations, pauses, ums and ah’s, corrections, false starts, repetitions, interjections, stuttering and slips of the tongue” (Garnham 1985: 206). The features of speech which Garnham includes in this category are substantially those avoided by Goffman’s lecturer, or cautioned against in manuals about how to speak in public (Bellenger 1979: 73).

The structuralist bias of traditional linguistic analysis gives little consideration to linguistic features like disfluencies, which tend to be dismissed as “the greasy parts of speech” (Goffman 1964: 61). However, a number of studies in applied linguistics highlight these “temporal variables” of oral production (Raupach 1980: 269-270) as a quantitative/distributional basis for evaluation of fluency. The perspective is often to compare different levels of competence, for example in a first and second language (ibid., Towell et al. 1996, Onnis 1999) or at different stages in second language acquisition (Towell et al., op. cit.).

1 “Temporal variables” do not provide an evaluation of fluency only by default (i.e. absence of disfluencies). They also focus on such parameters as speech rate and length of uninterrupted segments (“runs”).
In research on interpreting, disfluencies are part of Kopczynski’s (1981) system of error assessment, used with some modifications in Vik-Tuovinen’s (1995) longitudinal study of four interpreting students.

2.2. Silent and filled pauses

Technically, the two features Garnham refers to as “pauses” and “ums and ahs” are classed as “silent” and “filled” pauses respectively (Duez 1982: 13-14). Duez’s inclusion of filled pauses in a composite class of “non silent pauses”, including false starts and repetitions, highlights their status as disfluencies; silent pauses, by contrast, do not necessarily reflect hesitation and fumbling – indeed, Duez’s analysis of political speech shows that they can be used to rhetorical effect (ibid.). This stylistic effect of silent pauses is identified in an earlier study by Clemmer, O’Connell and Loui (1979), who match expert assessments of oral readings with the readers’ pausing patterns. The role of silent pauses in comprehension is similarly commented on by Holmes (1984), who evaluates pausing in improvised stories.

The functional distinction between silent and filled pauses is relevant to public speaking. Bellenger (op. cit.) recommends judicious distribution and timing of the former, which he simply calls “pauses”; control of the latter, referred to as “bruits”, should be far more vigilant – though their total exclusion would be unnatural.

Interestingly, some earlier studies on pauses (Henderson, Goldman-Eisler, Skarbek 1965; Henderson, Goldman-Eisler, Skarbek 1966; Goldman-Eisler 1967) do not systematically state whether the focus is on silent or filled pauses. These studies identify pausing as a feature of cognitive activity during linguistic production. Their theoretical premise is that oral production requiring little or no effort of formulation, like aloud reading and – according to the source concerned – simultaneous interpretation (Goldman-Eisler 1967: 125), does not entail the same alternation of pauses and fluency as impromptu speech. The functional distinction between silent and filled pauses is not explicitly discussed, nor is the concept of pauses as a help to the listener.

The status of silent pauses is varied, in that they may be stops for breath, deliberate pauses for emphasis or hesitations. Deese (1980) considers that a skilful speaker probably makes hesitations coincide with silent pauses at natural syntactic breaks in the utterance. A parallel to this is the way that good reading aloud discreetly takes advantage of appropriate breaks in the text, both for stylistic effect and to give the reader an idea of what comes next. What is important is to recognise that silent pauses are not necessarily disfluencies,
while conspicuous filled pauses almost certainly can be classed as such in the context of professional public speaking (including conference interpretation).

Given this distinction, combined data for both filled and silent pauses arguably afford insight into how obtrusively or discreetly speakers hesitate. The present study therefore includes both types of pause.

3. The study: objectives and methods

Comparative study of temporal variables in speech suggests that even high levels of competence in a foreign language do not imply the same degree of fluency as in native speech (Towell et al., op. cit.; Onnis, op. cit.).

Control of temporal variables thus seems to offer an appropriate perspective from which to evaluate interpretation in the B language. Fluency is admittedly not the most important criterion by which to evaluate interpretation. Nevertheless, the user surveys referred to above suggest that its role in the overall impact of the target speech is far from negligible. Its relevance to the “querelle du A et du B” is thus not to be dismissed.

The present study of pauses as a comparative index of fluency examines the hypothesis that the target speech is more fluent in the A language than the B language. Subjects’ explanations of major pauses are also examined, to identify any trends in terms of factors they perceive as relevant to fluency.

Subjects were all final year students at the S.S.Li.M.I.T., Forlì, with English as their B language. Each recorded two short consecutive interpretations, one in Italian and one in English. To ensure comparability, the same recorded input speeches were used for all participants. Apart from a short written briefing, they received no prior information on the texts before listening to each speech.

