The filler EH in Swedish

Horne, Merle

Published in:
Working Papers

2006

Citation for published version (APA):
The Filler $EH$ in Swedish

Merle Horne
Dept. of Linguistics and Phonetics, Centre for Languages and Literature, Lund University
merle.horne@ling.lu.se

Abstract
Findings from a pilot study on the distribution, function and phonetic realization of the filler $EH$ in interviews from SweDia2000 interviews are presented. The results show that $EH$ occurs almost exclusively after function words at the beginning of constituents. The phonetic realization of $EH$ was seen to be of three basic forms: a middle-high vowel (e.g. [e], [ɛ], [ɔ]), a vowel+nasal (e.g. [ɛm], [ɛm], [ɛn]), and a vowel with creaky phonation (e.g. [ɔ], [ŋ]). The vowel+nasal realization occurs as has been shown for English before other delays and is associated with planning of complex utterances. Since creaky phonation is associated with terminality, the creaky vowel realization of $EH$ could be interpreted as signalling the juncture between the filler and an upcoming disfluency.

1 Introduction
The filler, or ‘filled pause’ $EH$ has often been termed a ‘disfluency’ (e.g. Shriberg, 2001), since it constitutes a delay in the flow of speech associated with referential meaning. However, since it can often be assigned pragmatic functions, such as signalling an upcoming focussed word (Bruce, 1998), or need on the part of the speaker to plan or code his/her speech and thus a desire to hold the floor, $EH$ can also be considered to be an integral part of the linguistic system (see e.g. Allwood (1994), and Clark & Fox Tree (2002) who refer to it as a ‘word’). In a study on English, Clark & Fox Tree (2002) found that its realization as $Uh$ signals a minor delay in speaking, whereas $Um$ announces a major delay in speaking.

A number of studies on Swedish have reported some characteristics of $EH$ in different speaking styles, but none have focussed on the variation in the phonetic realization of $EH$ as far as I know. Hansson (1998), in a study on the relationship between pausing and syntactic structure in a spontaneous narrative, found that the filled pauses (n=22) in her material occurred at clause boundaries after conjunctions and discourse markers and before focussed words. Lundholm (2000) in a study on pause duration in human-human dialogues found that the filler $EH$ (n=55) in authentic travel-bureau dialogues occurred in turn non-initial position and had a duration similar to silent planning pauses (mean = 340 ms). Eklund (2004) in a number of studies on simulated human-human and human-machine dialogues found that the most common position of $EH$ (n=2601) was utterance-initial before another disfluency and that it was most often followed by $jag$ ‘I’, and $det/den$ ‘it’. The filled pauses were found to have a mean duration of about 500 ms, thus considerably longer than those found by Lundholm (2000) in authentic task dialogues.

2 Current study
The present study has been carried out to pursue the investigation of $EH$ in spontaneous data to get some better idea as to its distribution, function, and phonetic realization in authentic interviews where the speech is basically of a monologue style. Spontaneous speech from the
SweDia 2000 interview material was used for the study (<http://www.swedia.nu/>). The speech of two female speakers from Götaland (a young woman from Orust and an older woman from Torsö) was transcribed and all EH fillers were labeled.

3 Results
3.1 Distribution of EH
The spontaneous SweDia data showed that EH occurs principally in non-utterance-initial position. There were only two cases of EH in utterance-initial position in the data studied and their mean duration was 899 ms.

EH occurs almost exclusively as a clitic to a preceding function word: 127 of the 137 instances of EH were cliticized to a preceding function word. The most frequent function words preceding EH were the coordinate conjunctions och ‘and’ and men ‘but’ which often have discourse functions, e.g. introducing topic continuations, new topics, etc. 52 cases of EH occurred after these two function words. Och EH ‘and UH’ was the most common function word+filler construction and was often (in 30 of 38 cases) preceded or followed by an inhalation break, a clear indicator of a speech chunk boundary (see Horne et al., 2006). The second most frequent function word category preceding EH was the subordinate conjunction att ‘that’ which also sometimes functions as a discourse marker introducing a non-subordinate clause. 24 instances 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=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.

