Cognition 108 (2008) 837–842

Contents lists available at ScienceDirect

Cognition

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

Brief article

How speakers interrupt themselves in managing problems in speaking: Evidence from self-repairs

Mandana Seyfeddinipur [a,c,*], Sotaro Kita [b], Peter Indefrey [c,d]

[a] Department of Psychology, Stanford University, Jordan Hall, Building 420, Stanford, CA 94305-2130, USA
[b] School of Psychology, Birmingham University, Birmingham, UK
[c] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[d] F. C. Donders Center for Cognitive Neuroimaging, Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 8 May 2007
Revised 1 May 2008
Accepted 5 May 2008

Keywords:
Self-monitoring
Disfluency
Language production
Conversation

ABSTRACT

When speakers detect a problem in what they are saying, they must decide whether or not to interrupt themselves and repair the problem, and if so, when. Speakers will maximize accuracy if they interrupt themselves as soon as they detect a problem, but they will maximize fluency if they go on speaking until they are ready to produce the repair. Speakers must choose between these options. In a corpus analysis, we identified 448 speech repairs, classified them as major (as in a fresh start) or minor (as in a phoneme correction), and measured the interval between suspension and repair. The results showed that speakers interrupted themselves not at the moment they detected the problem but at the moment they were ready to produce the repair. Speakers preferred fluency over accuracy.

1. Introduction

Putting ideas into words proceeds through stages from conceptual planning to articulatory encoding (Levelt, 1989). Things can go wrong at any of these stages, and speakers monitor their speech to identify these problems. When they do detect a problem, they must decide whether or not to interrupt their speech to repair it, and if so, when.

Deciding how to handle problems in speaking is constrained by several demands. One demand is to be accurate, to provide correct information. Another is to be fluent, to produce utterances in a timely fashion (Clark, 1996, 2002). Speakers have to balance these competing demands. If they interrupt themselves the moment they detect a problem, they may have to pause to plan the repair. And by pausing, they risk losing the floor, appearing ineloquent or opting out (Clark & Wasow, 1998). But if speakers do not interrupt themselves right away, they may say something inaccurate, potentially leading to a misunderstanding. Different accounts have been given of how speakers balance these competing demands.

Speakers who prefer accuracy should suspend speaking as soon as they detect a problem in their production. This assumption underlies the Main Interruption Rule Hypothesis (MIR): speakers interrupt their entire speech production upon detecting trouble (Levelt, 1983, 1989; Nooteboom, 1980). Speakers interrupt instantly for outright errors (*left* instead of *right*), suspending within a word. But they delay the interruption to complete the current word under articulation when the word is merely inappropriate (e.g., *the blank uh white circle*), or when it is correct (*left side uh right side*; *side* is correct, but *left* is not). Speakers plan and process the following repair only after the suspension, during the so-called cut-off-to-repair interval (see Fig. 1). We will refer to the planning and processing of a repair as *replanning*.

According to MIR, speakers start replanning only after suspension. But if replanning takes time, speakers can never resume with a cut-off-to-repair interval of zero ms. And yet there are cases of repairs with zero ms cut-off-to-repair intervals (Blackmer & Mitton, 1991; Oomen & Postma, 2002). In these cases, speakers must have replanned part or the entire repair before they suspended speech.

* Corresponding author. Tel.: +1 951 905 4058.
E-mail address: mandsey@stanford.edu (M. Seyfeddinipur).

