

Running head: ATTENTION AND HESITATIONS IN SPEECH

Attention orienting effects of hesitations in speech: Evidence from ERPs

Philip Collard, Martin Corley and Lucy J. MacGregor

University of Edinburgh

School of Philosophy, Psychology and Language Sciences

David I. Donaldson

University of Stirling

Department of Psychology

Philip Collard

Psychology

School of Philosophy, Psychology and Language Sciences

University of Edinburgh

Edinburgh EH8 9JZ, UK

(tel) +44 131 651 1304; (fax) +44 131 650 3461; P.Collard@sms.ed.ac.uk

### **Abstract**

Filled-pause disfluencies such as *um* and *er* affect listeners' comprehension, possibly mediated by attentional mechanisms (Fox Tree, 2001). However, there is little direct evidence that hesitations affect attention. The current study used an acoustic manipulation of continuous speech to induce attention-related ERP components (Mismatch Negativity [MMN] and P300) during the comprehension of fluent and disfluent utterances. In fluent cases, infrequently occurring acoustically manipulated target words gave rise to typical MMN and P300 components when compared to non-manipulated controls. In disfluent cases, where targets were preceded by natural sounding hesitations culminating in the filled pause *er*, an MMN (reflecting a detection of deviance) was still apparent for manipulated words, but there was little evidence of a subsequent P300. This suggests that attention was not reoriented to deviant words in disfluent cases. A subsequent recognition test showed that non-manipulated words were more likely to be remembered if they had been preceded by a hesitation. Taken together, these results strongly implicate attention in an account of disfluency processing: Hesitations orient listeners' attention, with consequences for the immediate processing and later representation of an utterance.

### **ERP evidence for the attention orienting effects of hesitations in speech**

Disfluency is common in spontaneous speech (Fox Tree, 1995). Listeners encounter disfluency with such regularity that its effects on speech processing are of natural interest to those who research spoken-language comprehension. Converging lines of evidence show that disfluency can affect the way in which an utterance is understood. For example, hesitations in speech affect the confidence that listeners have in speakers' knowledge (Brennan & Williams, 1995), and disfluent corrections of a message may leave a lingering representation of the original content (Ferreira, Lau, & Bailey, 2004). Hesitations also affect syntactic representation, marking breaks in syntactic structure at phrase boundaries (Bailey & Ferreira, 2003).

But what happens at the point at which a disfluency has been encountered? Research addressing this question has tended to focus on hesitation-type disfluencies because these are often associated with local markers, such as elongations to words (e.g., *thee*) and filled pauses (e.g., *um*, *uh* or, in British English, *er*). Recently, Corley, MacGregor and Donaldson (2007) used Event-Related Potentials (ERPs) to demonstrate an immediate effect of hesitations while listening to spoken utterances such as (1) and (2).

(1) Everyone's got bad habits and mine is biting my [er] nails.

(2) Everyone's got bad habits and mine is biting my [er] tongue.

Using the N400 effect as an index of integration difficulty, they compared listeners' responses to unpredictable (difficult to integrate) words (2) against predictable words (1) in fluent contexts, and in disfluent contexts where the critical words were preceded

by hesitations. The magnitude of the N400 (predictability) effect was significantly reduced for disfluent utterances, showing a clear effect of hesitations on listeners' language processing. Importantly, the N400 differences were associated with representational differences: listeners were more likely to remember words which had been preceded by a hesitation in a forced-choice recognition task. One possible account of these findings is based on linguistic prediction, or expectancy. There is increasing evidence that listeners make online predictions during language comprehension (e.g., Altmann & Kamide, 1999; Van Berkum, Brown, Zwitterlood, Kooijman, & Hagoort, 2005). Furthermore, eyetracking evidence suggests that hesitations marked by prolongations such as *thee* and filled pauses such as *uh* may lead listeners to update their predictions about upcoming words (Arnold, Tanenhaus, Altmann, & Fagnano, 2004). Specifically, Arnold et al., (2004) showed that following hesitation, listeners were more likely to predict the upcoming mention of a discourse-new object, albeit from a limited set of candidate referents. In the absence of sufficient information from the environment regarding possible speech referents, hesitation may cause a reduction in the extent to which specific predictions are made, leading to the N400 attenuation observed by Corley et al. (2007).

