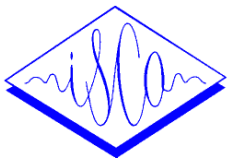

Proceedings of DiSS'05, Disfluency in Spontaneous Speech Workshop.
10–12 September 2005, Aix-en-Provence, France, pp. 151–155.

Disfluencies as a window on cognitive processing.

An analysis of silent pauses in simultaneous interpreting

Myriam Piccaluga[*], *Jean-Luc Nespoulous*[**], *Bernard Harmegnies*[*]

[*] Université de Mons-Hainaut, Mons, Belgique
[**] Université de Toulouse-Le Mirail, France

Abstract

The paper focuses on silent pauses observed in the productions of subjects involved in simultaneous interpreting tasks. Four bilingual subjects with various degrees of expertise in interpreting and various degrees of mastery of the languages involved (French and Spanish) have been recorded while interpreting utterances of French and Spanish talks. The source discourses had been perturbated by changes both in speech rates (by time compression) and in auditory quality (by addition of a parasiting noise). On the basis of acoustical analyzes performed on the subjects' productions, statistical analyzes focus both on the number and on the duration of the observed pauses. This double approach enables investigations of the kind of cognitive disturbances caused by the independent variables and allows further speculation on the semiology of the pauses durations.

1. Introduction

It is well known that a vocal signal used by a speaker in order to communicate is not produced at a constant rate but, on the contrary, involves many solutions of continuity. These disfluencies may take many forms (repetition, self-correction, vocalic lengthening, etc.). In this study we will focus on phenomena of that kind that are characterized by a break in the phonic flow: in other words, we will focus on *silent pauses* (as opposed to 'ums', hesitation phenomena and other 'filled pauses').

Authors who have investigated this domain generally agree that these breaks in the flow of speech take various forms. Some pauses result from simple physiological necessity (breathing) or articulatory requirements (linked to the realization of certain phonemes: the silence that precedes the release of a plosive, for example); some have a semantic function, and contribute to the discourse strategy of the speaker (delimiting sense groups, for example); still others are linked neither to articulatory mechanisms nor to segmentation in sense units, are not part of the communicative intention of the speaker, but stem from difficulties in speech production which can arise at any point in the speech production cycle (while planning the utterance or accessing the lexicon or even during phonetic implementation, etc.).

These interruptions, whatever their form and cause, are therefore inherent in all ordinary speech production. However, in situations where speech production is combined with another task, breaks in the flow of speech may be linked to imperatives other than those which usually lie behind speech production.

Simultaneous interpreting is an especially interesting task, when seen from this point of view, as it combines simple production of the target language with the preliminary tasks (which appear to an unsuspecting onlooker to be performed simultaneously) of understanding the source language and translating from source to target. Thus, for the interpreter, the production of speech signal is only the last link in a chain of complex cognitive processes.

During the interpreting process, various factors can affect the pauses, or cause them to appear when, under ordinary conditions, they would not be observed. Silences may be linked to difficulties in comprehending the initial message, and to various processes, such as searching for an equivalent translation of the source term in the target language, and indeed difficulty in expressing the concepts in the target language. In other words, it would be reasonable to think that the proliferation and/or lengthening of these pauses is linked to difficulties of various kinds in performing the complex task of interpreting.

This may seem obvious, but it must be observed that few of the researchers who have taken an interest in the process of interpreting have carried out systematic or sustained studies of pauses. Some studies of pauses are carried out on the basis of purely subjective analysis [2]. On the other hand, when a study is based on acoustic analysis, a summarizing approach is usually taken, designed to measure global indices such as, for example, the ratio of the total duration of pauses to total phonation time [1,6,8,9], and extensive studies like those carried out on other speech situations [3,4] are sorely lacking in this field.

In this article we shall employ a dual analytical approach (centered both on pauses frequencies and duration) which will aim to target accurately the characteristics of pauses which appear in the speech of subjects involved in an interpretation task, with the goal of shedding light on the cognitive functions which are active while the subject is interpreting.

2. Method

2.1. Subjects

There were four subjects. All are from the Barcelona area, where they had been living for several decades. They were all female, had a good command of French and Spanish, and spoke both languages regularly. However, for all four subjects, Spanish was the dominant language. Aside from these similarities, the subjects differed in their interpreting expertise and in their command of their languages.

