LUND UNIVERSITY

Attitude reports in spontaneous dialogue

uncertainty, politeness and filled pauses

Horne, Merle

Published in:
From Quantification to Conversation

2012

Document Version:
Early version, also known as pre-print

Link to publication

Citation for published version (APA):
Horne, M. (2012). Attitude reports in spontaneous dialogue: uncertainty, politeness and filled pauses. In S. Larsson, & L. Borin (Eds.), *From Quantification to Conversation: Festschrift for Robin Cooper on the occasion of his 65th birthday* (pp. 297-308). College Publications.

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

To appear in Borin, L. & Larsson, S. (Eds). *From quantification to conversation: Studies in theoretical and computational linguistics.*

Attitude reports in spontaneous dialogue: uncertainty, politeness and filled pauses

Merle Horne

1. Attitude reports in dialogue

Integrating information in dialogue interaction is a complex process that involves multimodal reasoning. Building on work in situation semantics (Barwise and Perry 1983), Cooper and colleagues (e.g. Cooper 1996; Cooper and Ginzburg 1996) have developed a model for the analysis of dialogue meaning which incorporates different information sources. One aspect of dialogue semantics that has received particular attention is the role of information states and their relationship to the analysis of propositional attitudes. Attitude reports are important in the analysis of spontaneous discourse and in practical applications such as automatic speech recognition and understanding. They give important information as to the speaker's attitude to his/her utterance in relation to the listener. The present contribution focuses on attitude reports in Swedish spontaneous speech.

1.1. Attitude reports in spontaneous speech: pragmatic particles

When investigating the meaning structure of spontaneous speech, it becomes clear that the expression of attitude reports related to mental states like *believe* and *know* is not always coded exclusively by attitudinal predicates like *believe, doubt,* etc. In spoken spontaneous dialogue, where the speaker and addressee(s) interact in real time, propositional attitudes are often modulated by other means than just lexical predicates. For example, pragmatic mood-related particles such as Swedish *liksom* 'like' in spoken language (see Anward 1999; Andersen and Fretheim 2000, Aijmer 2004) add a modulating, interactive dimension to the speaker's propositional attitude:

(1) *Man* ***tror liksom*** *att man nått botten, men det blir bara sämre och sämre*

'You think like you've reached the bottom, but it only gets worse and worse' (*www.slowfat.com/pblog.php?dec=7*)

1.1.1. Hedges, uncertainty and politeness

Pragmatic particles, in combination with attitudinal predicates like *believe* are related to the immediate interactive dimension of dialogue semantics. The mood-related pragmatic particle *liksom* 'like' is what has been referred to as a *hedge* (Lakoff 1972), a mechanism used to modulate the assertive force of an utterance. Hedging has been observed to be an important communicative device used to soften the modal attitude or illocutionary force of a statement. In Japanese, for example, hedges have been shown to be realized both lexically (e.g. with verbs and adverbs) and through non-lexical hedging devices (e.g. syntactic and prosodic devices) (Nittono 2003). Glottalized filled pauses can thus be interpreted as a kind of hedging mechanism related to politeness phenomena and modality. Hedging functions as a means for realizing what has been termed *positive politeness*, i.e. a strategy for maintaining a positive relationship or *face* between parties (Brown and Levinson 1987). Hedges can in other words be associated with face-saving, politeness and indirectness. The propositional attitude *uncertainty* is related to politeness in the sense that a speaker can be uncertain as to how to best formulate an utterance so as not to lose face (Aijmer 2004). The uncertainty attitude of the speaker uttering (2) is expressed through the use of the glottalized filled pause *EH~* (glottalization will be represented as (~) :

(2) *Jag ser ju att EH~ PAUSE tiden räcker ju liksom inte till ändå för mig*
'I see you know that UH~ PAUSE the time just you know like isn't long enough anyway for me'

It can be assumed that it is not the truth value of the proposition that the speaker is uncertain about in this case, but rather the optimal way of coding the message for the addressee who she has never spoken to before. The other interactional devices used by the speaker, the pragmatic particles *ju* 'you know' (twice) and *liksom* 'like', are further evidence for the very listener-oriented speech coding that the speaker is involved in.