Both speeches were from conference recordings, kindly made available by the organisers of the events concerned. The recording in English was the opening of a speech on British attitudes to Europe, given to an audience of Italian students by an English professor. The Italian source speech was again the start of a lecture, in this case by a journalist speaking to a group of industrialists about the international impact of the 1973 oil crisis.

While it is difficult to ensure a uniform level of difficulty in different speeches, the two were reasonably comparable in the following respects:

a) theme: contemporary history;
b) audience with no specialist knowledge of the subject;
c) no particular difficulties in terms of extralinguistic knowledge (Gile 1995: 216);
d) “off the cuff” delivery;
e) no slides or overheads used;
f) duration (3’50” and 3’40”’);
g) density of information;
h) mean speed (127 words per minute in English, 119 w.p.m. in Italian).2

Immediately after each consecutive interpretation, the recording was played back to the subject and comments were sought on major hesitations or clusters of brief pauses. Separate explanations could obviously not be sought for every minor individual pause, in some cases lasting less than 0.10 sec. Given that the consecutive interpretation generally took from 3 to 4 minutes and the average frequency of pauses was about 25 per minute (sometimes far more), subjects could not realistically be expected to offer detailed retrospective analysis of even the briefest and most innocuous.

Subjects’ explanations of pauses were collected by stopping the tape after each prominent or sustained hesitation, so that the subject could answer the question: “Is there any particular reason for this hesitation?” The question was deliberately left open, even when it seemed that there might be an evident reason (e.g., difficulty in finding a target language equivalent for a word or expression), to avoid “prompting” subjects. Leading questions, like “Did you have a lexical problem there?”, were avoided. Information collected in this way provided the basis for the analysis of how subjects perceived their pauses.

Recordings of all interpretations were transferred on to the hard disk of a Macintosh iMac™, using a software for visualisation and editing of audio files (SndSampler 3.7.1™, © Alan Glenn, Midland Mi, USA). The software was then used to divide the recordings into 20 second audio files and convert these into oscillograms on which each pause could be highlighted and matched with the corresponding sound recording. This allowed measurement of pause duration in hundredths of a second, using a sampling frequency of 44 kHz.

Times for silent and filled pauses were calculated separately. Any “mixed” pauses, comprising both silent and filled segments, were considered as filled pauses. Once all pauses had been inventoried, total pause duration and average pausing time per minute were calculated for each interpretation.

Average pausing times per minute in the two languages were compared by t tests for paired data.

Subjects’ explanations of their pauses were also inventoried and sorted into five categories. These were: (1) difficulties of formulation (lexis, grammar); (2) difficulty with notes (e.g., indecipherable symbol); (3) logical doubts (e.g., “Does this comment make sense?”); (4) no apparent reason perceived by the subject; (5) others (e.g., thinking about previous difficulties).

2 Comparison of speech rate in different languages on the basis of a word count is admittedly crude (Pöchhacker, 1993), but is sufficient to ensure that overall speed of delivery is reasonably comparable.
For each subject, values of the various classes of explanation were calculated as percentages of the total count. For example, subject 1 gave a total of 32 and 38 explanations regarding pauses in Italian and English respectively. Of these, the scores for difficulties in formulation were 15 in Italian and 12 in English. The percentage values were thus 46.88% (15/32) for the former and 31.58% (12/38) for the latter. The purpose of converting scores into percentages was to provide a common denominator for the statistical analysis, since the total number of pauses commented on varied from subject to subject.

Like pause times, percentages of the different classes of hesitation in the two languages were compared by t tests for paired data.

4. Results and Discussion

a) Comparison of pauses in Italian and English

Average pause times per minute are shown in Table 1. While group means afford only a limited basis for comparison, they show that: (i) pausing accounts for an appreciable proportion of overall speaking time; (ii) values in English are appreciably higher than in Italian, particularly for filled pauses and for total pause time.