Two of the subjects ('*int1*' and '*int2*') were professional interpreters, with less than 5 years' and 20 years' professional experience respectively. One subject ('*stud*') was in her third year of translation and interpreting studies. The other ('*biling*') had no experience of interpreting. Moreover, their command of French (and therefore their degree of bilingualism) was variable: the two subjects with the most balanced bilingualism were *int2* (very early contact with French, which she used very frequently with family) and *biling* (a teacher of French as a foreign language in a Spanish university); the two other subjects use French only for professional purposes, in interpretation situations.

2.2. Source corpora and linguistic combinations

Each subject was asked to perform six interpreting tasks, each consisting of interpreting a conference speech which was originally given on the floor of the European Parliament and then re-recorded in a laboratory by expert native speakers. The speeches were on issues of general policy and did not involve specialized vocabulary. Each subject had to interpret three speeches in each of the two combinations (French to Spanish and Spanish to French).

2.3. Disruptions introduced into source corpora

Disruptions were introduced into each text in the laboratory. These were: firstly, a local alteration in the speed of the source speech (increasing the rate through a reduction of 80%, 70% or 60% of the total duration of speech without modifying the characteristics of F_0); and secondly, local addition of noise interference (0 dB, 3 dB or 6 dB with respect to the average level of the source discourse).

Each source speech was divided into one-minute portions, without the subject's knowledge; portions 2, 4 and 6 were affected by disruptions, while the other portions were unaffected. Noise interference and temporal compression were applied to each disrupted portion. The levels of independent variables *compression* and *noise interference* were organized according to a Latin square design, which was designed to balance out the associations between levels of the independent variables, and to neutralize the possible semantic effect of the content of the speech.

2.4. Acoustic processing

The subjects performed their interpreting tasks at the Faculty of Translation and Interpreting, Barcelona Autonomous University. The results were analyzed acoustically in the phonetics laboratory at the University of Mons-Hainaut. All acoustic analysis was carried out using Multispeech software from Kay Elemetrics. Pauses were detected using a pitch analysis algorithm, with manual corrections based on interactive listening and examination, at relevant points, of narrow-band spectrograms. The data collected were converted into Excel spreadsheets. For the purposes of the present study, we will only include pauses which could legitimately be considered [4, 5, 7, 10] as neither physiological nor articulatory in nature; that is, pauses longer than 200 ms.

2.5. Statistical analysis

The quantitative analysis that we performed aimed to test the hypothetical effect of our independent variables ('IV') and of their interactions. In each case, we investigated the effect of the IVs *noise interference*, temporal *compression* and linguistic *combination* as well as *subject*. Note that the IV *subject* is considered to be fixed, in the same way as the other variables. This is due to the fact that the selection of subjects was not done randomly; on the contrary, it was the result of close study with the aim of identifying individuals who were prototypical of the groups they represented. This inferential analysis therefore is not intended to lead to any form of generalization to the universe of the subjects, but is designed to explore and shed light on the validity, for future research, of the individual characteristics of these subjects which led us to choose them for this study.

We shall rely on variance analysis to analyze the interval data. However, in so far as the distributions of pause duration are highly dissymmetrical, we will apply an inverse hyperbolic tangent transformation, with a view to giving the dependent variable the Gaussian features required by variance analysis. For numerical analysis of data, we rely on a fully saturated hierarchical loglinear model, which enables us to process nominal data in a conceptual framework which is close to that of variance analysis.

3. Results

3.1. Number of pauses

As table 1 shows, two IVs appear to have a significant effect[1]: *subject* and *combination*. Conversely, neither of the two IVs related to disruption appear to have significant effects.

Of the interactions, only two are clearly significant: the interaction of *subject* and *combination* on the one hand, and the interaction of *noise interference* and *compression* on the other.

Table 1: results of loglinear analysis applied to all pauses longer than 200 ms.

