Cognition 108 (2008) 601–607

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

# Concurrent processing of words and their replacements during speech

Robert J. Hartsuiker [a,*], Ciara M. Catchpole [b], Nivja H. de Jong [c], Martin J. Pickering [b]

[a] Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium
[b] Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK
[c] Amsterdam Center for Language and Communication, University of Amsterdam, Spuistraat 134, 1012 VB Amsterdam, The Netherlands

ARTICLE INFO

*Article history:*
Received 1 February 2007
Revised 21 April 2008
Accepted 27 April 2008

*Keywords:*
Speech production
Speech monitoring
Replanning
Self-interruption
Self-correction
Picture naming

ABSTRACT

Two picture naming experiments, in which an initial picture was occasionally replaced with another (target) picture, were conducted to study the temporal coordination of abandoning one word and resuming with another word in speech production. In Experiment 1, participants abandoned saying the initial name, and resumed with the name of the target picture. This triggered both interrupted (e.g., *mush-* ...scooter) and completed (*mushroom* ...scooter) productions of the initial name. We found that the time from beginning naming the initial picture to ending it was longer when the target picture was visually degraded than when it was intact. In Experiment 2, participants abandoned saying the initial name, but without resuming. There was no visual degradation effect, and thus the effect did not seem to be driven by detection of the stopping cue. These findings demonstrate that planning a new word can begin before the initial word is abandoned, so that both words can be processed concurrently.



## 1. Introduction

People often need to undo one action and replace it with a different action. They do so when they realize they have just made or are making a mistake or when the environment changes in a way that renders the initial action unsatisfactory. In psycholinguistic research, the former case has been studied extensively (e.g., Levelt, 1983). It occurs when speakers realize that their utterance is erroneous or insufficiently informative to their addressees. But the latter case is also very common and occurs when the event that is being described changes in a relevant respect. For instance, a sportscaster might start saying that Horse *A* is leading (and this is true when he or she starts planning the utterance), but abandon the utterance when Horse *B* overtakes Horse *A*.

This article contrasts two accounts of the temporal coordination of abandoning a word and resuming with its replacement, in situations when the context changes. On one account, the resumption is planned after the abandoned word is stopped (Levelt, 1989). Alternatively, planning the resumption might begin when the abandoned word is still being uttered (Hartsuiker & Kolk, 2001), so that for a while the words are processed concurrently.

To detect the need to abandon utterances, speakers must monitor their own speech (e.g., Blackmer & Mitton, 1991; Hartsuiker & Kolk, 2001; Hartsuiker, Pickering, & De Jong, 2005; Levelt, 1983, 1989; Postma, 2000). They can monitor both overtly produced speech (external-channel monitoring) and the speech plan before articulation (internal-channel monitoring). The self-monitoring system inspects speech for linguistic well-formedness, but also for appropriateness in the context (e.g., Is the utterance specific enough for the interlocutor to discriminate among potential referents in the context? Is it still sensible given that some new event has just occurred?) and intelligibility to the interlocutor (Levelt, 1989). After detecting trouble, the monitor can give the signal to abandon ongoing speech and plan a resumption (see Hartsuiker & Kolk, 2001).

* Corresponding author. Address: Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium. Tel.: +32 9 2646436; fax: +32 9 2646496.
*E-mail address:* robert.hartsuiker@ugent.be (R.J. Hartsuiker).

0010-0277/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2008.04.005

When the monitor gives this signal, speech does not stop immediately; stopping any action takes time (Logan & Cowan, 1984). This time-to-stop is estimated at about 150–200 ms (Hartsuiker & Kolk, 2001; Levelt, 1989; Slevc & Ferreira, 2006). According to Levelt's theory of monitoring, the signal halts every component of the language production system at roughly the same time. A subsequent *editing phase* (often filled with *uh* or *um*) is used for the planning of the resumption. Thus, the production of the abandoned word and planning the resumption are sequential processes. In contrast, Hartsuiker and Kolk's computational model of self-monitoring assumes that the stop signal and the planning of the resumption both begin immediately on the detection of trouble. Thus the production of the abandoned word and the planning of the resumption occur concurrently, from when the trouble is detected to when speech is stopped. This article attempts to arbitrate between these *sequential* and *concurrent* accounts.

Evidence against a sequential account comes from studies that considered the time from the interruption of speech to its resumption (i.e., interval 2–3 in Example 1).