It is clear that disfluency can affect linguistic processes, such as prediction, but such processing differences may in turn be predicated on other mechanisms, such as attention. Compared to the most likely continuation of an utterance (fluent production of the next word), disfluency introduces novelty. Such novelty might occupy attention and hence limit the processing of the following part of the utterance. Alternatively, the novelty could enhance attention to, and facilitate the processing of, subsequent words. The existing data seem to support the latter alternative. In a word monitoring task, for example, Fox Tree (2001) found that participants identified targets more quickly

following a hesitation including *uh*, which she attributed to heightened attention. An effect of disfluency on attentional processes might also account for findings that listeners respond more quickly to disfluent instructions (Brennan & Schober, 2001) and are more likely to remember words that follow a disfluent hesitation (Corley et al., 2007). Importantly, once attention is directed at an utterance, standard predictions as to what may follow may be affected.

Given these suggestions, the aim of the present study is to investigate directly the contention that attention for subsequent material is affected by disfluent hesitation in speech, using an ‘oddball’ ERP paradigm. In such experiments, listeners are occasionally presented with stimuli that are physically deviant from more frequent standard stimuli, for example with respect to pitch or amplitude. The deviant stimuli elicit a cascade of neural events related to their detection and the orientation of attention towards them. The ERP effects commonly elicited by such oddball stimuli are the Mismatch Negativity (MMN) and members of the P300 family of components, such as the P3a and P3b. The MMN, an early (100-250 ms post stimulus) centro-frontal negative difference wave (Schröger, 1997) appears to index neural processes involved in identification of deviance in the acoustic environment and can be modulated by highly focused attentional states (Alho, 1995). Occurring after the MMN at around 300 ms post stimulus, the frontally maximal P3a and the subsequent parietally maximal P3b are positive components typically associated with identification of, and attentional orientation to, deviant stimuli, and with the subsequent induced memory updating (Polich, 2004). Modulation of these attention related ERP components following hesitations would provide strong evidence that the hesitations alter the attentional state of listeners.

In the current study, participants listened to recorded utterances containing infrequent changes to the auditory characteristics of single words. Half of the time, the manipulated words followed hesitations. These were marked by natural changes to the speech, such as elongations to words within the hesitation (e.g., *thee*), and the filled pause *er*. The acoustic changes were designed such that the manipulated words would be physically deviant from the acoustic regularities set up by the preceding speech, but did not alter the linguistic content of the utterances. Because the deviant words were infrequent and therefore novel with respect to their contexts, they would be expected to induce equivalent attention-related ERPs in both fluent and disfluent conditions, unless, as we predicted, the attentional state of listeners was affected by preceding disfluency. If hesitations result in changes to the processing of subsequent words (indexed by alterations to the ERP signal) then we might expect some longer lasting changes to the representation of these words. Following Corley et al. (2007), we assessed this in a surprise recognition memory test.

## **Method**

### **Participants**

Twelve native English speakers participated in the experiment (7 male; mean age 23 years; range 17-36). All were right handed and reported no known neurological impairment. Informed consent was obtained in accordance the University of Stirling Psychology Ethics Committee guidelines. Participants were given financial compensation and course credit where applicable.

## Materials

The stimuli consisted of 160 pairs of recorded utterances taken from Corley et al. (2007; an example is given in 1 above) which ended with a highly predictable target word (mean cloze probability 0.84). Fluent and disfluent versions of utterances were recorded by a native British English speaker who was instructed to produce the utterances as naturally as possible. Disfluent versions incorporated a hesitation before the utterance-final word which included signs of disfluency that were natural to the speaker, such as prolongations to preceding words (e.g., the prolonged definite article *thee*) and culminated in a filled-pause *er*. Utterances were recorded with a pseudo-target ‘pen’ so that there were no acoustic cues to the upcoming word. Targets were recorded in separate carrier sentences and spliced onto the fluent and disfluent utterances, resulting in acoustically identical targets across the fluent and disfluent contexts. An additional 80 unrelated filler utterances were recorded. These were of a similar length to the experimental utterances. Half contained various types of disfluency, including hesitations marked by filled pauses, and disfluent repairs at varying positions within the utterances. Using the 320 experimental recordings, 320 additional stimuli were created by acoustically manipulating the target words to make them acoustically deviant. To do this, we applied an equalisation pattern that was biased to the mid-range frequencies from the target word onset until the end of the utterance and resulted in an amplification of 2.8dB across all frequencies except for the 125-1000Hz range. In this range we applied a bell curve-like pattern which ranged from 2.8dB to 18dB and peaked at 500Hz. The salient effect of the manipulation was to make the speech sound momentarily compressed, not unlike speech over a poor telephone line.

Four versions of the experiment were created, each containing 40 fluent normal, 40 disfluent normal, 40 fluent manipulated, and 40 disfluent manipulated recordings.