0010-0277/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2008.05.004

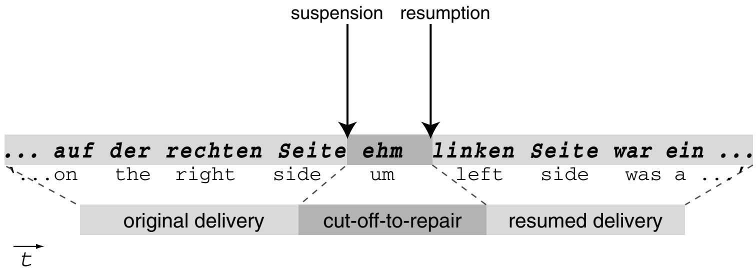

Fig. 1. Schematic representation of the structure of a disfluent utterance (based on Clark, 1996 and Blackmer & Mitton, 1991). In the given example the speaker said the erroneous word, *rechten* ('right'). He suspends his original delivery after *Seite* ('side'). The time interval following suspension is called cut-off-to-repair interval. Sometimes speakers pause or utter *uh/um*. Sometimes there is no such interval but speakers resume their delivery immediately. The moment the speaker begins to resume with the repair is called resumption.

The MIR can account for at least some zero ms cut-off-to-repair intervals if it is assumed that the interruption process itself takes time: there is an interval between the moment one *initiates* the interruption process and the moment one *suspends* speaking (Hartsuiker & Kolk, 2001). The idea is that when speakers detect a problem, they simultaneously initiate both the interruption process and replanning. If speakers complete replanning in the same time it takes to interrupt speech, they can start uttering the repair with a cut-off-to-repair interval of zero ms.

The process of interruption, however, takes only about 150–200 ms (Hartsuiker & Kolk, 2001; Levelt, 1989), so MIR can account only for repairs of small units (e.g., phonemes). Speakers need far more time to plan a major repair like a fresh start (e.g., *this house had um the entrance was big*), because major repairs require the generation of entirely new conceptual and syntactic representations. But if the production of a single word takes at least 600 ms from conceptualization to articulation (Indefrey & Levelt, 2004), the 150–200 ms window for the interruption process seems too short for a speaker to generate a fresh start in that time.

An alternative hypothesis is that speakers prefer fluency over accuracy. They interrupt not at the moment they detect a problem, but when they have a solution for the problem (Blackmer & Mitton, 1991). This way speakers minimize the cut-off-to-repair interval and resume speaking in a timely fashion (Catchpole, Hartsuiker, & Pickering, 2003). We term this the Delayed Interruption for Planning Hypothesis (DIP). According to DIP, when speakers detect trouble, they do not interrupt but start replanning while continuing to speak according to their original plan. Only once they have completed replanning do they interrupt their original delivery and utter the repair. Such interruptions can lead to suspensions either within or after a word (within- or after-word suspensions), depending on the timing of the interruption and ongoing speech. If speakers manage to align the timing of interruption and replanning optimally, they can utter the repair with a zero ms cut-off-to-repair interval.

The assumption behind DIP is that speakers can replan as they continue to speak based on the materials from the original utterance still in the formulator or articulatory buffer. If they cannot complete replanning before they run out of prepared material, they have to suspend afterword, assuming the smallest unit of buffering is the phonological word (Levelt, 1989). Then they have to complete the remainder of replanning during the cut-off-to-repair interval.

The goal of the current study is to investigate whether speakers prefer fluency (DIP) or accuracy (MIR). The hypotheses differ in how much replanning time they allow before suspension.[1] We can therefore test them using a simple logic: the more speakers can replan before suspension, the less replanning they have to do after suspension. The time spent replanning after suspension should be reflected in the cut-off-to-repair interval. We can infer how much time speakers have for replanning before suspension under each hypothesis from the way speech is suspended: within-word or after-word.

According to MIR, speakers have only short time for replanning before suspension. This time is fixed and cannot exceed certain upper bounds. When speakers detect a problem in the current word, they simultaneously initiate both interruption and replanning. The interruption process requires 150–200 ms to suspend overt speech, so speakers can replan during that interval (Hartsuiker & Kolk, 2001). If speakers detect the problem in the word early enough, that will lead to a within-word suspension. But if they happen to detect the problem toward the end of the word, it can lead to an after-word suspension instead. In this case, speakers also have 150–200 ms available for replanning before suspension.

