THE LITTLE VOICE IN YOUR HEAD: ERROR-BASED INVESTIGATIONS OF ABSTRACTED AND ARTICULATED INNER SPEECH

BY

GARY MICHAEL OPPENHEIM

B.A., Grinnell College, 2001

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts in Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2009

Urbana, Illinois

Master's Committee:

Professor Gary S. Dell, Chair
Professor J. Kathryn Bock

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*), two speech-error phenomena that have been localized to an articulatory-feature processing level and a lexical-phonological level, respectively. In the first chapter, we examine 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. In the second chapter, we focus on the role of articulation and motor imagery in inner speech, comparing inner speech without articulatory movements to articulated (mouthed) inner speech. As before, the speech errors occurring during unarticulated inner speech did not exhibit the phonemic similarity effect—just the lexical bias effect. In contrast, speech errors that occurred in articulated inner speech exhibited the phonemic similarity effect and lexical bias effect, similar to our previous findings for overt speech. The results constrain our previous interpretation and are interpreted as support for a *flexible abstractness* account of inner speech. That is, articulatory information is incorporated into inner speech if it is available, but is not essential to the phenomenon. This conclusion has ramifications for issues related to the embodiment of language and speech and for the theories of the processes involved in speech production.

To my mother and father, whose love and support has
guided me to this point and inspired me ever onwards.

ACKNOWLEDGEMENTS

I am greatly indebted to my advisor, Gary S. Dell, who co-authored the two articles included in this document. I also thank J. Kathryn Bock, and the members of the Dell and Bock labs for their comments on this work and other contributions. Thanks also my family, friends, and housemates for their support and feedback throughout this process. This research was funded in part by National Institutes of Health Grants DC 00191 and HD 44458.

TABLE OF CONTENTS

CHAPTER 1: INNER SPEECH SLIPS EXHIBIT LEXICAL BIAS, BUT NOT THE PHONEMIC SIMILARITY EFFECT[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

---

[1] This chapter represents a collaboration with my advisor, Gary Dell. It has been published as follows:

> Oppenheim, G. M., & Dell, G. S. (2008). Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition, 106*(1), 528-537.

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, in press; Dell, 1986; 1990; Humphreys, 2002; Hartsuiker, Anton-Mendez, Roelstraete, & Costa, 2006; Hartsuiker, Corley, & Martensen, 2005; 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, Roelofs, & Meyer, 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) = 3.27, nonlexical (WREATH) = 3.26; Outcomes: lexical similar (LEAF) = 2.59, lexical dissimilar (BEEF) = 2.47; nonlexical similar (LEATH) = 0.09; nonlexical dissimilar (BEETH) = 0.0)

The preliminary experiment demonstrated significant lexical bias and phonemic similarity effects (and no interaction) on the totals of overt onset errors (Figure 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.

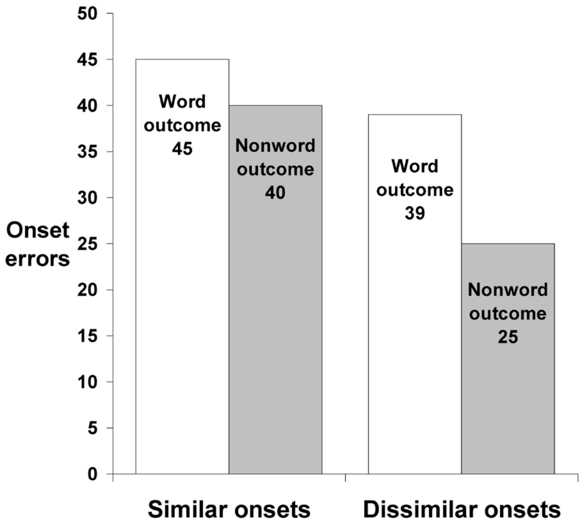

Figure 1. Examining onset errors on the first word of each pair (including complete exchanges, partial exchanges, and anticipations), the preliminary experiment demonstrated significant lexical bias (*p*=.0304) and phonemic similarity (*p*=.0173) effects, with no indication of an interaction (*p*=.4229). Additional details about this preliminary experiment are available from the authors upon request.

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

Methods

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.

Materials

32 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.

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.” Experimenter: “LEAF instead of what?” Participant: “I said LEAF instead of REEF”). Participants’ utterances were digitally recorded and transcribed both on- and off-line.

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.

Results/Discussion

Errors were recorded on 1217 of the 6144 recitations. 193 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).

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 (Figure 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.

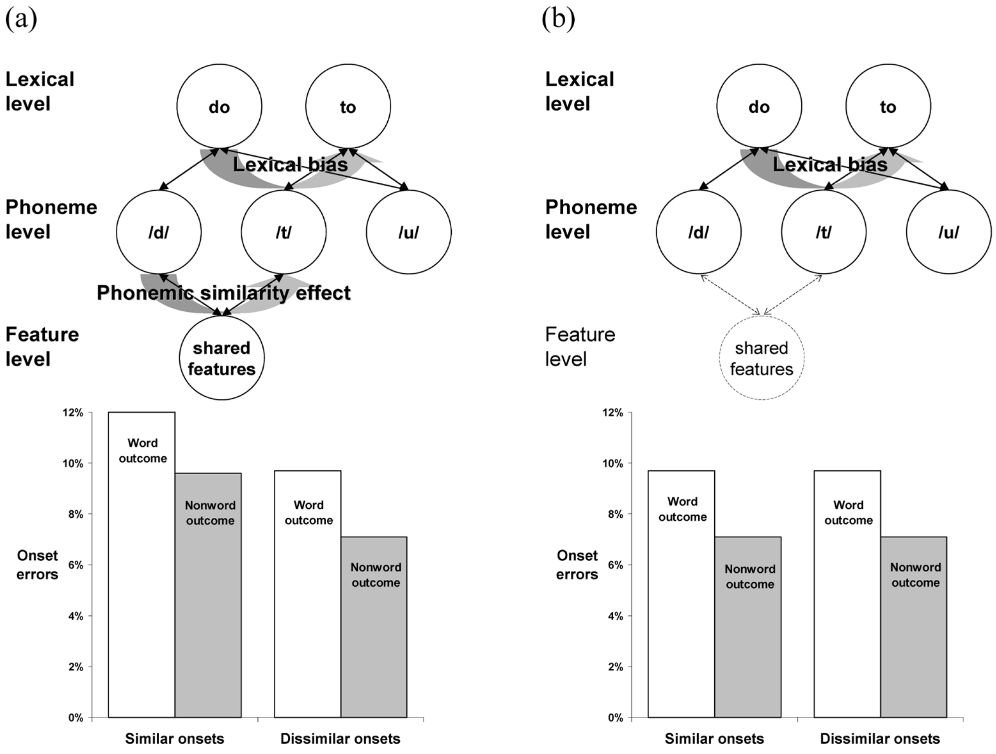

Figure 2. Model predictions for inner speech errors, based on the (a) Unimpoverished and (b) Surface-Impoverished hypotheses. Activated features feed back to connected phonemes, increasing the probability that a similar phoneme will be selected. (a) *With feature-level activation,* the model predicts 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.)