1.2. Attitude reports and prosody: uncertainty, politeness and intonation

In addition to lexical predicates and pragmatic particles, attitude reports can further be realized by prosodic means. For example, an intonation contour with a final rise in declaratives has been associated with an attitude of *uncertainty* on the part of the speaker (Pierrehumbert and Hirschberg 1990; Ohala 1983; Gussenhoven 2004; House 2003).

Speaking with a high level of fundamental frequency has also been associated with politeness according to the *Frequency Code* (Ohala 1983; Gussenhoven 2004). The example in Figure 1 is an assertion with a high-rising question-like intonation which could have been uttered as a hedging answer to a question such as: *Where would you like to live?* This intonation pattern signals an attitude of uncertainty on the part of the speaker, both as regards the utterance's assertive/illocutionary force, i.e. the degree of speaker commitment to the truth associated with the proposition, as well as the speaker's attitude as regards the hearer's background knowledge, e.g. uncertainty as regards the addressee's knowledge as to the whereabouts of Ottawa.

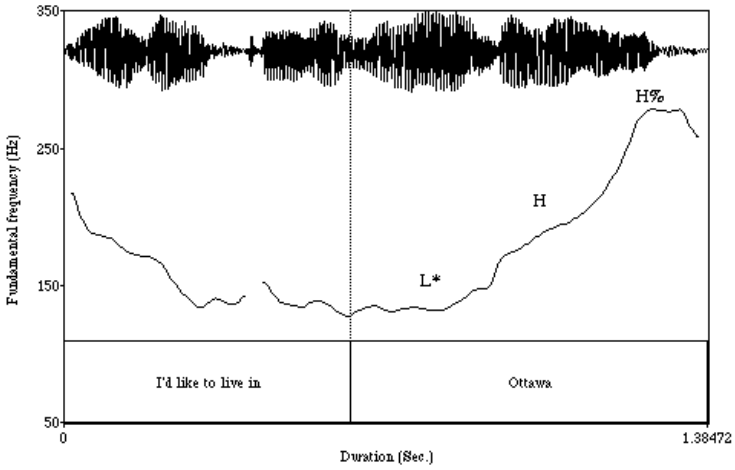

Figure 1. Wave form and fundamental frequency (pitch) curve for the utterance *I'd like to live in Ottawa* with a final question-like intonation: a rising pitch accent (L*+H) on *Ottawa* and a high phrase accent (H%).

2. Uncertainty, politeness and glottalized filled pauses

In the present contribution, drawing on data from Swedish spontaneous speech, it will be illustrated how another voice-related cue, the parameter *glottalization* or *creaky voice* used in the production of filled pauses in spontaneous speech can also be interpreted as signalling the speaker attitude *uncertainty*.

2.1. The filled pause *EH*

In speech technology studies, the filled pause or filler *EH* 'UH' has often been termed a *disfluency* (e.g. Shriberg 2001), since it constitutes a delay in the flow of speech associated with lexical meaning. Despite their lack of referential or lexical meaning, filled pauses have, nevertheless, been associated with clear discourse functions, such as signaling an upcoming focussed word (Strangert 1991; Bruce 1998), or the intention to continue speaking. Since *EH* can therefore be associated with relatively clear discourse or pragmatic meanings, it has sometimes been classified as a *word,* an *interjection* (Allwood et al. 1990; Clark and Fox Tree 2002).

