Brief article

Inner speech slips exhibit lexical bias, but not the phonemic similarity effect ☆

Gary M. Oppenheim *, Gary S. Dell

Beckman Institute, Department of Psychology, University of Illinois at Urbana-Champaign, 405 North Matthews Avenue, Urbana, IL 61801, USA

Abstract

Inner speech, that little voice that people often hear inside their heads while thinking, is a form of mental imagery. The properties of inner speech errors can be used to investigate the nature of inner speech, just as overt slips are informative about overt speech production. Overt slips tend to create words (*lexical bias*) and involve similar exchanging phonemes (*phonemic similarity effect*). We examined these effects in inner and overt speech via a tongue-twister recitation task. While lexical bias was present in both inner and overt speech errors, the phonemic similarity effect was evident only for overt errors, producing a significant overtness by similarity interaction. We propose that inner speech is impoverished at lower (featural) levels, but robust at higher (phonemic) levels.
© 2007 Elsevier B.V. All rights reserved.

1. Introduction

Most people hear a little voice inside their head when thinking, reading, writing, and remembering. This voice is inner or internal speech, mental imagery that is generated by the speech production system (Sokolov, 1972). Inner speech is the basis of rehearsal in short-term memory (e.g. Baddeley, Thomson, & Buchanan, 1975) and some phonological influences in reading and writing (e.g. Hotopf, 1980). It may even play a role in auditory hallucinations in schizophrenia (e.g. Ford & Mathalon, 2004).

We produce inner speech the same way that we speak, except that articulation is not present (Levelt, Roelofs, & Meyer, 1999). We hear the speech in our mind, though, through an *inner loop* that transmits the speech plan at the phonetic (e.g. Levelt, 1983, 1989) and/or phonological (e.g. Wheeldon & Levelt, 1995) level to the speech comprehension system. The existence of this inner loop gives a good account of our ability to monitor our planned speech for errors (Hartsuiker & Kolk, 2001; Postma, 2000; Roelofs, 2004; Slevc & Ferreira, 2006).

Inner speech is characterized by slips of the "tongue" that can be internally "heard", despite the absence of sound or significant movements of the articulators (Hockett, 1967). Inner slips that are reported during the internal recitation of tongue twisters are similar to overt errors made when the same material is spoken aloud (Dell & Repka, 1992; Postma & Noordanus, 1996). This fact alone makes credible the view that overt errors are not really slips of the tongue. Rather, they are slips of speech planning, a process that occurs both during inner and overt speech.

The properties of inner slips can be used to investigate inner speech, just as overt slips are informative about overt production. Here, we compare inner and overt errors to investigate the processing levels in production and how these differ between inner and overt speech. The phenomena that we are concerned with are the *lexical bias* and *phonemic similarity* effects. Lexical bias is the tendency for phonological errors to create words (e.g. REEF LEECH → LEAF REACH) over nonwords (e.g. WREATH LEAGUE → LEATH REEG) (Baars, Motley, & MacKay, 1975; Costa, Roelstraete, & Hartsuiker, 2006; Dell, 1986, 1990; Humphreys, 2002; Hartsuiker, Corley, & Martensen, 2005, 2006; Nooteboom, 2005a). This effect has been attributed to either the interactive flow of activation between lexical and phonological levels (Dell, 1986) or a prearticulatory editorial process that suppresses nonword utterances (Baars et al., 1975; Levelt et al., 1999). The phonemic similarity effect is a tendency for similar phonemes to interact in slips. For example, the likelihood of REEF LEECH slipping to LEAF REACH is greater than that of REEF BEECH slipping to BEEF REACH, because /r/ is more similar to /l/ than it is to /b/. This effect has often been demonstrated in natural error analyses (MacKay, 1970; Shattuck-Hufnagel & Klatt, 1979) and in at least one experimental manipulation (Nooteboom, 2005b). Explanations for the effect posit a role for sub-phonemic features in the relevant representations (e.g. Dell, 1986).

