SILENT PAUSES AND DISFLUENCIES IN SIMULTANEOUS INTERPRETATION: A DESCRIPTIVE ANALYSIS

By
Benedetta Tissi
Freelance Conference Interpreter

Literature on simultaneous interpretation (SI) contains many recommendations addressed to interpreters on the need to avoid pauses and interruptions in order to produce a more fluent high-quality delivery. Yet not many scholars have gone deeper into this subject and questions still remain open as to whether such occurrences depend on the interpreter, situational constraints of SI or the source-text (ST).

In order to start investigating this subject ten interpretation students were asked to translate simultaneously two text excerpts of a spontaneous speech containing pauses, hesitations and interruptions. It is the aim of this study to analyse whether and to what extent the presence of such occurrences in the ST affects the interpreter’s comprehension and delivery. Occurrences in both ST and target-texts (TT) were described in an attempt to produce a taxonomy which can be specifically applied to SI; pauses and interruptions were not seen as mere interruptions of fluency; whenever possible, their communicative value and/or tactical use were stressed.

1. SI as a composite form of communication

1.1. The implications of orality

Interpretation has been defined as a form of oral translation (Riccardi 1999: 161). Even though this seems quite obvious, it may be useful to point out the implications of orality for SI both as a process and as a form of communication. Interpretation studies on orality have dealt with its transience (Seleskovitch 1978: 14; Kopczynski 1980: 24; Gile 1995b: 68-69 in Straniero 1999: 109) or with its features at syntactic and semantic level, while the interactive dimension of oral communication as a pragmatic language variety has been completely neglected (Straniero 1999: 109). Since SI is essentially a form of communication, an approach focusing on this dimension would be suitable. One distinctive feature of spoken language is its limited planning, leading to fragmentary speech with false starts and repeats, Ricardi (1997: 62-63). Besides, spoken language is characterized by its paraverbal and non-verbal dimensions which have to be taken into account in SI, as Stenzl (1983: 40) writes:
While in translation the message is conveyed entirely by graphic means, interpretation involves not only linguistic elements and what they convey, but also intonation, voice quality, changes in pitch and loudness, pauses and non-linguistic elements […], which can all contribute to the message and may have to be verbalised by the interpreter.

The presence of prosodic and non-linguistic elements in the ST is important, but so too is their verbalisation by the interpreter. As a matter of fact those elements play a double role: in the ST for the interpreter’s comprehension and in the TT for the listener’s comprehension and the quality of the interpretation. Most scholars consider only the occurrence of non-fluencies in the interpreter’s delivery, without examining how the presence of such items in the ST could possibly influence his/her comprehension and thus presentation. Recommendations or pieces of advice by professionals or experts on the latter subject are very frequent; Viaggio (1992: 311) writes:

The interpreter […] must be duly conversant with the uses of oral speech – first and foremost intonation and pause management. […] He should be trained in maximizing the use of extra-linguistic clues and intonation in order to save breath; for instance, conveying modal information suprasegmentally.

Straniero (1999: 110) also points out the potential of a strategic use of prosody. The importance of correct prosody management in SI is widely accepted and has been included in the criteria for SI quality assessment (Kurz 1993; Viezzi 1999). Pöchhacker (1994) has developed a grid of prosodic parameters which have to be related both to ST and TT in order to achieve a well-balanced qualitative evaluation of an interpretation. One of the very few studies dealing with the ST-TT relationship from a prosodic point of view is Déjean Le Féal’s (1978). The author investigates the difficulties experienced by interpreters with a read ST in comparison to spontaneous ST, due to what she calls manque d’idéation chez le locuteur (lack of ideation by the speaker). This is typical of speakers reading their speeches and takes the form of weakened prosodic prominence and longer speech bursts between pauses, making ST comprehension by the interpreter more difficult. Déjean Le Féal’s work (1978) is significant because it is, to quote Stenzl (1983: 27), “a first approach to an interpretation-specific text typology”. Déjean Le Féal focuses on the influence of the ST on the interpreting process rather than on its product, which has yet to be analysed.
1.2. The implications of simultaneity

One of the typical constraints of SI is simultaneity of listening and speaking, which entails adaptation to the characteristics of the speaker’s delivery. This has to be taken into account in an attempt to analyse the pragmatic dimension of SI, as it involves prosody and the presence of pauses and interruptions in the interpreter’s delivery. Déjean Le Féal (1978), for example, attributes the lack of regularity and the occurrence of pauses in interpreters’ speeches to the need to anticipate ST items first, and check their conformity with the original speech afterwards. Though her hypothesis is confirmed by Stenzl (1983: 38) and Čeňková (1989: 53) it has to be considered that anticipation is just one out of a range of possible SI strategies (Kalina 1998; Riccardi 1999) and can therefore be seen only as a partial cause of irregular or fragmentary elocution. Kalina (1992: 253) sees another obstacle to the interpreter’s comprehension which can affect his/her delivery in what she defines as “the prolonged presence of the source text road signs”. Gile (1995: 97) confirms her view adding that simultaneity can sometimes make semantic and syntactic choices easier for the interpreter. In a similar perspective other SI researchers even suggest that pauses and non-fluencies could be exploited in a strategic way (Gringiani 1994: 38). Besides using them for monitoring his/her own anticipations, as mentioned above, the interpreter can produce filled micropauses (micropauses remplies) (Čeňková 1989) in order to slow down his/her delivery and concentrate on listening. Another strategy to cope with complex or temporally undetermined syntactic structures is what Setton (1999: 50) calls waiting, namely the insertion of short pauses at grammatical boundaries, in order to gain time without giving the listener the impression of omitting parts of the original message. It still has to be verified if a tactical use of non-fluencies can be systematically related to the presence of particular occurrences in the ST. For this purpose the ST-TT relationship in SI has to be examined in a situational and functional perspective.