2.1.1. Distribution of EH in spontaneous dialogue

Using spontaneous speech in spontaneous interview situations involving speakers who had never met before (data from *SweDia* database), Horne (2006) showed that *EH* occurred in this material almost exclusively as a clitic to a preceding function word or discourse marker: 127 of the 137 instances of *EH* occurred in this environment. *EH* occurred most often after *och* 'and' and *men* 'but' which often function as discourse markers, e.g. introducing new topics, topic continuations, etc. 52 cases of *EH* occurred after these two function words. *Och EH* 'and UH' was the most common function word + *EH* construction. The second most frequent function word category preceding *EH* was the subordinate conjunction *att* 'that' which also can function as a discourse marker introducing a non-subordinate clause. 24 instances of *EH* occurred after *att*. The other instances of *EH* were found after the following function words: preposition (n=13), articles (n= 9), pronouns (Subject) (n= 9), basic verbs or auxiliary verbs (n=9), demonstrative article (n=5), indefinite adjective (n=3), subordinate conjunction (other than *att* 'that') (n=2), Negation (n=1). Content words preceded *EH* in only 7 cases. Finally, there was 1 case where *EH* was a repetition. Thus, the dominating discourse context for filled pauses is after a discourse marker, in clause-initial position. *EH* is thus intimately associated with a context where it can be assumed to be related to the coding of clauses in working memory.

2.1.2. Phonetic realization of EH

Three basic realizations of the filled pause *EH* were observed in Horne (2006): a middle-high front or central vowel: e.g. [ɛ] (see Fig. 2), a nasalized

vowel or vowel + nasal consonant: e.g. [əm] (see Fig. 3), and a glottalized or creaky vowel: e.g. [ə̰] (see Fig. 4).

The vowel realizations of *EH* were the most frequent (n=61) and had a mean duration of 268 ms and a SD of 136 ms. The nasalized or vowel+nasal realizations were second in frequency (n=43), and had a mean duration of 436 ms and a SD of 185 ms. These showed a distribution like the vowel+nasal fillers in English that Clark and Fox Tree (2002) analyzed, i.e. they were always followed by other kinds of 'delays', sometimes several in sequence as in Figure 3 with SWALLOW, SMACK, INHALE following *EH*.

The glottalized realizations of *EH* were the fewest (n=31) had a mean duration of 310 ms and a SD of 150 ms. Their duration thus overlaps with the durations of the vowel and vowel+nasal realizations. Unlike the vowel+nasal realizations, the only other delay that was observed to follow the glottalized filled pause was a silent pause.

2.3. Glottalization and dialogue interaction

Glottalized filled pauses in dialogue are in some sense unexpected, since *EH* is often assumed to be a signal that the speaker wants to hold the floor, whereas glottalization or creaky voice, on the other hand, is assumed to be

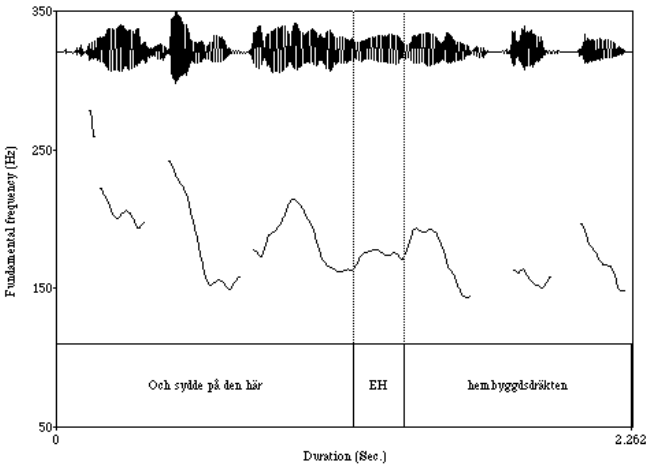

Figure 2. Example of the realization of *EH* as the middle high vowel [ə] in the utterance *och sydde på den här EH hembyggdsdräkten* 'and sewed this UH folk dress'

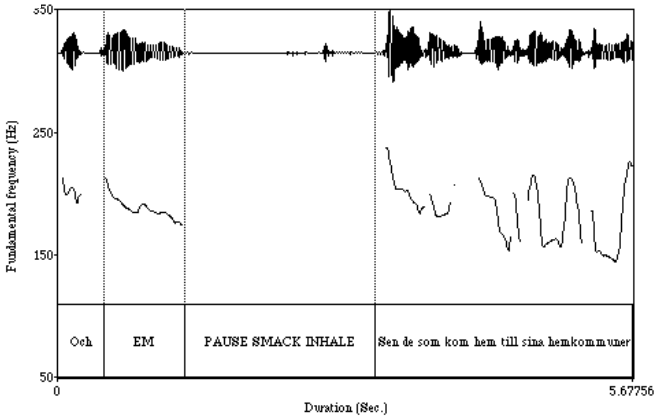

Figure 3. Example of the realization of *EH* as a vowel+nasal [ɛm] before a relatively long delay in speech: *Och EM SWALLOW SMACK INHALE sen de som kom hem till sina hemkommuner* 'and UM SWALLOW SMACK INHALE then the ones that came back to their home municipalities…

