"(Dis)fluency across spoken and signed languages: applications of an interoperable annotation scheme"

Crible, Ludivine ; Dumont, Amandine ; Grosman, Iulia ; Notarrigo, Ingrid

Abstract
Natural languages, whether spoken or signed, are characterized by the presence of so-called disfluencies or “fluencemes” (Götz 2013) which reflect the online nature of production and comprehension processes. Fluencemes are related to the complex constructs of fluency and disfluency, which are commonly associated with continuity and interruption, as illustrated in Ejzenberg’s (2000: 287) definition: “the degree to which speech is articulated smoothly and continuously without any ‘unnatural’ breakdowns in flow”. Fluency is also fundamentally context-bound: the production and perception of fluencemes strongly depend on distributional (frequency, position, combination) and contextual factors (interactional features), so much so that the same fluenceme can either be fluent or disfluent depending on its local and situational environment. This ambivalence, acknowledged in our use of the term (dis)fluency, motivates a componential approach to the phenomenon that decomposes a h...

Document type: Communication à un colloque (Conference Paper)

Référence bibliographique
Crible, Ludivine ; Dumont, Amandine ; Grosman, Iulia ; Notarrigo, Ingrid. (Dis)fluency across spoken and signed languages: applications of an interoperable annotation scheme. International Conference on Fluency & Disfluency Across Languages and Language Varieties (Louvain-la-Neuve, Belgium, du 15/02/2017 au 17/02/2017).
Natural languages, whether spoken or signed, are characterized by the presence of so-called disfluencies or “fluencemes” (Götz 2013) which reflect the online nature of production and comprehension processes. Fluencemes are related to the complex constructs of fluency and disfluency, which are commonly associated with continuity and interruption, as illustrated in Ejzenberg’s (2000: 287) definition: “the degree to which speech is articulated smoothly and continuously without any ‘unnatural’ breakdowns in flow”. Fluency is also fundamentally context-bound: the production and perception of fluencemes strongly depend on distributional (frequency, position, combination) and contextual factors (interactional features), so much so that the same fluenceme can either be fluent or disfluent depending on its local and situational environment. This ambivalence, acknowledged in our use of the term (dis)fluency, motivates a componential approach to the phenomenon that decomposes a holistic impression of “speed and effortlessness” (Chambers 1997: 535) into a typology of fluencemes which, once combined, allow the researcher to evaluate the relative fluency of the speaker or signer. Starting from Shriberg’s (1994) seminal work, an increasing body of corpus studies have indeed adopted this componential view, though with slightly different definitions, methods and research agendas, across a variety of languages (e.g. English in Besser & Alexandersson 2007, Portuguese in Moniz 2013 or French in Pallaud et al. 2013).

Following this line of research, this paper presents an annotation scheme that aims to offer an exhaustive yet flexible coverage of fluencemes, and that is applicable to both spoken and signed languages. It was designed and tested on a variety of speech situations uttered in spoken French, native and learner English as well as in Belgian French Sign Language (henceforth LSFB). The fluencemes included in the model can be divided in four groups: (1) simple fluencemes – they include pauses (filled, unfilled, hand stops), palm-ups, discourse markers, explicit editing terms, false-starts and incomplete truncations; (2) compound fluencemes, a category that covers repetitions (identical, modified, framing) and substitutions (morphosyntactic, propositional); (3) insertions (lexical, parenthetical) and deletions; and (4) “diacritics” (the term is taken from Shriberg 1994), that add contextual information to an existing fluenceme (misarticulation, lengthening, embedding, re-ordering or syntactic completion). The main innovative aspects of this annotation scheme involve, first, its applicability to multilingual and multimodal corpora (especially to signed languages), and, second, its technical format that enables the handling of complex embedded structures on a single layer of annotation thanks to a bracketing and numbering system. Another major, more theoretical, characteristic of the scheme concerns the ambivalent (fluent vs disfluent) nature of fluencemes: not only are the consensual hesitations and interruptions included in the typology, but typically “fluent” devices are also included, that is, the strategic uses of fluencemes for rhetorical or structuring effects, as in examples 1 and 2 below.

1. we have to have (0.980) um life saving floats we have to have (0.280) life buoys we have to have (0.470) bilge pumps (Backbone corpus, “bb_en002_amphibioustours”)
2. AVOIR PT:DET DIRE BIEN AVOIR PT:DET ENVIRON (LSFB corpus, CLSFBI3705)
   there are competent [interpreters] there are weaker [interpreters]
These examples contain several occurrences of pauses and hand stops, as well as “fluent” modified repetitions (“we have to have” in English, AVOIR PT:DET in LSFB) which are not corrective reformulations but rather create an enumeration or a contrast, respectively, thus exploiting the linguistic material in the local context for strategic purposes. This, we argue, allows the researcher to refrain from early or arbitrary judgments at this stage of the analysis.

The replicability of the annotation was assessed by an inter-annotator agreement analysis on a 7,000-word sample of radio interviews in spoken French, where we found substantial to almost perfect kappa-scores between two annotators: $\kappa = 0.67$ including disagreement on detection of fluencemes; $\kappa = 0.82$ when excluding disagreement on detection and categories with less than 10 occurrences. These results are very encouraging and indicate that the protocol can reliably be used in multilingual and multimodal corpora.

The potential of this annotation model is then illustrated in the genre of interviews in six corpora, namely Backbone (Kohn 2012), the French Corpus of Humorist Speech (C-Humour, Grosman 2016), the Louvain Corpus of Annotated Speech - French (LOCAS-F, Degand et al. 2014), the Louvain International Database of Spoken English Interlanguage (LINDSEI, Gilquin et al. 2010), the Corpus LSFB (Meurant 2015), and VALIBEL Database (Dister et al. 2009). The most frequent fluencemes were compared across the different languages and modalities, revealing whether signers and speakers of different languages and modalities resort to the same devices in a similar interaction. While spoken languages (i.e. native French, and native and learner English) show very similar quantitative behaviors (absolute and relative frequency of fluencemes), LSFB appears to make use of a different top set of fluencemes. For instance, modified repetitions are prevalent in LSFB and far more frequent than in spoken languages which prefer unfilled pauses: this result might be an indication of different strategies in coping with language production. In a last step, the study zooms in on the clustering tendencies of palm-ups and filled pauses (such as *uh* and *uhm* in English), investigating whether and under what conditions these two categories can be considered as multimodal equivalents, as is sometimes claimed in the literature (e.g. Locker McKee & Wallingford 2011; van Loon 2012). More specifically, palm-ups and filled pauses are systematically compared in terms of their proportion in isolation vs. in combination with other adjacent fluencemes and, if so, what and where these adjacent fluencemes appear (i.e. in initial, medial or final position in the sequence).

The method and analysis reported in this paper contribute to the growing interest for multimodal data, providing further insights into, on the one hand, the potential equivalence of fluencemes in the spoken and signed modalities and, on the other, how similar speaking situations might breed different outputs depending on (the nature of) the spoken language (French/English, native/learner). Lastly, this multimodal and multilingual annotation scheme successfully addresses a number of methodological issues, namely universality and interoperability, in view of further enhancing our understanding of the complex notion of (dis)fluency.

References