(1) $_1$Mush$_2$-, $_3$Chair

In a corpus of naturalistic speech (obtained from a radio talk show), Blackmer and Mitton (1991) found that the interruption-to-resumption interval was less than 100 ms in almost 50% of their data, and was sometimes 0 ms. Oomen and Postma (2001) reported similar results in a picture naming experiment that elicited interruptions and resumptions. These results are difficult to reconcile with the sequential account, because the resumption would be planned during the interruption-to-resumption interval. It is unlikely that speakers can plan a word in less than 100 ms (e.g., it takes at least 170 ms to produce a prepared syllable; see Hartsuiker & Kolk, 2001).

In contrast, Hartsuiker and Kolk's (2001) concurrent account can accommodate these findings, because the resumption has been wholly or partly prepared when speech is halted. In their simulations, model-generated distributions of interruption-to-resumption intervals corresponded well with data from Oomen and Postma (2001).

However, Hartsuiker and Kolk (2001) considered only speech errors and their repairs and not other cases where one needs to replace one word with another. It is possible that when an error occurs, planning processes are disordered, and hence error data are unrepresentative of error-free speech. For example, planning two words concurrently might even cause speech errors. Similarly, the observation of extremely short intervals between interruption and resumption in Blackmer and Mitton's (1991) study appeared to be restricted to errors; for other types of replacements (e.g., when a word was insufficiently appropriate) much longer times were observed (i.e., 200–300 ms longer on average). Finally, note that Blackmer and Mitton's and Oomen and Postma's (2001) data on the time-course of interruption and resumption come from studies in which the experimenters had no control over each interruption and resumption (e.g., about the moment the speaker realized the need to resume).

Therefore, this study tests a further prediction of the concurrent account. If continued production of the to-be-abandoned word and planning of its replacement take place concurrently, then the resumption planning could influence the production of the abandoned word. In particular, an increase in the difficulty of resumption planning may increase the time to interrupt or complete the abandoned word. This is because two concurrent tasks may compete for a limited set of cognitive resources so that an increase in difficulty of one task slows down performance on the other task (e.g., Ferreira & Pashler, 2002; Meyer & Kieras, 1997; Pashler, 1994). In contrast, the sequential model does not predict any influence from the resumption on the duration of the abandoned word.

To test these predictions, we conducted an experiment that elicited abandoned words and resumptions (Experiment 1) and a control experiment that elicited abandoned words, but not resumptions (Experiment 2). Experiment 1 used a replanning paradigm in which speakers begin naming one picture (the initial picture), but that picture is quickly replaced by another picture (the target picture) on a small proportion of trials (following Hartsuiker et al., 2005). Our goal was to make speakers abandon their response to the initial picture and name the target picture instead. The majority of trials (87.5%) were no-change trials, so that participants would not anticipate the change on the change trials. In the change trials, the SOA was −300 ms (i.e., initial picture occurred 300 ms before target picture). This value was chosen so that participants would often not be able to stop themselves from saying at least part of the initial picture's name.

Hartsuiker et al. (2005) manipulated the semantic and phonological relatedness between initial name and the target name. In all conditions, responses were observed in which participants interrupted within the initial name (Example 1), and responses in which the initial name was said completely (Example 2).

(2) $_1$Mushroom$_2$, $_3$Chair

The time to initiate the resumption (i.e., the time from target picture presentation to the point marked "3" in the examples) was affected by both semantic and phonological relatedness to the initial name, showing that the planning of resumptions is affected by their relationship to the abandoned word. Thus, Hartsuiker et al. demonstrated an influence of the abandoned word on resumption production.

The present study tested whether the resumption can influence the production of the abandoned word. We manipulated whether the target picture on change trials was easy or hard to describe. To do this, we presented pictures that either had intact contours or were visually degraded by deleting 50% of the contours. Naming degraded pictures takes longer than naming intact pictures (Meyer, Sleiderink, & Levelt, 1998). According to a sequential model, the resumption is planned only when speech stops (whether within a word or at its end). Thus, the time from initial word onset to interruption (i.e., 1–2 in Examples 1 and 2) should be unaffected by whether the target pictures were visually degraded (making resumption comparatively difficult) or not (making resumption comparatively easy).