Each target word occurred only once in each version of the experiment. Two copies of each of the 80 fillers were added to each set, resulting in a total of 320 recordings of which 80 ended in deviant target words. Thus the overall deviant to normal utterance ratio was 1 in 4, ensuring that manipulated stimuli remained relatively novel ‘oddballs’ throughout the experiment.

### **Procedure**

The experiment comprised two sections. In the first, participants listened to the 320 experimental utterances and fillers. Materials were presented in a random order via computer loudspeakers in two blocks lasting around 20 minutes each, and separated by a break of a few minutes. Participants were instructed to listen to the recordings as if they were part of a normal conversation, but were not given any other task. They were not told specifically about the presence of the disfluencies or acoustically manipulated words, but were told that occasionally, the sound editing quality would drop, which they should ignore.

Electroencephalogram (EEG) was recorded from 61 silver/silver-chloride electrodes embedded in an elasticized cap at standard 10-20 locations (Jasper, 1958), using a left-mastoid reference. Electro-oculograms (EOGs) were collected to monitor for eye-movements. EEG and EOG were amplified (bandpass filtered online, 0.01 – 40 Hz) and continuously digitized (16 bit) at 200Hz. Electrode impedances were kept below 5K $\Omega$ . Epochs were created from the EEG (150ms before the onset of the target words to 800ms after the onset) and these data were re-referenced offline to the average of the left and right mastoid electrodes, baseline corrected (relative to the average over the pre-stimulus interval) and smoothed over 5 points. Before averaging into ERPs, individual epochs were screened for drift of  $\pm 75\mu\text{V}$  over 500ms (amplitude difference between first and last data point of each epoch), and for artefacts of  $\pm 75\mu\text{V}$ . The



screening process resulted in the loss of 10.47% of epochs, with no significant variation in rejections between conditions [ $F(3,33) = 1.756$ ]. Average ERPs were formed time locked to the onset of target words for each participant (minimum of 16 artefact free trials were required for inclusion).

In the second section of the experiment participants performed a surprise recognition memory test for the material that they had heard. The 160 utterance-final (previously heard) target words were presented visually interspersed with 160 frequency-matched foil words, which had not been uttered at any previous point during the experiment. After a 500ms fixation cross, each word was presented for 750ms, followed by a blank screen for 1750ms. Participants were instructed to decide whether each word had occurred at any previous point during the experiment and respond ‘old’ or ‘new’ via a button-box placed in front of them. Responses which took longer than 2500ms were discarded.

## Results

ERPs associated with the onsets of deviant target words were compared to ERPs to non-manipulated standard controls for fluent and disfluent conditions. Because pre-stimulus baselines in fluent and disfluent utterances were different (including an *er* for disfluent cases), effects related to the acoustic manipulations were analysed separately for fluent and disfluent conditions.

---

Insert Figure 1 about here

---

Figure 1 shows the distribution of the oddball effects over 100-400ms. In fluent utterances, deviant words elicit an early negativity with an initial left hemisphere bias (100-150ms) which spreads laterally into a very typical MMN distribution (150-200ms). A large positive difference appears fronto-centrally at the midline (250-300ms) and develops into a widespread centroparietally maximal positivity (300-400ms). This pattern represents a typical P300 complex.

In disfluent utterances, effects are much smaller and less widespread. There is some indication of early negativity at the midline fronto-centrally (100-150ms) which becomes lateralised with a right hemisphere bias (150-200ms). No frontocentral positivity is apparent although a less focal and greatly diminished centroparietal positivity can be seen later (300-400ms).

Figures 2 and 3 show the waveforms of the MMN and P300 effects at electrodes used in the statistical analyses (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4), for fluent and disfluent utterances respectively. In fluent utterances (figure 2), deviant stimuli give rise to midline dominant MMN and P300 effects. There is clear indication of a P3a-like early frontal component (as with the topographic depiction of the data; figure 1). Data from disfluent utterances are presented on the same scale (figure 3) and show oddball effects which are much smaller in magnitude.

ERPs were quantified by measuring the mean voltages for deviant and standard targets over two time windows, consistent with the MMN (100-200ms) and the P300 (250-400ms), for fluent and disfluent utterances separately. Greenhouse-Geisser corrections to degrees of freedom were applied and corrected F and p values are reported where appropriate.

---

Insert Figure 2 about here

---

Analyses used three-way ANOVAs with factors of deviance (infrequent deviant, standard), location (electrodes F, C and P) and laterality (electrodes 3, z and 4).

For the fluent conditions, in the MMN time window, results showed a significant main effect of deviance [ $F(1,11) = 13.152$ ,  $\eta_p^2 = .545$ ,  $p = .004$ ] indicating that deviant stimuli elicited a widespread negativity across the scalp (mean voltages of  $-1.701\mu\text{V}$  and  $-.118\mu\text{V}$  for deviant and standard stimuli respectively). No other effects involving the factor of deviance reached significance [ $F_s < 2.170$ ].