We use the lexical bias and phonemic similarity effects to probe inner speech. Will inner slips exhibit these effects and, if so, how will they compare in magnitude to the effects in overt speech? There are three possibilities:

Unimpoverished hypothesis. Inner speech is planned exactly as normal speech, except that the articulators are not moved (e.g. Dell, 1978; Levelt, 1989). If so, the lexical bias and phonemic similarity effects will be equally strong in overt and inner speech.

Surface-impoverished hypothesis. Inner speech is impoverished at a surface level, having weakened or absent lower-level representations (e.g. featural level). For example, Dell and Repka (1992) claim inner speech inconsistently activates phonological nodes, but is lexically intact. Wheeldon and Levelt's (1995) conclusion that the inner loop perceives holistic phonological segments is also consistent with the surface-impoverished hypothesis. More generally, Chambers and Reisberg (1985) claim that mental imagery's representations are semantically interpreted instead of being composed of raw sensory information. If speech imagery (i.e. inner speech) is similar, it should emphasize deep rather than surface information.

Because lexical bias requires the activation of deeper lexical representations, whereas the phonemic similarity effect is based on surface featural representations, the surface-impoverished hypothesis predicts preserved lexical bias, but a weakened phonemic similarity effect, in inner speech.

Deep-impoverished hypothesis. Inner speech represents speech sounds or gestures, and not higher level information. This hypothesis is rooted in conceptions of a short-term memory comprised of auditory or articulatory representations, rather than lexical and semantic representations (e.g. Baddeley, 1966). If inner speech is like this, then phonemic similarity should affect inner slips, but higher levels (lexical bias) should not.

The experiment reported in this paper used tongue-twister recitation to create both overt and inner slips. Internal recitation of tongue twisters is an effective way of producing inner slips (Dell & Repka, 1992; Postma & Noordanus, 1996), and the reported slips are often identical to those that occurred during overt recitation.

To create the materials for this experiment, we first did a preliminary experiment that generated only overt errors. We used the classic Baars et al. (1975) SLIP procedure to elicit onset errors in two-word CVC targets that manipulated slip outcome lexicality and onset phoneme similarity in 32 sets of four matched target word pairs (e.g. REEF LEECH → leaf reach; WREATH LEECH → leath reach; REEF BEECH → beef reach; WREATH BEECH → beeth reach). We manipulated phonemic similarity by changing the second onset of the pair (e.g. /l/) from a phoneme that differed from the first (/r/) by one feature to one that differed by two features (/b/). Outcome lexicality was manipulated in the first word of each target pair by a minimal change to its coda (/č/ to /θ/). The first word in each pair was identical within a condition of outcome lexicality (REEF, lexical outcome), and the second word was identical within a condition of phonemic similarity (LEECH, similar condition). The second slip-outcome (reach) was identical for all pairs within a set. Since word frequency affects phonological errors (Dell, 1990), the first word of each critical pair was controlled for target and slip-outcome log frequency (Kučera & Francis, 1967): Targets: lexical (REEF) = 2.59, nonlexical (WREATH) = 2.47; Outcomes: lexical similar (LEAF) = 3.27, lexical dissimilar (BEEF) = 3.26; nonlexical similar (LEATH) = 0.09; nonlexical dissimilar (BEETH) = 0.0.
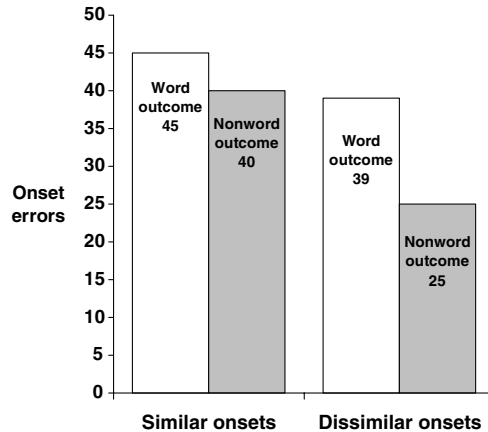
Fig. 1. Examining onset errors on the first word of each pair (including complete exchanges, partial exchanges, and anticipations), the preliminary experiment demonstrated significant lexical bias ($p = .0304$) and phonemic similarity ($p = .0173$) effects, with no indication of an interaction ($p = .4229$). Additional details about this preliminary experiment are available from the authors upon request.