1.3. The ST-TT relationship

Pöchhacker (1994: 205) closely examined this subject questioning TT-autonomy. In his view, particular features like the presence of both speaker and interpreter, the same communicative situation, traces and intrusions of the ST in the interpreter’s delivery speak in favour of an interdependence between the two texts. An implicit confirmation comes from Riccardi (1999: 161), who states that in SI language and text serve interpretation, having no autonomous purpose. So far, the ST-TT relationship has been analysed merely with regard to the verbal component of the text, namely through error grids based on the principle of informative equivalence between ST and TT (Barik 1969; Gerver 1974;
Lambert 1982 in Stenzl 1983). Mazzetti (1998: 4) deplores the scarce attention paid to the ST-TT correlation from the non-verbal point of view and states the need to acknowledge that successful interpretation depends not only on TT quality but also on ST language and presentation. Following Pöchacker’s (1994: 129) recommendation to single out as analytical criteria prosodic features which can be found both in ST and TT, this study aims at examining pauses and different types of interruptions, allowing both a quantitative and qualitative description.

2. Pauses and interruptions as elements of linguistic production and in SI research

The heterogeneous character of studies on this subject has already been mentioned and can be noticed even on an examination of the definitions and classifications of the occurrences. Some experts divide pauses into individual and functional pauses, others describe them as silent and filled pauses. Another quite vague question concerns the relationship between pauses and hesitations. In order to cast light on these issues, studies on the subject will be presented following the different approaches. It has to be stressed that in many of them non-fluencies are not the object of the study but only a useful tool to investigate psychological and cognitive mechanisms in communication and/or SI.

– Pauses as traces of cognitive activity

In the 1950s Goldman-Eisler found that the distribution of pauses in speech was not accidental, and she distinguished breathing pauses from hesitation pauses. On the basis of her first studies Goldman-Eisler (1958) concluded that the first lexical item after a silent pause is more difficult to predict than any lexical item uttered in a fluent context. Silent pauses of this kind are produced in order to gain time during the process of linguistic production. In a later work by the author (1961) the incidence of pauses is related to the cognitive effort required by the linguistic activity carried out. In this context she demonstrates that the number of pauses diminishes with the progressive automatisation of the task.

Maclay & Osgood (1959 in Martin 1967) link up with Goldman-Eisler’s early studies and propose a first classification of pauses and three types of interruptions, stressing their different functions. Silent pauses, filled pauses and repeats are used to take time for the choices required during language codification, whereas false starts are devices to correct what has been said immediately before. A later study by Tannenbaum, Williams & Hillier (1965) confirmed Goldman-Eisler’s (1958) findings about the low predictability of lexical words after a pause, but found that the word before a pause is equally difficult to predict. The explanation therefore lies in the different types of pauses
analysed in the two studies. Goldman-Eisler (1958) concentrated on silent pauses, whereas Tannenbaum et al. (1965) examined different types of occurrences and came to the following conclusion: silent and filled pauses are devices to take time before an increase of information, while repeats and false starts are produced to temporise before a correction. Therefore the less predictable word is located after silent and filled pauses but before repeats and false starts. This leads to the conclusion that an integrated perspective considering the type of non-fluency, its duration and localisation has to be applied.

– Pauses as functional items in a language system
Osgood (1954 in Suci 1967) was the first to assume the existence of functional units for information transmission. On this premise Suci (1967) claimed that pauses could be the boundaries of such functional units, by virtue of their non-casual distribution in speech (Goldman-Eisler 1958). He therefore defined minimal language units as items resisting progressive fracturing and demonstrated how speech segments between pauses fulfil this condition. He subsequently examined whether these psychological units were based on syntactic structure. After having carried out a series of experiments Suci concluded that as there is no correspondence between these units and traditional syntactic structure, pauses must reflect a different sort of structural organisation. Moreover, it had to be borne in mind that there are individual differences in the structuring of verbal material. Keseling (1992) too, examined pauses in this perspective and came to the conclusion that pauses are indeed functional elements, but they are not subject to a fixed system of rules.

– Pauses in communication
Pauses fulfil many roles in oral communication. The most visible is their influence on elocution speed: the higher the number of pauses, the lower the elocution speed. Pauses also contribute to the disambiguation of syntax (Mazzetti 1998) as in the case of compound words or word lists. Moreover, they contribute to discourse segmentation and help give prominence to particular text samples, drawing the listener’s attention to certain elements and making it easier for him/her to understand the message. In the latter case pauses have a predominantly stylistic and rhetorical function. As already stated, the use and the incidence of pauses are strongly characterized by the speaker’s individuality, both from a physiological and an emotional point of view.

– Pauses and speech reception
The listener’s reception of pauses and interruptions has not been closely examined yet, and the few scholars who have dealt with this subject support
diametrically opposite views, considering pauses as an obstacle for speech reception, as irrelevant for speech decoding or as elements making comprehension easier. Goldman-Eisler (1968: 14) is an exponent of the first group, claiming that:

a large proportion of pauses in spontaneous speech does not fit in with the linguistic structure and does not seem to serve communication, indeed it may at times impede rather than facilitate decoding.

Martin (1967), on the contrary, assumed that the speaker’s hesitations do not play a relevant role for the listener. He analysed how different subjects reproduce utterances they have just heard and observed that while in the speaker’s production pauses are mainly caused by hesitation, in the subjects’ reproductions they tend to coincide with the boundaries of grammatical units. In this way his starting point is confirmed and it is proved that the listener reorganises verbal material while decoding it. Keseling (1992) summarised the position of the third group, stating that pauses mark coherent passages to the receiver. At first sight the three positions appear mutually exclusive, but on a closer examination the different statements appear to be related to different types of pauses. The first and the second position refer to hesitation pauses, which are characteristic for linguistic production but have no function for reception. Indeed, in some cases they may compromise decoding. The third statement, on the other hand, refers to pauses with a signalling or even emphatic function.