In the P300 time window there was a significant effect of deviance [ $F(1,11) = 51.080$ ,  $\eta_p^2 = .823$ ,  $p < .001$ ] reflecting a positivity associated with deviant words that was widespread across the scalp (mean voltages of  $4.390\mu\text{V}$  and  $.325\mu\text{V}$ , for deviant and standard stimuli respectively). Significant deviance by laterality [ $F(2,22) = 10.045$ ,  $\eta_p^2 = .477$ ,  $p = .001$ ] and deviance by location by laterality [ $F(4,44) = 7.920$ ,  $\eta_p^2 = .419$ ,  $p < .001$ ] interactions indicate that the deviance effect was larger over midline sites, and that this midline bias was largest at frontal and posterior sites. No other effects involving the factor of deviance reached significance [ $F_s < 1.668$ ].

---

Insert Figure 3 about here

---

For the disfluent conditions, in the MMN time window, there was a significant deviance by location interaction [ $F(2,22) = 4.950$ ,  $\eta_p^2 = .310$ ,  $p = .017$ ] indicating negativity associated with deviant words that was confined to frontal and central sites (mean voltages of  $.776\mu\text{V}$ ,  $.100\mu\text{V}$ ,  $-.864\mu\text{V}$  for frontal, central and posterior sites respectively for the standard stimuli and  $-.266\mu\text{V}$ ,  $-.908\mu\text{V}$ ,  $-.859\mu\text{V}$  for the deviant stimuli). No other effects involving deviance reached significance [ $F_s < 2.092$ ].

In the P300 time window there was a significant deviance by location interaction [ $F(2,22) = 6.033$ ,  $\varepsilon = .553$ ,  $\eta_p^2 = .354$ ,  $p = .028$ ] indicating positivity associated with deviant words that was confined to posterior sites (mean voltages of  $1.619\mu\text{V}$ ,  $.606\mu\text{V}$ ,  $-.171\mu\text{V}$  for frontal, central and posterior sites respectively for the standard stimuli and  $.836\mu\text{V}$ ,  $1.084\mu\text{V}$ ,  $.974\mu\text{V}$  for the deviant stimuli). No other effect involving deviance reached significance [ $F_s < 2.024$ ].

These analyses demonstrate robust and typical MMN and P300 effects for acoustically deviant words in fluent stimuli. In disfluent contexts, the early negativity and later positivity are much weaker and less widespread, and there are some distributional differences between fluent and disfluent ERPs. However, the antecedents and gross topographies of the effects support an interpretation of MMN followed by P300 complex in each case. We therefore conducted a further analysis to compare effect sizes across fluent and disfluent conditions. Because the disfluent condition gave rise to interactions between deviance and location in both the MMN and P300 windows, location was also included as a factor in these comparisons. Each analysis was conducted on the deviance effect (ERPs to deviant items minus standard ERPs) using the factors of fluency and location, with the same electrode set as the previous analyses, collapsed across laterality. In the MMN time window, there were no significant effects involving fluency [ $F_s < 2.804$ ]. This is perhaps surprising in light of Figure 1, which corresponds to a mean difference between conditions of  $.898\mu\text{V}$  across electrodes. In the P300 time window, a large difference between the fluent and disfluent conditions (mean of  $4.434\mu\text{V}$  and  $.187\mu\text{V}$  for fluent and disfluent respectively) was confirmed [ $F(1,11) = 32.484$ ,  $\eta_p^2 = .747$ ,  $p < .001$ ]. The interaction of fluency and location was not significant [ $F(2,22) = 1.476$ ].

A final consideration was addressed using an additional analysis which examined the responses to disfluent items over time. By comparing responses during the first and second halves of the experiment, we were able to establish that the responses to deviant items following a hesitation did not differ over the course of the experiment, either for the MMN ( $F_s < 1.433$  for all effects involving half) or the P300 (deviance by half:  $F(1,11) = 2.187$ ; other  $F_s < 1.035$ ).

The second analysis focused on performance in the recognition task. As in Corley et al. (2007), the probability of correctly identifying words heard in the comprehension block of the experiment was quantified with stimulus identity treated as a random factor.<sup>1</sup> Overall, 57% of the previously-heard words were correctly recognized (false alarm rate 18%). Figure 4 shows the recognition probability of utterance-final words by fluency and deviance.