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 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.

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? J*ournal of Experimental Psychology: Human Perception and Performance*, *11*, 317-328.

Costa, A., Roelstraete, B., & Hartsuiker, R. J. A. (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, S.C.: 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.

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. In *To honor Roman Jakobson* (Vol. 2, pp. 910-936). The Hague: Mouton.

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

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* (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/CRC.

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.

CHAPTER 2: MOTOR MOVEMENT MATTERS: THE FLEXIBLE ABSTRACTNESS OF INNER SPEECH[2]

Introduction

Embodiment, in the domain of language processing, is usually taken to be about whether meaning is sensory-motor in nature, specifically in terms of engaging sensory or motor simulations of the events signified by linguistic referents (*e.g.* Barsalou, 1999; Lakoff, 1987; Pulvermüller, 2005). For instance, understanding the word 'reach' may require basic visual, auditory, proprioceptive, and motoric circuitry to simulate the act of reaching so that the main difference between actually reaching and merely understanding the word *reach* is an apparent lack of motor movement. This perspective is consistent with the discovery of somatotopically relevant "mirror neurons" (e.g. Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006), and is specifically supported by studies that show response-time compatibility effects between the meaning of a linguistic stimulus and the required motor response (*e.g.* Glenberg & Kaschak, 2002) and by findings of trace amounts of semantically relevant muscle activity (*e.g.* Foroni & Semin, 2009) during the processing of action verbs. Thus, the processing of language meaning appears to invoke motor simulations of the semantically represented actions.

But there is another question that arises concerning embodiment and language: to what extent do internal representations of speech have sensory or motor components? The motor theory of speech perception (Gallantucci, Fowler, & Turvey, 2006; Liberman, Dellatre, & Cooper, 1952; Liberman & Mattingly, 1985) is a classic example of a theoretical stand on this

---

[2] This chapter represents a second collaboration with my advisor, Gary Dell. It has been submitted for publication as follows:

> Oppenheim, G. M. & Dell, G. S. (submitted). Motor movement matters: the flexible abstractness of inner speech.

question; it claims that listeners perceive a syllable by internally simulating that syllable's production. A second example, and the one that we investigate here, concerns the nature of inner speech—the silent, internal speech that accompanies many cognitive activities (*e.g.* Dell, 1978; Dell & Repka, 1992; Postma & Noordanus, 1996; Oppenheim & Dell, 2008; Vygotskiĭ, 1965; as contrasted with auditory verbal imagery, *e.g.* McGuire *et al.*, 1996; Shergill *et al.*, 1999; Smith, Reisberg, & Wilson, 1992). Although there is no overt articulation or auditory consequence of inner speech, one can nevertheless ask to what extent this process is a motoric one.

There are two contrasting views on this question. The *motor simulation view* (*e.g.* Dell, 1978; Nooteboom & Quené, submitted; Postma & Noordanus, 1996; cf. Reisberg, Smith, Baxter, & Sonenshine, 1989) holds that inner speech is a relatively complete motoric simulation of the articulation process. The only difference is that the articulators do not move, perhaps because they are inhibited. The *abstractness view* (*e.g.* Caplan, Rochon, & Waters, 1992; Dell & Repka, 1992; MacKay, 1992; Macken & Jones, 1995; Oppenheim & Dell, 2008; Vygotsky, 1962; Wheeldon & Levelt, 1995), on the other hand, holds that inner speech is the consequence of the activation of abstract linguistic representations. For instance, inner speech may only involve the sequencing of phonemes that are not specified in terms articulatory features.

Six arguments for the abstractness of inner speech

As reviewed below, most of the evidence supports the abstractness view.

1.) Inner speech is faster than overt speech. Although the durations for tasks involving inner and overt recitation of particular words are highly correlated, overt speech tasks take longer, suggesting that inner speech production is abbreviated in some manner (*e.g.* MacKay, 1981; 1992). While this finding does not require that inner speech uses abstracted motor representations, it does suggest that inner speech lacks a full specification of articulatory properties.

2.) Inner speech uses less brain than overt speech. Similarly, while neuroimaging studies have shown that inner speech production activates many of the same brain areas as overt speech (*e.g.* Paulesu, Frith, & Frackowiak, 1993; Yetkin *et al.*, 1995), these areas tend to be less active in inner speech, suggesting that processing in these areas is not as complete or reliable as in overt speech. Specifically, inner speech involves less activation of brain areas thought to subserve the planning and implementation of motor movements (*e.g.* Barch *et al.*, 1999; Palmer *et al.* 2001; Schuster & Lemieux, 2005; see also McGuire *et al.*, 1996 and Shergill *et al.*, 1999, for neuroimaging work that differentiates between inner speech and auditory verbal imagery). These physiological observations certainly do not compel an abstractness view, since less motor activation may still be sufficient to produce motor simulations. But, to the extent that activation of motor and premotor areas corresponds to the psycholinguistic concepts of articulatory planning or simulation, the reduced activation for inner speech suggests that it is characterized by degraded representations of motor movements.

3.) Inner speech does not require articulatory abilities. The ability to overtly articulate a word is not required for successful use of inner speech. Anarthric patients, who have brain lesions that disrupt overt articulation (*e.g.* BA 1, 2, 3, 4, 6), nevertheless show indirect signs of intact inner speech (Baddeley & Wilson, 1985; Vallar & Cappa, 1987). Similarly, localized magnetic interference (*i.e.* rTMS) can disrupt healthy participants' overt speech while leaving their inner speech seemingly intact (Aziz-Zadeh, Cattaneo, Rochot & Rizzolatti, 2005). These dissociations suggest that inner speech does not need to be articulatory.