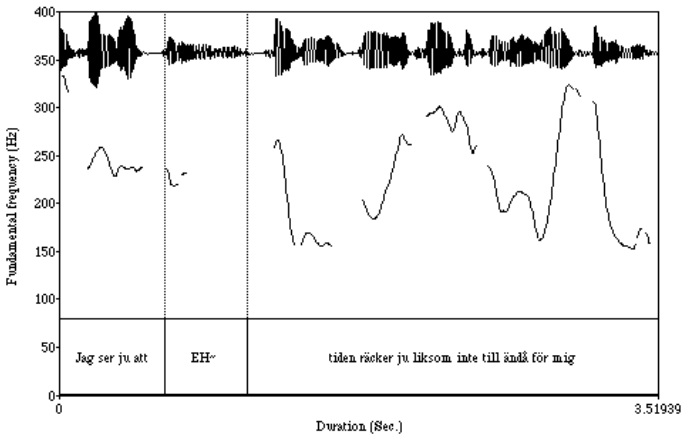

Figure 4. Example of the realization of *EH* as the creaky/glottalized vowel [ə̰] in the utterance: *Jag ser ju att EH tiden räcker ju liksom inte till ändå för mig* 'I see you know that UH time just like isn't enough you know anyway for me'.

a signal of finality (Ladefoged, 1982). Ogden (2001) shows for example, that in Finnish, creaky voice in utterance-final position functions in dialogue to signal the end of a speaker's turn. Grivičić and Nilep (2004) have observed creaky voice quality in the pronuncation of English *yeah* as feedback in telephone speech and suggest that it is a signal on the part of the listener for a change of the discourse topic. Duncan and Fiske (1977) observed that when creaky voice is coupled with low pitch, it can signal the end of a conversational turn.

2.3.1. Glottalization and attitude reports: uncertainty

Associated with a filled pause, however, glottalization or creak seems to have other functions than signaling turn change (Shriberg 2001). Bear et al. (1992), Lickley (1994), Nakatani and Hirschberg (1994) have observed that glottalization is not uncommon before a speech repair. Glottalized *EH* could, therefore, be interpreted as a juncture signal for an upcoming pause in speech that is *not* to be interpreted by the listener as a TRP (Transition Relevance Place). When it occurs *within* a non-completed turn, glottalization seems to function as a signal that the speaker wishes to continue but is somewhat uncertain as to how to encode the upcoming speech:

(3) *men den såg ju inte ut EH~ det var någon*
'but it did not look like UH~ it was somebody'

2.3.2. Glottalized filled pauses as procedural markers

Glottalized filled pauses thus can be considered to be procedural markers of speaker attitude. They occur in environments where the speaker seems to be uncertain about how he or she should code a proposition for a listener who the speaker is not well acquainted with, such as in the example in (4):

(4) *Jag ser ju att EH~ PAUSE tiden räcker ju liksom inte till ändå för mig*
'I see you know that UH~ PAUSE the time you know just isn't like long enough anyway for me'

Notice in this example (see also Figure 4) involving the filled pause *EH~*, the cooccurrence of the 'softener' *liksom* 'like' as well as the intensifying positive feedback seeking particle *ju* 'you know' (2 occurrences), which

constitute further evidence of the interactional, hedge-related interpretation of the glottalized filled pause. The speaker's fundamental frequency is also relatively high in the context of the filled pause. This high pitch constitutes additional support for a hedge interpretation of the glottalized filler, since, as mentioned above, high F0, according to the *Frequency Code,* is associated with uncertainty and politeness. The uncertainty, however, is not related to the truth value of the proposition being coded but rather to procedural encoding of the message (Blakemore 1987), i.e. it is related to the speaker's mental state as to how the message should be encoded in the most relevant way given the speaker-addressee relationship.