In contrast, the concurrent model predicts that planning of the resumption begins as soon as the monitor detects the change in picture. Because planning the resumption requires processing resources, it should interfere with the concurrent production of the initial word. This prediction is independent of whether the initial word is interrupted (Example 1) or not (Example 2), because the model assumes that the resumption starts from change detection, which occurs before the actual interruption and also before the end of initial word completion (where the interruption process was presumably too slow to stop the word). Therefore, this model predicts that the onset-to-interruption interval (i.e., 1–2 in Examples 2 and 3) should be longer when the target pictures are visually degraded (because resumption should be comparatively difficult) than when they are not (because resumption should be comparatively easy). Experiment 2 presented the same conditions, but participants were now told not to resume with the name of the target picture. In this way, we controlled for possible effects of visual degradation on the perception detection of the change.

## 2. Experiment 1

### 2.1. Method

#### 2.1.1. Participants

Thirty-one University of Edinburgh students (aged 17–30) participated for payment. All were native speakers of English.

#### 2.1.2. Materials

We selected 32 pictures as initial pictures and 32 different ones as target pictures (see Appendix). There were two versions of each target picture: an *intact* version showing the normal contours and a *degraded* version in which 50% of the contours were erased (Fig. 1). The target pictures were taken from Meyer et al. (1998).[1] Pairs of initial and target pictures were unrelated in phonology or meaning. The names of the initial pictures had a mean frequency of 33.1 per million (range: 6.5–199.1) in the Celex database (Baayen, Piepenbrock, & Gulikers, 1995), and a mean length of 6.03 phonemes (range: 4–7). The names of target pictures had a mean frequency of 81.2 per million (range: 1.4–384.7) and a mean length of 4.53 phonemes (range: 3–7). Additionally, we selected 56 filler pictures similar to the target pictures.

The experimental materials were combined into a pseudo-random master list of 256 items, which included 32 critical pairs of initial and target pictures, the 32 initial pictures in isolation, the 32 target pictures in isolation, 32 fillers that were presented four times each, and 24 fillers that were presented twice each. The list was constrained so that experimental trials were separated by at least three filler trials. From this master list, two counterbalanced lists were derived so that in each list, half the target pictures occurred in the degraded condition and half in the intact condition, and that across lists, each picture occurred once in each condition. Each list was divided into two blocks, and two further lists were created by swapping the order of the blocks. Each participant was presented with one list.

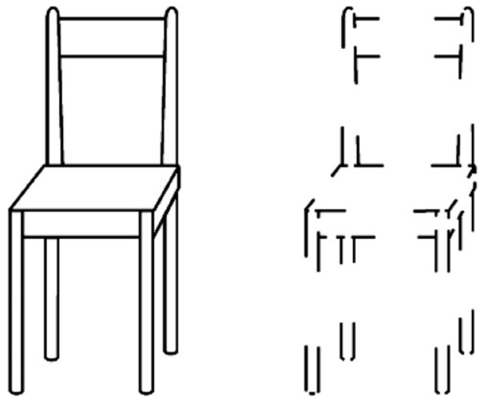

**Fig. 1.** Example of a target picture in the intact and degraded conditions.

#### 2.1.3. Procedure

The participants were first familiarized with the names of all 160 pictures. Each picture was presented on a computer screen with its name underneath. Participants were asked to name the picture aloud and progressed to the next picture by pressing a key.

In the experimental phase, participants were instructed to name all the pictures that appeared on the screen. The instructions emphasized the importance of speaking slowly and clearly. If two pictures appeared after each other, the participants were to try to stop naming the first picture and name the second picture. In all trials, a fixation cross appeared in the center of the screen for 2500 ms. In the no-change trials, the target picture subsequently appeared for 500 ms. In the change trials, the initial picture appeared for 300 ms, and was replaced by the target picture, which appeared for 500 ms. Inter-trial interval was 3300 ms after a change trial, and 3000 ms after a non-change trial.

#### 2.1.4. Apparatus

The experiment was implemented in E-prime (Schneider, Eschman, & Zuccolotto, 2002a, 2002b) and speech was recorded on DAT-tape via a high-quality directional microphone. Using the phonetic software package PRAAT (Boersma & Weenink, 1992), we manually determined the naming latency, moment of interruption or completion of the first name, and moment of resumption for the experimental trials. We also determined the naming latency for target pictures appearing in isolation.