Source of Variation	Degrees of Freedom	Chi-square	Signif.
subject	3	54.927	0.000
combination	1	17.878	0.000
noise interference	2	3.518	0.172
compression	2	0.026	0.987
subject*combination	3	13.802	0.003
subject*noise interference	6	10.967	0.089
combination*noise interference	2	0.507	0.776
subject*compression	6	2.251	0.895
combination*compression	2	0.747	0.689
noise interference*compression	4	13.866	0.008

A descriptive study of the differences shows that the significance of the *subject* variable can be principally attributed to the difference between *int2* (who was remarkable for the large number of pauses she made) and the three other subjects, who recorded lower numbers of pauses and for each of whom the number of pauses was approximately equal.

The significant difference attributable to the effect of the *combination* variable is linked to the fact that the number of pauses was generally lower in the French to Spanish combination than in Spanish to French. In other words, the subjects paused less often when they were speaking the dominant language.

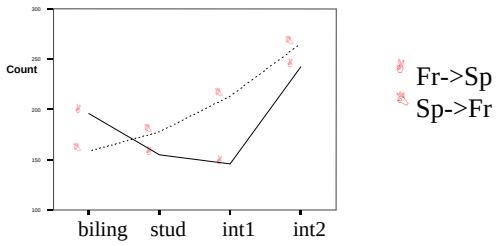

Figure 1: number of pauses by *subject* and by *combination*

Analysis of the dual interaction *subject*combination* (see fig.1) shows that the difference in numbers of pauses between the two combinations varies according to subject. This number is particularly low in *int2*, particularly high in *int1*, and

[1] Due to lack of space and a desire for clarity, we will confine ourselves, in this study, to the analysis of the main effects of these independent variables and to the dual interactions between them; analysis of higher-level interactions did not offer substantial additional information.

intermediate in the two other subjects. It is notable that *biling* is the only subject who paused more often in Spanish than in French.

The study of the interaction between *noise interference* and *compression* shows that, generally, the number of pauses increases as a function of the amount of noise interference, but this increase varies considerably as a function of the rate of compression: there is a very clear increase when the level of compression is high or medium.

3.2. Duration of pauses

Variance analysis shows a significant effect for the *subject*, *combination* and *noise interference* IVs. Conversely, no significant effect was observed when considering the effect of (temporal) *compression* in isolation. A significant effect can be observed for the interactions *subject*combination*, *compression*combination* and *compression*noise interference*.

Table 2: results of variance analysis applied to all pauses longer than 200 ms in duration.

Source of Variation	Degrees of Freedom	F	Signif.
subject	3	9.735	0.000
combination	1	16.819	0.000
noise interference	2	8.896	0.000
compression	2	1.557	0.211
subject*combination	3	9.967	0.000
subject*noise interference	6	0.124	0.993
combination*noise interference	2	2.816	0.060
subject*compression	6	1.235	0.285
combination*compression	2	5.292	0.005
noise interference*compression	4	4.020	0.003

In other words, the VI temporal *compression* may well have no effect on the duration of pauses when considered in isolation, but its effects are felt when its interaction with *linguistic combination* and *noise interference* is considered.

It can be observed that in the group as a whole, the mean duration of pauses is 768 ms. From this point of view the differences between subjects are great, as the mean pause duration for *int2* (620 ms) differs by more than 300 ms from that observed for *stud*. The lowest durations were observed in *biling* and *int2*, and the greatest in *stud* and *int1* (see fig. 2)

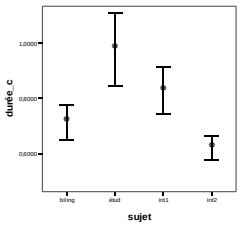

Figure 2: mean and standard deviation of duration of pauses according to *subject* (left to right: *biling, stud, int1* and *int2*)

In the Spanish to French combination, the subjects tended generally to produce shorter pauses (670 ms on average) than in the French to Spanish combination(865 ms on average), as shown in fig. 3a.

Under the influence of an increase in noise interference, the average pause duration tended to increase. Fig. 3b shows that from *noise interference* level 1 to level 2, a reduced increase of the order of 30 milliseconds can be observed. Conversely, between levels 2 and 3, the difference was of the order of 200 ms (706 ms compared to 902 ms on average).