4.) Articulatory suppression does not (necessarily) eliminate inner speech. A motor simulation view of inner speech requires that its production engages articulatory resources, implying that articulatory suppression (*e.g.* in the form of concurrent articulation tasks, such as repeating "tah tah tah") should impair the performance of tasks that require inner speech. However, the predicted impairment has not consistently been observed in phoneme monitoring tasks, in which the pronunciation of a word must be internally generated and monitored for the presence of a target phoneme. While Smith, Reisberg, and Wilson (1992) reported that articulatory suppression impaired phoneme monitoring accuracy for orthographically presented words, Wheeldon and Levelt (1995) demonstrated only a very limited cost from suppression when subjects monitored their Dutch translations of auditorily presented English words for phoneme targets. Thus, while speakers may employ fine-grained articulatory simulations (*e.g.* when processing orthographic stimuli), these may not be required for all inner speech tasks.

5.) Inner speech practice does not (necessarily) transfer to overt speech performance. If inner and overt speech involve comparable planning processes, one should expect that practicing an utterance in inner speech would improve overt performance, and vice versa. Such transfer definitely occurs when the purpose of the internal practice is to rehearse conceptual or lexical information. For instance, MacKay (1981) had participants practice translating a sentence from German to English, using either inner or overt speech. Both yielded equivalent improvements in subsequent overt performance. A more everyday example: mentally rehearsing a shopping list is quite effective for later reproduction of the list. Inner-speech practice is less effective, though, when the overt-speech task is articulatorily challenging. Dell and Repka (1992) found that while overt practice of tongue twisters improved the accuracy of both inner and overt recitation of those tongue twisters, inner practice improved inner, but not overt, performance. This asymmetry suggests that inner speech may fail to engage the articulatory planning that makes tongue twisters twist tongues. Together the studies suggest that inner speech involves relatively intact higher-level representations (*e.g.* on the word or message level), but degraded lower-level representations (*e.g.* sub-phonemic features).

6.) Phonological errors in inner speech do not show a phonemic similarity effect. Oppenheim and Dell (2008) compared self-reported speech errors made during the overt or inner recitation of tongue-twisters. Phonological errors in both conditions tended to produce more words than nonwords (*i.e.* lexical bias, *e.g.* REEF→LEAF is more likely than WREATH→LEATH), suggesting intact lexical-phonological processes in both inner and overt speech. But only the overt speech errors tended to involve similarly articulated phonemes (*i.e.* the phonemic similarity effect, *e.g.* REEF→LEAF is more likely than REEF→BEEF). It is well known that speech errors are strongly affected by phonemic features (e.g. Goldrick, 2004; MacKay, 1970), and hence the absence of an effect of shared features on the inner speech errors suggests that inner speech involves phonemic, but not subphonemic (e.g. articulatory) representations. This result further implies a distinction between a processing level concerned with lexical-phonological representations (present in inner speech) and a post-lexical level at which featural information is relevant (present in overt speech). In fact, such a distinction has been well supported by studies of errors by individuals with brain damage (e.g. Goldrick & Rapp, 2007) and is central to one well known theory of production (Levelt, Roelofs, & Meyer, 1999).

Challenges for the abstractness view (i.e. support for the motor simulation view of inner speech)

However, the abstractness view cannot offer a full account of even those results we have discussed above. For instance, the regularity with which the timing of inner and overt speech tasks ordinally track each other (*e.g.* silent sentence rehearsal: MacKay, 1981; 1992; reading: Abramson & Goldinger, 1997) suggests that inner speech preserves many aspects of overt speech. Moreover, the inner speech activation of brain areas involved in motor planning (*e.g.* Barch *et al.*, 1999; Yetkin *et al.*, 1995), although small, suggests at least some degree of motor simulation. In fact, for decades, cognitive neuroscientists were sufficiently convinced of the correspondence between inner and overt speech that they regularly used inner speech as a proxy for overt speech in neuroimaging tasks, as a strategy to avoid technical limitations (e.g. Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997).

Furthermore, while the Oppenheim and Dell's (2008) finding that inner slips are insensitive to phonemic similarity result is suggestive, it contradicts the conclusions of some other studies of errors in inner speech. In studies comparing inner and overt performance on tongue twisters, Meringer and Meyer (1895, cited in MacKay, 1992), Dell (1978), and Postma and Noordanus (1996) all reported remarkably similar error distributions in inner and overt speech. Postma and Noordanus, in particular, contrasted the errors that subjects reported while reciting tongue twisters in inner speech, mouthed speech (*i.e.* inner speech with silent articulatory movements), noise-masked overt speech (*i.e.* overt speech without auditory feedback), and normal overt speech. Crucially, participants reported similar numbers and types of errors (*e.g.* phoneme anticipations, perseverations, etc.) across the first three conditions, with higher reporting rates only in the normal overt speech condition. This pattern supports the motor simulation view, suggesting that unarticulated inner speech and normal overt speech engage similar planning processes – right down to the level of individual motor movements – with any apparent differences being attributable to auditory error detection. In addition, the specific claim of Oppenheim and Dell (2008) that inner speech does not involve featural information is, itself, disputed. For instance, Brocklehurst and Corley (2009) recently reported finding a phonemic similarity effect in inner speech errors that did not significantly differ from that in overt speech, thus arguing that inner speech does incorporate featural information. Our studies, reported here, will be directly relevant to this controversy.

Finally, analyses of quickly interrupted speech errors (*e.g.* REEF→LE…; *e.g.* Nooteboom & Quené, 2008) – which must have been interrupted before initiating articulatory movements (Levelt, 1983) – demonstrate that some amount of detailed articulatory information must be available prior to overt speech. Specifically, these analyses indicate that errors in which a target phoneme is replaced by a dissimilar one are, under some conditions, more likely to be quickly interrupted than those in which the replacement is more similar to the target If, as posited by the phonological monitoring account (Levelt, 1983), these errors are detected by monitoring inner speech, then inner speech would have to include detailed subphonemic representations in order for its monitoring to produce similarity effects for interrupted overt errors.

Synthesis: A flexible abstractness account of inner speech

Acknowledging both the support for the abstractness view and a number of challenges, we introduce the *flexible abstractness view*: Given that we can produce both overt speech (clearly a motoric act) and we can also produce fairly abstract inner speech, we seem to have the ability to control the extent of motoric expression in speech production. Given this control, we ought to be able to vary the extent to which inner speech is motoric. Perhaps such variation can explain the variability in the extent to which we see evidence of articulatory effects in inner speech.

We test this hypothesis by comparing reported speech errors in two forms of inner speech, normal unarticulated inner speech, and mouthed (*i.e. silently articulated*) speech. The former, we hypothesize, is more abstract than the latter. As was done in Oppenheim and Dell (2008), we examine lexical bias and the phonemic similarity effect for phonological errors in inner speech, as a way to gauge the involvement of higher- and lower-level processes in the two forms of inner speech. If the flexible abstraction view holds true, then adding the motor movements to inner speech should increase motor planning, thereby restoring the phonemic similarity effect.

