

Hesitation Markers and Audience Design: Position Matters

Kerstin Fischer^{1*} and Nathalie Schümchen²

¹*Department of Design and Communication, University of Southern Denmark, Sonderborg kerstin@sdu.dk*

²*Department of Design and Communication, University of Southern Denmark, Kolding nats@sdu.dk*

In this study, we investigated the effect of framing a speech event as an instance of teaching in comparison to delivering a talk on the perception of hesitation markers. We expected that, given the interpersonal and discourse structuring functions of hesitation markers and their role in marking important information for the listener, people would judge hesitation markers more positively if they occurred in pedagogical contexts. That is, when a speaker ostensibly designs their utterances for the particular communication partner, hesitation markers should be evaluated more positively. The findings of our questionnaire study of the speech of six TED talkers support these hypotheses, but significantly more so for hesitation markers before important words inside the clause. Thus, position has an influence on the effects of hesitation markers with respect to audience design.

INTRODUCTION

While hesitation markers are often treated as an unwanted speech behavior that should be avoided as much as possible [1], they actually fulfill numerous useful functions in interactions. In particular, they function to indicate ongoing thought processes [2], which helps structure the information and make one's thought processes transparent and thus accessible to the partner; furthermore, hesitation markers indicate ad hoc production in comparison to canned, prefabricated speech, which, in turn, serves a social-interactive function.

We therefore hypothesized that if listeners focus on the degree to which a speaker designs their utterance for the respective audience, for instance, in order to teach important information, listeners would appreciate the functions of hesitation markers and not associate them with nervousness or lack of knowledge [3].

METHOD

The questionnaire study was carried out using a between-subject design. In order to study the effect of the framing of the speech event, we developed a questionnaire in which people were either told that they would hear excerpts from 'great teachers' or from 'great speakers'. Then, participants heard short audio files of 11-21secs extracted from six TED talks, which they had to rate according to the speaker's perceived traits. Subsequently, participants had to answer comprehension questions concerning the sentences they had heard before, since previous work suggests that hesitation markers contribute to better comprehension and memory [4].

Stimuli Creation

We selected three male and three female TED talkers from a variety of disciplines, where topics range from robotics to sociology. We selected three initial hesitation markers

in discourse structuring functions and three medial ones that occur before important words. In one condition, participants heard the original version of the stimulus (utterance including 'uh'), in another condition, the hesitation marker was edited out, and in a third condition, the hesitation marker was edited out and replaced by silence. Each participant was presented with six stimuli, each uttered by a different speaker: two original stimuli including 'uh', two without hesitation, and two with silence instead of 'uh'.

Questionnaire

All stimuli were integrated into an online survey using LimeSurvey. The survey started out with a welcome text followed by demographic questions. In order to test our hypothesis that hesitation markers serve an important function regarding addressee orientation, the audio stimuli were framed in one of two ways: participants were told to listen either to "great speakers" or to "great teachers". The framing was reinforced by asking the participants to rank their expectations toward what either a good speaker or teacher is supposed to be good at, such as *Speaks fluently, Preparedness, Intelligence, Friendliness, High education, Focuses on current task*. These attributes were selected on the basis of known preconceptions about hesitation markers.

Thus, two differently framed questionnaires that each included two utterances of each condition in random order were designed (resulting in 90 possible combinations). The dependent measures are participants' responses to questions about pragmatic function, but also about the suspected degree of audience design exhibited by the speaker. Therefore, the participants had to answer two sets of questions after each audio clip. The first set of questions addressed to which extent the speaker is perceived as trying to get something across, is involved, wishes the listener to really understand, and is perceived as friendly, likeable and polite (among other categories). The second

set of questions concerned participants' expectations about good speakers and good teachers.

The survey was sent out via the crowdsourcing platform "Prolific". We decided to only recruit native speakers of English.

RESULTS

In total, 223 participants filled out our survey. 47 participants were excluded due to incomplete surveys, completion durations under minimum time, or L1 other than English, which left us with 176 participants, whose mean age is 37 years (range 16-73), and evenly distributed across gender (81 female, 94 male, 1 other). 19 are students, 36 hold an MA degree or higher, 42 have completed high school, 49 professional training, and 30 replied 'other'.

A first analysis of the data shows significantly more positive ratings for utterances with hesitation markers marking important words as well as more prefaced important words with 'uh' were rated as significantly less nervous and unconcentrated. With regard to the framings, analyses reveal significant differences in the ranking of important attributes. 35.8 % of all participants in condition 1 put *speaks fluently* on the first rank while the majority in condition 2 (45.1 %) prioritized *knowledge about topic* highest. In both conditions, *high education* ranks lowest (condition 1: 74.4 %; condition 2: 48.8).

These different expectations correlate significantly with participants' ratings of the speakers as honest and friendly when hesitation markers occurred clause initially, and as significantly more educated, honest, knowledgeable, nervous and wanting to get their point across when they used hesitation markers before important words. However, while these correlations are all significant, correlations are low, ranging between $r=0.09$ to 0.25 .

Because of the considerable differences based on position, an independent-samples t-test was conducted to compare the subjective ratings of hesitation markers dependent on their positioning. There are consistent highly significant differences in the scores for initially positioned hesitation markers and hesitation markers prefacing important words; in particular, speakers who use hesitation markers in medial position before important words are rated significantly higher with respect to how important it is to them whether the listener understands them, how important it is to them to get their point across, the degree with which they take their partner into account and with respect to how intelligent, educated, knowledgeable and prepared they are, but also concerning nervousness and lack of concentration.

A chi-square test of the relation between hesitation markers' position and the participants' performance at comprehension questions shows that significantly more participants gave the correct answers to comprehension questions when the hesitation marker occurred before the

important word (62.8%). Only half of the participants (48.5%) answered correctly when hesitation markers were in initial position ($p=0.003068$; $X^2=8.7665$).

DISCUSSION

The results on the effects of the framing show that people's expectations have an influence on their evaluation of hesitation marker occurrences. However, these effects are much more pronounced for utterance-medial occurrences before important words; the results show consistently more positive evaluations for 'uh' in this position, even though it is probably much more salient there [5]. The analysis of comprehension effects also favors hesitation markers in medial position.

ACKNOWLEDGMENTS

The authors would like to thank the Danish Council of Independent Research for their support of the project *Improving Second Language Pedagogy at the Prosody-Grammar-Pragmatics Interface Using Human-Robot Interaction*.

REFERENCES

- [1] Clark, Herbert H. & Fox Tree, Jean E. (2002): Using *uh* and *um* in Spontaneous Speaking. *Cognition* 84: 73-111.
- [2] Fischer, K. 2000. From Cognitive Semantics to Lexical Pragmatics: The Functional Polysemy of Discourse Markers. Mouton de Gruyter.
- [3] Fox Tree, J. E. (2001): Listeners' uses of *um* and *uh* in speech comprehension. *Memory and Cognition* 29: 320-326.
- [4] Fox Tree, Jean E. (2002): Interpreting Pauses and *Ums* at Turn Exchanges. *Discourse Processes* 34,1: 37-55.
- [5] Niebuhr, O. & Fischer, K. (2019): Do not hesitate – unless you do it shortly or nasally. *Proceedings of Interspeech* 2019.