#### 2.1.5. Scoring

Three-hundred and four responses (30.6%) were discarded because of errors in naming the initial or target picture or because of disfluencies (e.g., *um*, repetitions) before naming. Errors occurred more frequently in the degraded condition (correct: 327, error: 169) than in the intact condition (correct: 361, error: 135); [$\chi^2(1) = 5.17$, $p < .05$]. The remaining 688 responses were divided into three

[1] We thank these authors for providing us with the set of degraded and intact pictures.

categories: interrupted initial name (e.g., *Mush-*, Chair), completed initial name (e.g., *Mushroom*, Chair), and initial name skipped (e.g., chair). Only the first two categories were included in analyses of time intervals.

## 2.2. Results

### 2.2.1. Naming target pictures in isolation

To ensure that visual degradation of our stimuli slowed down naming latencies, we first considered the naming latencies when the target pictures were presented in isolation. As expected, degraded pictures were named more slowly than intact pictures (Table 1). The 52 ms difference was significant in paired-samples *t*-tests that treated participants, $t_1(28) = 3.75$, $p < .001$, and items, $t_2(31) = 2.70$, $p = .01$, as a random variable.

### 2.2.2. Change trial data

The number of responses of each type is listed in Table 2. There was no significant difference in response frequency by condition [$\chi^2(2) = .06$, $p = .97$], and this was also true when only interrupted and completed responses were considered [$\chi^2(1) = .004$, $p = .95$].

We excluded response times more than three standard deviations from the grand mean (onset of initial name: 2.3%, initial name onset to interruption/completion: 0.5%, interruption/completion to resumption: 2.5%; onset of target name 1.5%). The mean naming latencies for the remaining data are listed in Table 1. The analyses treated degradation (degraded vs. intact) and response type (interrupted vs. completed initial name) as within-participants and within-items variables. Because response type was not under experimental control, and there were highly divergent number of observations per cell, analyses that respectively average over items and participants are inappropriate. As in Hartsuiker et al. (2005), we therefore used linear mixed effect (multilevel) models (Pinheiro & Bates, 2000), so that every observation counted, and so that one can simultaneously generalize to participants and items.

Analyses on initial name onset latency revealed no main effects of degradation or response type (*F*s < 1) and no interaction, $F(1, 554) = 1.61$, $p = .21$. Additional analyses that treated log frequency and length in phonemes of both the initial and the target picture names as covariates showed the same pattern of significance. Inclusion of covariates did not improve model fit.

The time from initial name onset to the point of interruption/completion was longer for completed than interrupted responses, $F(1, 565) = 96.63$, $p < .0001$, and importantly, longer for degraded than intact pictures, $F(1, 565) = 8.04$, $p < .005$. There was no interaction ($F < 1$). When the covariates were included, there was a slight improvement in model fit (without covariates: $R^2 = 0.43$, with covariates: $R^2 = 0.45$), but importantly, the pattern of significance remained the same. Additionally, the length of the initial name significantly predicted this interval, $F(1, 561) = 12.98$, $p < .001$: the longer the initial name, the longer speakers took to interrupt or complete it.

The time from the point of interruption/completion to the target name onset was longer for interrupted than completed responses, $F(1, 553) = 80.31$, $p < .0001$. There was no degradation effect ($F < 1$) and no interaction ($F < 1$). This pattern was unaffected by the inclusion of covariates (which slightly improved overall model fit: without covariates: $R^2 = 0.26$; with covariates: $R^2 = 0.31$). This interval was significantly predicted by both initial name frequency, $F(1, 549) = 7.63$, $p < .01$, and target name frequency, $F(1, 549) = 24.99$, $p < .0001$: the more frequent the initial name or the target name, the shorter the interval.

For completeness sake, we also conducted analyses on target name onset latencies. These analyses revealed no main effect of degradation, $F(1, 559) = 1.47$, $p = .23$, no main effect of response type, $F < 1$, and no interaction, $F(1, 559) = 1.21$, $p = .28$. Including log frequency and length in phonemes of the initial and target picture as covariates did not change the pattern of results, but slightly improved model fit (without covariates: $R^2 = 0.22$; with covariates: $R^2 = 0.24$). Additionally, the time to target name onset was predicted by initial picture frequency, $F(1, 555) = 6.75$, $p < .01$, and target picture frequency, $F(1, 555) = 8.94$, $p < .003$.