– Classification attempts
It has already been mentioned that classifications are very heterogeneous in this field. Since it would be impossible to explain them all, only those considered relevant to the aim of the study will be described in this section. The first categorisation was Maclay & Osgood’s (1959 in Tannenbaum, Williams and Woods 1967), which divided occurrences into silent pauses, filled pauses, repeats and false starts and was adopted in many later studies on the subject (Tannenbaum, Williams & Hillier 1965; Tannenbaum, Williams & Woods 1967; Martin 1967; Duez 1982). Hieke (1981: 148) maintained the same occurrences but criticised what he defines Maclay & Osgood’s “concatenated approach”: “in this standard classification system, items receive joint attention only if they are in close proximity with each other”. He pleads for an “integrated approach” analysing interruptions and errors in a broader perspective. In his view hesitations are traces of two particular aspects of language production: prevention and correction of errors, which form the categories of “stalls” and “repairs” respectively. Repeats are divided into “prospective repeats” to take time for speech planning and “retrospective repeats” to correct errors or recreate
Silent Pauses and Disfluencies in Simultaneous Interpretation

a link with already uttered items. Hieke’s new taxonomy can be summarised as follows:
– stalls: silent pauses, filled pauses, prospective repeats and syllable lengthening;
– repairs: false starts, retrospective repeats, restoration of links.
Hieke (1981: 150) also introduced a completely new conception of hesitations:

Hesitations [...] form an integral part of speech production in the positive sense, a view quite in opposition to the attitude that there is fluency on one hand and hesitancy on the other. Not only are hesitations a normal component of fluency if they occur in moderation, but now pauses and the other hesitations can actually be considered wellformedness phenomena rather than disfluencies, at least as far as they serve as devices by the speaker to produce more error-free, high-quality speech.

This view is shared by Magno Caldognetto, De Zordi & Corrà (1982) whose categorisation is taken as starting point for the taxonomy adopted in this study. Occurrences are gathered under the hyperonym “non-fluencies” and divided into silent pauses and disfluencies, the latter including various types of items with different functions. The scheme below gives an overview of the different categories; those which have been maintained in this analysis will be defined in part 3.

<table>
<thead>
<tr>
<th>NON-FLUENCIES</th>
</tr>
</thead>
<tbody>
<tr>
<td>SILENT PAUSES</td>
</tr>
<tr>
<td>– initial pause</td>
</tr>
<tr>
<td>– juncture pause</td>
</tr>
<tr>
<td>– clause-internal pause</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

All non-fluencies can be found in all spontaneous speech, varying according to the speaker, his/her cognitive activity and a series of socio-linguistic variables. It is interesting to investigate the typology of non-fluencies in a complex task like SI.

– The double role of pauses in SI
One of the very few studies on pauses in SI is Čeňková’s *L’importance des pauses en interprétation simultanée* (1989). First of all she confirms that parallel listening and speaking by the interpreter is a given, though related to the speaker’s elocution speed and the number and duration of pauses in his speech.
She stresses the double role of pauses, namely in the ST to facilitate decoding and in the TT for segmentation by the interpreter. Her experiments confirm the importance of pauses for message segmentation but also the difficulty experienced by the interpreter in exploiting them because of their short duration. Čechková introduces a new definition of pauses, which are understood not only as interruptions in the flow of speech but also as significant variations in intonation and rhythm or even the juxtaposition of two semantically independent items. The pause is no longer an objectively quantifiable silence but rather an item depending on the interpreter’s and the listener’s perception.

3. Experimental study

The experiment, carried out at the SSLMIT of the University of Trieste, consisted of the SI into Italian of two tape-recorded spontaneous speeches delivered in German.

3.1. Aim of the study

This study aims at systematically analysing non-fluencies in ST and TT from the quantitative and qualitative point of view, in order to investigate possible correlations between the occurrences in the texts and to propose a SI-specific functional taxonomy of non-fluencies.

3.2. Materials

The STs adopted for the experiment were two excerpts from the tape recording of a speech delivered by an Austrian politician during a round-table meeting in Trieste on December 15th 1999. The excerpts lasting 3.18 and 5.85 minutes respectively, are both spontaneous speeches. As the incidence of non-fluencies is closely linked to on-line planning, which is typical of spontaneous discourse and SI (Déjean Le Féal 1978: 85) and given the aim of the study, it was considered essential for ST and TT to be comparable in this respect. Riccardi (1997: 67) pointed out the implications of a spontaneous ST for its interpretation, namely that it enables the interpreter to follow the development of the text and to make use of analogous strategies in his/her interpretation. As the texts were neither technical nor specialised, the participants were not given any information in advance.

After having delivered their SI, participants were asked to fill up a questionnaire concerning their impressions about elocution speed, fluency and the incidence of pauses in the ST as well as the occurrence of non-fluencies in their own deliveries.
3.3. Participants

The participants were ten interpretation students who had successfully completed their curriculum of SI exams from German into Italian. All students were Italian native-speakers chosen to avoid non-fluencies due to imperfect command of the language.

The experiment was carried out with students because, as stressed by Riccardi (1997), their deliveries tend to be much closer to the ST than those of professional interpreters. One could therefore suppose that the former are more influenced by the presence of non-fluencies in the ST. In order to point out differences in the degree of adhesion to the ST and possible implications for the occurrence of non-fluencies, a professional with three years’ experience was asked to interpret the ST under the same conditions.

3.4. Technical equipment

The deliveries of the 11 subjects were recorded with a Philips AAC 500 double-track machine.

Silent pause durations were measured with the Wave Studio software by Creative, version 4.06. It must be pointed out that even though a computer was used measurements had to be carried out manually, which inevitably precludes absolute precision.

3.5. Methods

Before describing the procedure of the experiment some methodological problems will be pointed out and the explanation of the relevant solutions provided.

The first problem concerned the measurement of elocution speed, namely the choice between words and syllables per time unit. Even though perfect correspondence can not be assured in either case, the latter solution was chosen because syllables are more language-independent than words. Having to compare elocution speed in two different languages, syllables appeared more suitable (Pöchhacker 1994: 132).