Methods

Participants recited tongue-twister phrases that manipulated the opportunity for errors to exhibit both the phonemic similarity effect and the lexical bias effect. Each recitation trial used either silent unmouthed inner speech, or silent mouthed inner speech. All conditions were manipulated in a within-participant and within-item-set fashion for maximum power.

Participants

Eighty 18- to 30-year-old Champaign-Urbana residents participated in exchange for cash or course credit. 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.

Materials

Thirty-two matched sets of four-word tongue-twister phrases were devised, as illustrated in Table 2. These were, for the most part,[3] the same stimuli used in Oppenheim and Dell (2008), designed to test for lexical bias on the third word and a phonemic similarity effect on the second and third words. Phoneme similarity was manipulated by changing the onsets of the first and fourth words (e.g. /l/) from a phoneme that differed from the second and third (/r/) by one articulatory feature to one that differed by two features (/b/). Outcome lexicality was manipulated in the third word of each target set by a minimal change to its coda (/č/ to /θ/). The second word in each set was identical in all conditions, and the third word was identical within a condition of outcome lexicality (REEF, lexical outcome). Since word frequency affects phonological errors (Dell, 1990; Kittredge, et al., 2008), the third word of each set was controlled for target and slip-outcome log frequency (Kučera & Francis, 1967): Targets: lexical (REEF) = 3.38, nonlexical (WREATH) = 3.28; Outcomes: lexical similar (LEAF) = 2.73, lexical dissimilar (BEEF) = 2.63; nonlexical similar (LEATH) = 0.09; nonlexical dissimilar (BEETH) = 0.0).

---

[3] We replaced one item set, a w / m(similar) / r(dissimilar) phoneme contrast. This set showed a strong reverse similarity effect in Oppenheim & Dell (2008)' overt condition, suggesting a possible limitation of our three-feature metric for assessing similarity (cf. Frisch, 1996). However, we did include this item as a 33rd trial for each subject. Substituting that data for data from our replacement item would not change the outcome of our statistical analyses; if anything, it would strengthen our effects.

Table 2. A matched set of four-word tongue-twisters.

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

These phrases were placed into counterbalanced lists, yielding four 32-item lists with eight phrases of each condition in each list. Within each list, half of the phrases in each condition were marked to be 'imagined' while mouthing and half were marked to be 'imagined' without mouthing; the order of these overtness conditions was pseudorandom and fixed. A second version of each of these four lists then reversed the mouthed/unmouthed pattern, resulting in a total of eight lists, with each participant assigned to one.

Procedure

Each trial consisted of an overt study phase followed by a silent testing phase. At the start of the study phase, a phrase appeared in the center of a 17" computer screen, in white 18-point Courier New font on a black background. After three seconds, a quiet 1-Hz metronome began. Participants then recited the phrase aloud four times, in time with the metronome, pausing between repetitions, and then pressed the spacebar to continue, signaling that they had memorized the phrase. The metronome then stopped and the screen went blank for 200 ms.

Then the test phase began. A picture appeared in the center of the screen, cueing the subject to imagine saying the phrase while either mouthing (a mouth) or not mouthing (a head), and a faster (2-Hz) metronome began 500 ms later. The phrase reappeared in a small, low-contrast font at the top of the screen; participants were instructed that they could check this between recitations, but should avoid looking at the words otherwise. Participants now attempted the phrase four times, pausing four beats between attempts. For the unmouthed condition, participants were instructed to imagine saying the phrase, in time with the metronome, without moving their mouth, lips, throat, or tongue. The mouthed condition was identical, except that participants were now instructed to silently articulate while imagining the phrase. As they did so, participants were instructed to monitor their inner speech in both conditions, stopping to report any errors immediately, and precisely specifying both their actual and intended 'utterances' (*e.g.* "Oops, I said LEAF REACH instead of REEF LEECH"). After completing the four fast attempts, participants pressed the spacebar, and the next trial began 200 ms later.

To ensure procedural consistency within and across subjects, each participant was trained through two demonstration trials and four practice trials (two mouthed and two unmouthed). In the rare case that a participant's reporting of an error was unclear, the experimenter prompted them for more information whenever it was possible to do so in a timely manner (*e.g.* Participant: "Oh, I said LEAF." Experimenter: "LEAF instead of what?" Participant: "I said LEAF instead of REEF"). Error reports were digitally recorded and transcribed both on- and off-line.

Analyses

Any observed or self-reported deviations from the instructed procedures (*e.g.* mouthing in the unmouthed condition, or reporting errors imprecisely, as "Um, I said LEAF," or "Oops, I said… oh, nevermind,") were dealt with by excluding the affected trials (80 trials, < 1%). Entire participants (n=17) were replaced if this meant excluding more than 25% (*i.e.* 4 trials) of their data from any one condition or more than 6.25% (*i.e.* 8 trials) of their data overall. Each of the remaining trials was categorized as follows, based on self-reports:

1. *Target errors* were exactly those utterances where one onset in the phrase replaced the other without creating other errors. For instance, replacing REEF with LEAF (a simple onset phoneme anticipation) was considered a target error,

whereas replacing REEF with LEE, LEAD, LEECH, LEAFS, or LEATH was not.

2. *Competing errors* included any errors, other than the target errors, where the target onset could be construed as having been replaced by the other onset in the phrase. For instance, replacing REEF with LEE, LEAD, LEECH, LEAFS, or LEATH would all be considered competing errors.
3. *Other contextual word errors* were cases in which a participant reported misordering the second and third words or their codas (e.g. REED→REEF or REEF→REED). These are not directly relevant to our hypotheses, but account for a substantial proportion of the non-target errors.
4. *Miscellaneous errors* included all other errors, such as noncontextual phoneme errors, vowel or coda errors (*e.g.* REEF → RIFE, REEF → REAL), disfluencies (*e.g.* reported as, "I just stopped instead of saying the third word,"), and other multi-phoneme or word-level errors (*e.g.* REEF → RIND, REEF → FROG).
5. *No errors reported* were those trials where no errors were reported.

Since participants had been instructed to stop and report any errors immediately, in the rare event that a subject reported multiple errors in a single attempt, only the first of these errors was recorded.