## 2.3. Discussion

As predicted by the concurrent account, the duration of the abandoned word increased when the replacement picture was visually degraded and therefore harder to name. A manipulation check (i.e., naming degraded and intact pictures in isolation) confirmed that these degraded pictures were indeed harder to name. There was no significant effect of degradation on the time interval between interruption/completion and resumption, although descriptively

**Table 1**
Mean naming latency in ms for the initial and target pictures by response type and degradation (standard deviations in brackets) in Experiment 1

| Response interval | Degraded | Intact |
|---|---|---|
| Target in isolation | 704 (203) | 652 (197) |
| Onset of initial name | | |
| Initial name interrupted | 595 (118) | 589 (110) |
| Initial name completed | 600 (121) | 612 (155) |
| Initial name to interruption/completion | | |
| Initial name interrupted | 287 (114) | 273 (99) |
| Initial name completed | 413 (90) | 402 (90) |
| Interruption/completion to target name | | |
| Initial name interrupted | 225 (176) | 208 (160) |
| Initial name completed | 75 (116) | 71 (123) |
| Onset of target picture to target picture name | | |
| Initial name interrupted | 830 (220) | 802 (227) |
| Initial name completed | 807 (222) | 801 (238) |

**Table 2**
Frequency of each response type for the degraded and intact conditions in Experiment 1

| Response | Degraded | Intact |
|---|---|---|
| Initial name interrupted | 111 | 120 |
| Initial name completed | 176 | 195 |
| Initial name skipped | 40 | 46 |

*Note.* Responses containing errors were excluded.

this interval was longer in the degraded condition. It is thus possible that the additional time to name degraded pictures not only influenced the time-course prior to the point of interruption/completion (as only the concurrent account predicts), but also the timing subsequent to that point (which is consistent with both the concurrent and sequential accounts).

The time intervals were predicted by several covariates. The interval from the initial name onset to the point of interruption/completion was predicted by the length of the initial name. This is hardly surprising: in the case of completed productions, words with longer names will take longer to produce; in the case of interrupted production, words with longer names offer more opportunity to interrupt at a later point in time but still within the word. The interval from the point of interruption/completion to target name onset was predicted by the frequency of both initial name and resumption, and so was the interval from target picture appearance to target name onset. The effect of resumption frequency is unsurprising, as frequency effects in picture naming are well-attested (e.g., Jescheniak & Levelt, 1994). The effect of initial name frequency is again compatible with our concurrent account: as the abandoned word and the resumption are processed concurrently, increased ease in processing the abandoned word (because of frequency) can free up resources for processing the resumption. Therefore, the resumption can be produced somewhat earlier.

Finally, we observed that the interval from the point of interrupting or completing the initial name to the target name onset was nearly 150 ms shorter when the initial name was completed than when it was interrupted. It is possible that interrupting itself incurs a processing cost, for example because it may be time-consuming to first deactivate the speech motor system (in order to interrupt) and then to reactivate it (in order to resume), but this finding is also consistent with the concurrent account. On that account, if the speaker completes the initial name and concurrently plans the resumption, there should be sufficient time to have the resumption (almost) ready at the end of the initial word. Indeed, in this situation speakers resumed in under 75 ms. However, if the speaker interrupts the initial name, it is less likely that planning the target name will be nearly completed. In this situation, speakers took over 200 ms to resume. The sequential account has no obvious explanation of this large difference.

Thus, our main finding of a longer duration of abandoned words in the degraded condition supports the concurrent account, and all the additional findings discussed above are at least consistent with such an account. However, we need to consider an alternative explanation of our main finding. On this explanation, the increased duration of the initial word in the degraded condition in comparison to the intact condition is due to difficulties in visually processing the target pictures. Specifically, because object recognition is more difficult in the degraded condition, it may take the participants longer to detect the need to abandon the initial name. In order to evaluate this alternative explanation, we conducted a further experiment, in which participants were told to abandon naming the initial picture whenever a second picture appeared, but, importantly, were not told to name the second picture. If the alternative explanation is correct, this experiment should replicate the effect of degradation on the duration of the abandoned word, because it should take longer to detect the stopping cue in the degraded than the intact condition. But if the concurrent account is correct, there should be no effect of degradation, because there is no need to process the name of the second picture concurrently with the name of the first picture.

## 3. Experiment 2

### 3.1. Method

#### 3.1.1. Participants

Twenty-six further University of Edinburgh students participated for payment. All were native speakers of English.