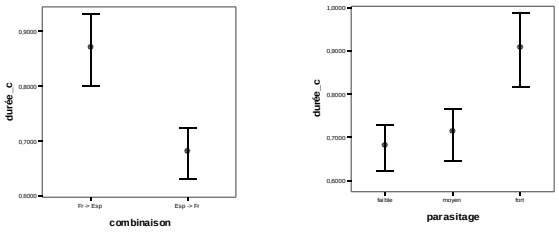

Figure 3: a (left) :mean and standard deviation of duration of pauses according to linguistic *combination* (left to right: *Fr->Sp* and *Sp->Fr*) and b (right) to *noise interference* (right: *low, medium, high*)

Analysis of the dual interaction *subject*combination* (fig. 5) confirms the previously noted difference in profile of the four subjects. However, this shows that, from subject to subject, the difference in pause duration between the two combinations manifests itself in very different ways. Thus, the mean durations for *int2* are approximately equal (the difference being around 30 ms). *Biling* shows greater differences but these is still only of the order of 122 ms on average. Conversely, *int1* and *stud* show much greater differences between combinations (of the order of 200 ms and 500 ms respectively).

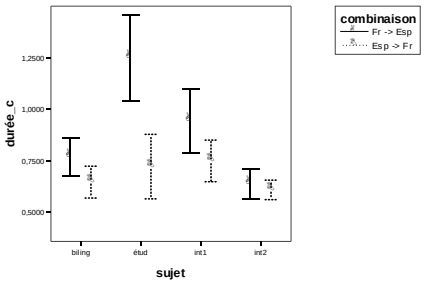

Figure 5: mean and standard deviation of pause duration according to the interaction *subject*combination* (left to right: subjects *biling, stud, int1* and *int2*)

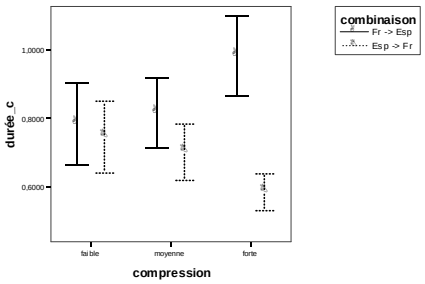

Figure 6: mean and standard deviation of pause duration according to the interaction *compression*combination* (left to right: *low, medium* and *high* compression)

Under the joint effect of *compression* and *combination*, average pause duration varies in opposed directions. Thus, a consistent increase in duration as a function of an increase in compression rates in the French to Spanish combination can be observed (784 ms, 816 ms and 982 ms respectively). Conversely, in the Spanish to French combination, a consistent decrease is observed (746 ms, 702 ms and 585 ms).

In any case, analysis shows that, whatever rate of compression is applied, the duration corresponding to the lowest level of noise interference is lower than the duration

corresponding to the highest level of noise interference. However, regular gradation is only observable under high levels of compression.

4. Conclusion

The two types of analysis that we carried out (on number and on duration of pauses) showed in each case that dependent variables are sensitive to the IVs used in our analysis. However, the combination of approaches (number and duration of pauses) seems to be particularly relevant, in so far as it enables richer observations to be made. Thus it can be observed that the subject who had the greatest linguistic and interpreting expertise (*int2*) was manifestly superior in terms of the number of pauses. But analysis of duration shows that this subject also produced the shortest pauses. It can also be seen that if the three other subjects could not be differentiated by how many pauses they made, differ one from another in the durations of their pauses. In contrast, it was observed that *stud*, who produced the lowest number of pauses, recorded the longest average pause duration.

We also saw that pauses were more numerous in the Spanish to French combination than in French to Spanish; these were of longer duration in the French to Spanish combination than in the Spanish to French combination. These observed tendencies are, however, sensitive to the *subject* effect: the three subjects having some degree of contact with the world of professional interpreting had higher scores when speaking French than when speaking Spanish, but the opposite was recorded in the subject with no interpreting experience (*biling*). The number of pauses is revealed to be sensitive to the level of interpreting expertise, while the pause duration seems to be sensitive to linguistic expertise: the subjects recorded different pause durations in each of the two combinations, but it was also noted that for *biling* and *int2* these differences were only minor, while for the other two subjects the differences in duration were greater.

The relevance of joining these approaches together is confirmed by the emergence of an effect that noise interference has on duration, while this IV appeared to have no effect on the number of pauses.