Thus, we propose that glottalized filled pauses signal the speaker attitude *uncertainty* related to the assertive force of the utterance, more specifically as to *how* the message should be coded linguistically in a dialogue situation where the speaker is not sure of the optimal way to code a message. The data we have used to illustrate this assumption come from dialogues where the speaker and listener have never met before and therefore have very little common ground. We assume the attitude associated with glottalized filled pauses is related to that associated with the pragmatic particle *liksom* 'like'. However, glottalized *EH~* is not associated with uncertainty as regards the referential meaning of the proposition but rather to uncertainty regarding what the speaker determines to be the optimal way of coding the utterance being planned. This attitude can be thought to arise due to the paucity of the common ground established in the current discourse. It is associated with the interactive dimension of the discourse, in the sense that it involves the speaker's degree of awareness of the need to code his or her utterance in a way that best suits the context, given the background knowledge that the addressee has or seems to have. Since it is a kind of uncertainty related directly to speech planning/performance in a dialogue setting, it is therefore a procedural related attitude report.

Nonglottalized *EH*, on the other hand, is not readily associated with this uncertainty. Nonglottalized filled pauses occur in contexts generally associated with prominence in discourse such as signalling an upcoming focussed word (new information) (Strangert 1991). In this type of context, the speaker is not perceived as being uncertain about how the utterance is to be coded. Notice that in the example in Figure 2, *och sydde på den här EH hembyggdsdräkten* 'and sewed this UH folk dress' one could not add the hedge *liksom: *och sydde på den här hembyggdsdräkten liksom* '*and sewed like this folkdress', since the presence of the demonstrative determiner *den här* 'this' indicates that the referent *hembyggsdräkten* is known to the speaker and the listener. One could, however, have said *och sydde på en hembyggdsdräkt liksom* 'and sewed like a folkdress' where the indefinite article *en* 'a' indicates a referent, which is new in the discourse and thus

assumed by the speaker to be relatively inaccessible for the listener. The use of a glottalized *EH~* at the beginning of the clause would make the interactive nature of the utterance more apparent and the speaker attitude more uncertain as regards the procedural encoding of the utterance: *och EH~ sydde på en hembyggdsdräkt liksom* 'and UH~ sewed like a folkdress'.

3. Summary and conclusion

The glottalized realization of the filled pause *EH* has been hypothesized to have a function related to that of the 'hedging' particle *liksom* 'like'. It is assumed to be an interaction-related hedge signal like *liksom* in the sense that it is intimately associated with the speaker's cognitive activity in speech planning, i.e. it is a procedural cue to the speaker's attitude.

The data discussed here suggest that glottalized filled pauses would appear to be related to attitude reports that involve uncertainty and positive politeness. Glottalized *EH* is thus assumed to constitute a procedural cue to the speaker's mental state as regards the most relevant way to code the current message, given the dialogue situation where the common ground for speaker and addressee is minimal. The analysis is based on data from discourse situations where the speaker and listener have never met before.

The uncertainty attitude associated with glottalized filled pauses is supported by the fact that other cues to uncertainty and politeness are sometimes present in the utterances containing a glottalized *EH*, for example, pragmatic particles such as the hedge *liksom* 'like' as well as a relatively high fundamental frequency. These cues have independently been associated with the same speaker attitude of uncertainty as well as to politeness. In order to strengthen the present analysis, however, more data from more dialogue situations has to be analyzed and compared.

Acknowledgements

This research was supported by a grant from the Swedish Research Council *(VR)*.

References

Aijmer, Karin
2004 Pragmatic markers in spoken interlanguage. *Nordic Journal of English Studies* 3: 173–190.

Allwood, Jens, Joakim Nivre, and Elisabeth Ahlsén
1990 Speech management – On the non-written life of speech. *Nordic Journal of Linguistics* 13: 3–48.