The preliminary experiment demonstrated significant lexical bias and phonemic similarity effects (and no interaction) on the totals of overt onset errors (Fig. 1). Its materials then formed the basis of the tongue twisters for the main study. We created four-word tongue twisters whose last two words were the critical pairs from the previous experiment, and whose first two words came from a preceding "interference pair", which had been used in the preliminary experiment to increase the chance of a slip. For example, in the preliminary experiment the critical pair REEF BEECH from the lexical/dissimilar condition was preceded by the interference pair, BEAN REED. Putting them together makes the new test sequence: BEAN REED REEF BEECH. The other three conditions were assembled in a similar manner (Table 1). Thus, phonemic similarity was manipulated by changing the onsets of the first and fourth words across conditions and holding the second and third words constant; this strategy allows direct comparison of slips on REED REEF, for example, when the surrounding words have dissimilar (/b/) onsets to slips on the same words when the surrounding words have similar (/l/) onsets. Outcome lexicality was manipulated on the third word: slips of REEF to the words LEAF or BEEF can be compared to slips of WREATH to the nonwords LEATH or BEETH.

Table 1
A matched set of four-word sequences

	Similar onsets				Dissimilar onsets			
Word outcome	lean	reed	reef	leech	bean	reed	reef	beech
Nonword outcome	lean	reed	wreath	leech	bean	reed	wreath	beech

2. Methods

2.1. Participants

Forty-eight 20- to 30-year-old Champaign-Urbana residents received $10 for participating. All had normal or corrected-to-normal vision and hearing and were American English speakers who had not learned any other languages in the first five years of their lives.

2.2. Materials

Thirty-two matched sets of four-word sequences were devised as described above. Sequences were placed into counterbalanced lists, yielding four 32-item lists with eight sequences of each condition in each list. Within each list, half of the sequences in each condition were marked to be recited aloud and half were marked to be 'imagined'; the order of these overtness conditions was pseudorandom and fixed. A second version of each of these four lists then reversed the overtness pattern, resulting in a total of eight lists.

2.3. Procedure

The procedure for each sequence consisted of a study phase followed by a testing phase. Each sequence was presented in the center of a 17″ computer screen, in white 18-point Courier New font on a black background. Three seconds after the sequence appeared, a 1-Hz metronome began to play at a low volume. Participants then recited the sequence aloud four times, in time with the metronome, pausing between repetitions, and then pressed the spacebar to continue. The metronome then stopped and the screen went blank; by this point the participants should have memorized the sequence. After 200 ms, a cue to recite either aloud (a mouth) or internally (a head) appeared in the center of the screen. A half second later a faster (2-Hz) metronome began and the sequence reappeared in a small, low-contrast font at the top of the screen; participants were instructed that they could check this in between recitations, but should avoid looking at the words during their recitations. Participants now attempted to recite the sequence four times, pausing four beats between recitations and stopping to report any errors immediately. Error reports were to include both actual and intended 'utterances' (e.g. "Oops, I said LEAF REACH instead of REEF LEECH"). After completing the four fast repetitions, participants pressed the spacebar, whereupon the display went blank, the metronome terminated, and the next trial began after a 200 ms delay.

Each participant was assigned to one of the eight lists. During four practice trials (two inner and two overt), participants were encouraged to really imagine saying the word sequences without moving their mouths (on inner trials), and to immediately stop and report any errors that they made during the fast recitations. In the rare case that a participant's reporting of an error was unclear, the experimenter prompted the participant for more information (e.g. Participant: "Oh, I said LEAF." Experi-

menter: "LEAF instead of what?" Participant: "I said LEAF instead of REEF"). Participants' utterances were digitally recorded and transcribed both on- and off-line.

2.4. Analyses

All relevant errors were replacements of an onset by the other onset in the sequence. Only onset replacements on the third word were counted (e.g. REEF → LEAF, WREATH → LEATH, REEF → BEEF, and WREATH → BEATH) in tests of lexical bias, because this was the word in which outcome lexicality was manipulated. We counted onset replacements on both the second and third words for tests of phonemic similarity. As explained earlier, these two words are exactly balanced between the similar and dissimilar conditions.

