Hesitations in read vs. spontaneous French in a multi-genre corpus

Jean-Philippe Goldman [1,3]*, Mathieu Avanzi* [2]*, Antoine Auchlin* [1]

[1] Department of Linguistics, University of Geneva, Switzerland
[2] Chaire de Linguistique française, Neuchâtel University, Switzerland
[3] TCTS Labs, University of Mons - UMONS, Belgium

`Jean-Philippe.Goldman@unige.ch; Mathieu.Avanzi@unige.ch; Antoine.Auchlin@unige.ch`

Abstract

This study is a part of an on-going work whose goal is the prosodic characterization of various speaking styles in a multi-genre 70-minutes French corpus as well as the development of prosodic automatic detection tools. In this corpus, a manual annotation prominences and disfluencies like hesitations and syntactic ruptures is used to show evident phonological aspects of hesitation in regard to quality, pause position and proximity to syntactic rupture.

Index Terms: hesitation, filled pause, vowel lengthening, spoken French, disfluencies

1. Introduction

Since seminal work by [1] it is accepted that hesitation comprises filled pauses (like "euh", "um" or "er"), lengthened syllables that do not correspond with a final phrase boundary, false starts, non semantic repetitions and non syntactic pauses.

In French, **hesitations** are generally expressed by a quasi-lexicalized central vowel [2] (this hesitation marker is transcribed "euh" and can be considered the English equivalent of "um" or "er"), or by an extra-lengthened vowel (the so-called "**filled pauses**"[3]).

Acoustically, scholars consider that both hesitation markers and filled pauses have the same properties. Usually, there are described as being strongly lengthened in contrast with stressed syllables (lengthening being a consistent property of French accent, [4]), as having a constant vocalic quality (the spectral envelope deformation which accompanies the vowel is very little), and associated with a level or slightly falling contour ([3] [5] [6] [7]).

The distribution of hesitation vs. vowel lengthening is considered as depending on the context. The analysis of a 70 minutes corpus of spontaneous classroom interactions (involving 11 speakers) has led [6] to show that the presence of a "euh" is more frequent after a lexical word or after a functional word ending with a CVC syllable. On the other hand, vowel lengthening is frequent for functional words ending with a (C)V syllable.

Scholars also agree to say that hesitation markers do not play the same function when they occur before or within a speech unit. According to [3] [8] [9], hesitations within a speech segment (81% in Campione's corpus [3]) indicate that the prosodic unit is not finished and is being continued; hesitations occurring before a new segment starts (19% in Campione's corpus) are employed by speakers to keep the floor and prevent any turn transition.

Finally, the communicative context also plays a part in the frequency of hesitation marks. [8] summarizes several studies showing that the frequency and mean duration of hesitation marks increase according to the degree of spontaneity of a message.

2. Material

This study is based on the C-PROM corpus [10][11], a freely available 70-minute corpus comprising 7 speaking style samples of 10 minutes each (Reading (LEC), Political Address (POL), Conference (CNF), News Broadcast (JPA), Radio Interviews (INT), Map Task (ITI) and Narratives (NAR)). Each style is represented by 2 to 7 speakers and all recordings have been automatically segmented in phones, syllables and orthographic words using Praat [12] and the EasyAlign tool [13].

Two annotators manually checked the segmentation and followed a strict protocol described in [10] to annotate the corpus at a syllabic level for prominences and delivery phenomena, like hesitations (labeled as "h"), post-tonic syllabic schwa (as in "c'est dinguE" [sɛdɛ̃gə], which can be perceived as hesitant or not) and unaccented post-tonic syllables (appendices). Eventually, an automatic grammatical annotation, manually checked, provided part of speech information at word level [14].

3. Quantitative description of hesitations

The whole corpus comprises 17,820 articulated syllables, 1,737 silent pauses and 22 overlap, paraverbal or noise segments. Among the articulated syllables, 579 are annotated as hesitations, either "euh" markers or lengthened vowels, and 444 were annotated as syntactic interruption (i.e. repetitions or false starts). 117 syllables are labeled both hesitation and interruption. In this section, we describe the hesitations according to their type, syntactic location and combination with silent pauses.

3.1. Frequency and type by speaking style

Table 1 shows the number and percentage of hesitations per speaking style. As expected [8], spontaneous styles have far more hesitations than read or prepared speech (JPA, LEC and POL). Within spontaneous speech, CNF and INT qualify for semi-spontaneous speech as speakers talk about a well prepared topic. The rate of hesitation is below true spontaneous speech (ITI & NAR). As for read speech, JPA has a surprisingly a few hesitations. This can be accounted for by two reasons: a high articulation rate (5.87syll/sec) and breaking news that force the speaker to improvise.

The multi-tier annotation allows distinguishing between "euh" and other vowel elongations thanks to the word tier. The rightmost column in Table 1 shows the percentage of "euh" hesitations relatively to the number of hesitations by style. Results show that this proportion is rather regular throughout the styles (from 55% to 77%). The sparse data for LEC and POL hesitations prevent from taking into account the proportion for these 2 speaking styles

Table 1 Number of syllables and hesitations by speaking style, and number of "euh". In parenthesis, percentage of hesitation by style and percentage of "euh"'s relatively to the number of hesitations

	Speak. style	#syll	#h (%)	#euh (%/#h)
spon-taneous	ITI	2234	151 (6.8%)	91 (60%)
	NAR	2670	187 (7.0%)	118 (63%)
semi-spont.	CNF	3132	103 (3.3%)	79 (77%)
	INT	2615	77 (2.3%)	44 (55%)
read	JPA	3165	60 (1.9%)	34 (57%)
	LEC	1829	0 (0%)	0 (0)
	POL	2175	1 (0.1%)	1 (100)
total		17820	579(3.14)	365(65)

3.2. Hesitations and discourse-syntactic rupture

The 579 hesitations can be categorized according to their type ("euh" vs non-"euh") and their direct proximity to a syntactic rupture as in Table 2.

Table 2 Contingency table for "uh"/non-"euh" vs. rupture/non-rupture of 575 hesitations

	rupture	non-rupture	Total
"euh"	28	337	365
non-"euh"	89	125	215
Total	117	462	579

Syntactic ruptures correspond to interruptions (with or without reparation) or non semantic repetitions. The greater proportion of non-"euh" for ruptures (8% of "euh" hesitations, 42% for non-"euh") is significant (χ^2=99.09; p<0.001). In other words, "euh" markers tend not to co-occur with a syntactic rupture.

3.3. Position of hesitation toward adjacent pause

Table 3 displays the number of hesitation with an immediate pause before of after (and in parenthesis the number of "euh" for each category). No statistical difference was found for the presence of a pause before and/or after a hesitation or before and/or after a "euh".

Table 3 Number of hesitations with a pause before and/or after an hesitation (In parenthesis, the number of "euh" for this category)

	No pause before	Pause before	Total
No pause after	324 (188)	104 (67)	428 (255)
Pause after	120 (85)	31 (25)	151 (110)
Total	444 (273)	135 (92)	579 (365)

Nevertheless the proportion of "euh" with a pause after (110/151=72%) is significantly greater than the proportion of "euh" without a pause after (255/428=59%) (χ^2=7.87; p<0.001) but no regarding preceding pauses, showing some recurrent patterns of hesitations and pauses within spontaneous speech.

4. Discussion

This preliminary study looked at some aspects of hesitation in regard to hesitation quality, pause position and proximity to syntactic rupture. Some significant evidences showed specific phonological properties. The next step is twofold: 1. to point out other similar parameters that would clarify our intuition about 2 types of "euh", like position within intonation and/or syntactic unit (more precisely short "euh" near unit beginning and long "euh" at unit boundaries) and 2.a to have a more phonetic look to acoustic properties of such phenomena with relative prosodic parameters such as the ones used in automatic prominence detection as in [11].

5. References

[1] Maclay, H. & Osgood, C.E., Hesitation Phenomena in Spontaneous English Speech, Word, 1, 19-43, 1959.

[2] Vasilescu I., Candea M., Adda-Decker M. Perceptual salience of language-specific acoustic differences in autonomous fillers across eight languages, *Proceedings of Interspeech 2005*, Lisbonne

[3] Campione, E. and Véronis, J., "Pauses and Hesitations in French Spontaneous Speech", Proc. DISS'05, Aix-en-Provence, 42-46, 2006.

[4] Delattre, P. 1966. L'accent final en français : accent d'intensité, accent de hauteur, accent de durée », in P. Delattre, *Studies in French and comparative phonetics*, La Haye, Paris : Mouton, 66-72.

[5] Guaïtella, I. Hésitations vocales en parole spontanée : réalisations acoustiques et fonctions rythmiques, *Travaux de l'Institut de Phonétique d'Aix*, vol.14, pp. 113-130, 1991.

[6] Candea, M. *Contribution à l'étude des pauses silencieuses et des phénomènes dits « d'hésitation » en français oral spontané. Etude sur un corpus de récits en classe de français*, Ph.D Université Paris III , 2000

[7] Corley M. & Stewart, O. W. Hesitation Disfluencies in Spontaneous Speech: The Meaning of *um, Language and Linguistics Compass* 2/4 (2008): 589–602.

[8] Duez D. Signification des hésitations dans la parole spontanée, 2001.

[9] Lacheret-Dujour, Anne (2003), La prosodie des circonstants en français parlé, Société linguistique de Paris, Leuven-Paris, Editions Peeters.

[10] Goldman Jean-Philippe, Auchlin Antoine, Roekhaut Sophie, Simon Anne-Catherine, Avanzi Mathieu, *Prominence perception and accent detection in French. A corpus-based account*, In: Proceedings Speech Prosody 2010, Chicago, 11-14 mai 2010, 2010, p. 100575

[11] CPROM : http://sites.google.com/site/corpusprom/

[12] Boersma, Paul & Weenink, David (2010). Praat: doing phonetics by computer [Computer program]. Version 5.1.34, retrieved 31 May 2010 from http://www.praat.org/

[13] EasyAlign Goldman http://latlcui.unige.ch/phonetique

[14] Goldman, J.-Ph., A. Auchlin, S. Roekhaut, A.C. Simon & M. Avanzi. 2010. Prominence perception and accent detection in French. A corpus-based account, Proceedings of Speech Prosody 2010 (Chicago, Illinois, USA, May 10, 2010)