Given the structure of the materials, our analyses focused on just the second and third word of each item set. This strategy avoids potential confoundings arising from asymmetrical phoneme confusability (*e.g.* the probability of /s/-> /š/ is greater than that of /š/-> /s/, Stemberger, 1991) by holding the target onsets constant across all conditions. Thus, tests of phoneme similarity effects count errors on both the second and third word. Tests involving lexical bias, however, are restricted to target errors on the third word, which is the only part of the word set in which we systematically manipulated outcome lexicality (*e.g.* REEF → LEAF, WREATH → LEATH, REEF → BEEF, and WREATH → BEATH).

We computed the proportions of trials that contained target errors, and report them with the raw error counts below. Any trials ending before the critical word was attempted – that is, before the second word in the similarity analyses, and before the third word in the lexical bias analysis – were not included. Thus, converting the counts into proportions adjusts for the possibility that some utterances were interrupted before the critical words were attempted. Such opportunity-adjusted proportions were computed separately for each condition, participant (for the by-participant analyses) and item set (for the by-items consideration), and serve as the input for our statistical analyses.

Analyses used Wilcoxon signed-rank tests, with 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 at $\alpha=0.05$ 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 dissenting”). All planned tests of lexical bias, phoneme similarity effects, and similarity-by-mouthing interactions are directional, based on the findings of Oppenheim and Dell (2008) and related effects in the overt speech literature. We also report p-values from directional tests when describing non-significant unexpected effects that nonetheless have an implied directionality (e.g. phonemic similarity in unmouthed inner speech; the implied direction is that errors are more likely in the similar condition). This treatment is conservative as our theoretical perspective predicts null effects in these cases

Results / Discussion

Participants reported errors on roughly a quarter of the 10240 recitations. Consistent with previous reports (*e.g.* Postma & Noordanus, 1996), we found similar error totals for mouthed

(1301 errors) and unmouthed (1241) inner speech, suggesting that the mouthing manipulation did not greatly affect the overall probability of error production, detection, or reporting in this experiment. Moreover, errors showed similar distributions across word positions in the mouthed and unmouthed conditions (Figure 3). Such similarity suggests that we can now interpret any differences in target error rates as reflecting underlying processes of inner speech production and comprehension or monitoring rather than merely differences in the base rates of error production.

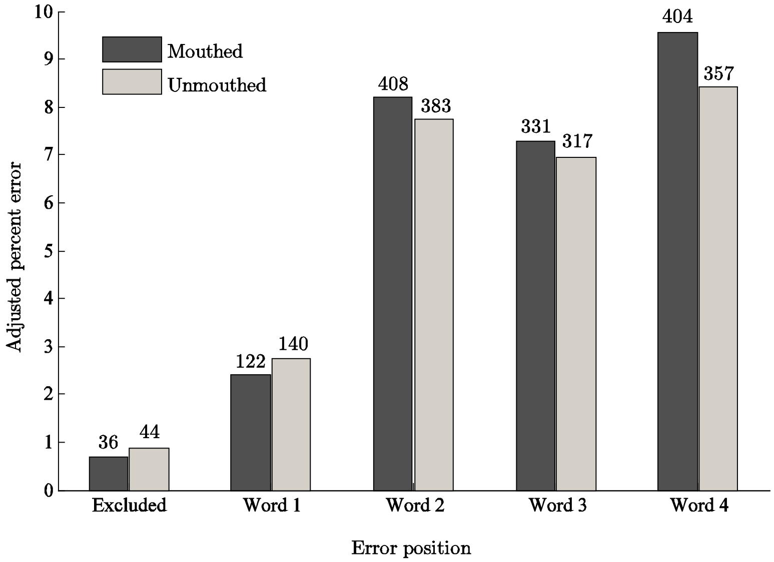

Figure 3. Total adjusted mouthed and unmouthed percent error collapsed across similarity and lexicality conditions, plotted by position within the phrases. Each column shows the opportunity-adjusted error rate for that position, and the number above it gives the unadjusted error count. Although excluded trials and errors on words 1 and 4 were not included in our other statistical analyses, we report aggregate error counts here in order to demonstrate general similarities in terms of error distributions over word positions.

Phonemic similarity effects

So how does silent articulation affect the phonemic similarity effect? To address this question, we restrict our focus to the 366 targeted onset anticipations and perseverations on words two and three (Table 3).

Table 3. Trials featuring an error on word two or three, for examination of phoneme similarity effects. Each cell lists the number of trials resulting in a certain outcome, followed by the percentage of the relevant trials that this represents. Only target errors were statistically examined, but we include other counts here for informational purposes.

	Mouthed		**Unmouthed**	
Outcome	**Similar**	**Dissimilar**	**Similar**	**Dissimilar**
Target errors	123 (4.98%)	75 (3.01%)	92 (3.74%)	76 (3.07%)
Competing errors	39 (1.58%)	29 (1.17%)	30 (1.22%)	27 (1.09%)
Other contextual word errors	120 (4.86%)	135 (5.42%)	115 (4.68%)	118 (4.77%)
Miscellaneous errors	100 (4.05%)	118 (4.74%)	114 (4.64%)	128 (5.17%)
No errors reported	2087 (84.53%)	2132 (85.66%)	2106 (85.71%)	2126 (85.90%)

Confirming a prediction of our flexible abstraction view of inner speech, only mouthed inner speech elicited a phonemic similarity effect (mouthing condition by similarity interaction, p

$< .03$, 1/5 item sets dissenting). Target errors involving similar phonemes were significantly more likely in the mouthed condition ($p<0.002$, 0/5 sets dissenting), but not in the unmouthed condition ($p>0.16$, 3/5 sets dissenting). This interaction suggests that, although inner speech can operate on a more abstract form-based level, it can also incorporate lower-level articulatory planning. The lack of a phonemic similarity effect in unarticulated inner speech replicates our previous findings and reinforces the view that inner speech can have degraded subphonemic representations under some conditions.

The presence of the effect in mouthed speech, in addition to supporting the flexibility of the extent to which such representations are used, also argues against an alternative explanation for the lack of a phonemic similarity effect found in unmouthed inner speech. This alternative is that the inner speech errors involving similar phonemes are difficult to detect because there is no sound in unmouthed inner speech. Hence, errors in the similar-phoneme condition would be under-reported and the overall effect of similarity is diminished. In our mouthed speech condition, there is no sound, and yet the similarity effect was robust (5% similar slips, 3% dissimilar slips), and was the same size as that found in Oppenheim and Dell's (2008) overt-speech condition (4% similar slips, 2% dissimilar slips). Thus, the presence of auditory information about the slip is not required for the similarity effect to be obtained. (Below, we test for, and eliminate, a more sophisticated form of this alternative, in which the similarity effect is attributed to a monitoring and repair process that works more effectively on mouthed slips.)