Speakers may have slightly more time for replanning before suspension when they delay interruption to complete the word they are currently saying. If we assume that the average word is about 400 ms long (Levelt, 1989) and that speakers detect the problem at word onset, they may have up to 250 ms (=400–150) more for replanning before suspension. So, when speakers suspend their speech *after* a word, they have on average slightly *more* time for replanning before suspension, and so they should need *less* time for replanning after suspension, during the cut-off-to-repair interval. MIR's first prediction, therefore is, that the

[1] We cannot observe the replanning time before suspension directly, because it depends on the moment of error *detection* rather than moment the error surfaces in speech. Thus replanning time before suspension is not reliably reflected in the interval from error onset in speech to cut-off (error-to-cut-off interval).

Table 1
Overview of the predictions for each hypothesis

Dependent variable	Prediction by MIR	Prediction by DIP
Cut-off-to-repair interval	after-word ⩽ within-word	after-word > within-word
Cut-off-to-repair interval	minor repair < major repair	within-word: minor repair = major repair after-word: minor repair < major repair
Repairs that can follow a within-word suspension with a zero ms cut-off-to-repair interval	minor repairs	minor repairs & major repairs

Note that "after-word" and "within-word" in the prediction columns refers to repairs with an after-word suspension and repairs with a within-word suspension, respectively.

cut-off-to-repair interval should be shorter for after-word suspensions than for within-word suspensions.

Second, according to MIR, speakers have only a short fixed interval available for replanning before suspension. This time may be long enough for minor repairs such as phonetic changes, but it is too short to plan major repairs such as fresh starts. For these, speakers have to complete replanning during the cut-off-to-repair interval. If we assume that it takes longer to replan major repairs than minor repairs, then MIR predicts that the cut-off-to-repair interval should be longer for major repairs than for minor repairs, regardless of how speakers suspend their speech.[2]

Third, according to MIR, when speakers suspend speech within-word, they must have replanned for only a short time before suspension. In this short time, they can replan a minor repair, but not a major repair. So for within-word suspensions, MIR predicts that the cut-off-to-repair interval can be zero for minor repairs, but not for major repairs.

DIP makes very different predictions. According to DIP, the time speakers have for replanning before suspension depends on the amount of material buffered by formulation and articulation processes. So whenever speakers suspend speech within-word, that indicates that they have finished replanning and interrupted before running out of buffered material. Depending on how successfully they align suspension and resumption, they can resume either right away or after a short interval. In contrast, whenever speakers suspend speech after-word, that indicates that they did *not* complete replanning before suspension, but had to cease speaking because they ran out of buffered material. In this case, they have to complete replanning during the cut-off-to-repair interval. So DIP predicts that the cut-off-to repair interval should be longer when speakers suspend after-word than when they suspend within-word.

Second, DIP predicts that the cut-off-to-repair interval should be longer for major repairs than minor repairs, but only when speakers suspend speaking after-word. When they suspend speaking after-word, they must have run out of prepared material and need to complete replanning after suspension. This period should therefore be longer for major than for minor repairs. But when speakers suspend speaking within-word, they may have just finished replanning, and so the cut-off-to repair interval should be comparable for major and minor repairs.

Third, DIP predicts that when speakers suspend speaking within-word, either a major or a minor repair can follow a zero cut-off-to-repair interval.

For an overview of the predictions see Table 1 below.

2. Method

2.1. Data

We compiled a corpus of German speech disfluencies by videorecording participants as they described to an interlocutor the layouts of their homes or apartments (Linde & Labov, 1975; Ulmer-Ehrich, 1982). Altogether, recordings lasted 96.3 min.

2.2. Participants

Participants were 6 male and 6 female native German speakers, all students at the Freie Universität Berlin or the Universität Mainz.

2.3. Coding

We coded all overt repairs. We defined an overt repair as a disfluency containing some indication of a speech suspension (e.g., a glottal stop, laryngalization, fillers, or silent pause greater than 200 ms) and a resumption in which there was a modification of the original delivery.