3.2 Phonetic realization of EH in non-initial position
Three basic realizations of the filler EH have been observed: (1) a middle-high front or central vowel: e.g. [e], [ɛ], [ɔ] (see Fig. 1), (2) a nasalized vowel or vowel+nasal consonant: e.g. [ɛn], [ɔm], [ɔ n] (see Fig. 2), (3) a glottalized or creaky vowel: e.g. [ə], [ɛ] (see Fig. 3).

Figure 1. Example of the realization of EH as the middle high vowel [ɔ].

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 & Fox Tree (2002) analysed, i.e. they were
always followed by other kinds of ‘delays’, sometimes several in sequence as in Figure 2 with SWALLOW, SMACK, INHALE following EH.

**Figure 2.** Example of the realization of EH as a vowel+nasal [em]. Notice the other delays (SWALLOW, SMACK, INHALE) following [em].

The creaky vowel realizations of EH were the fewest (n=31) and 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 creaky filler was a silent pause. Creaky fillers are in some sense unexpected, since EH is often assumed to be a signal that the speaker wants to hold the floor, whereas creak, on the other hand, is assumed to be a signal of finality (Ladefoged, 1982). Nakatani & Hirschberg (1994), however, have observed that glottalization is not uncommon before a speech repair, and thus the creaky EH could, therefore, be interpreted as a juncture signal for an upcoming disfluency. Observation of the SweDia data shows in fact that the creaky realizations have a tendency to occur before disfluencies, as in the following examples: a) *men den såg ju inte ut [ə] det var någon ‘but it did not look [ə] it was somebody’, b) *å då var det en [ə] en kar som heter Hans Nils som blev ordförande ‘and then there was a [ə] a guy named Hans Nilsson who became chairman’. Creaky fillers also occur in environments where the speaker seems to be uncertain or have problems in formulating an utterance, e.g. för
att då blev det ju så att PAUSE [ʔ] PAUSE Johannesberg det skulle ju läggas ner ‘since it happened that PAUSE [ʔ] PAUSE Johannesburg it was going to be shut down’.

4 Summary and conclusion
This study on the distribution, function and phonetic realization of the filler EH has shown that the occurrence of EH in the SweDia spontaneous speech studied here is mostly restricted to a position following a function word at the beginning of an utterance. This supports and generalizes the findings of Hansson (1998) and Lundholm (2000) who found the filler EH most often in utterance internal position after conjunctions/discourse markers in spontaneous speech, both of a monologue and dialogue type. This differs from the findings for simulated task-related dialogues in Eklund (2004), where the filler EH occurred almost exclusively in utterance-initial position. This difference is most likely due to the simulated nature of the speech situation where the planning and coding of speech is more cognitively demanding.

As regards the phonetic realization of the filler EH, the patterning in Swedish is seen to be partially similar to the findings of Clark & Fox Tree (2002) for English: A vocalic realization of EH occurs before shorter delays in speech whereas a vowel+nasal realization correlated with relatively longer delays in speech. The duration of the vocalic realizations in the present data (mean = 268 ms) corresponds rather well with the median duration for EH found by Lundholm (240 ms) in spontaneous dialogues; thus, we would expect that the fillers in her data were realized mainly as a vowel such as ([e], [e], [ə]). A third realization, with creaky phonation, whose distribution overlaps with the other two realizations would appear to be associated with relatively more difficulty in speech coding; the creaky phonation, associated with termination, perhaps signals that the speaker has problems in formulating or coding his/her speech, and was observed to sometimes occur before repairs and repetitions. More data is needed, however, in order to draw more conclusive results.

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

References
SweDia 2000 Database: http://www.swedia.nu/.