The main difficulty concerning silent pause measurement consisted in establishing a minimum threshold for measurement which would be possible with non-specialised technical equipment. In SI literature the following thresholds can be found: 0.18 seconds (Duez 1982), 0.25 seconds (Duez 1982, Goldman-Eisler 1968 and 1972), 0.3 seconds (Tannenbaum, Williams & Wood 1967). For this study a minimum value of 0.25 seconds was chosen in order automatically to exclude interruptions due to articulatory constraints (see
Goldman-Eisler 1968: 12). With the purpose of minimising the margin of error, silent pauses were classified into 9 categories per interval of 0.25 seconds. The upper threshold was established with reference to the longest pause in the ST, namely 2.4 seconds, included in the interval 2.25-2.50 seconds. The interpreters’ deliveries were characterised by much longer pauses; consequently two more categories were added: 2.5-5 seconds and 5 seconds or more. In this case larger intervals were chosen, considering that in presence of such atypical values for non-mediated spontaneous speech, 0.25-second differences were almost insignificant.

3.6. Procedure

The deliveries of the participants were transcribed as faithfully as possible, listening to the recordings. In case of doubts an expert was asked for advice. Afterwards, silent pauses were measured and their durations in brackets were inserted in the transcriptions (see appendix). Lastly, non-fluencies were counted and divided according to the categories described below. For the purpose of a qualitative analysis a few ST samples containing significant non-fluencies were chosen, and their counterparts in the TTs were described and scanned for non-fluencies.

3.7. Analysis scheme

<table>
<thead>
<tr>
<th>SILENT PAUSE</th>
<th>NON-FLUENCIES</th>
<th>DISFLUENCIES</th>
</tr>
</thead>
<tbody>
<tr>
<td>- grammatical and/or communicative pauses</td>
<td>FILLED PAUSES</td>
<td>INTERRUPTIONS</td>
</tr>
<tr>
<td>- non-grammatical pauses</td>
<td>- vocalised hesitations</td>
<td>- repeats</td>
</tr>
<tr>
<td></td>
<td>- vowel and consonant lengehings</td>
<td>- restructuring</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- false starts</td>
</tr>
</tbody>
</table>

This taxonomy is based on the categorisation by Magno Caldognetto, De Zordi & Corrà (1982), revised in order to meet SI-specific features. As can be seen in the scheme, non-fluencies were divided into two macrocategories which can easily be distinguished at the level of perception, namely silent pauses and disfluencies, the latter including filled pauses and interruptions. The three categories of silent pauses, filled pauses and interruptions are made up of different types of occurrences. With regard to Magno Caldognetto, De Zordi &
Corrà’s (1982) taxonomy some changes were introduced. They are illustrated below, together with the categories of the new classification proposal:

A) Silent pauses

These are the only occurrences corresponding to an interruption in language production and can therefore be measured objectively. Each silence between two articulated sequences has been considered a silent pause. The initial pause, namely the silence before starting with the linguistic task, was left out, considering that SI requires by definition a certain time-lag which can not be considered a non-fluency. The criterion of pause omission was introduced, as it was considered that the absence of pauses necessary for correct text segmentation could hinder comprehension (see also Mazzetti 1998).

Silent pauses were simply divided into grammatical and/or communicative pauses and non-grammatical pauses. In comparison with the original classification, communicative pauses were introduced by virtue of the important role they play in spontaneous speech and SI, giving prominence to discourse items.

B) Filled pauses

a) Vocalized hesitations

This class includes all vocalized expressions of hesitation, which have been transcribed as üh, ühm, mm for German and eh, ehm, mm for Italian, regardless of their duration. As the interpreted versions contained many brief hesitations sounding like a schwa, they were transcribed as ə in order to distinguish them from eh. Glottal clicks and guttural sounds were also included in this category.

b) Vowel and consonant lengthenings

Although not included in the reference taxonomy, the category of vowel and consonant lengthenings is not new in studies on non-fluencies (see Magno Caldognetto, Vagges & Job 1983). These occurrences were included in the taxonomy because they are typical of spontaneous speech and by virtue of their high incidence in the interpreters’ deliveries.

C) Interruptions

Interruptions include many types of occurrences and sometimes it may be difficult to decide in which category they have to be ranged. In such cases an expert was asked for advice.
a) Repeats

These occurrences include non-semantic repetitions of a phrase, word, or even part of a word. All repetitions with a stylistic or rhetorical function were of course excluded from this category.

b) Restructuring

Restructuring can be defined as an utterance rectifying what the speaker has just said (phrase, word or part of a word). This category includes corrections of phonological lapse as well as formulation or content errors.

Another subgroup includes structure reformulations, that is to say when the speaker decides to express the same concept with a new formulation.

Finally, there is the sub-category of syntactic mixtures, occurring when the speaker juxtaposes two syntactically incompatible structures. The difference between structure reformulations and syntactic mixtures is that in the latter case the speaker does not explicitly try to rectify his/her previous utterance.

c) False starts

False starts occur when the speaker interrupts an utterance and begins a new one without having completed it.

One of the guidelines for the application of this scheme was Hieke’s integrated approach (1981), which consists of the analysis of the broad context in which non-fluencies occur. In this way attention is not only focused on the immediate surroundings of each occurrence but it is linked to the whole sentence structure and the communicative context.