---

Insert Figure 4 about here

---

A 2-way ANOVA with factors of fluency and deviance revealed a significant interaction between the two factors [ $F(1,147) = 5.382$ ,  $\eta_p^2 = .035$ ,  $p = .022$ ]. For standard stimuli, a pairwise comparison of recognition probabilities for words which had been heard in fluent or disfluent contexts showed a significant difference [ $t(147) = 2.114$ ,  $\eta_p^2 = .030$ ,  $p = .036$ ], suggesting that acoustically normal words were more likely to be recognized following disfluency. Conversely, there was no difference in the recognition probabilities for deviant words [ $t(147) = 1.083$ ].

## Discussion

Large deflections in the ERPs were observed when participants encountered infrequently-occurring acoustically deviant words in standard fluent speech. Given their polarities, distributions, timings and antecedent conditions, it is clear that the ERP deflections correspond to the typical neural signatures of attention capture and orientation, the MMN and P300. When the same deviant words were encountered following a hesitation, there was some evidence for MMN and P300-like effects in the appropriate time windows. However, compared to the fluent case, amplitudes were greatly reduced, and distributions were less widespread.

Polich (2004) provides a model of ERPs elicited by auditory deviance. In his model, the MMN is associated with the detection of deviance by attentional systems. The P300 is driven by the novelty of the stimulus, and is associated with orientation of attention towards deviant stimuli (frontal P3a component) and subsequent memory-updating processes (parietal P3b component). The reduction of the observed ERP effects following disfluency in the present study provides *prima facie* evidence that hesitation affects the listener's attentional system. Moreover, the reduced response to novelty suggests that when the acoustically deviant words were encountered, attention was already oriented towards the speech, consistent with previous claims that hesitations heighten attention.

At first glance, these findings are reminiscent of results from attentional blink paradigms (e.g., Raymond, Shapiro, & Arnell, 1992). In attentional blink experiments, participants are less likely to detect a second target stimulus after a first, to which attention has presumably been oriented; this is accompanied by a reduced P300 to the second target (Vogel, Luck, & Shapiro, 1998). However, there are three reasons to

suggest that the present findings cannot be accounted for in terms of an attentional blink. First, Corley et al. (2007) have demonstrated that the N400 effect related to low cloze probability words is attenuated following hesitations. There is no equivalent N400 attenuation in the attentional blink paradigm (Vogel et al., 1998). Second, in attentional blink paradigms, the attentional attenuation tends to be maximal about 300ms after the onset of the initial orienting target (at a lag of 3 items, 100ms/item; Vogel et al., 1998). In the present study, the mean delay between the onset of the *er* and that of the target word was 598ms (SD 103ms) (this is a low estimate for the time between events because signs of disfluency such as word prolongations sometimes occurred before the *er*).

The third and most important reason for rejecting an attentional blink account comes from the recognition task. Hesitations cause subsequent (acoustically normal) target words to be *more* likely to be later recognised, in direct contrast to what would be predicted if hesitation induced an attentional blink. This increase replicates the finding of Corley et al. (2007) that differences in the processing of fluent and disfluent utterances lead to long-term differences in the representations of those utterances, and further suggests that despite the acoustic manipulations necessary for the purposes of the present study, participants were engaged in comparable language processing. Salient (here, deviant) items were recognised equally often whether they had originally been encountered in fluent or disfluent utterances, possibly ascribable to a ceiling effect, given the numbers of stimuli and time between encoding and recognition of up to 55 minutes. Taken together, the results of the present study suggest that hesitations orient listeners' attention to the ongoing utterance. In contrast to attentional blink studies, attention is not 'occupied' by hesitation, rather it is heightened so that listeners specifically attend to (and subsequently recognise) the words which follow. If the

subsequent word is acoustically deviant, the standard MMN and P300 responses to deviance are attenuated, because attention is already oriented to the disfluent utterance. This provides a straightforward account for the increased likelihood of recognition following hesitations, as well as for the facilitated reaction times for targets that have been found in earlier studies (Brennan & Schober, 2001; Fox Tree, 2001).

Previous accounts of disfluency processing have either focused on changes in attention (Fox Tree, 2001) or changes to linguistic mechanisms (Arnold et al., 2004) that occur when hesitations marked by filled pauses are encountered. However, these accounts are not mutually exclusive. Hesitations may induce a low-level response that heightens listeners' attention, and this may in turn affect linguistic processes which alter the linguistic availability of subsequent material. Clearly, such an account would require elaboration: For example, it is presently unclear whether listeners' "heightened attention" is speech-specific or represents a more general state of arousal. Such issues remain questions for future research. The importance of the present study is that it provides clear evidence that attention *is* affected by hesitation in an utterance, either concomitantly with, or as a precursor to, linguistic processes.