We computed the proportions of trials that contained target errors, and report them along with the count data below. These proportions were computed separately for each condition and each participant (for the by-participant analyses) and for each item set (for the by-items consideration).

Analyses used Wilcoxon signed-rank tests, using a continuity correction (Sheskin, 2000), an adjustment for tied ranks (Hollander & Wolfe, 1973) and a reduction of the effective n when differences between paired observations were zero (e.g. Gibbons, 1985; Sheskin, 2000).

We reject or fail to reject the null hypothesis based on the by-participants analyses but, to document the consistency of the effects across item groups for each contrast, we also examined the 5 item sets with the largest differences in either direction. Where null hypotheses are rejected, we report the number of those sets in which the difference was not in the overall direction (e.g. as, "1 out of 5 sets in the opposite direction"). All planned tests of lexical bias and phonemic similarity main effects are one-tailed as these effects are well known in the literature. Any tests of interactions, though, are two-tailed as there is no firm basis for an expected direction.

2.5. Results/discussion

Errors were recorded on 1217 of the 6144 recitations. One hundred and ninety three of these recitations contained at least one expected onset substitution in word positions 2 and 3, appropriate for analysis of phonemic similarity effects, and 125 contained target errors in word position 3, useful for the analysis of lexical bias.

The results replicated the findings of the preliminary experiment for overt speech, but suggest differences for inner speech. Overall, more word- (84 errors [proportion of relevant trials that were erroneous = 2.96%]) than nonword-outcome (41 [1.45%]) slips were produced ($p = .0024$; 0 out of 5 item sets in the opposite direction). This main effect of lexical bias held true for both overt speech (48 [3.33%] to 21 [1.48%]; $p = .0089$; 0 out of 5 item sets in the opposite direction) and inner speech (36 [2.59%] to 20 [1.42%]; $p = .0089$; 0 out of 5 item sets in the opposite direction) conditions. There was no detectable interaction between lexical bias and overtness ($p = .5542$),

or between lexical bias and phonemic similarity in either overt ($p = .7361$) or inner speech ($p = .5993$).

Examination of word positions two and three for phonemic similarity showed that overt slips more often involved similar (66 [4.39%]) than dissimilar (36 [2.40%]) phonemes ($p = .0353$; 1 out of 5 item sets in the opposite direction), but slips in inner speech exhibited no such phonemic similarity effect (39 [2.58%] to 52 [3.49%]; $p = .8483$). This phonemic similarity by overtness interaction was significant ($p = .0234$; 0 out of 5 item sets in the opposite direction).

3. General discussion

The principal findings are easy to state. The lexical bias and phonemic similarity effects are robust in overt speech; they were demonstrated in two experimental paradigms. In a direct comparison between inner and overt slips, lexical bias was present in both, but the phonemic similarity effect was only present with overt slips.

The straightforward interpretation of these results is that inner speech is impoverished. Either inner speech's generation or its interpretation by the comprehension system (or both) lacks representations that support the phonemic similarity effect. These findings are contrary to the hypothesis that inner speech is the product of an articulatory/acoustic system with no contact with lexical information (the deep-impoverished hypothesis). They are, instead, consistent with the surface-impoverished hypothesis, in which featural, but not lexical, representations are weakened. We should note that our claim of surface-impoverishment may not hold true for the sort of inner speech that Levelt (1989) describes as the basis for monitoring in overt speech production. In fact, Levelt's inner speech may well be more fully specified, because of impending overt articulation.

The spreading-activation model of Dell (1986) can be used to simulate the data (Fig. 2). If access to features is blocked, the model's error rate in similar conditions equals that of the dissimilar conditions. The loss of the phonemic similarity effect does not affect lexical bias, however. Blocking the features in the model therefore mimics the inner speech condition, while leaving them accessible simulates the overt condition. More generally, a plausible account of the data is that lexical bias and phonemic similarity effects are generated by a hierarchical speech production/perception system with lexical bias mediated by access to lexical representations and phonemic similarity mediated by featural representations, and that inner speech lacks the latter more than the former.