Lexical bias

Both mouthed and unmouthed target errors showed significant lexical bias, consistent with the assertion that both types of inner speech engage higher-level (*e.g.* phonemic and lexical) representations (Table 4). Recall that outcome-lexicality was specifically manipulated on Word 3 of each phrase. Target errors here produced more words than nonwords in both articulatory conditions (mouthed: p<0.006, 1/5 sets dissenting; unmouthed: p<0.03, 1/5 sets dissenting). No significant interactions between lexical bias and similarity (mouthed: p>0.18; unmouthed: p>0.25) or modality (p>0.11) emerged. These results replicate Oppenheim and Dell's (2008) finding of lexical bias for phonological errors in unarticulated inner speech, extending it to articulated inner speech as well. Moreover, the lack of an interaction between articulation and lexicality contrasts with the presence of an interaction between articulation and phonemic similarity, reinforcing our interpretation that unarticulated inner speech is specifically impaired in terms of lower-level articulatory representations.

Table 4. Trials featuring an error on word three, for examination of lexical bias effects. Each cell lists the number of trials resulting in a certain outcome, followed by the percentage of the relevant trials that this represents. Only target errors were statistically examined, but we include other counts here for informational purposes.

	Mouthed				Unmouthed			
	Similar		Dissimilar		Similar		Dissimilar	
Outcome	**Word**	**Non-word**	**Word**	**Non-word**	**Word**	**Non-word**	**Word**	**Non-word**
Target errors	49 (4.31%)	24 (2.12%)	31 (2.71%)	16 (1.41%)	35 (3.04%)	19 (1.71%)	24 (2.09%)	19 (1.68%)
Competing errors	5 (0.44%)	10 (0.88%)	5 (0.44%)	8 (0.71%)	3 (0.26%)	10 (0.90%)	8 (0.70%)	6 (0.53%)
Other contextual word errors	17 (1.50%)	20 (1.77%)	12 (1.05%)	17 (1.50%)	13 (1.13%)	10 (0.90%)	21 (1.83%)	11 (0.97%)
Miscellaneous errors	33 (2.90%)	25 (2.21%)	38 (3.32%)	21 (1.85%)	28 (2.43%)	41 (3.68%)	33 (2.87%)	36 (3.17%)
No errors reported	1032 (90.85%)	1053 (93.02%)	1060 (92.50%)	1072 (94.53%)	1071 (93.13%)	1034 (92.82%)	1064 (92.52%)	1062 (93.65%)

Error generation versus error repair accounts of lexical bias

Finding lexical bias in both inner speech conditions supports the view that higher-level influences on inner speech errors are present, regardless of the abstractness of the inner speech. We can also use these data to help understand what causes lexical bias in general. According to one class of theories, there are more word-outcome slips than nonword-outcome slips because more word outcome slips are directly generated by the phonological encoding process. For example, interactive theories of production hold that phoneme strings that form words are more likely to be encoded than those that do not because a top-down spread of activation from the lexical level biases the activation of phonemes so that they correspond to a single word (e.g. Dell, 1986; 1990; Nozari & Dell, 2008). Alternately, many theories of speech production posit that errors may be intercepted and repaired before a speaker becomes aware of them, or even before the error itself is even spoken (e.g. Levelt, Roelofs, & Meyer, 1999). By some accounts, potential nonword errors are more likely to be caught and repaired than potential word-outcome errors (*e.g.* Baars *et al.*, 1975, Hartsuiker, Corley, & Martensen, 2005; Hartsuiker, Antón-Méndez, Roelstraete, Costa, 2006; Nooteboom, 2005; Nooteboom & Quené, 2008). If so, then lexical bias could simply reflect the workings of a speech monitor that repairs speech errors with varying degrees of effectiveness (*e.g.* Garnsey & Dell, 1984). In sum, lexical bias could arise either because word-outcome slips are more likely to be generated (direct generation account), or because nonword-outcome errors are more likely to eliminated by a post-generation monitoring and repair process (biased repair account).

We can use our data to distinguish the contributions of these two mechanisms by modeling one version of the proposed repair process (Figure 4). This exercise follows Nooteboom and Quené (2008)'s insight that the contributions of monitoring to error patterns can be examined by considering errors other than target errors, and by treating the various kinds of errors as the outcome of a multinomial process. Specifically, Nooteboom and Quené assumed that the monitor may repair an utterance incorrectly, creating certain error types that would not otherwise occur in great quantities. Here, we assume that an error initially generated as, for instance, WREATH→BEETH (the Target error) may be repaired so that it is ultimately spoken (or in the case of inner speech, reported) as something else. These 'Competing errors' could correspond to a word from earlier in the tongue-twister (e.g. WREATH→ BEECH), some other word (e.g. WREATH → BEET), or a non-word (e.g. WREATH→BEESH). The occurrence of these competing errors can thus be attributed to the monitoring/repair process. By comparing how they, along with target errors and correct trials, are influenced by the lexical-bias manipulation, we can draw some conclusions about the mechanisms underlying the lexical bias effect.

We limit our free parameters by restricting this analysis to Target errors, Competing errors, and Correct trials. The analysis assumes that there are two stages to production: the initial generation of a planned utterance, and the monitoring/repair of that utterance. At each stage, multiple events can happen, so that the various results of a particular trial can be represented as branches of the tree shown in Figure 4. During the first stage, we assume that participants activate and select phonemes (*e.g.* as represented in Dell's 1986 model) so as to generate a target slip (e.g. WREATH→ BEETH) with probability *s*. The rest of the time (*i.e.* with probability (1-*s*)), the generation process is assumed to be correct (WREATH→WREATH). Then, all of these potential utterances are evaluated at the monitoring/repair stage. All correct utterances pass the monitor and are then reported (or spoken if the task were to produce overt speech) as correct. Of the initially generated target errors (*s*), some are missed by the monitor, with probability (1-*r*), and are reported as target errors (thus total probability for target errors = *s**(1-*r*)). The monitor 'repairs' the rest with probability *r*, yielding a total proportion of *s*r* repaired utterances. We assume that these repairs produce either correct utterances (*(s*r)/2*) or competing errors (*(s*r)/2*) with equal[4] probability (thus yielding total probabilities for correct utterances = (1-*s*)+(*s***r*)/2 and competing errors = (*s***r*)/2).

---

[4] For our purposes, the precise ratio of repaired-as-correct to repaired-as-competing-errors does not matter. This ratio must only be constant across similarity conditions.