<table>
<thead>
<tr>
<th>subject</th>
<th>pauses (s.p.m.) in Italian:</th>
<th>pauses (s.p.m.) in English:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>filled</td>
<td>silent</td>
</tr>
<tr>
<td>1</td>
<td>6.02</td>
<td>7.66</td>
</tr>
<tr>
<td>2</td>
<td>8.84</td>
<td>8.92</td>
</tr>
<tr>
<td>4</td>
<td>7.28</td>
<td>8.17</td>
</tr>
<tr>
<td>5</td>
<td>12.83</td>
<td>2.54</td>
</tr>
<tr>
<td>6</td>
<td>2.05</td>
<td>10.59</td>
</tr>
<tr>
<td>7</td>
<td>10.41</td>
<td>5.69</td>
</tr>
<tr>
<td>8</td>
<td>6.40</td>
<td>9.34</td>
</tr>
<tr>
<td>9</td>
<td>8.37</td>
<td>2.55</td>
</tr>
<tr>
<td>10</td>
<td>2.92</td>
<td>11.69</td>
</tr>
<tr>
<td>12</td>
<td>5.11</td>
<td>3.66</td>
</tr>
<tr>
<td>13</td>
<td>6.57</td>
<td>3.92</td>
</tr>
<tr>
<td>14</td>
<td>6.01</td>
<td>2.85</td>
</tr>
<tr>
<td>15</td>
<td>9.31</td>
<td>7.13</td>
</tr>
<tr>
<td>mean:</td>
<td>7.17</td>
<td>6.47</td>
</tr>
</tbody>
</table>

Table 1: Seconds per minute of pause time in consecutive interpretations by 15 final year students.
With regard to the first point, speaking time net of pauses can be calculated as a percentage of total speaking time. This percentage, called the P/T (Phonation/Time) ratio, is about 75% and 66% of the group means for Italian and English respectively. While these values may seem to reflect limited phonation in relation to pauses, they are consistent with other P/T ratio data.³ For example, Onnis’s (1999) study of an oral narrative task by English native speakers considered “fluent” in Italian identifies a P/T ratio of about 65% in both languages.

In terms of the comparison between languages, statistical tests identify significantly higher filled pause times in English than Italian. The same is true of total pauses (silent and filled pauses together), while silent pauses alone do not differ significantly between the two languages (t = 1.710; sig. = 0.109). Significant differences are summarised in Table 2.

<table>
<thead>
<tr>
<th>pair</th>
<th>t</th>
<th>sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>total pauses in English - total pauses in Italian</td>
<td>5.812</td>
<td>0.000</td>
</tr>
<tr>
<td>filled pauses in English - filled pauses in Italian</td>
<td>4.173</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 2: Significant differences in seconds per minute of pause time (paired t tests)

The appreciably higher incidence of pauses in English than Italian reinforces the hypothesis that output is more fluent in the A language (Italian) than the B language (English).

The statistical analysis also shows a significant negative correlation between silent and filled pauses in English (Pearson correlation coefficient = -0.633, p = 0.05). This means that the more subjects use one, the less they use the other. In this case, filled pauses are far more prominent. The group’s management of hesitation in English is thus not consistent with the standard of Goffman’s lecturer, whose ability to keep hesitation discreetly backstage leaves little room for “ums and ah’s”.

b) Subjects’ explanations of their pauses in Italian and English

The occurrences of the various categories of explanation for pauses are summarised in Table 3. Significant differences in percentages for the various

³ A caveat in this respect is that some authors calculate the ratio by differentiating only silent pauses from speaking time, while others also subtract filled pauses.
classes in the two languages, based on t tests for paired samples, are listed in Table 4.

<table>
<thead>
<tr>
<th>explanations for pauses</th>
<th>Italian</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>formulation</td>
<td>146</td>
<td>174</td>
</tr>
<tr>
<td>notes</td>
<td>116</td>
<td>106</td>
</tr>
<tr>
<td>logic</td>
<td>60</td>
<td>44</td>
</tr>
<tr>
<td>no reason</td>
<td>42</td>
<td>67</td>
</tr>
<tr>
<td>others</td>
<td>19</td>
<td>21</td>
</tr>
<tr>
<td>total</td>
<td>383</td>
<td>412</td>
</tr>
</tbody>
</table>

**Table 3:** Total explanations for pauses

<table>
<thead>
<tr>
<th>pair</th>
<th>t</th>
<th>sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>formulation in English - formulation in Italian</td>
<td>1.772</td>
<td>0.098</td>
</tr>
<tr>
<td>logic in English - logic in Italian</td>
<td>-2.379</td>
<td>0.034</td>
</tr>
<tr>
<td>no reason in English - no reason in Italian</td>
<td>3.040</td>
<td>0.009</td>
</tr>
</tbody>
</table>

**Table 4:** Significant differences in subjects’ explanations of pauses

The raw data and statistical analysis highlight a number of trends.

A first observation concerns the relative weight subjects give to linguistic and extralinguistic factors as determinants of hesitation. Table 3 shows that the combined totals for “notes” and “logic” are more or less as high as the figures for “formulation” (indeed, they are higher if Italian alone is considered). In other words, subjects at this stage in their training perceive that hesitation in consecutive interpretation stems as much from problems in rereading notes and resolving logical inconsistencies as from strictly language-related difficulties.