4. Results and discussion

In accordance with the aim of this study the discussion will be carried on in two directions: ST-TT comparison on the one hand and the description of common trends and differences in the interpreters’ deliveries on the other. Whenever possible the findings are displayed in tables. Because of the limited sample examined, the average value for TT1 and TT2 can not always be considered significant, especially when there are great differences among the interpreters. Therefore it has sometimes been omitted, while TT minimum and maximum values for each item are always reported in brackets.
A) General text features

<table>
<thead>
<tr>
<th>Text features</th>
<th>ST1</th>
<th>TT1</th>
<th>ST2</th>
<th>TT2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text duration (min)</td>
<td>3.18</td>
<td>3.15</td>
<td>5.85</td>
<td>5.80</td>
</tr>
<tr>
<td></td>
<td>(3.02 – 3.22)</td>
<td>(3.02 – 3.22)</td>
<td>(5.68 – 5.87)</td>
<td>(5.68 – 5.87)</td>
</tr>
<tr>
<td>Silent pause duration (min)</td>
<td>0.57</td>
<td>0.97</td>
<td>1.60</td>
<td>1.94</td>
</tr>
<tr>
<td></td>
<td>(0.57 – 1.31)</td>
<td>(0.57 – 1.31)</td>
<td>(1.30 – 2.57)</td>
<td>(1.30 – 2.57)</td>
</tr>
<tr>
<td>Elocution speed</td>
<td>267</td>
<td>222</td>
<td>208</td>
<td>200</td>
</tr>
</tbody>
</table>

Table 1

As shown in table 1, ST1 duration and TT1 average duration are almost the same, whereas TT1 values for each subject vary from 3.02 to 3.22 minutes. The lowest finding shows a 0.16 minute divergence from the average, compared to the 0.04 of the highest value. For the second text, too, the duration of the ST is slightly higher than the TT average. Here again the highest TT value (5.87 min) is closer to the average than the lowest (5.68 min). Shorter TT may be due to a shorter speech, to shorter and/or fewer pauses or to omissions with respect to the original text. In order to interpret these findings correctly pause durations have to be considered, too. Total pausing time is longer in the interpreters’ texts both for T1 and T2. As a first consequence it can be stated that interpreters spoke for less time than the speaker, although the findings about silent pause incidence will also have to be considered for an exhaustive analysis. In this context it has to be stressed that only one of the interpreters shows comparable trends in T1 and T2, which leads to the conclusion that the students’ pausing modalities are not deliberate but rather the result of comprehension or reformulation difficulties related to features of the ST form (and/or content). The third feature considered beside text and pause duration was elocution speed. Its values are 222 syllables per minute in TT1 against 267 syll/min in ST1. The slower rhythm of the interpreter’s deliveries is related to the longer average duration of silent pauses. In T2 elocution speed is 208 syll/min in ST and 200 syll/min on average in TT. Here values are much closer, but again the ST is delivered at a higher speed than the TT. For a correct interpretation of these findings it must be considered that ST and TT are produced in two different languages, which may at least partially explain some of the divergent results. As far as TTs are concerned, extreme values for each item indicate that there are great individual differences among the interpreters’ productions. It should be tested whether the recipients of SI prefer fewer and longer pauses or a series of shorter occurrences which could facilitate the segmentation of the message. Though it goes beyond the aim of this study to answer this question, the importance of an integrated perspective considering the interaction of different variables should be stressed.
B) Silent pauses

<table>
<thead>
<tr>
<th>Silent pauses</th>
<th>ST1</th>
<th>TT1</th>
<th>ST2</th>
<th>TT2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of pauses</td>
<td>56</td>
<td>46.8</td>
<td>130</td>
<td>86.4</td>
</tr>
<tr>
<td>Average duration (sec)</td>
<td>0.609</td>
<td>–</td>
<td>0.739</td>
<td>–</td>
</tr>
<tr>
<td>Longest pause (sec)</td>
<td>1.653</td>
<td>–</td>
<td>2.400</td>
<td>–</td>
</tr>
</tbody>
</table>

Table 2

Table 2 summarizes silent pause values. In both texts the average number of pauses in the TT is lower than that in the ST: 46.8 versus 56 in T1 and 86.4 versus 130 in T2. The values of the second text show a greater difference which clearly emerges if the number of pauses per minute is calculated for each text: 14.86 pauses/min (TT1) versus 17.61 (ST1) and 14.90 (TT2) versus 22.22 (ST2) respectively. In spite of a significant difference between the STs, the TT values are almost identical, which implies that regardless of the number of pauses per minute in the original texts, the interpreter produced the same amount of utterances per minute in both TTs. It can be observed that the highest number of pauses in TT is only slightly higher than that of the speaker in the first text (57 vs. 56 pauses), whereas it is substantially lower (106 vs. 130 occurrences) in the second. This finding confirms the previous conclusion that the speaker produced more pauses than the interpreters, which can be attributed to the fact that he was delivering a spontaneous speech. Further explanation is provided by an analysis of the average pause duration, which is higher in the interpreters’ deliveries. Even if no average value could be calculated because of the great differences among the participants, it clearly emerges that the interpreters’ average pauses are much longer than the speaker’s, reaching 2.156 and 2.448 seconds in TT1 and TT2 respectively. The lowest value, on the other hand, almost coincides with that of the ST for T1 (0.605 sec vs. 0.609 sec) and lies above it for T2 (0.833 sec vs. 0.739). The findings are even more striking if the longest pause values for each text are compared. Here again, it was not possible to calculate an average value because of very divergent values. Both in TT1 and TT2 even the lowest maximum pause duration is considerably higher than the longest occurrence in the relevant ST: 2.326 sec vs. 1.653 sec for T1 and 3.537 sec vs. 2.4 sec for T2. The highest values in the TTs exceed 12 and 11 sec respectively. This difference between original texts and interpreted versions is illustrated in the figure below.
Figure 1: Pause distribution in ST and TT 1 and 2

Figure 1 reproduces the distribution of silent pauses along duration intervals from 0.25 sec to 5 sec and over. In spite of a slight difference for ST2 there is a parallel trend in all the texts: there is a high incidence of occurrences in the intervals from 0.25 to 1.25 sec and a much lower one for the following intervals until 2.50 seconds. This finding leads to the conclusion that normal pauses fall within 0.25 and 1 sec, both for spontaneous speech and SI. The difference between the two modalities lies in the remaining two classes, which have been created ad hoc for TT occurrences. The significant number of pauses they contain is a consequence of SI-specific constraints and of the fact that the participants were students, still experiencing difficulties in attention-sharing and comprehension.