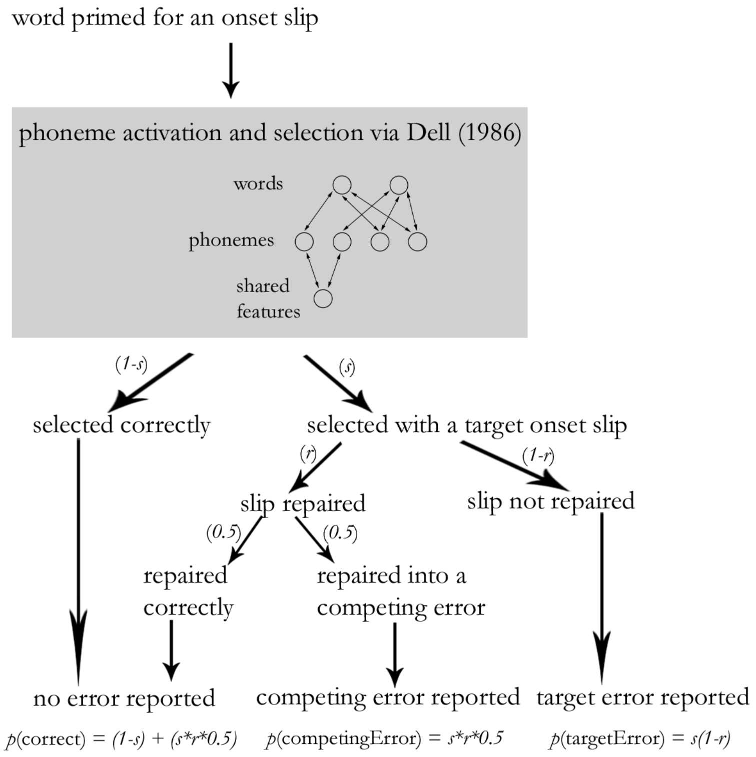

Figure 4. Flow chart model of the role of monitoring and repair processes in creating reported distributions of target errors, competing errors, and non-error trials. Figures in parentheses give the probability of taking each tree branch. Equations at the bottom specify the expected proportions of each outcome, in terms of the *s* and *r* parameters.

With our observed outcomes in each condition (from Table 4), we can then solve the equations in Figure 4, thereby estimating the proportion of target slips initially generated (*s*), and the proportion of those slips that are repaired into other responses (*r*). The direct generation account of lexical bias predicts that more word-outcome slips are initially generated, corresponding to larger *s* values in the word than nonword condition. The biased repair account predicts that fewer word-outcome slips are repaired, corresponding to *smaller r* values in the word than nonword condition. The resulting estimates are listed in Table 5.

Table 5. Estimates of target slip generation and repair rates, based on the equations given in Figure 4 and the data in Table 4.

	Mouthed				**Unmouthed**			
	Similar		**Dissimilar**		**Similar**		**Dissimilar**	
Estimate	**Word**	**Non-word**	**Word**	**Non-word**	**Word**	**Non-word**	**Word**	**Non-word**
Target slip generation rate (s)	5.4%	4.0%	3.7%	2.9%	3.7%	3.7%	3.6%	2.9%
Error repair rate (r)	16.9%	45.5%	24.4%	50.0%	14.6%	51.3%	40.0%	38.7%

The *s*-values, indexing the target *slip* generation rates, show that, in three of the four conditions, more slips are generated in the word-outcome than in the nonword-outcome condition. This pattern therefore supports a direct generation account of the lexical bias effect.

In addition, though, the *r*-values, indexing the slip *repair* rates, tend to support the biased repair account of lexical bias. They generally show a reverse lexical bias (with the unmouthed dissimilar condition as an exception), suggesting that nonword errors are more likely to be repaired. This result emerges because the competing errors are generally more numerous in the nonword conditions, which is consistent with the conclusion that some degree of lexical bias arises because nonword outcomes are more likely to be altered. Thus, our analysis indicates that error generation and repair biases both contribute to the observed lexical bias effects. This conclusion is quite in line with recent studies of overt speech errors that have concluded that interactive feedback, which generates lexical bias directly, and a monitoring/repair system that is sensitive to lexical status are jointly responsible for the observed lexical bias effect (Hartsuiker et al., 2005; Nooteboom & Quené, 2008). We should acknowledge, though, that this analysis is only suggestive. There are too few competing errors for robust inferential statistical testing.

Generation versus repair accounts of the phonemic similarity effect

We can apply this same model to our phonemic similarity data, to distinguish between the direct generation and biased repair accounts of these effects. As with lexical bias, the direct generation account holds that similar-phoneme slips outnumber dissimilar-phoneme slips because more similar-phoneme slips are directly generated by the phonological encoding process. For instance, similarly articulated phonemes are thought to share lower-lever representations (*i.e.* articulatory features) that spread activation back to connected phonemes, thus biasing phoneme selection (*e.g.* Dell, 1986; Oppenheim & Dell, 2008). In contrast, the biased repair account holds that potential errors involving dissimilar phonemes are more likely to be caught and repaired than potential similar-phoneme errors because they are easier for the monitor to 'hear' (*e.g.* Nooteboom, 2005; *cf.* Hartsuiker, 2006). Thus phonemic similarity effects could also result from either the direct generation of more similar-phoneme slips, or the biased repair of more dissimilar-phoneme slips.

As before, we test these accounts by plugging our observed outcomes in each condition (from Table 3) into the equations in Figure 4, and solving for *s* (the proportion of target slips initially generated) and *r* (the proportion of target errors that are repaired into other responses). The resulting estimates are listed in Table 6.

Table 6. Estimates of target slip generation and repair rates, based on the equations given in Figure 4 and the data in Table 3.

	Mouthed		**Unmouthed**	
Estimate	**Similar**	**Dissimilar**	**Similar**	**Dissimilar**
Target slip generation rate (s)	8.9%	6.8%	5.9%	5.8%
Error repair rate (r)	38.8%	39.5%	43.6%	41.5%

The slip generation rates (*s*) show the same articulation-by-phonemic similarity interaction that we reported for target errors, thus supporting the direct generation account of the phonemic similarity effect. But the *r*-values show very little difference across similarity conditions, suggesting that biased repair contributes little, if anything, to our observed similarity effect in the mouthed condition or the lack of an effect in the unmouthed condition. These results obtain, in large part, because competing errors show the same phonemic similarity effect as target slips, rather than complementary effect predicted by the biased repair account. Therefore, we can conclude that while a biased repair process may contribute to the lexical bias for phonological errors (Table 3), it is not noticeably biased with respect to phonemic similarity and contributes exceedingly little to our observed similarity effect and its interaction with mouthing.

