

Phonetic and functional features of lexicalized pauses in Italian

Loredana Schettino¹, Violetta Cataldo¹

¹Department of Humanities Studies, University of Salerno, Italy

Abstract

This research explores lexicalized filled pauses in Italian tourist guides' speech, addressing the existing correlation between their formal, mainly phonetic, features and pragmatic functions. The study reveals that pauses' F₀ contour correlates with their main functions, allowing to distinguish pauses working as controlled focusing devices, from more explicitly unintentional hesitation pauses.

Key words: Speech, Disfluencies, Lexicalized Filled Pauses, Discourse Markers

Introduction

The spontaneous human speech comes with a number of phenomena affecting its 'fluency', namely pauses, repairs, repetitions, whose incidence may vary depending on the type of speech and its planning degree. It is by now recognized that some of such 'disfluencies' are charged with communicative values. In particular, pauses provide the listener with valuable meta-information about the ongoing speech (e.g. about discourse planning, its structure, speakers' modal disposition). Though, there is no evidence of speakers' control on their production (see Eklund 2004, Corley and Stewart 2008). Two forms of disfluencies are usually identified: repairs, something in the utterance is retraced and altered; hesitations, as temporary interruptions and delays in flowing speech (Lickley 2015). Hesitation pauses may be subdivided in unfilled (silent) pauses and filled pauses. Among filled pauses, those with a proper lexical form, here referred to as 'lexicalized' filled pauses (such as discourse markers), may be distinguished from the 'non-lexicalized' filled pauses (prolongations, vocalizations and the like).

This exploratory study aims to investigate disfluencies falling into the set of pauses. It follows a previous investigation on Silent Pauses (SP), Filled Pauses (FP, such as 'ehm' or 'eeh'), Vocalized Filled Pauses (VFP, segmental prolongations) (see Cataldo *et al.* in press) and extends the object of study to Discourse Marker Filled Pauses (DMFP, such as *diciamo* 'let's say', *quindi* 'so'). DMFPs are here identified as polyfunctional lexical elements, external to propositional content, which, given their semantic meaning, develop pragmatic functions due to the context they are uttered in (Bazzanella 2006).

To this end, the existing correlation between DMFPs' formal, mainly phonetic, features and pragmatic functions is addressed. Such studies might find application in automatic speech processing (see Origlia *et al.* 2018).

Methodology

For this study, a dataset was considered of the main audio-visual corpus of visits led by expert guides in San Martino's Charterhouse (Naples). The dataset consists of circa 35' semi-spontaneous speech by three female tourist guides (G01, G02, G03). The selected data were annotated using ELAN, for a multilevel annotation of disfluencies, and PRAAT for the description and analysis of their phonetic realization.

DMFPs were analyzed according to the following parameters:

Pragmatic functions – 1) Interactional Function (IF), related to discourse planning within the interaction; 2) Meta-Textual Function (MTF), related to structuring and emphasizing information in speech; 3) Cognitive Function (CF), dealing with ongoing cognitive processes related to the propositional content and the illocutionary force (based on Bazzanella 2006)¹.

Phonetic features – DMFP F₀ profile (rising, falling, flat, valley, peak); mean pitch in Hz (of the Tonal Unit, TU, and of the DMFP occurrence); creaky vocalizations; DMFPs distribution within the TU.

Results

The incidence of the 639 DMFPs on the guides' speech is quite even across the speakers (one item uttered per about 9 words and about 17 per minute).

Pragmatic functions. Figure 1 shows DMFPs are mainly adopted as focusing devices, textual markers and/or boosters. In fact, out of 463 focusing devices, about the half is also boosters and a quarter carries out a text marker function as well. Despite this common tendency, more specific individual preferences reflect the idiosyncrasies found in the previous study.

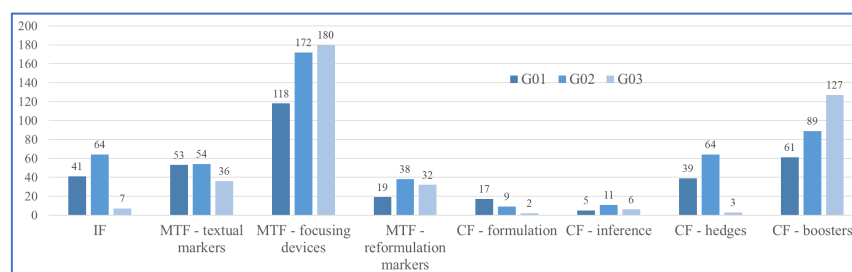


Figure 1. DMFP occurrences per speaker and per function

Phonetic and functional features of lexicalized pauses in Italian

G01 prefers an ‘on the fly’ strategy, employing a number of formulation pauses and hedges, fillers, and phatic devices; G02 tends to avoid non-lexicalized pauses altogether through ‘juxtaposition’ of utterances and uses a high number of lexicalized pauses instead; G03 adopts a more controlled, ‘rhetorical’ style, using DMFPs as focusing devices and boosters mainly.