Another criterion of analysis was the distinction between grammatical/communicative pauses and non-grammatical pauses. The following results emerge from a ST-TT comparison: 57% grammatical/communicative pauses in ST1 versus a TT1 average of 52% and 58% against 57% in ST2 and TT2 respectively. On the whole, the percentage of “correct” pauses is slightly lower in the interpreted texts than in the ST, which is probably due to a lack of control of linguistic production typical of SI.

C) Disfluencies
a) Filled Pauses

<table>
<thead>
<tr>
<th>Filled pauses</th>
<th>ST1</th>
<th>TT1</th>
<th>ST2</th>
<th>TT2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocalised hesitations</td>
<td>12</td>
<td>–</td>
<td>13</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0 – 34)</td>
<td></td>
<td>(1 – 53)</td>
</tr>
<tr>
<td>Vow. and cons. lengthenings</td>
<td>1</td>
<td>57.5</td>
<td>2</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(18 – 99)</td>
<td></td>
<td>(66 – 161)</td>
</tr>
</tbody>
</table>

Table 3

Table 3 shows that the number of vocalised hesitations is almost the same in both ST1 and ST2, meaning that the number of occurrences per minute is much
lower in the second text: 0.45 versus 3.77 in the first. The incidence of these occurrences in the TT is very heterogeneous, ranging from no occurrences to 34 in TT1 and from 1 to 53 in TT2. Since it would have been pointless to calculate average values, occurrences were ranged in 3 classes: 0 to 10 items, 11 to 20 and 20 or more. For both TTs most of the interpreters’ deliveries came into the first class, whereas only a few participants produced a high number of vocalised hesitations. As can be seen in table 5, the participants produced a similar number of utterances in both texts, which confirms that the incidence of these occurrences is highly individual.

The picture concerning vowel and consonant lengthening is quite heterogeneous, too: only 1 and 2 occurrences respectively in the ST against average values of 57.5 in TT1 and 112 in TT2. The patent difference between ST and TT may also be due to particular language features and to the individuality of the speaker. Here again it may be useful to have a closer look at values in the interpreters’ deliveries, ranging from 18 to 99 in TT1 and from 66 to 161 in TT2. But even in-between occurrences are very divergent, as shown in table 5. The significant presence of syllable lengthenings in interpreted texts may also be due to their tactical use described by Čeňková (1989), mentioned in part 1. In her view syllable lengthenings allow the interpreter to concentrate on listening, or to bridge a gap while waiting for new material to translate. Of course, the application of such strategies is very subjective, as confirmed by the significant differences between the participants. Finally, another particularity has to be stressed: if supported by intonation, syllable lengthenings may have a function of discourse scansion. After a lengthening at the end of a word (which expresses hesitation), the lengthening of the tonic vowel of a following word has the function of drawing attention back to an important informative element (often coinciding with a lexical item). Such vowel lengthenings have a communicative value.

b) Interruptions

Table 4 illustrates the incidence of different types of interruptions. ST1 and ST2 present 8 and 9 repeats respectively, whereas occurrences range from 0 to 3 in TT1 and from 1 to 8 in TT2. The difference between ST an TT is striking, especially in the first text, but also in the second, since 7 participants produced from 1 to 4 occurrences only (see table 5). The reason for such divergent findings is that the STs are spontaneous speeches with on-line planning. TTs are of course planned on-line too, but in SI some parts of the process, such as lexical or semantic choices, are facilitated (Gile 1995). As he/she keeps a certain time-lag from the original discourse, the interpreter recognises repeats in the ST and avoids reproducing them. One could even assume that such occurrences
facilitate the interpreter’s task, since the repeated item does not overload his/her memory and delays the following words, allowing him/her to concentrate on delivery.

Restructuring is typical for spontaneous speech and is even more frequent in SI, where the interpreter does not fully control the progression of discourse, as he/she gets knowledge of it in a fragmentary way. On the other hand, time-lag should provide enough time for the interpreter to recognise such occurrences and avoid reproducing them. Restructurings are limited both in ST and TT and there is no link between the occurrences in the two modalities. There are 5 occurrences in ST1 and 7 in ST2, while their interpreted versions include on average 5.8 and 10.8 occurrences respectively. Here again, great differences can be noticed among the interpreters: from 0 to 10 restructurings in the first text and from 1 to 22 in the second. Because of the great variety of possible restructurings only a detailed qualitative analysis could explain their causes and evaluate their implications for text fluency.

False starts are the class with the lowest number of disfluencies. There are no occurrences at all in the ST and from 0 to 1 and 0 to 2 in the TTs. From a closer look at table 5, it can be seen that only 3 and 4 subjects respectively produced false starts in their deliveries. Only one interpreter made false starts in both T1 and T2, which indicates that such occurrences are not an individual characteristic but rather the result of difficulties in coping with the original text. In SI a false start can be the consequence of comprehension difficulties or of a wrong anticipation. Generally speaking, false starts are the result of a lack of control of linguistic production. This is confirmed by the frequent co-occurrence of false starts with other non-fluencies. Nevertheless, only a deeper analysis of each case can provide precise answers.

<table>
<thead>
<tr>
<th>Interruptions</th>
<th>ST1</th>
<th>TT1</th>
<th>ST2</th>
<th>TT2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeats</td>
<td>8</td>
<td>–</td>
<td>9</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0–3)</td>
<td></td>
<td>(1–8)</td>
</tr>
<tr>
<td>Restructurings</td>
<td>5</td>
<td>–</td>
<td>7</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0–10)</td>
<td></td>
<td>(1–22)</td>
</tr>
<tr>
<td>False starts</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0–2)</td>
<td></td>
<td>(0–1)</td>
</tr>
</tbody>
</table>

Table 4

D) Common trends in TTs

While describing the incidence of the different occurrences in the TTs, the gap between the interpreters’ lowest and highest values has been stressed and great individual variations have been mentioned. Table 5 summarizes the values of
each interpreter for the most significant categories. The values contained in it confirm the presence not only of significant differences but also of a kind of parallel behaviour by the interpreters in both T1 and T2. Parallel trends of this kind can be observed especially in the highest and lowest silent pause duration and the number of vocalised hesitations and syllable lengthenings. A perusal of table 5 shows further similarities of this kind, which are certainly an interesting starting point for further investigation on the subject.