It is important to recognize that inner speech is a product of perception as well as production. We know its properties by our perception of it. Consequently, the impoverishment at the featural level could, logically, be caused during production, perception, or both. If production is responsible, features may be absent from the inner speech production code (e.g. as in Wheeldon & Levelt's, 1995, phonological code, or our simulation) and hence no effect of shared features occurs in errors. If the perceptual system is responsible, there are at least two possibilities. For one, the features could be generated, but poorly perceived. For example, it may be hard
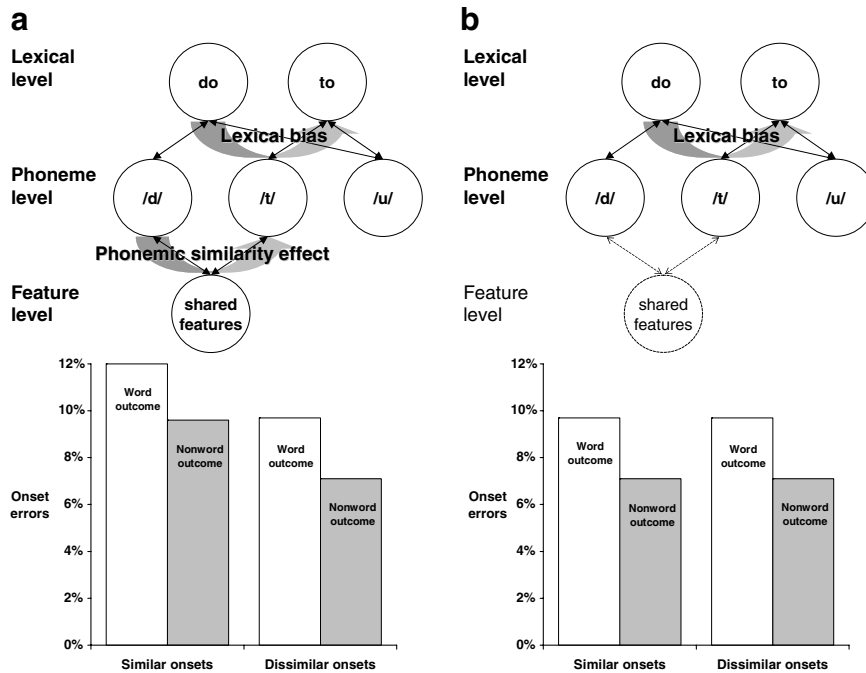
Fig. 2. Model predictions for inner speech errors, based on the (a) Unimpoverished and (b) Surface-Impoverished hypotheses. Activated features feed back to connected phonemes, increasing the probability that a similar phoneme will be selected. (a) *With feature-level activation*, the model predicts an error distribution in inner speech that is identical to that in overt speech. (b) *Without feature-level activation*, no phonemes receive feedback activation from the features, and so both similar and dissimilar items are treated as if they were dissimilar. Lexical bias occurs in both conditions due to activation feeding back from phonemes to words. (Connection weights = 0.2, decay = 0.4, activation spreading period = 4 time steps, standard deviation of activation noise = 0.68; additional details are available upon request.)

to internally "hear" the all of the features, and so slips involving similar phonemes might not be detected. Or, instead, the features could be present in the production system, but their effects on slips may not be transmitted to the perceptual system. Our experiment does not distinguish among these possibilities. A corollary to this caveat is that, although we simulated the experiment with an interactive model that attributed the impoverishment of inner speech to the production system, its findings do not compel interactive explanations for the error phenomena or the conclusion that the impoverishment is solely within the production system.