#### 3.1.2. Materials and procedure

The same materials and procedure were used as in Experiment 1. The only difference was the task, which was now to stop speaking whenever two pictures appeared after each other.

#### 3.1.3. Scoring

We discarded 42 responses (5%) because of errors in naming the initial picture, because of disfluencies before naming, or because of responses in which the second picture was (partially) named. Error frequency in the degraded condition (correct: 391, error: 25) did not differ significantly from that in the intact condition (correct: 416, error: 17; $\chi^2(1) = 1.54$, $p = .21$). The remaining 807 responses were divided into three categories: no response, interrupted initial name, and completed initial name.

### 3.2. Results and discussion

The number of responses of each type is listed in Table 3. There was no significant difference in response frequency by condition when all three response categories were considered, $\chi^2(2) = 4.26$, $p = .12$, and no effect when only interrupted and completed responses were considered, $\chi^2(1) = 0.21$, $p = .65$.

We excluded response times more than three standard deviations from the grand mean (onset of initial name: 1.5%, initial name onset to interruption/completion: 1.0%). The mean naming latencies for the remaining data are listed in Table 4.

Analyses of initial name onset latency revealed an effect of response type, $F(1, 350) = 17.76$, with interrupted responses having longer latencies than completed responses.

**Table 3**
Frequency of each response type for the degraded and intact conditions in Experiment 2

| Response | Degraded | Intact |
|---|---|---|
| Initial name interrupted | 112 | 104 |
| Initial name completed | 93 | 77 |
| No response | 186 | 218 |

*Note*. Responses containing errors were excluded.

There was no effect of degradation ($F < 1$), and no interaction, $F(1, 350) = 1.69$, $p = .19$. A further analyses that treated log frequency and length in phonemes of both the initial and the target picture names as covariates showed the same pattern of significance (but slightly improved model fit: without covariates: $R^2 = 0.32$; with covariates: $R^2 = 0.33$). None of the covariates significantly predicted this interval (all $p$'s > .38).

The time from initial name onset to the point of interruption/completion was longer for completed than interrupted responses, $F(1, 352) = 206.36$, $p < .0001$. Importantly, there was no effect of degradation, $F(1, 352) = 1.03$, $p = .31$, and no interaction ($F < 1$). A further analysis that treated log frequency and length in phonemes of both the initial and the target picture names as covariates showed the same pattern of significance (and slightly improved model fit: without covariates: $R^2 = 0.47$, with covariates: $R^2 = 0.51$). None of the covariates significantly predicted this interval (all $p$'s > .14).

This experiment showed that the second picture's visual degradation had no effect on the point of interruption/completion when participants simply had to abandon the naming of the initial picture (rather than to abandon and resume with the name of the second picture). This result clearly argues against an account of the results in Experiment 1 according to which degradation affects detection of the stopping cue. In conjunction, the two experiments support the concurrent hypothesis, whereby the processing of an abandoned word and planning the resumption can proceed in parallel, so that difficulty in resumption can affect the duration of the abandoned word.

## 4. General discussion

Experiment 1 found that when speakers begin to produce one picture name, but then abandon this name and resume with another picture name, the duration of the abandoned word was longer when the replacement picture was visually degraded than when it was not. Experiment 2 showed that this effect was not due to increased difficulty in detecting the cue to abandon the initial name: when the task was only to abandon the initial name (without resumption) the degradation effect disappeared. Taken together, these results argue for an account in which the resumption is planned concurrently with the abandoned word, so that difficulty in planning the resumption can lengthen the abandoned word. In contrast, the results are inconsistent with the sequential account, in which resumption planning waits until the abandoned word is interrupted or completed.

The initial naming latencies were similar in both experiments (approximately 600 ms). But in contrast to Experiment 1, Experiment 2 revealed an effect of response type on these latencies, with interrupted responses having slower naming latencies than completed responses. Presumably, if speakers initiate a response very quickly, the stopping process is sometimes too slow to interrupt word-internally. Note that a similar effect of response type was also observed in Experiment 1 (but not Experiment 2) of Hartsuiker et al. (2005).

In both experiments, the interval from name onset to the point of interruption/completion was longer for completed than interrupted responses. It is interesting to note that the overall durations of both these intervals were considerably (roughly 100 ms) shorter in Experiment 2 than Experiment 1. We take this as further support of a concurrent account: in Experiment 2 there is no concurrent planning of a resumption, which frees cognitive resources for the processing of the initial word.

