Supplementary Motor Area Activation in Disfluency Perception. An fMRI Study of Listener Neural Responses to Spontaneously Produced Unfilled and Filled Pauses

Robert Eklund1 & Martin Ingvar2,3
1 Department of Culture and Communication, Linköping University, Linköping, Sweden
2 Department of Clinical Neuroscience, Karolinska Institute, Stockholm, Sweden
3 Osher Center for Integrative Medicine, Karolinska Institute, Stockholm, Sweden

ABSTRACT
Spontaneously produced Unfilled Pauses (UPs) and Filled Pauses (FPs) were played to subjects in an fMRI experiment.

For both stimulus increased activity was observed in the Primary Auditory Cortex (PAC). However, FPs, but not UPs, elicited modulation in the Supplementary Motor Area (SMA), Brodmann Area 6.

Our results provide neurogenic confirmation of the disfluency effect—between UPS and other kinds of speech disfluency and could also provide a partial explanation for the previously reported beneficial effect of FPs on reaction times in speech perception.

Our results also have potential implications for the two of the suggested functions of FPs: the "floor-holding" and the "help-me-get-back-hypothesis".

Keywords: speech disfluency, filled pauses, unfilled pauses, speech perception, spontaneous speech, fMRI, Auditory Cortex, PAC, Supplementary Motor Area, SMA, Brodmann Area 6, etc.

INTRODUCTION
Almost no one is completely fluent when speaking.

Speech disfluency has been studied since at least the 1930s. Disfluencies have commonly been regarded as "performance errors" in speech production.

However, several studies indicate that certain kinds of disfluencies can have beneficial effects on listener perception (Tracy & Watson, 2011; Bar, Brey, & Schindler, 2012; Fortune & Bakker, 2004; For Tare, 2001; pitch, 1980).

The most common voice disfluency is the filled pause (FP), "et cetera". The reported average frequency of filled pauses (FPs) ranges from 1.5 to 7.6% at word level (Eklund, 2010).

Neuroscientific studies on disfluency have traditionally consisted of psycholinguistic studies, starting with Kutas and Hillyard (1980) on the N400 component.

The present study uses functional Magnetic Resonance Imaging (fMRI) to analyse the effect of authentic disfluencies to study the effect of unfilled and filled pauses on brain processing.

METHOD (1)
Subjects
The subjects were 15 healthy adults (9 females, 6 males) aged 18 to 35 years.

The fMRI system has high temporal resolution, but the fMRI scanner used has high temporal resolution, but the fMRI scanner used was not a high-resolution scanner. The fMRI scanner used was an off-the-shelf scanner.

Equipment
The fMRI scanner used was the General Electric 1.5 T Signa EXC Freefield Head (Signa). The fMRI scanner used was an off-the-shelf scanner.

The stimulus data used were excerpts from a human-human dialog speech data created by Eklund (2001: 1-108). Subjects were asked to play the role of travel agents listening to customers reading travel bookings over the telephone, including a task block (Eklund, 2004b: 105).

From the original data set, four repetitions were chosen (SPGR) and a continuous epoch, 240 ms, was used for the stimuli. The stimuli were 400 ms in length.

Stimulus data is shown in Table 1 below.

ANALYSES AND RESULTS (1)
Using Fluent Speech (FS) as the baseline condition, the following three conditions were analyzed:

(1) Filled Pauses > Fluent Speech
(2) Unfilled Pauses > Fluent Speech
(3) Filled Pauses > Unfilled Pauses

The results were calculated with a false positive few type II error rate p < 0.05. They were further divided with a cluster level threshold of 10 contiguous voxels. No activation in Brodmann Area 22, associated with semantic processing, was observed.

COMMENTS ON DATA AND ANALYSES
1. We used fMRI to study disfluency perception, not aEEG with its focus on temporal aspects of speech perception.
2. We investigated perceptual modulation caused by FPs, not speech rhythm (orients, disfluency) or general cognitive processing.
3. Unlike previous studies where the auditory stimuli were produced by speaking using an off-the-shelf scanner, we used ecologically valid stimulus data.

ANALYSES AND RESULTS (2)
(1) UPS and FPs modulated the Primary Auditory Cortex (PAC). That heightened attention influences auditory perception has been shown (Fellows et al., 2004; and more) and the observed attention-enhancement function of FPs could possibly help explain the shorter reaction times to linguistic stimulus that follow FPs as reported by e.g. Fox (2001; 1996). Note, however, that UPS also exhibited PAC, but the observed "attention-enhancement function of FPs could possibly help explain the shorter reaction times to linguistic stimulus that follow FPs as reported by e.g. Fox (2001; 1996)."

Both UPS and FPs lead to heightened attention.

(2) Filled Pauses activated (pre-)motor areas, (e.g. BA6). (Not observed in UPS)

DISCUSSION AND CONCLUSIONS
(1) "Floor-holding" hypothesis: FPs at first proposed by McCarthy and Ojemann (1999) that FPs are used to keep the floor in polylogy.

→ Our results imply that FPs could possibly be used to keep the floor in polylogy.

(2) "Help-me-out" hypothesis: as suggested in and by Wilkes-Gibbs (1988). FPs are expected to act as a "semicomplementary" signal asking for interlocutor help in conversation.

Our results imply that FPs might serve this function well.

SUMMING IT UP
Our results - conceptually speculative - suggest that FPs, but not UPS, activate motor areas in the listener brain.

Both FPs and UPS activate PAC, which lends support to the attention-heightening hypothesis that has been forwarded in the literature and would now seem clear that it is not the break in the speech stream per se that causes this activation, since UPS seemingly do not have this effect.