Neural correlates of the processing of unfilled and filled pauses

Robert Eklund1, Peter Fransson2, Martin Ingvar1,3
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. While both stimuli resulted in increased activity in the Primary Auditory Cortex (PAC) and the Supplementary Motor Area (SMA), FPs were associated with increased activation in BA6, and the Supplementary Motor Area, Brodmann Area 6.

This observation provides neurocognitive confirmation of the oft-reported difference between FPs and other kinds of speech disfluency and also could provide a partial explanation for the previously reported beneficial effect of FPs on reaction times in speech perception.

The results are discussed in the light of the suggested role of FPs as floor-holding devices in human polysyllabics.

Keywords: speech disfluency, filled pauses, unfilled pauses, speech perception, spontaneous speech, MRI, Auditory Cortex, PAC, Supplementary Motor Area, SMA, Brodmann Area 6, BA6

INTRODUCTION

Almost no one is completely fluent. The most common voiced disfluency is the filled pause (FP), "uh." The reported average frequency of filled pauses (FPs) ranges from 1.9 to 7.6% (Eklund, 2010).

Studies indicate that certain kinds of disfluencies can have beneficial effects on listener perception (Fraunorf & Watson, 2011; Barr & Seyfeddinipur, 2010; Ferreira & Bailey, 2001; Fox Tree, 2003; Reich, 1990).

Neurocognitive studies have traditionally focussed on electrophysiological studies, beginning with Kutas and Hillyard (1980) on N400 component.

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

METHOD (1)

Subjects

The subjects were 16 healthy adults (9 F/7 M) ages 22–54 (mean age 40.3, standard deviation 9.3). All subjects volunteered to participate in the experiment. All subjects were right-handed as determined by the Edinburgh Handedness Inventory (Oldfield, 1971). All subjects received higher education. After a full explanation of the experimental design and informed consent was obtained from all subjects, a written participation remuneration was also administered.

Equipment

The fMRI scanner used was the General Electric 1.5 T Signa HDX 3.0-Tesla scanner at Karolinska Institute, Stockholm, Sweden. The coil used was a General Electric Starlark bird-head head coil (3.5 T).

Stimulus data

The stimulus data were taken from the human-human dialog speech data described in (Eickhoff, 2004; 2004-103-106). Subjects were asked to pay the role of talker agents listenting to customers needing travel bookings over the telephone, following a task script (Eickhoff, 2004-103-106). From the original data set, four speakers were chosen (2M/2F) and a number of sentences were elicited that were fluent except that they were filled by pauses (30%). A number of non-fluent sentences were also elicted, resulting number of both UPs and FPs roughly correspended to reported incidences of UPs and FPs in spontaneous speech.

Stimulus data are shown in Table 1.

ANALYSES AND RESULTS (1)

Using Fluent Speech (FS) as the baseline condition, the following three contrasts were analyzed:

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

The results were calculated with a False Discovery Rate (FDR) at p < 0.05 (Genovese, Lazar & Nichols, 2002) with a cluster level threshold of 10 contiguous voxels.

No activation in Brodmann Area 22, associated with semantic processing, was observed.

METHOD (2)

Experimental design

The stimulus file described above were used in an event-related design where single-word auditory stimuli were presented on a computer screen. The subjects were briefed whether they were still awake/focused on the task, breathing naturally, and to concentrate as all to allow for reliable BOLD acquisition. FPs and UPs were modeled as events in SPM and were convolved using the Haemodynamic Response Function (HRF) in SPM.

Experimental setting

The subjects lay supine head first in the scanner with earplugs to protect them from scanner noise and headphones with the sound played to them. The perceived sound level was fairly sufficient and no subjects reported having any problems hearing what was said. Head movement was constrained using foam wedges and/or tape.

Experimental instructions

The subjects were instructed to listen carefully to what was said, as if they were the addressed travel agent, but that they were not expected to react verbally to the alternatives or any other way, only that they needed to pay attention to the information provided by the client. All subjects understood the procedure without any confusion.

Analysis and data acquisition

Stimulus data

The fMRI scanner used was the General Electric 1.5T Excite HD Twinspeed scanner at Karolinska Institute, MR-center, Stockholm, Sweden. The coil used was a General Electric Starlark bird-head head coil (3.5 T).

Stimulus data were taken from the human-human dialog speech data described in (Eickhoff, 2004-103-106). Subjects were asked to pay the role of talker agents listenting to customers needing travel bookings over the telephone, following a task script (Eickhoff, 2004-103-106). From the original data set, four speakers were chosen (2M/2F) and a number of sentences were elicited that were fluent except that they were filled by pauses (30%). A number of non-fluent sentences were also elicited, resulting number of both UPs and FPs roughly corresponded to reported incidences of UPs and FPs in spontaneous speech.

Stimulus data are shown in Table 1.

ANALYSES AND RESULTS (2)

Results are shown in Table 2 and Figure 1 below:

DISCUSSION AND CONCLUSIONS

(1) UPS and FPs modulated the Primary Auditory Cortex (PAC). That heightened attention influences auditory perception has been shown (Pettiti et al., 2004 and more) and the observed attention-heightening function of FPs could possibly help explain that shorter reaction times to linguistic stimuli that follow FPs as reported by e.g. Fox Tree (2001, 1995). Note, however, that UPS also exhibited PAC activation.

Both UPSs and FPs led to heightened attention

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

Known already since Bricken (1940) that SMA is active in the processing of speech, and several later studies have confirmed both SMA and pre-SMA play a role in speech production (e.g. Goldberg, 1988; Alario et al., 2006) and speech perception (e.g. Jacobson, 2006; Wilson et al., 2006) and that FPs seem to “kick-start” speech production.

POTENTIAL IMPLICATIONS

(1) “Floor-holding” hypothesis of FPs, as first proposed by e.g. Jacobson (1940) that FPs are used to keep the floor in polyphony

Our results imply that FPs act as counter-productive for this.

(2) “help-me-out” hypothesis, as suggested in Clark and Wilkes-Gibbs (1986): FPs can be used as a (semi-definite) signal asking for interlocutor help in conversation

Our results imply that FPs might serve this function well

SUMMING IT UP

Our results suggest that FPs – unlike UPs – and FPs – activate motor areas in the listening brain. However, both FPs and UPSs activate PAC, which lends support to the attention-heightening hypothesis that has been forwarded in the literature.

It would also 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.