2.4. Cut-off-to-repair interval

To measure cut-off-to-repair intervals, we identified the moments of suspension and resumption for each repair with the annotation tool MediaTagger (Brugman & Kita, 1995). Because we based our timing on video files, we were limited to a 40 ms increment size (a video frame). Whenever the interval between suspension and resumption was shorter than 40 ms, we assigned a cut-off-to-repair interval of 40 ms (one frame).

2.5. Suspension type

We coded each suspension as within-word or after-word following Levelt (1983).

[2] This prediction also assumes that speakers are equally likely to delay interruption till after-word for minor and major repairs. If they were more likely to delay for major repairs than for minor repairs, then the predicted effect would be weakened (or in the extreme case, even reversed) because of the gained extra replanning time.

Table 2
Overview of resumption type, example, and repair type

Resumption type	Example	Repair type
Fresh start	wenn man links in ehm vorm Haus war eine Garage ('when one left into um in front of the house was a garage')	Major
Addition	ging nochmal son langer Flur son ganz schmaler langer Flur ('went again sucha long hallway such an entirely narrow long hallway')	Minor
Deletion	auf der ganz linken auf der linken Seite ('on the completely left on the left side')	Minor
Substitution	auf der rechten eh linken Seite ('on the right uh left side')	Minor
Mixed	und hatte en grossen Kamin eh en Eckkamin ('and had a big fireplace uh a corner fireplace')	Minor

2.6. Repair type

We classified each resumption as adding, deleting, substituting an element or a mix thereof; or abandoning of the original delivery with a fresh start (Clark, 1996; see Seyfeddinipur, 2006). We classified each resumption as a minor repair if it repeated old material and as a major repair if it did not (see Table 2).

2.7. Reliability

Two trained raters, both blind to the hypotheses, coded the repairs. The second rater independently transcribed and coded the suspension and repair type of 15% of the coded disfluencies randomly selected from the total corpus ($N = 1202$). The raters agreed on 89% of the suspension types and on 74% of the repair types. This percentage is comparable to the 76% agreement reported in Blackmer and Mitton (1991) and the 73% agreement reported in Levelt (1983). Disagreements were resolved by discussion.

2.8. Results

Participants produced a total of 597 overt repairs. We excluded 87 of them because the repair was unclassifiable.[3] We also excluded one participant who did not provide any major repairs with withinword suspensions and another whose overall mean cut-off-to-repair interval exceeded the group mean by more than two standard deviations. That left 448 overt repairs for analysis (see Table 3).

Table 3
Number of major and minor repairs following within-word and after-word suspensions

Repair type	Suspension type	
	Within-word	After-word
Minor repair	100	148
Major repair	43	157

We carried out a repeated measures 2×2 ANOVA on cut-off-to-repair intervals (dependent variable) with suspension type (within-word, after-word) and repair type (major, minor) as factors (see Appendix for mixed effect regression analysis). The analysis provides three lines of support for DIP over MIR.

First, according to DIP, the cut-off-to-repair interval should be longer following after-word than within-word suspensions, but according to MIR, it should be the opposite. The data support DIP: cut-off-to-repair intervals were significantly longer following after-word suspensions (430 ms) than following within-word suspensions (164 ms), $F(1,9) = 14.86$, $MSE = 29.71$, $p < .01$.

Second, following MIR, the cut-off-to-repair intervals should be longer for major than for minor repairs independent of whether the suspension was within-word or after-word. According to DIP, in contrast, the cut-off-to-repair interval should be longer for major than for minor repairs, but *only* for after-word suspensions. Cut-off-to-repair intervals were significantly longer for major repairs (378 ms) than for minor repairs (217 ms), $F(1,9) = 8.35$, $MSE = 19.34$, $p < .05$. Importantly, in support of DIP, the difference between major and minor repairs was larger for after-word than for within-word suspensions, $F(1,9) = 5.251$, $MSE = 23.35$, $p < .05$ (see Fig. 2).