<table>
<thead>
<tr>
<th>Occurrence/Subject</th>
<th>N.</th>
<th>Max.</th>
<th>n. voc.</th>
<th>n. syll.</th>
<th>Repeats</th>
<th>Restructuring</th>
<th>False starts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Int. 1</td>
<td>TT1</td>
<td>47</td>
<td>8.433s</td>
<td>1</td>
<td>45</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Int. 1</td>
<td>96</td>
<td>6.052s</td>
<td>1</td>
<td>107</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>TT1</td>
<td>51</td>
<td>5.655s</td>
<td>1</td>
<td>55</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>TT2</td>
<td>93</td>
<td>7.743s</td>
<td>4</td>
<td>111</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Int. 2</td>
<td>TT1</td>
<td>47</td>
<td>7.808s</td>
<td>13</td>
<td>37</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>TT2</td>
<td>89</td>
<td>8.264s</td>
<td>11</td>
<td>73</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Int. 3</td>
<td>TT1</td>
<td>42</td>
<td>4.363s</td>
<td>28</td>
<td>48</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>TT2</td>
<td>82</td>
<td>5.515s</td>
<td>33</td>
<td>108</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Int. 4</td>
<td>TT1</td>
<td>52</td>
<td>2.326s</td>
<td>12</td>
<td>65</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>TT2</td>
<td>92</td>
<td>3.537s</td>
<td>28</td>
<td>122</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>Int. 5</td>
<td>TT1</td>
<td>37</td>
<td>12.643s</td>
<td>2</td>
<td>18</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>TT2</td>
<td>63</td>
<td>11.342s</td>
<td>5</td>
<td>66</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Int. 6</td>
<td>TT1</td>
<td>57</td>
<td>2.388s</td>
<td>12</td>
<td>99</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>TT2</td>
<td>92</td>
<td>4.460s</td>
<td>13</td>
<td>147</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Int. 7</td>
<td>TT1</td>
<td>57</td>
<td>6.039s</td>
<td>0</td>
<td>85</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>TT2</td>
<td>106</td>
<td>4.673s</td>
<td>8</td>
<td>144</td>
<td>5</td>
<td>22</td>
</tr>
<tr>
<td>Int. 8</td>
<td>TT1</td>
<td>34</td>
<td>7.676s</td>
<td>34</td>
<td>88</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>TT2</td>
<td>67</td>
<td>8.697s</td>
<td>53</td>
<td>161</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>Int. 9</td>
<td>TT1</td>
<td>44</td>
<td>3.750s</td>
<td>2</td>
<td>35</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>TT2</td>
<td>84</td>
<td>7.360s</td>
<td>5</td>
<td>80</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 5

5. Conclusions

From the results of the analysis and description of silent pauses and disfluencies in ST and TT the following conclusion can be drawn: occurrences in ST and in the interpreters’ deliveries are certainly related. Yet the data collected confirm that the influence is not as direct as one could assume. TTs contain fewer, but altogether longer silent pauses than ST, as confirmed by their average duration and particularly by the longest pause values in the interpreters’ speeches. In addition, TTs have a slightly higher number of grammatical pauses, while the only parallel trend for these occurrences in ST and TT seems to be their distribution within the different duration-intervals. Both ST and TT have a much
Silent Pauses and Disfluencies in Simultaneous Interpretation

higher incidence of occurrences in the intervals from 0.25 to 1.25 seconds. Unlike the original texts, TTs show a remarkable incidence of pauses from 2.5 to 5 and from 5 seconds up, which do not even exist in ST. Another difference concerns vowel and consonant lengthenings, which are much more numerous in the interpreted texts. The values concerning interruptions indicate the prevalence of repeats in ST, while false starts occur only in TTs. As far as filled pauses and restructuring are concerned it is difficult to point out general trends because of marked differences among the interpreters. Altogether this heterogeneous picture speaks against the presence of a systematic correspondence of non-fluencies in ST and TT. Occurrences in the participants’ productions were very often due to problems of technique, namely unbalanced attention sharing, which may compromise the listening phase and lead to a fragmentary delivery. Some difficulties can be attributed to the fact that the subjects were students, but the presence of many non-fluencies in the ST surely does not make their task easier. Even if there is no formal correspondence, some occurrences in TT are caused by the need to wait for new items, delayed because of a speaker’s pause and/or interruption. Besides, the fact that there is no systematic correspondence between occurrences in the original text and its interpreted version does not mean that there is no link at all. The description of text samples has shown that certain non-fluencies in ST may cause hesitation in the speaker which, though expressed through a different type of occurrence, is still clearly related to the ST item. In many cases it is useful to keep a certain time-lag from the original text in order to reformulate or even summarise the message, a strategy adopted here by the professional interpreter (although it must be stressed that differences between this participant’s and the students’ deliveries concerned informative completeness rather than the presence/absence of disfluencies). A finding emerging from a comparison among TTs is the strong subjectivity in the production of non-fluencies by each participant. Further investigation is required to find out if a sort of individual pattern exists in the use of silent pauses and disfluencies. Many findings support this hypothesis: the maximum and minimum values for pause duration, longest pause, filled pauses and syllable lengthenings are reported by the same participants in both TT1 and TT2.

Another important result of this study is the communicative, sometimes even strategic use of some non-fluencies by the interpreters. The most common examples are: silent or filled pauses before a correction, which give salience to the rectified item, lengthenings on the tonic vowel, contributing together with intonation to draw attention to the lexical item and finally retrospective repeats, re-establishing a connection with an interrupted utterance.