## References

- Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear and Hearing, 16*, 38-51.
- Altmann, G. T. M. & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*, 247-264.
- Arnold, J.E., Tanenhaus, M.K., Altmann, R.J., & Fagnano, M. (2004). The old and thee, uh, new: Disfluency and reference resolution. *Psychological Science, 15*, 578-582.
- Bailey, K. G. D. & Ferreira, F. (2003). Disfluencies affect the parsing of garden-path sentences. *Journal of Memory and Language, 49*, 183-200.
- Brennan, S.E. & Schober, M. F. (2001). How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language, 44*, 274-296.
- Brennan, S.E. & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language, 34*, 383-398.
- Corley, M., MacGregor, L. J., & Donaldson, D. I. (2007). It's the way that you say it: Disfluency in speech affects the comprehension process. *Cognition, 105*, 658-668.
- Ferreira, F., Lau, E.F., & Bailey, K.G.D. (2004). Disfluencies, parsing, and tree-adjointing grammars. *Cognitive Science, 28*, 721-749.
- Fox Tree, J. E. (1995). The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language, 34*, 709-738.

- Fox Tree, J. E. (2001). Listeners' uses of um and uh in speech comprehension. *Memory and Cognition*, 29, 320-326.
- Jasper, H. H. (1958). Report to the committee on methods of clinical examination in electroencephalography. Appendix: the ten-twenty system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, 10, 370-375.
- Lang, A. (2000). The information processing of mediated messages: A framework for communication research. *Journal of Communication*, 50, 46-70.
- Polich, J. (2004). Neuropsychology of P3a and P3b: A theoretical over view. In N.C. Moore & K. Arikan (Eds.), *Brainwaves and Mind: Recent developments*. (pp 15-29) Kjellberg Inc: Wheaton, IL.
- Pritchard, W. S. (1981). Psychophysiology of P300. *Psychological Bulletin*, 89, 506-540.
- Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, 18, 849-860.
- Rugg, M. D., Mark, R. E., Gilchrist, J., & Roberts, R. C. (1997). ERP Repetitions effects in indirect and direct tasks: effects of age and interim lag. *Psychophysiology*, 34, 572-586.
- Schröger, E. (1997). On the detection of auditory deviants: A pre-attentive activation model. *Psychophysiology*, 34, 245-257.
- Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43, 443-467.

Vogel, K. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1656–1674.

**Author Note**

We thank the members of the Psychological Imaging Laboratory at the University of Stirling for support throughout the experiment, and the University of Edinburgh Disfluency Group and three anonymous reviewers for useful feedback on earlier drafts of this paper.

This research was partially supported by the College of Humanities and Social Science at the University of Edinburgh (P.C.), a Staff Research Development Grant (M.C.), the Economic and Social Research Council (L.J.M.), and the Biotechnology and Biological Sciences Research Council (D.I.D.).

### Footnote

1) Traditional adjustments for individual error-rates, such as  $d'$ , are inappropriate, since the properties of 'old' stimuli are determined by their context of occurrence and hence there are no comparable categories of 'new' stimuli. Using stimulus identity as a random factor ensures that per-participant biases to respond 'old' or 'new' are controlled for across the experiment.

Twelve target words were inadvertently repeated in the experiment, resulting in 148 distinct targets. Removing data from the repeated targets did not affect the outcome of the ANOVA, but the fluency effect for standard stimuli became marginal [ $t(135) = 1.993, \eta_p^2 = .027, p = .055$ ].

### Figure Captions

*Figure 1:* Topographic maps (anterior up; electrodes shown as black dots) illustrating the mean distributions of the deviance effects (deviant minus standard ERPs) over 100-400ms (in 50ms time windows) for fluent (top) and disfluent (bottom) utterances.

*Figure 2:* Grand average ERPs for deviant (continuous lines) relative to standard (dotted lines) target words in fluent utterances (positive up). Waveforms show data from left, midline and right electrodes at frontal, central and parietal sites (from left to right and top to bottom: F3, Fz, F4, C3, Cz, C4, P3, Pz, P4).

*Figure 3:* Grand average ERPs for deviant (continuous lines) relative to standard (dotted lines) target words in disfluent utterances (positive up). Waveforms show data from left, midline and right electrodes at frontal, central and parietal sites (from left to right and top to bottom: F3, Fz, F4, C3, Cz, C4, P3, Pz, P4).

*Figure 4:* Recognition probabilities for utterance-final words that were originally presented as acoustically deviant or standard stimuli, in fluent or disfluent contexts (error bars represent one standard error of the mean).