Indeed, the cut-off-to-repair interval following after-word suspensions was 300 ms longer for major repairs (580 ms, $SD = 344$) than for minor repairs (280 ms, $SD = 106$), (planned comparisons t-test with Bonferroni adjustment, $t(9) = 2.705$, $p < .05$). In contrast, the cut-off-to-repair interval following within-word suspensions did not differ for major (175 ms, $SD = 113$) and for minor repairs (154 ms, $SD = 69$), ($t(9) = .578$, n.s.).

Third, as predicted by DIP, within-word suspensions were followed within 0–40 ms cut-off-to-repair intervals by major and minor repairs. These "zero" cut-off-to-repair intervals accounted for 19% of the minor repairs and 9.5% of the major repairs (corresponding to 74.3% and 25.7% of within-word suspensions with 0–40 ms cut-off-to-repair intervals).

3. Discussion

Do speakers strive for accuracy, interrupting themselves the moment they detect a problem in what they are saying (MIR), or do they strive to maintain fluency, interrupting themselves only once they have planned the repair (DIP)? We report three main findings that support DIP over MIR: in handling problems in speaking, speakers prefer fluency over accuracy.

First, when speakers suspended their speech within-word, the cut-off-to-repair interval was shorter than when they suspended after-word. Speakers apparently needed less time for replanning after suspension following within-word than after-word suspensions. This finding fits DIP but not MIR. It suggests that when speakers suspended within-word they did not interrupt immediately but replanned for a longer time than when they suspended speech after-word.

[3] Unclassifiable repairs were mostly cases in which a word was interrupted so early (in the middle of or after the first phoneme or morpheme) that we could not reconstruct what the speaker was about to say, hence it was not possible to classify the repair (see Seyfeddinipur, 2006).

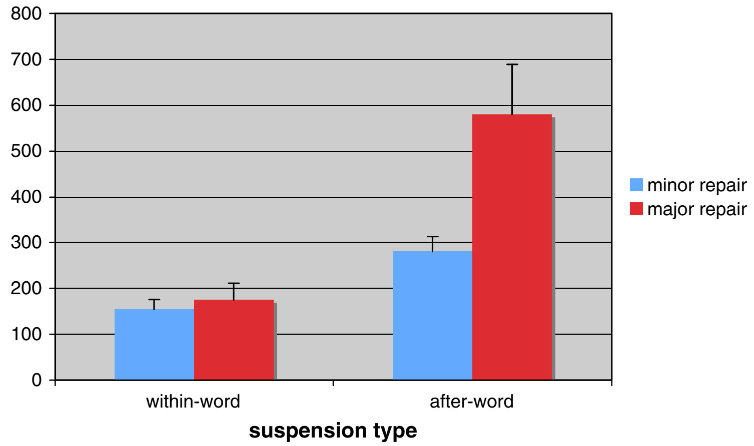

Fig. 2. Duration of cut-off-to-repair intervals (ms) for within-word vs. after-word suspensions and for major vs. minor repairs.

Second, when speakers suspended within-word, the cut-off-to-repair interval did not differ for major and minor repairs. This, too, is in line with DIP. The cut-off-to repair interval was longer for major than for minor repairs only when participants suspended speech after-word. So by the time speakers suspended within-word, they had already completed most of their replanning, and that held just as often for major as for minor repairs. In contrast, when speakers suspended after-word, they still had most of their replanning to do during the cut-off to repair interval.

Third, when speakers suspended speech within-word, they produced both major and minor repairs without any cut-off-to-repair interval. As predicted by DIP, speakers delayed interrupting speech until they completed their replanning for major and minor repairs. This way they could resume with the repair right after suspension. Immediate major repairs are a problem for MIR (Hartsuiker & Kolk, 2001), because it allows only a very short time interval for replanning before suspension, an interval insufficient to plan a new utterance from scratch.