An important matter with reference to non-fluencies in SI is how they are perceived by the listener. As Viezzi (1999: 150) observes, only an in-depth
analysis of the communicative situation and the receivers’ characteristics and
expectations will allow the interpreter to use the best strategies for a successful
interpretation in this sense.

It is therefore impossible at this stage to identify clear trends in the incidence
of non-fluencies in the interpreters’ productions. Further experiments with larger
samples will have to be carried out in order to draw significant conclusions.
Besides numerical findings this study offers some indications for continued
exploration of non-fluencies as relevant strategic tools for the achievement of
the goals of a particular communicative situation. The following approach by
Enkvist & Björklund (1989: 325 in Pöchhacker 1994: 136) could be extended to
SI research:

We are [...] looking at hesitations and structure shift, not as instances of
regrettable human imperfection but rather as important devices helping
people to manage in spite of the strains inherent in real-time processing
of discourse. They are worth studying as integral parts of spoken
communication and discourse which help speakers to maintain cohesion
and coherence and to adapt their text to the requirements of receptors and
situations.

Appendix

Following symbols were used in the transcription:

x: vowel or consonant lengthening
( ) silent pause duration in seconds
* glottal click
[ ] paralinguistic remark (cough, etc.)
äh, ähm brief vocalised hesitation (like a schwa)
mm vocalised hesitations in the German text
vocalised hesitation in both German and Italian text
eh, ehm vocalised hesitations in the Italian text
Ø clause or sentence omission in the interpreter’s text
xxx repeat (of a letter, word or phrase)
xxx correction
xxx double correction (counted as 2 corrections)
xxx false start

SPEAKER: [...] dieses Element wird meines Erachtens überschätzt (1.068) es ist
ein Element unseres Erfolgs es ist aber nicht das Element das entscheidende
Element sondern das ist ein (0.985) ein ein ich würde sagen ein ein
Bausteinssystem ein Modul (0.461) in einem Erfolgskonzept das (0.457) äh nicht ausschließlich auf der Frage der Zuwanderungspolitik aufbaut [...] 
INT 1: [...] l’elemento dell’immigrazione della politica dell’immigrazione è stato sopravvalutato è uno degli elementi ma non certo quello fondamentale (3.452) si tratta di: una compone:nte di un sistema modula:re potremmo dire ma (1.817) non è certo: (0.440) il punto fondamentale:le [...] 
INT 2: [...] un altro eleme:nto un altro moti:vo della no della forza del nostro partito (0.318) è: (0.731) rappresentata: (4.035) dal fatto che: (5.655) abbiamo: (0.763) assunto: (0.350) un elemento: (0.286) una posizione: particolare in merito all’immigrazione (3.590) la nostra politica: (0.636) sull’immigrazione non è nulla di spettacolare:re [...] 
INT 3: [...] è: uno degli elementi del nostro elemen uno dei nostri (0.287) ehm degli'elementi che ci ha permesso di vi:ncere le elezioni non l’unico ma uno degli’elementi (1.928) abbiamo: (1.409) propo:sto (0.916) un'a nostra politica che non si concentra solame:nte sulla politica d’immigrazione:ne [...] 
INT 4: [...] questo (0.746) è un elemento molto importante della nostra politica ma non l’elemento più importa:nte (1.347) direi: che ehm abbiamo elaborato un mod:ulo per quanto riguarda l’immigrazione che non si concentra solamente sull’immigrazione sulla politica dell’immigrazione (0.385) come ta:le [...] 
INT 5: [...] anche se secondo me questo motivo viene (0.365) purtroppo sopravvaluta:to non è di certo l’elemento decisivo è solo uno dei tante eleme:nti (1.254) io dire:i che: è solamente un: modulo (0.317) nel modul nel nostro concetto di succe:sso (2.326) comunque la nostra politica: α degli’immigrati viene spesso sopravvalu:ta [...] 
INT 6: [...] questo è un elemento molto importa:nte (1.484) è uno degli’elementi che ha contribuito alla nostra vittoria ma non l’elemento decisivo fondamentale (4.040) dunque è uno degli’elementi che ha: (0.470) costituito (0.460) il nostro successo ma non l’elemento decisivo [...] 
INT 7: [...] è un elemento del nostro successo sicuramente ma non è l’unico: (1.662) anzi (1.306) direi che si tratta (0.373) di un mo:du:lo ne in un conce:itto di: successo (0.484) che: (1.170) perché la nostra politica non si interessa solo * non è: solo politica d’immigrazione: [...] 
INT 8: [...] ovvero questo eleme:nto secondo me è sopravvalu:ta:to (0.680) è un elemento del nostro succe:ss:so ma non è l’unico eleme:nto l’elemento decisi:vo (2.446) direi (0.567) che (0.731) si ha è una struttu:ra che si compone di tanti pe:zzi (1.286) e che non: (0.290) consiste solo nella problematica dell’immigrazione [...] 
INT 9: [...] qui tutti questi eleme:nti sono quelli da prendere in considerazione:ne (0.531) per: eh giustificare (0.258) il: mia vittoria non è s:olo quest’ultimo elemento è un:o degli eleme:nti eh della mia vitto:ria (4.089) e quindi non ci sì è concentrati solamente sulla questio:ne della: politica d’immigrazione: [...]

Silent Pauses and Disfluencies in Simultaneous Interpretation 123
INT 10: [...] (0.792) quest’ultimo elemento per no secondo me è stato sopravvalutato:to è uno degli’elementi che ha portato (1.105) a votare:rci (0.927) fa parte (0.568) di un progetto (0.718) di successo del nostro parti:to che non si basa soltanto sulla politica dell’immigrazione:ne [...]  
INT 11: [...] questo elemento quindi a mio avviso viene sopravvalutato (0.816) è sicuramente un elemento del nostro successo ma non è l’elemento fondamentale (3.784 *) diciamo: che la nostro: concetto di successo è stato costituito da vari (0.278) a elemen:ti (0.297) tra cui l’immigrazione non è stato: sicuramente l’aspetto più: importante [...]  

References  