Our study provides the first experimental test of Hartsuiker and Kolk's (2001) concurrent account. The findings suggest that their account applies to cases where the trouble is not a speech error, but rather a word that was initially correct, but which ceases to be correct as a result of a change in the environment. This generalization is important, because speech error repair may differ from other types of resumptions, in that speech errors reflect disordered planning processes.

A remaining issue is the precise mechanism by which resumption difficulty affects the duration of the abandoned word. We have suggested that the production of the abandoned word and planning of the resumption compete for a limited pool of cognitive resources, so that a greater demand for these resources from one process (resumption planning) leaves the other process with fewer resources, and therefore processing takes longer. But it is also possible that speakers strategically prolong an abandoned word when it is more difficult to resume. This fits the claim made by Fox Tree and Clark (1997) and Clark and Fox Tree (2002) that speakers can strategically lengthen certain function words (e.g., saying *the* as *thee*) or disfluencies (lengthening *uh* or *um*) to indicate to addressees that they have upcoming production difficulties.

In conclusion, our experiments demonstrate that planning a resumption affects the duration of an abandoned word. This confirms an account on which the resumption and the abandoned word are processed concurrently (Hartsuiker & Kolk, 2001), but argues against an account in which these processes take place sequentially (Levelt, 1989).

**Table 4**
Mean naming latency in ms for the initial and target pictures by response type and degradation (standard deviations in brackets) in Experiment 2

| Response interval | Degraded | Intact |
|---|---|---|
| Onset of initial name | | |
| Initial name interrupted | 609 (93) | 598 (91) |
| Initial name completed | 561 (84) | 577 (90) |
| Initial name to interruption/completion | | |
| Initial name interrupted | 184 (70) | 180 (74) |
| Initial name completed | 292 (77) | 283 (75) |

## Acknowledgements

This research was sponsored by Grant R000223812 from the Economic and Social Research Council (United Kingdom) awarded to Rob Hartsuiker and Martin Pickering and a British Academy Research Readership awarded to Martin Pickering.

## Appendix.

Names of initial pictures:

window, finger, mountain, chicken, button, orange, glasses, apple, cricket, camel, lion, sandwich, basket, hammer, barrel, pencil, rabbit, suitcase, needle, monkey, ruler, candle, onion, ladder, lemon, arrow, mushroom, helmet, kettle, tiger, spider, cannon.

Names of target pictures:

door, foot, arm, wall, top, mouth, chair, bottle, star, tie, nose, boat, bone, hat, steps, trousers, sofa, whistle, mouse, spoon, comb, sausage, sack, axe, rake, vase, pear, saw, scissors, sledge, scooter, pliers.

## References

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The Celex lexical database (CD-ROM): Linguistic data consortium*. Philadelphia, PA: University of Pennsylvania.

Blackmer, E. R., & Mitton, J. L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition, 39*, 173–194.

Boersma, P., & Weenink, D. (1992). *Praat*. Available from http://www.fon.hum.uva.nl/praat.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and *um* in spontaneous speaking. *Cognition, 84*, 73–111.

Ferreira, V. S., & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 1187–1199.

Fox Tree, J. E., & Clark, H. H. (1997). Pronouncing "*the*" as "*thee*" to signal problems in speaking. *Cognition*, 151–167.

Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology, 42*, 113–157.

Hartsuiker, R. J., Pickering, M. J., & De Jong, N. H. (2005). Semantic and phonological context effects in speech error repair. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 921–932.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word-frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 824–843.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition, 14*, 41–104.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91*, 295–327.

Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition, 66*, B25–B33.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance. 1. Basic mechanisms. *Psychological Review, 104*, 3–65.

Oomen, C. C. E., & Postma, A. (2001). Effects of time pressure on mechanisms of speech production and self-monitoring. *Journal of Psycholinguistic Research, 30*, 163–184.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220–244.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed effects models in S and S-Plus*. New York: Springer.

Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition, 77*, 97–131.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002a). *E-Prime user's guide*.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002b). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools Inc.

Slevc, L. R., & Ferreira, V. S. (2006). Halting in single word production: A test of the perceptual loop theory of speech monitoring. *Journal of Memory and Language, 54*, 515–540.