Figure 1:

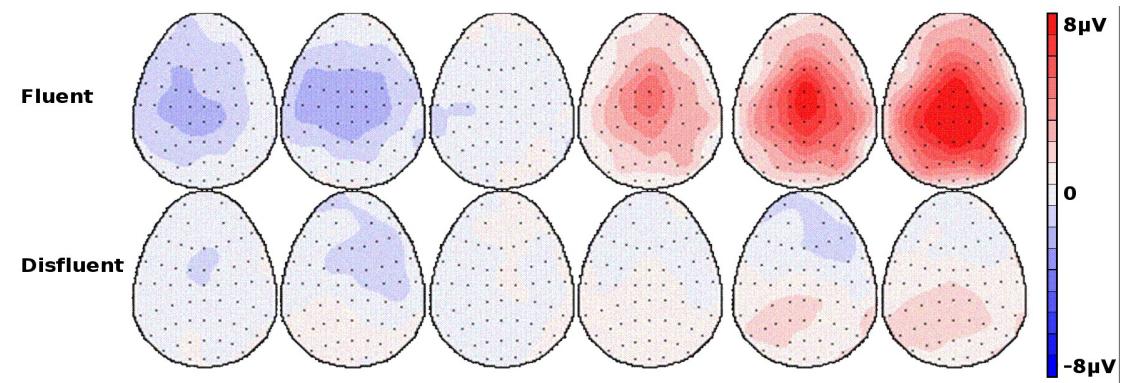


Figure 2:

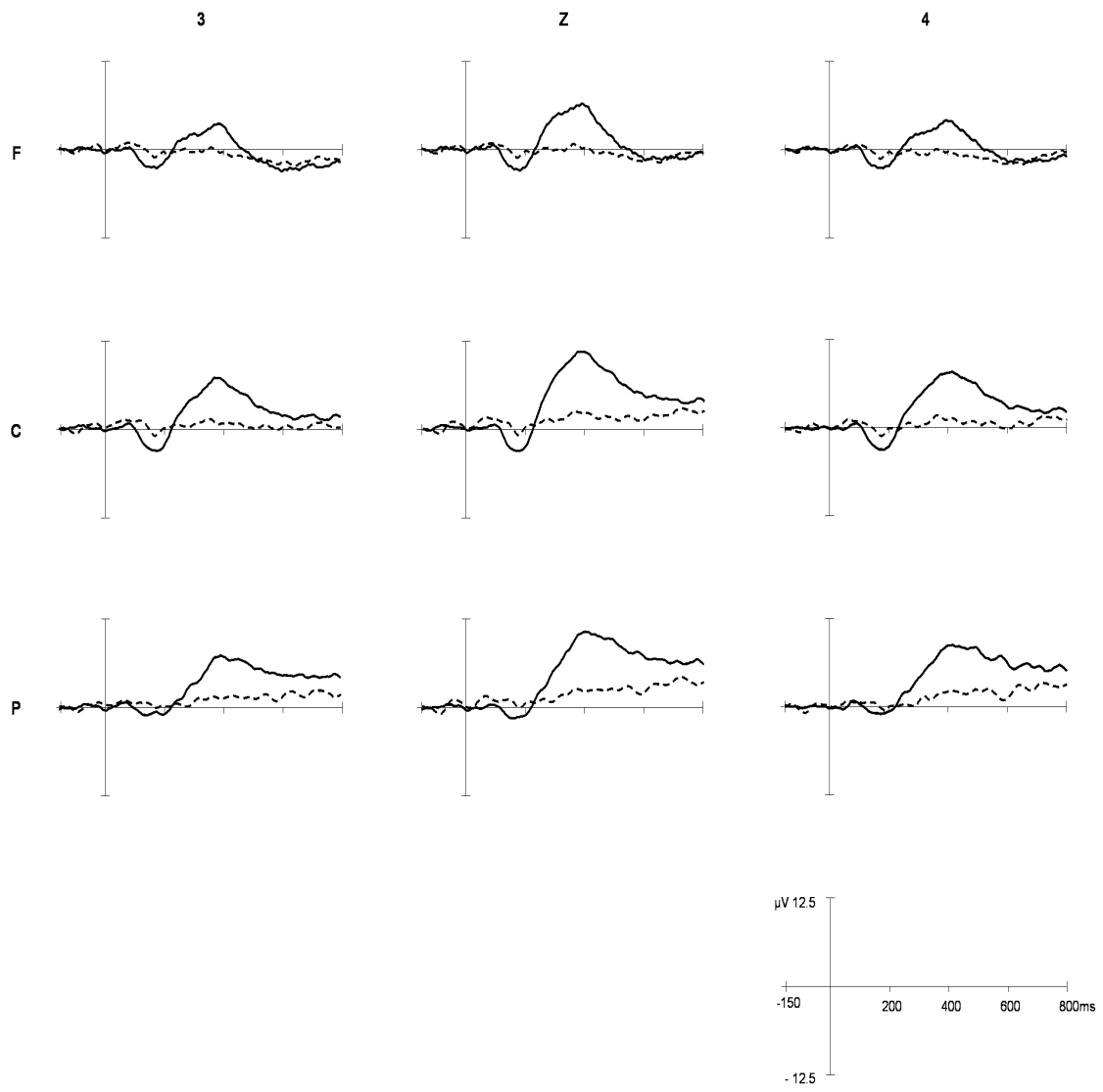




Figure 3:

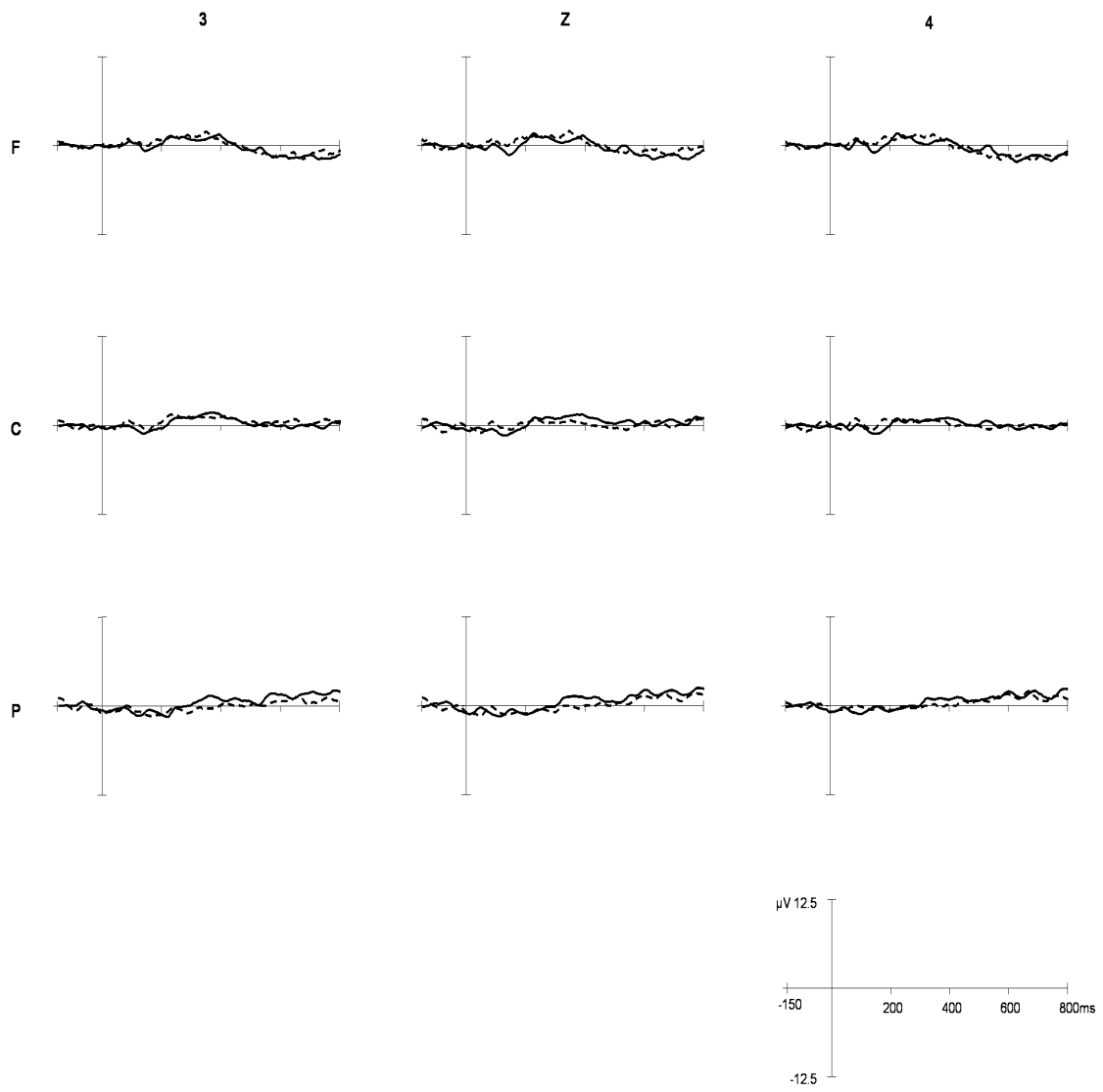


Figure 4:

