Toward a semantics for French short positive feedback utterances

Laurent Prévot

Aix Marseille Université & CNRS, Laboratoire Parole et Langage
5 Av. Pasteur, 13100 Aix-en-Provence, FRANCE
`laurent.prevot@lpl-aix.fr`

Although feedback utterances are ubiquitous in dialogue and identified as a crucial aspect of language interaction, existing semantic/pragmatic proposals do not offer a comprehensive model for them. The present abstract argues for a two-step formal model using fairly standard semantics for lexical item entries and attempting to derive the pragmatic communicative functions from the semantics thanks to rich context modeling. The objective is not only the formal model *per se* but the possibility of using it as the backbone for a more empirical approach, in the spirit of (Gravano et al., 2012) or (Neiberg et al., 2013) but in which the usefulness of the semantic layer will be investigated.

Phenomena targeted Following the original proposal of (Yngve, 1970), we take *back-channel utterances* as utterances produced on the back channel of the conversation. If they were produced on the main channel, they will disrupt the flow of the speaker at that moment. Following (Bunt, 1994), we take *feedback utterance* as an utterance through which a dialogue participant provides information about his processing of the partner's previous utterances. This includes information about perception, interpretation, evaluation (agreement, surprise,...) or dispatch (fulfillment of a request,...). The topic of this paper are the *positive* feedback.[1]

Objectives and related work A crucial objective for our formal model is to help make more precise the interaction between the different modalities involved. Our starting point is a model in which all the feedback utterance instances associate a base form,[2] a prosodic form[3] and more acoustic-phonetic properties. When visual channel is involved, gestures or facial expressions can be combined and/or constitute another type of base forms. In previous empirical work (Stolcke et al., 2000; Gravano et al., 2012; Neiberg et al., 2013) all the instances received a communicative function. However, little has been said precisely in term of semantics. In the formal work of (Bunt, 2012), there is room for semantics but the proposal kept this part relatively abstract. Consequently, this work also aims at reducing the current gap between theses data-drivenstudies and more theoretical contributions such as (Ginzburg, 2012) or (Lascarides and Asher, 2009).

(Stolcke et al., 2000; Gravano et al., 2012; Neiberg et al., 2013) are all proposing some conclusions about the impact of the different features in their learning or classifying systems but the results are semantically shallow. In (Stolcke et al., 2000) and (Gravano et al., 2012), we only learn (on this aspect) that the stronger clues are the tokens identity which is the shallow way to get to semantics. (Gravano et al., 2012) shows the relevance of positional features which are shallow discourse features but do not enter in their analyses. (Neiberg et al., 2013) has the more surprising results that the base form is not relevant and it is phonological operations and prosody that are contributing the more to the communicative function.[4]

Providing a semantics for lexical forms, for phonological prosodic forms (contours) and for facial expressions will allow to study precisely how their interact and whether they behave more like in a compositional or a constructional fashion. This question will be addressed within and across modalities. More precisely, it is hypothesized that having a formalization of the pragmatic impact (in a given context) of a given lexical/prosodic/gesture association and a formal semantics for each of these elements we will be able to understand how they combine. A preliminary and easier to answer question in which semantics can help also is whether all these forms are compatibles. Predicting incompatibilities from our

---

[1]Of course, the polarity of the item considered is not a good clue for the positive discourse function since, for example, *'no'* is regularly used as positive feedback targeting negative utterances.

[2]Here we restrict the study to a closed list of lexical items and their combinations or repetitions.

[3]One issue is however that phonological categories prove to be very difficult to annotate on these rather reduced forms (D'Imperio et al., 2013).

[4]However, this could be due to functions there are looking at and to the way the specifically select the data subset for their experimental study.

model will be an interesting intermediate check for our approach.

Selection of forms The selection of the forms studied is straightforward. There are the most frequents forms found in our French spoken corpora. The seven tokens *ouais, oui, d'accord, voilà, okay, mh, ah* represents almost 10% of the total number of tokens in a MapTask Corpus (Gorish et al., 2014) and nearly 6% in a French conversational corpus (Prévot et al., 2013). Other potential feedback items are very far in term of frequency from the set we plan to scrutinize in this work.