Andersen, Gisle, and Thorstein Fretheim (eds.)
2000 *Pragmatic Markers and Propositional Attitude*. University of Bergen/University of Trondheim Pragmatics and Beyond NS 79. Amsterdam: John Benjamins.

Anward, Jan
1999 Allt du önskar kan du få? Om SAG och talspråket. *Svenskans Beskrivning* 24: 329–353.

Barwise, Jon, and John Perry
1983 *Situations and attitudes*. Cambridge, Mass.: MIT Press.

Bear, John, John Dowding, and Elizabeth Shriberg
1992 Integrating multiple knowledge sources for detection and correction of repairs in human-computer dialog. *Proceedings of the 30th Annual Meeting of the Association for Computational Linguistics*, 56–63.

Blakemore, Diane
1987 *Semantic Constraints on Relevance*. Oxford: Blackwell.

Brown, Penelope, and Stephen Levinson
1987 *Politeness. Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Bruce, Gösta
1998 *Allmän och Svensk Prosodi*. Lund: Dept. of Linguistics and Phonetics, Lund University.

Clark, Herb, and Jean Fox Tree
2002 Using *uh* and *um* in spontaneous speech. *Cognition 84*: 73–111.

Cooper, Robin
1996 The attitudes in Discourse Representation Theory and Situation Semantics. In *Logic, Language and Computation*, Vol. 1, Jerry Seligman and Dag Westerståhl (eds.), 137–150. Stanford: CSLI Publications.

Cooper, Robin, and Jonathan Ginzburg
1996 A compositional situation semantics for attitude reports. In *Logic, Language and Computation,* Vol. 1, Jerry Seligman and Dag Westerståhl (eds.), 101–113. Stanford: CSLI Publications.

Duncan, Starkey, and Donald Fiske
1977 *Face to Face Interaction: Research Methods and Theory*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Grivičić, Tamara, and Chad Nilep
2004 When phonation matters: the use and function of *yeah* and creaky voice. *Colorado Research in Linguistics* 17 (http://www.colorado.edu/ling/CRIL/Volume17_Issue1/index.htm) .

Gussenhoven, Carlos
2004 *Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.

Horne, Merle
2006 The filler *EH* in Swedish. In Proceedings of Fonetik 2006, *Working Papers* (Dept. of Linguistics and Phonetics, Lund University) 52: 65–68.

House, David
2003 Perceiving question intonation: the role of pre-focal pause and delayed focal peak. *Proceedings of the 15th International Conference of the Phonetic Sciences* (Barcelona): 755–758.

Ladefoged, Peter
1982 The linguistic use of different phonation types. *University of California Working Papers in Phonetics* 5: 28–39.

Lakoff, George
1972 Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic* 2: 458–508.

Lickley, Robin
1994 *Detecting disfluency in spontaneous speech*. Ph.D. dissertation, University of Edinburgh.

Nakatani, Christine, and Julia Hirschberg
1994 A corpus-based study of repair cues in spontaneous speech. *Journal of the Acoustical Society of America* 95: 1603–1616.

Nittono, Miharu
2003 *Japanese hedging in friend-friend discourse.* Ph.D. dissertation, Columbia University.

Ogden, Richard
2001 Turn transition, creak and glottal stop in Finnish talk-in-interaction. *Journal of the International Phonetic Association* 31: 139–152.

Ohala, John
1983 An ethological perspective on common cross-language utilization of f0 in voice. *Phonetica* 41: 1–16.

Pierrehumbert, Janet, and Julia Hirschberg
1990 The meaning of intonational contours in the interpretation of discourse. In *Intentions in Communication*, Philip Cohen, Jerry Morgan and Martha Pollack (eds.), 271–311. Cambridge, Mass.: MIT Press.

Shriberg, Elizabeth
2001 To 'errrr' is human: ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association* 31: 153–169.

Strangert, Eva
1991 Phonetic characteristics of professional news reading. Proceedings from Fonetik 1991, *PERILUS* XIII: 39–42. Stockholm: Stockholm University.