

CoFee - Toward a multidimensional analysis of conversational feedback, the case of French language

Laurent Prévot, Roxane Bertrand

Laboratoire Parole et Langage, CNRS & Aix-Marseille Université, Aix-en-Provence, France

laurent.prevot@lpl-aix.fr, roxane.bertrand@lpl-aix.fr

Abstract

Conversational feedback is mostly performed through short utterances such as *yeah*, *mhmm*, *okay* not produced by the main speaker but by one of the other participants of a conversation. Such utterances are among the most frequent in conversational data. They also have been described in psycho-linguistic models of communication as a crucial communicative tool for achieving coordination or alignment in dialogue. The newly funded project described in this paper addresses this issue from a linguistic viewpoint by combining fine-grained corpus linguistic analyses of semi-controlled data with formal and statistical modeling. The impoverished aspect of the linguistic material present in these utterances allows for a truly multi-dimensional analysis that can explain how different linguistic domains combine to convey meaning and achieve communicative goals.

Index Terms: Feedback, Backchannel, Semantics, Pragmatics, French Language

1. Objectives

The general objective of the CoFee¹ project is to propose a fine grained model of the form/function relationship concerning feedback behaviors in conversation. To succeed, we need to achieve:

- a fine-grained analysis of the different dimensions involved (prosody, lexical markers, acoustic non-verbal signals, facial expressions, head movements, gaze) ;
- a fine-grained analysis of the communicative functions related to feedback ;
- a rich characterization of two crucial contextual parameters: discourse context and production context ;
- the integration of these ingredients into a general model.

¹CoFee –Conversational Feedback: Multi-dimensional Analysis and Modeling– is a newly funded ANR (Agence Nationale pour la Recherche) 3-year project [2012-2015].

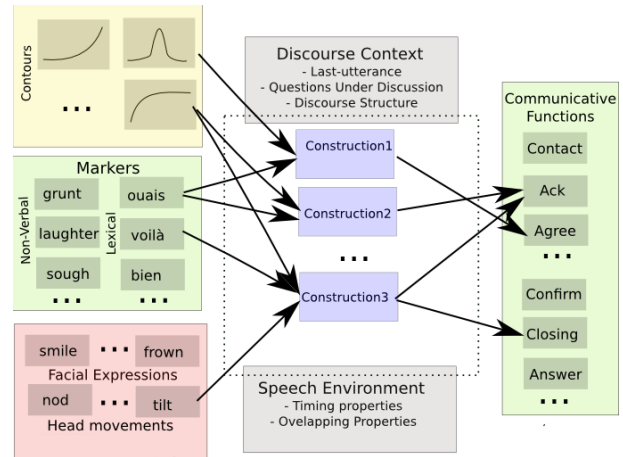


Figure 1: Sketch of the CoFee model

We consider that the truly multi-dimensional nature of the analysis proposed is an important and ambitious step for linguistic studies. Most of existing related work either focus on one domain and marginally integrates the other dimensions or constitutes a very shallow surface-based analysis grounded on a few features. Moreover, the integration of different situations of communication in such a precise study is also new and will allow to account for communicative situation variability from a more theoretical and experimental approach than what is done usually.

The present paper is structured as follow. We will start by better defining the object of our study *feedback items* in section 2. Then we will discuss some related work (Section 3) before presenting with some details of the data (Section 4) as well as the analysis and the modeling planned (Section 5). Finally we will briefly describe the current work and what will be presented at the workshop.

2. Definitions

CoFee is a study of the positive feedback items. This section is a clarification attempt, at least for the sake of the project.

- **Feedback:** In dialogue or conversation context, it can be associated with any evaluative communicative action about previously introduced material.
- **Backchannel:** Intuitively, backchannels are productions made by the participant holding the listener role.
- **Acknowledgment:** A positive feedback. Polarity is functional here since negative items can have a positive evaluation function.

2.1. Backchannels vs. acknowledgments

Acknowledgements and backchannels have sometimes been used as synonyms. Although these phenomena are frequently co-occurring, they constitute different aspects of verbal interaction.

The term *back-channel* was introduced by [1] and included a broad range of linguistic phenomena such as questions and short comments. The notion was broadened later to include other items such as verbalized signals, sentence completions, brief restatements, clarification requests,... Almost any communicative event can be a back-channel. Indeed, backchannels are sometimes described as production by the *listener* moving the definition issue to the speaker / listener definitions. This is however not as straightforward as it seems to be since listeners are commonly said to produce signals in the course of the communication. While [2] argues that participants tend to not overlap the production of their interlocutors thanks to an efficient *turn-taking* rule system, [3] shows that even if the turn-taking system is efficient it is not rare that participants speech overlaps.

Speaker/Listener distinction, and therefore backchannel definition combine both form and content issues. A participant that is not willing to take the turn should not produce utterances signaling his willingness to do so. Backchannels are typically briefs, low in intensity and may exhibit specific prosodic contours. Moreover, even if the listener desires to take the initiative, social rules (politeness) are forcing him to conform to turn-taking rules and therefore remain more or less in his listener role until the speaker yields the turn. At the content level, many productions can be back-channelled and only a few communicative acts (such as questions) tend to switch systematically the speaker/listener role.

Traditionally backchannels are divided between *continuers* and *assessments* [4]. Tottie [5] gives *continuers* a regulative function and *assessments* a supportive function. The former regulates the coming contribution of the interlocutor while the latter bring a supportive reaction to a previous contribution. Feedback is more clearly associated with the latter but it is difficult to systematically distinguish them and therefore most of the empirical studies are proposing to work on the phenomena as a whole.

2.2. Backchannels feedback

The difficulties mentioned above lead [3] to propose a *back-channel feedback* notion. Back-channel feedback:

- (i) respond directly to the content of an utterance of the other participant
- (ii) is optional
- (iii) does not require acknowledgment by the other participant.

However, these criteria concern only back-channelled feedback, not those occurring as part of a turn. Such feedback is rather common, specifically in task-oriented dialogues that require a detailed grounding of the information transmitted.

2.3. Sum-up

To sum-up, in COFEE we are examining positive feedback behaviors (mainly verbal behavior but also laughter and other communicative grunts [6]). Most of them are also back-channels but we do not exclude feedback items that are not back-channelled. The latter may be taken as an answer to a question but it is still optional, contrarily to answers. Based on earlier studies on French feedback [7, 8], the list of lexical items we are including in our study is : *oui* (yes), *ouais* (yeah), *mhmm*, *ok*, *d'accord* (right), *voilà* (that's it), *c'est ça* (that's it), *ah*, *bon* (well).

3. Related work

Among the more recent works, [9] proposed a broad study of the form/function relation for feedback. They use various features including acoustic and discourse ones. However “discourse” features are more shallow than the one we are planning to use (basically based on size of Inter Pausal Units –IPU– and of position of the item in the IPU). Moreover, we are also planning a more linguistic way for extracting the speech parameters than purely acoustic measurements. However, we will attempt to replicate many aspects of their study on our French corpora.

[10, 11] have a multi-dimensional model of communicative functions dealing with feedback behaviors. Here the modeling framework is very rich but as in Gravano and colleagues study the discourse and linguistic features used are very shallow since the goal was not to focus on feedback but on the identification of all communicative functions.

Formal semantics and pragmatics until recently remained away from feedback behaviors. Fortunately, in the recent years this field started to look more carefully at this issue. This movement can be traced back to [12] for at least a plea for a move in this direction. More complete framework allowing to work on feedback mechanisms in formal pragmatics are presented in [13] or [14].

4. Data and annotations

4.1. Corpora

Three corpora will be used in the course of the project:

- The Corpus of Interactional Data (CID) recorded by Roxane Bertrand and Béatrice Priego-Valverde [15] is a 8 hours (110K tokens) corpus composed of 8 conversations of 1 hour. It features a nearly free conversational style with only a single theme proposed to the participants at the beginning of the experiment. This corpus is fully transcribed and forced-aligned at phone level with signal. Moreover, it has been annotated with various linguistic information (Prosodic Phrasing, Discourse units, Syntactic tags,...) during the OTIM project [15]. (Visible at sldr.org/sldr000720/en)
- A 3h30 French MapTask created by Corine Astésano and Ellen Bard [16]. It has been recorded according to the original MapTask methodology. This corpus has been transcribed and aligned manually at utterance level. We are now planning an automatic phone alignment with the same methodology used in the previous project. (Visible at sldr.org/sldr000732/en)
- A French Negotiation Game Corpus that is currently under construction and that consist in negotiations games played by four participants. We are targeting a bigger corpus than the two others but not fully transcribed. We plan to transcribe only speech neighbouring feedback items which are less frequent in this setting than in the two previous ones. (Preview visible at sldr.org/sldr000773/en)

These three corpora constitute very different communicative situations and therefore cover an interesting range of functions feedback can play in dialogue.

4.2. Feedback Annotations

There are some rich annotation frameworks including feedback aspects such as [17, 11]. However given our focus on a few restricted forms we will only use part of these comprehensive frameworks. Moreover we inherit from previous annotations efforts. Namely for the CID corpus we already have some back-channels annotations performed. The categories annotated were: *continuer* (minimally takes note), *understanding* (understands), *assessment* (agrees with what has been said), and *evaluation* (evaluate and display an attitude about what has been said). Orthogonally turn-initiating and turn-ending features have been added. From another study, [7] we wanted to include (i) aspects related to the *confirmation* nature² of some feedback items and (ii) their discourse

²Related to allo-feedback in DIT scheme [11] for example.

structuring functions such as *closing* current discourse topic.

Perhaps the most original part of our annotation scheme is the annotation of the feedback scope. In [7], we identified 3 relevant scopes: *last utterance*, *last pair* or *wide scope*. In the corpus we used, this scope annotation was reasonably well annotated ($\kappa = 0.6$) and allowed us to specify the functions of some of the lexical items studied without having to rely on finer-grained functions.

5. Analysis and Modeling

The model we are aiming at combines a detailed multidimensional analysis of the forms involved, a deep modeling of the meaning of these forms and how these meanings are used to reach the communicative goals.

5.1. Analysis of the forms

About the first aspect we will perform both linguistic analysis (including in particular systematic prosodic analysis) and more acoustic measurements. About the prosodic aspects, a track we will follow is the Functional Data Analysis (FDA) such as proposed in [18]. We therefore adopt a really data-driven approach but guided by the linguistic analysis. FDA requires indeed to have some minimal hypotheses about the shapes of the contours before starting the purely statistical analysis that will distinguish several clusters of instances of contours. Concerning this level the goal is to delineate as precisely as possible the formal categories in all the dimensions considered (at least lexical item, prosodic contour, acoustic parameters).

5.2. Model of the functions

Concerning the functions, we consider that simply having a list of categories is not enough. To have a function one should be able to model the effects on context. Moreover, we are also interested into the meaning (if any) of the forms considered and how it is exploited by the participants to perform communicative actions. The set of communicative actions comes from the literature on these issues, in particular from the DIT+ framework [10]. However we would like to go step deeper by looking at how formal theory of dialogue [13, 10, 14] are handling these phenomena.

Despite differences with regard to the primitives and to representation tools, it is possible to list a few properties any semantic/pragmatic theory should feature for being able to deal with feedback items:

- Radical context dependence: Given the range of communication functions a simple word as *'yeah'* can rich in a conversation, it is clear that the theory has to be a theory of how the meaning of a new utterance is interpreted (and resolved) in a given

context ;

- Rich ontology of communication objects, the context in which utterances are resolved cannot simply be a flat representation of the actual world. Feedback has a meta-level nature, it is information about the information exchange, not about the content exchanged directly. Moreover feedback is also about processing of information by the speaker (cognitive realm) and about conventional rules of the exchange (social realm).