Phonetic features. As shown in Table 2A, most often DMFPs are realized with a flat F₀ profile or as a tonal trough within the TU (31%, 25%). Most noteworthy is the recurring F₀ ‘peak’ realization (30%), which appears to correlate with focusing devices². Considering focusing DMFPs (foc) as compared with the remaining DMFPs (non-foc.), a generalized linear mixed model testing the correlation between function (foc vs. non-foc) and F₀ profile showed that the latter is a significant predictor of the function ($\chi^2=20.74$, $p<.001$). Pairwise comparison among fixed levels then showed that ‘peak’ and ‘valley’ are significant predictors of, respectively, foc. and non-foc. function.

Such results confirmed also the mean pitch values registered for both DMFPs and their TU (Table 2B). Only the 194 DMFP peak occurrences are realized with mean pitch values about 25 Hz higher than those of their TU. Standard deviation high values depend on the inter-speaker variability (the maximum excursion in G03 is double the one in G02) and on the fact that the ‘peak’ doesn’t always reach the highest point in the TU.

Table 2. DMFP occurrences and percentages per F₀ profile and per speaker (A); mean pitch (Hz) of TUs and DMFPs, F₀ difference between DMFP peaks and TU mean pitch (Hz) and its standard deviation values (B).

A	F ₀ patterns					B	Mean pitch (Hz)			
	flat	valley	rising	falling	peak		tot DMFPs		peaks DMFP	
G01	84	33	2	7	47	TU DMFP		diff. st.dev.		
G02	96	79	26	25	47	G01	219,8	222,9	25,5	29,8
G03	19	47	4	21	100	G02	214,6	212	24,5	18,8
n. occ.	199	159	32	53	194	G03	215,3	223,5	23,9	35,5
% occ.	31%	25%	5%	8%	30%					

Discussion and conclusions

As for the pragmatic level, DMFPs are found to diverge from FPs and VFps. Given their multifunctional nature, the former set is highly used to perform meta-textual functions (most often as focusing devices) and cognitive functions (as mitigating or boosting devices), unlike non-lexicalized pauses, mostly used as a tool of the speech planning process.

As for their phonetic features, DMFPs occur with a wider range of possible realizations and show more modulated intonation patterns, as they have more segmental content, than the other voiced pauses. Furthermore, the ‘peak’ realizations, never found neither for FPs nor for VFPs, seem to distinguish structuring focusing DMFPs, working as deliberate rhetoric devices, from the other more explicitly unintentional lexicalized pauses (fillers and modal devices) showing a rather flat or valley pitch contour.

Notes

1. IF, i.e. speaker-oriented devices (turn-taking, fillers, attention-getting, phatic, checking comprehension, requesting agreement/confirmation, yielding the turn) and addressee-oriented devices (interrupting, back-channels, attention confirmed, phatic, acknowledgement, agreement/confirmation); MTF, i.e. textual markers (introduction, transition, list, digression, ending, quotation), focusing devices and reformulation markers (paraphrase, correction, exemplification); CF, i.e. markers of formulation/inference and modulation devices (hedges and boosters).
2. Ex: *i principi ispiratori della comunità sono legati a quello che è un altro ordine esistente in Europa ed è l'ordine benedettino* - ‘The community’s guiding principles are tied to **what it is** another existing order in Europe that is the Benedictine Order’.

Acknowledgements

Work funded by the Italian National Project PRIN ‘Cultural Heritage Resources Orienting Multimodal Experiences (CHROME)’ (#B52F15000450001).

References

- Bazzanella, C. 2006. Discourse Markers in Italian: towards a ‘compositional’ meaning. In Fischered, K. (ed.). *Approaches to discourse particles*, 449-464. Amsterdam, Elsevier.
- Cataldo, V., Schettino, L., Savy, R., Poggi, I., Origlia, A., Ansani, A., Sessa, I., Chiera, A. in press. Phonetic and functional features of pauses, and concurrent gestures, in tourist guides’ speech. *Studi AISV*.
- Corley, M., Stewart, O. W. 2008. Hesitation disfluencies in spontaneous speech: The meaning of um. *Language and Linguistics Compass*, 2(4), 589-602.
- Eklund, R. 2004. Disfluency in Swedish human-human and human-machine travel booking dialogues. Ph.D. dissertation, Linköping University Electronic Press.
- Lickley, R. J. 2015. Fluency and Disfluency. In Redford, M.A. (ed.). *The handbook of speech production*, 445-474. John Wiley & Sons.
- Origlia, A., Savy, R., Poggi, I., Cutugno, F., Alfano, I., D’Errico, F., Vincze, L., Cataldo, V. 2018. An Audiovisual Corpus of Guided Tours in Cultural Sites: Data Collection Protocols in the CHROME Project. *Proc. 2018 AVI-CH Workshop*, 1-4, Castiglione della Pescaia, Italy.