One potential alternative account might be that fresh starts (major repairs) are in fact not planned after the detection of a problem. Instead, they are cases in which speakers formulate two messages in parallel and then suddenly abandon the utterance under articulation in favor of the other, already encoded message. Under these circumstances, no or only little replanning is necessary and speakers could resume with a fresh start without any cut-off-to-repair interval. Similarly, parallel encoding might explain why cut-off-to-repair intervals were comparable between major and minor repairs following within-word suspensions.

This alternative account, however, has to be rejected for two reasons. First, it cannot explain why speakers' cut-off-to-repair intervals were longer for major than for minor repairs when speakers suspended after-word. If there was little or no replanning needed in fresh starts, cut-off-to-repair intervals should be the same for within- and after-word suspensions. Second, under this account, both encoded messages are correct, i.e. neither is erroneous. One utterance is suspended in favor of the other, presumably because the former is considered to be less appropriate than the latter. If the utterance is merely inappropriate, according to MIR, speakers should not suspend within but after-word.

According to our findings, major repairs require more replanning time than minor repairs, but the additional time is not reflected in the cut-off-to-repair interval following within-word suspensions. So speakers do not interrupt themselves the moment they detect a problem in what they are saying, but delay interruption to prepare the repair while continuing to speak.

However, speakers appear to risk confusing the listener because they make problematic information available for longer than necessary. This, it seems, might even cause listeners to correct the speaker or ask for clarification. But as long as speakers continue speaking, listeners are not likely to interrupt. And whenever speakers cannot repair fast enough and run out of prepared material, they can always use editing terms like 'no' or 'I mean' (Clark, 1994; Levelt, 1983). These terms provide listeners with accounts of the speakers' problem, minimizing the risk of misunderstanding. Also, speakers monitor listeners' facial expressions and can add further explanations on the fly when necessary (Clark & Krych, 2004). When listeners show their understanding in their following turn, speakers can decide if they need to reformulate. With such strategies, speakers can offset the cost of not interrupting a potentially problematic expression as soon as they detect it.

Although speakers appear to favor fluency over accuracy, there are probably circumstances in which they strive to maximize accuracy. Speakers may interrupt themselves right away when they detect an expression with socially drastic consequences. By contrast, they may not interrupt themselves on detecting minor phonological errors such as a 'cuf of coffee'. Speakers may base their decisions on a moment-by-moment evaluation of the social impact of the problem and of its impact on the flow of conversation.

In sum, although speakers may sometimes interrupt speech the moment they detect a problem, they do not do so by default. Instead, they prefer to interrupt only once they have planned the repair, so that they can keep the pause after suspension to a minimum. Minimizing gaps and silences caused by replanning is just one of the strategies speakers use to handle problems in speaking in conversation.

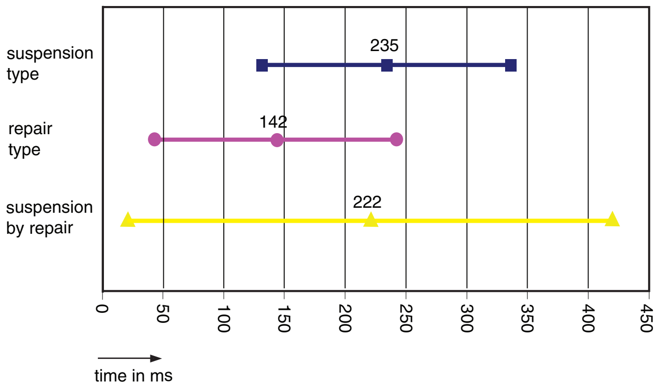

Fig. 3. MCMC point estimates (middle points) and HPD95 confidence intervals (outer points from the mixed-effect analysis.

Acknowledgements

The authors thank Dale Barr and Herb Clark for extensive discussion and comments on earlier drafts. They also thank Rob Hartsuiker and three anonymous reviewers for their suggestions on earlier versions of this paper.