Dynamic semantics and further work grounded in this paradigm all feature the first point while the second one is present in most of the works that have looked seriously at dialogue. For the formal modeling aspect of our work, we will focus on two semantic theories that have put dialogue on their agenda: SDRT (Segmented Discourse Representation Theory) from [13] and KOS from [14].

6. Current work

Our current work consist in building the data sets from the corpora and in finishing recording the third corpus. By the time of the workshop, we will have conducted some preliminary studies on a data subset. The study will include a FDA data analysis for at least two French lexical items in the CID corpus: *ouais* (yeah) and *voilà* (that's it). For this data subset, we will complete the annotations of the functions in order to have as small scale picture of our project to present during the workshop.

7. References

- [1] V. H. Yngve, "On getting a word in edgewise," in *Papers from the sixth regional meeting of the chicao linguistic society*, 1970, pp. 567–578.
- [2] H. Sacks, E. A. Schegloff, and G. Jefferson, "A simplest systematics for the organisation of turn-taking for conversation," *Language*, vol. 50, pp. 696–735, 1974.
- [3] N. Ward and W. Tsukahara, "Prosodic features which cue back-channel responses in english and japanese," *Journal of Pragmatics*, vol. 32, no. 8, pp. 1177–1207, 2000.
- [4] E. Schegloff, "Discourse as an interactional achievement: Some use of" uh-huh" and other things that come between sentences," *Georgetown University Round Table on Languages and Linguistics, Analyzing discourse: Text and talk*, pp. 71–93, 1982.
- [5] G. Tottie, "Conversational style in british and american english: The case of backchannels," *English corpus linguistics: Studies in honour of Jan Svartvik*. London & New York: Longman, pp. 254–335, 1991.
- [6] N. Ward, "Non-lexical conversational sounds in American English," *Pragmatics & Cognition*, vol. 14, no. 1, pp. 129–182, 2006.
- [7] P. Muller and L. Prévot, "An empirical study of acknowledgement structures," in *Proceedings of Diabrock, 7th workshop on semantics and pragmatics of dialogue*, Saarbrücken, 2003.
- [8] R. Bertrand, G. Ferré, P. Blache, R. Espesser, and S. Rauzy, "Backchannels revisited from a multimodal perspective," in *Proceedings of Auditory-visual Speech Processing*. Citeseer, 2007.
- [9] A. Gravano, J. Hirschberg, and Š. Beňuš, "Affirmative cue words in task-oriented dialogue," *Computational Linguistics*, vol. 38, no. 1, pp. 1–39, 2012.
- [10] H. Bunt, "Multifunctionality in dialogue," *Computer Speech & Language*, 2011.
- [11] V. Petukhova and H. Bunt, "Towards an integrated scheme for semantic annotation of multimodal dialogue data," in *Proceedings of the seventh international conference on language resources and evaluation*, 2010, pp. 2556–2563.
- [12] J. Allwood, J. Nivre, and E. Ahlsen, "On the semantics and pragmatics of linguistic feedback," *Journal of Semantics*, vol. 9, 1992.
- [13] A. Lascarides and N. Asher, "Grounding and correcting commitments in dialogue," *Journal of Semantics*, 2009.
- [14] J. Ginzburg, *The Interactive Stance: Meaning for Conversation*. Oxford University Press, 2012.
- [15] P. Blache, R. Bertrand, and G. Ferré, "Creating and exploiting multimodal annotated corpora: the toma project," *Multimodal corpora*, pp. 38–53, 2009.
- [16] C. Astésano, E. Bard, and A. Turk, "Structural influences on initial accent placement in french," *Language and Speech*, vol. 50, no. 3, pp. 423–446, 2007.
- [17] J. Allwood, L. Cerrato, K. Jokinen, C. Navarretta, and P. Paggio, "The MUMIN coding scheme for the annotation of feedback, turn management and sequencing phenomena," *Language Resources and Evaluation*, vol. 41, no. 3, pp. 273–287, 2007.
- [18] M. Gubian, L. Boves, and F. Cangemi, "Joint analysis of f0 and speech rate with functional data analysis," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. IEEE, 2011, pp. 4972–4975.