4. Conclusion

The little voice inside your head has much in common with articulated speech. Just like overt speech, inner speech has speech errors in it, and these errors exhibit

one of the most important error effects, lexical bias. But inner speech is also different from overt speech. Perhaps because inner speech lacks articulation, it is also impoverished at the featural level. Poor generation of features during the ''production'' of inner speech or poor sensitivity to features during its ''perception'' eliminated the effect of phonemic similarity on slips. Ultimately, we can understand inner speech as a form of mental imagery. Although images are much like the real thing, they are also more abstract (Pylyshyn, 1981) and less ambiguous (Chambers & Reisberg, 1985). In the speech domain, this translates into representations that emphasize lexical and segmental properties, rather than featural and articulatory ones.

References

Baars, B. J., Motley, M. T., & MacKay, D. G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior, 14*, 382–391.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology, 18*, 362–365.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short term memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 575–589.

Chambers, D., & Reisberg, D. (1985). Can mental images be ambiguous?. *Journal of Experimental Psychology: Human Perception and Performance 11*, 317–328.

Costa, A., Roelstraete, B., & Hartsuiker, R. J. (2006). The lexical bias effect in speech production: Evidence for feedback between lexical and sublexical levels across languages. *Psychonomic Bulletin & Review, 13*(6), 972–977.

Dell, G. S. (1978). Slips of the mind. In M. Paradis (Ed.), *The fourth LACUS forum* (pp. 69–75). Columbia, SC: Hornbeam Press.

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 93*, 283–321.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5*, 313–349.

Dell, G. S., & Repka, R. J. (1992). Errors in inner speech. In B. J. Baars (Ed.), *Experimental slips and human error: Exploring the architecture of volition* (pp. 237–262). New York: Plenum.

Ford, J. M., & Mathalon, D. H. (2004). Electrophysiological evidence of corollary discharge dysfunction in schizophrenia during talking and thinking. *Journal of Psychiatric Research, 38*, 37–46.

Gibbons, J. D. (1985). *Nonparametric methods for quantitative analysis* (2nd ed.). Syracuse, NY: American Sciences Press.

Hartsuiker, R. J., Anton-Mendez, I., Roelstraete, B., & Costa, A. (2006). Spoonish Spanerisms: A lexical bias effect in Spanish. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 949–953.

Hartsuiker, R. J., Corley, M., & Martensen, H. (2005). The lexical bias effect is modulated by context, but the standard monitoring account doesn't fly: Related beply to Baars, Motley, and MacKay (1975). *Journal of Memory and Language, 52*, 58–70.

Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology, 42*(2), 113–157.

Hockett, C. (1967). Where the tongue slips, there slip I. *To honor Roman Jakobson* (Vol. 2, pp. 910–936). The Hague: Mouton.

Hollander, M., & Wolfe, D. A. (1973). *Nonparametric statistical methods*. New York: Wiley.

Hotopf, W. H. N. (1980). Semantic similarity as a factor in whole-word slips of the tongue. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 97–109). New York: Academic Press.

Humphreys, K. R. (2002). Lexical bias in speech errors. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, Urbana, IL.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

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: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1–75.

MacKay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia, 8*, 323–350.

Nooteboom, S. G. (2005a). Listening to one-self: Monitoring speech production. In R. Hartsuiker, Y. Bastiaanse, A. Postma, & F. Wijnen (Eds.), *Phonological encoding and monitoring in normal and pathological speech* (pp. 167–186). Hove, UK: Psychology Press.

Nooteboom, S. G. (2005b). Lexical bias revisited: Detecting, rejecting and repairing speech errors in inner speech. *Speech Communication, 47*(1–2), 43–58.

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

Postma, A., & Noordanus, C. (1996). The production and detection of speech errors in silent, mouthed, noise-masked, and normal auditory feedback speech. *Language and Speech, 39*, 375–392.

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review, 88*(1), 16–45.

Roelofs, A. (2004). Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: Comment on Rapp and Goldrick (2000). *Psychological Review, 111*, 561–572.

Shattuck-Hufnagel, S., & Klatt, D. H. (1979). The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior, 18*, 41–55.

Sheskin, D. J. (2000). *Handbook of parametric and nonparametric statistical procedures* (2nd ed.). Boca Raton, FL: Chapman & Hall.

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.

Sokolov, A. N. (1972). *Inner speech and thought*. New York: Plenum.

Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language, 34*, 311–334.