Appendix A

Given the variable number of observations across conditions and across participants, we also analyzed the data using linear mixed-effect models (Baayen, Davidson, & Bates, in press; Hoffman & Rovine, 2007; Raudenbush & Bryk, 2002). In this analysis, the variables were coded using an effect coding scheme, such that significance tests for the parameter estimates would directly correspond to the significance tests for main effects and interactions in an ANOVA model (Cohen, Cohen, West, & Aiken, 2003). To test for simple effects, the model was re-fit with dummy coded variables. In mixed-effect models, degrees of freedom for parameter estimates can only be approximated (Raudenbush & Bryk, 2002). Rather than using the approximate degrees of freedom, p-values were calculated using the more conservative Markov-Chain Monte Carlo (MCMC) sampling procedure (Baayen, 2008; Baayen et al., in press). We will report the p-values as *pMCMC* to indicate that the values obtained were based on this procedure.

The analyses yielded the same pattern of results as the ANOVA: main effect of repair type: ($t = 2.76$, $pMCMC < .01$); main effect of suspension type ($t = 4.58$, $pMCMC < .01$); and a significant interaction between suspension and repair type ($t = 2.17$, $pMCMC < .05$). The effect of repair type depended on suspension type: longer cut-off-to-repair intervals in after-word suspension for major repairs (543 ms, SD = 666) than for minor repairs (290 ms, SD = 409), $t = 4.64$, $pMCMC < .01$. There was no difference in the cut-off-to-repair intervals following within-word suspensions for major repairs (197 ms, SD = 223) or for minor repairs (166 ms, SD = 211), $t = .36$, $pMCMC = .72$, n.s. (see Fig. 3 for MCMC point estimates and HPD95 confidence intervals (95 % confidence interval based on highest-posterior-density intervals from the MCMC simulations).

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics*. Cambridge: Cambridge University Press.
Baayen, R. H., Davidson, D. J., & Bates, D. M. (in press). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*.
Blackmer, E. R., & Mitton, J. L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition, 39*, 173–194.
Brugman, H., & Kita, S. (1995). Impact of digital video technology on transcription: A case of spontaneous gesture transcription. *KODIKAS/CODE: Ars Semiotica, An International Journal of Semiotics, 18*, 95–112.
Catchpole, C., Hartsuiker, R. J., & Pickering, M. (2003). Self-corrections in speech: Evidence against Levelt's Main Interruption Rule. *In Poster presented at the ninth conference on architectures and mechanisms for language processing*, Glasgow, Scotland.
Clark, H. H. (1994). Managing problems in speaking. *Speech Communication, 15*, 243–250.
Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
Clark, H. H. (2002). Speaking in time. *Speech Communication, 36*, 5–13.
Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology, 37*, 201–242.
Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language, 50*(1), 62–81.
Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, N.J.: Erlbaum.
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.
Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: Foundations and illustrative examples. *Behavior Research Methods, 39*(1), 101–117.
Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*, 101–144.
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.
Linde, C., & Labov, W. (1975). A site for the study of language and thought. *Language, 51*(4), 924–939.
Nooteboom, S. G. (1980). Speaking and unspeaking: Detection and correction of phonological and lexical errors in spontaneous speech. In V. A. Fromkin (Ed.), *Errors in linguistic performance* (pp. 87–95). New York: Academic Press.
Oomen, C. C. E., & Postma, A. (2002). Limitations in processing resources and speech monitoring. *Language and Cognitive Processes, 17*(2), 163–184.
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.
Seyfeddinipur, M. (2006). *Disfluency: Interrupting speech and gesture. MPI Series in Psycholinguistics, 39*. Max Planck Institute of Psycholinguistics, Nijmegen.
Ulmer-Ehrich, V. (1982). The structure of living space descriptions. In R. J. Jarvella & W. Klein (Eds.). *Speech, place, and action*. New York: Wiley.