The proposal Space constraints prevent a detailed presentation of the model, but overall we treat the lexical items in a relatively standard way either as propositional adverbs or type $\langle t, t \rangle$ for *oui, ouais* or as attitudinal operators of type $\langle \langle e, t \rangle, t \rangle$ for *d'accord, okay, voilà, ah*. *mh* is a special case since we consider that in default use it does not target propositional content and we will detail further its case here. Prosodic contours are also expected to act as operators on the propositional content that hold them.[5] Finally, gestures can be both modeled as propositional or attitudinal operators.

A key issue is to handle potential redundancy across modalities. The solution proposed at this stage is to combine the different attitudinal contributions through an unification-based mechanism. The issue become therefore to identify the relevant sets of dimension.

Illustration The first step of the work is to propose semantic entries for each of the forms considered. Lacking space for a definition of each attitude we try to provide explicit labels for a subset of attitudes.

Lexical items *ouais,oui*: $\lambda P.P$
ah: $\lambda P.x \, attitudeSet(x, P) \sqcup surprised$
ok,d'accord: $\lambda P.x \, attitudeSet(x, P) \sqcup agree$[6]
voilà: $\lambda P.x \, attitudeSet(x, P) \sqcup manifest$[7]

Prosodic contours Both contours types ($\sim$: high F0 standard deviation ; $\searrow$: Fall ; $\nearrow$: Rise) and their meaning are proposed, based on previous literature of several languages, for illustrating

the model and might not reflect what they will be ultimately in the French case.
$\sim$: $\lambda P.x \, attitudeSet(x, P) \sqcup surprised$
$\searrow$: $\lambda P.x \, attitudeSet(x, P) \sqcup grounded$
$\nearrow$: $\lambda P.x \, attitudeSet(x, P) \sqcup elicit$

Gestures Same comment as for prosody.
☺-NOD: $\lambda P.x \, attitudeSet(x, P) \sqcup grounded$
☺-SMILE: $\lambda P.x \, attitudeSet(x, P) \sqcup amused$

The research objective can be therefore formulated as characterizing the $\sqcup$ operations (including when attitudes are conflicting). As mentioned above, a first step consists in checking the incompatibilities. Here for example, *agree* is not compatible with *elicit* and therefore *d'accord* should not be compatible with $\nearrow$.

References

Bunt, H. (1994). Context and dialogue control. *Think Quarterly*, 3(1):19–31.

Bunt, H. (2012). The semantics of feedback. In *16th Workshop on the Semantics and Pragmatics of Dialogue (SEMDIAL 2012)*, pages 118–127, Paris (France).

D'Imperio, M., Petrone, C., and Prévot, L. (2013). Basic prosodic transcription of short french feedback utterances. In *Proceedings of Second Advancing Prosodic Transcription for Spoken Language Science and Technology*, Lisbon.

Ginzburg, J. (2012). *The Interactive Stance: Meaning for Conversation*. Oxford University Press.

Gorish, J., Astésano, C., Bard, E., Bigi, B., and Prévot, L. (2014). Aix map task corpus, the french multimodal corpus of task-oriented dialogue. In *submitted*.

Gravano, A., Hirschberg, J., and Beňuš, Š. (2012). Affirmative cue words in task-oriented dialogue. *Computational Linguistics*, 38(1):1–39.

Lascarides, A. and Asher, N. (2009). Grounding and correcting commitments in dialogue. *Journal of Semantics*.

Neiberg, D., Salvi, G., and Gustafson, J. (2013). Semi-supervised methods for exploring the acoustics of simple productive feedback. *Speech Communication*.

Prévot, L., Bigi, B., and Bertrand, R. (2013). A quantitative view of feedback lexical markers in conversational french. In *Proceedings of SIGDIAL*.

Stolcke, A., Ries, K., Coccaro, N., Shriberg, E., Bates, R., Jurafsky, D., Taylor, P., Martin, R., Ess-Dykema, C., and Meeter, M. (2000). Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational linguistics*, 26(3):339–373.

Yngve, V. H. (1970). On getting a word in edgewise. In *Papers from the sixth regional meeting of the chicago linguistic society*, pages 567–578.

---

[5]This is not a general proposal for the meaning of prosody.

[6]Actual grounding can be derived through the properties of the attitudes (eg. $grounded(x, P) \rightsquigarrow P \in gameBoard_x$)

[7]P is manifest for x