A second point is that the significant differences between languages in Table 4 show greater perception of language difficulties in English, while logical doubts are more prominent in Italian. The greater importance attached to language as the source of difficulty in English is consistent with the expected gap between command of expression in a foreign and a native language. Perception of logical doubts as more evident sources of difficulty in Italian than English suggests that the lesser demands on linguistic resources in the A language leave subjects more scope to focus on the nuances of content.

Thirdly, the finding that pauses considered as inexplicable are more prominent in English can be very tentatively interpreted as a consequence of the demands that control of expression makes on attentional resources. In other words, the perceived need for care with expression in English perhaps leads to lapses in overall coordination and control of delivery.
As a whole, the data on subjects’ perception of their pauses indicate that fluency depends on both linguistic and extralinguistic competences. The former, which subjects perceive as demanding greater attention in the B language, depend not only on linguistic knowledge but also on its availability (Gile 1995: 189 et seq.). This concept should not be ignored in any consideration of how far production in the target language can be appropriately described as automatic.

While the “théorie du sens” rightly highlights the importance of not limiting interpretation to a superficial rendering of the source speech’s wording, it can give the impression that formulation in the target language is an effortless adjunct to analysis of the input. Similarly, Weber (1989: 163) describes the consecutive interpreter’s target language output as practically automatic if the source speech has been well assimilated. By contrast, Thiéry (1981: 102) identifies a number of “opérations inhabituelles” that the consecutive interpreter must manage in the target speech (e.g., focus on the speech’s pragmatic impact, holding the floor in public). Similarly, the cognitive processes identified in Gile’s Effort Models include production among the non automatic efforts (Gile 1995: 97). This is consistent with Schmidt’s (1992: 376) statement that the phonological processing in speech can be automatised but higher level conceptual planning cannot. However, the distinction between automatic and non automatic components of speech production is to a certain extent “fuzzy”, in that “the controlled-automatic distinction should be viewed as a continuum rather than a dichotomy” (ibid.).

In the present study, subjects’ comments on expression difficulties in both the A and B languages suggest that target language formulation depends to a considerable extent on non automatic processes. Here, availability of relevant linguistic knowledge is probably an important factor in streamlining processes like lexical selection and reducing demands on attentional resources. This concept could explain in part why acquisition of fluency does not seem to progress uniformly or predictably as a function of overall linguistic knowledge.

Extralinguistic competences, like ability to listen analytically and to use notes (both while listening and during delivery of the target speech), arguably contribute just as much as language skills to the interpreter’s fluency. In addition, the use of strategies to manage difficulties should not be overlooked (Gile 1995: 129 et seq.). Though language teaching issues are not always relevant to interpreter training, classifications of language learners’ coping strategies identify a number of recurrent categories (e.g., omission, paraphrase, calques and borrowings, appeal to interlocutors) which overlap to a certain extent with those identified in interpreting (Ellis 1994: 397). Skill in unobtrusive use of such strategies is probably a major factor in fluency, whether in monolingual communication or in interpreting.
5. Conclusions

The findings of the present study afford tentative insight into trainee interpreters’ fluency in the A and B languages, as well as into factors on which this depends. However, the sample is relatively small and involves only two languages. Only with more extensive study on a broader range of language combinations will it prove possible to draw firmer conclusions.

In addition, the study should ideally be complemented by evaluation of fluency in native speakers of English with Italian as their B language. This would allow comparison of the two situations examined in the present study (source speeches in Italian and English, interpreted by subjects with a combination of “Italian A + English B”) with a further two variants (source speeches in Italian and English, interpreted by subjects with a combination of “English A + Italian B”). An investigation of this sort would make it possible to evaluate the possibility that the diverging pause profiles in the present study might be attributable to differences in rhythm between Italian and English, irrespective of their status as an “A” or “B” language.

On the basis of the present study, trainee interpreters’ fluency in the A and B languages seems to differ significantly. However, it is not suggested that this should dissuade them from using their B language. Improvement of fluency, with a view to closing the gap between the two languages, can be a challenging goal. An immediate lesson of the study is the scope for shifting the balance between silent and filled pauses, particularly in the B language. Close analysis of the problem triggers identified by individual subjects should also enable them to address their weaknesses by heightening sensitivity to which items of expression (e.g., links between speech segments, attenuation of over-bold statements) or general technique (e.g., legibility of notes) require attention.

Acknowledgment

I am grateful to Forlì colleague Giuseppe Nocella, without whose statistical expertise the data would have been far less amenable to interpretation and comment.

References