We observed, moreover, that the variable *compression* seems to have no direct effect on the number or duration of pauses. However, it would be false to assume that pauses are completely insensitive to variations in compression; no direct effect was observed, but several quite complex interaction effects were observed, which reveal only a general tendency for compression, in association with other variables, to have an indirect effect on the quantity of silence, particularly if the level of compression is high.

Data collected in this exploratory framework do not enable us to make inferences of any guaranteed validity. However, it may be observed that these data aid in gaining a better comprehension of the cognitive processes at work in this domain. The subject with the greatest linguistic and interpreting expertise notably made a large number of short pauses, which illustrates a markedly functional division of the source speech into regular chunks, which are then reproduced in the target language, separated by short pauses; conversely, the student subject, whose expertise both in language and interpreting was lacking, made fewer pauses but these were rather long, which did not accord with the optimal strategy displayed by the more experienced subject, and which probably illustrates disorganization in the production of the target language linked to various breakdowns in the interpretative process.

The general tendency for subjects to pause more often and for shorter periods when the speech is delivered in their dominant language could also be interpreted in terms of their having more effective chunking strategies for their stronger language; this could also be linked to difficulties in production of the target language, which is a hypothesis that could possibly partly explain the greater inter-combination contrast in subjects with less knowledge of French.

The effect of noise interference is clearly linked to comprehension phenomena, and it is useful to note that this effect mostly manifests itself in the appearance of pauses that are longer as a function of the level of noise interference.

Temporal compression effects were not observed; this could be explained either by the fact that an increase in source discourse speed had no consequences, or by the fact that, by choosing a threshold criterion of 200 ms, we prevented ourselves from observing pauses whose duration was shortened when the subject needed to speak more quickly. If this is the case, then these effects are linked to low-level production phenomena. Conversely, the contribution of the variable of compression to interaction effects (for high levels of compression) suggests difficulties of another kind, linked to comprehension and/or translation problems for such a fast source speech.

These observations could lead to further speculation about modeling the interpretative task, as they are rooted in analysis of low-level phenomena, but are aimed at understanding high-level phenomena.

From the point of view of basic research, the study of pauses in the task of interpreting is relevant in so far as it helps us work towards a better understanding of the cognitive processes in the interpreter. Moreover, from the point of view of applied research, this study is justified in that it is directly linked not just to the intelligibility of the speech produced by the interpreter, but also to the perception by the speaker of the interpreter's confidence, based on discontinuities in his/her speech; therefore, research of this kind contributes to the study of interpreting quality.

5. References

[1] Barik, Henry C. 1973. Simultaneous interpretation: Temporal and quantitative data, *Language and Speech*, 16, 237-270.

[2] Boschian Schiavon, Vesta. 1983. Velocità di parola e interpretazione simultanea, Università degli studi di Trieste: Monografie, N°20, Trieste.

[3] Campione, Estelle & Véronis, Jean. 2002. A Large-Scale Multilingual Study of Silent Pause Duration. In B. Bel & I. Marlien (Eds.), *Proceedings of the Speech Prosody 2002 conference*, 199-202.

[4] Campione, Estelle & Véronis, Jean. 2004. Pauses et hésitations en français spontané, *Actes des 25èmes Journées d'Etudes sur la PArole (JEP 2004*), 109-112.

[5] Candéa, Maria. 2000. *Contribution à l'étude des pauses silencieuses et des phénomènes dits "d'hésitation" en français oral spontané.* Thèse de doctorat, Université Paris III.

[6] Gerver, David. 1969. The effects of source language presentation rate on the performance of simultaneous conference interpreters, In E. Foulke (Ed.), *Proceedings of the 2nd Louisville conference on rate and/or frequency controlled speech*, University of Louisville, 162-184.

[7] Goldman-Eisler, Frieda. 1968. *Psycholinguistics: experiments in Spontaneaous Speech*, Academic Press, London and New York.

[8] Lee, Tae -Hyung. 1999. Speech proportion and accuracy in simultaneous interpretation from English into Korean, *Meta*, 44, 2, 261-267.
[9] Yagi, Sane M. 2000. Studying style in simultaneous interpretation, *Meta*, 45, 3, 520-547.
[10] Zellner, Brigitte.1994. Pauses and the Temporal Structure of Speech, In E. Keller (Ed.), *Fundamentals of speech synthesis and speech recognition*, 41-62.