General Discussion

Inner speech as verbal imagery

Our findings reveal three things about the nature of cognitive representations involved in inner speech. First, inner speech involves sublexical form-based representations such as phonemes. This point is clear from reports of phonological errors, and nonword errors in particular, which should not occur in substantial quantities without the use of phonological representations. Second, inner speech is, nevertheless, abstracted. It does not necessarily involve more fine-grained articulatory information to the same extent as normal overt speech. This point is supported by Oppenheim and Dell's (2008) report that phonemic similarity only affected slip rates in overt speech, and is reinforced by our current similarity by mouthing interaction for target errors – both suggest that fine-grained (*e.g.* featural) information figures less prominently in unarticulated inner speech. Finally, inner speech is a flexible form of imagery. It incorporates articulatory information when available, but when this information is not available, it operates on a more abstract level. This point is most clearly supported by our demonstration that phoneme similarity effects in inner speech are stronger when inner speech is silently articulated.

Inner speech as an embodied cognitive process

A central tenet of the embodiment hypothesis, as it is commonly specified (*e.g.* Barsalou, 1999; Wilson, 2002), asserts that offline cognition is based in sensorimotor processes. For instance, understanding the concept 'leaf' should entail sensorimotor simulation of visual, auditory, tactile, olfactory, and maybe even gustatory experiences with leaves, as well as consideration of the interactions that leaves afford. Similarly, imagining the word 'leaf' should necessarily entail a sensorimotor simulation of articulating the word, and feeling and hearing oneself articulate it (*e.g.* Wilson, 2001). If this were a complete, un-degraded simulation, then one should expect a high degree of similarity between unarticulated inner-speech and overt-speech phenomena and phenomenology.

And some similarity is clearly indicated by the data. The very fact that unarticulated inner speech has "slips" at all tells us that the real-time constraints on production that lead to overt slips operate when speech is internally generated. More than that, inner slips are, at least to a first approximation quite similar to overt slips. Dell (1978), for example, found that exactly the same errors occurred when tongue twisters were spoken or imaged. Postma and Noordanus (1996) showed similar distributions of error types reported across various speech modalities. Oppenheim and Dell (2008) and our present experiment showed that slips in overt speech and two kinds of inner speech exhibit the lexical bias effect. Moreover, our analysis of the role of monitoring/repair processes in creating lexical bias in inner speech led to a similar conclusion about that role that studies of overt speech errors did (Nooteboom & Quené, 2008). Thus, in many ways, unarticulated inner speech does seem to parallel overt speech.

However, our findings add to a growing body of literature that reveals where the parallels between inner and overt speech break down. For instance, if speakers repeatedly say a word such as “life” aloud, it will likely perceptually flip to “fly”. This verbal transformation effect does not occur in inner speech, if participants clamp their mouths shut or perform other tasks like chewing (Reisberg, Smith, Baxter, & Sonenshine, 1989). In addition, practicing tongue-twisters aloud improves performance in unarticulated inner speech, but the reverse does not hold true (Dell & Repka, 1992). Finally, in a previous study (Oppenheim & Dell, 2008), we demonstrated that the phoneme similarity effect for phonological errors characteristic of overt speech was significantly weaker in unarticulated inner speech. And our current findings show that this divergence can be attributed to the additional motor planning required for articulation.

We can therefore conclude that offline cognition is flexible in terms of the levels at which mental simulations are run. When inner speech proceeds normally – without articulation, but also without any other task occupying the articulators – the resulting verbal imagery shows little or no sign of simulating low-level motor movements. Adding articulation changes inner speech by integrating articulatory information. Thus, while inner speech – that is, verbal imagery – can reflect enacted low-level motor planning and sensory monitoring, this information is not essential to the phenomenon.

This flexibility is an important qualification to embodied theories of cognition. Verbal imagery, as a form of offline cognition, chiefly involves abstract representations, but may incorporate low-level sensorimotor information if available. In the absence of such low-level information, offline cognition may continue to involve simulations (as posited by embodied theories), but these simulations invoke only abstracted representations of sensorimotor processes.

To conclude, we note that these abstracted simulations should not be viewed as poor substitutes for the real thing. Unarticulated inner speech well represents the higher linguistic processing levels, as shown by the fact that errors in it exhibit lexical bias. This kind of verbal imagery should be quite effective for cognitive tasks in which the processing of meaning or lexical information is what is relevant (*e.g.* reading text, remembering a telephone number, rehearsing how you will begin a lecture). It is only when the purpose of inner speech is to actively represent or practice the articulatory details that the unarticulated inner speech will fail to do the job. Fortunately, the production system is sufficiently flexible that those details can be added without actual speaking aloud—as in our mouthed speech condition. Then, as we found, the resulting imagery well simulates real audible articulation.

References

Abramson, M., & Goldinger, S. D. (1997). What the reader's eye tells the mind's ear: Silent reading activates inner speech. *Perception & Psychophysics, 59*(7), 1059-1068.

Aziz-Zadeh, L., Cattaneo, L., Rochat, M., & Rizzolatti, G. (2005). Covert speech arrest induced by rTMS over both motor and nonmotor left hemisphere frontal sites. *Journal of Cognitive Neuroscience, 17*(6), 928-938.

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology, 16*(18), 1818-1823.

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*(4), 382-391.

Baddeley, A., & Wilson, B. (1985). Phonological coding and short-term memory in patients without speech. *Journal of Memory and Language, 24*(4), 490-502.

Barch, D. M., Sabb, F. W., Carter, C. S., Braver, T. S., Noll, D. C., & Cohen, J. D. (1999). Overt verbal responding during fMRI scanning: Empirical investigations of problems and potential solutions. *NeuroImage, 10*(6), 642-657.

Barsalou, L. W. (1999). Perceptual symbol systems. *The Behavioral and Brain Sciences, 22*(4), 577-609; discussion 610-60.

Brocklehurst, P. & Corley, M. (2009). Lexical bias and the phonemic similarity effect in inner speech. Paper presented at the *15th Annual Conference on Architectures and Mechanisms for Language Processing,* 7-9 September 2009, Barcelona.

Caplan, D., Rochon, E., & Waters, G. S. (1992). Articulatory and phonological determinants of word length effects in span tasks. *The Quarterly Journal of Experimental Psychology.A, Human Experimental Psychology, 45*(2), 177-192.

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

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283-321.

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.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5*(4), 313.

Foroni, F., & Semin, G. R. (2009). Language that puts you in touch with your bodily feelings: The multimodal responsiveness of affective expressions. *Psychological Science, 20*(8), 974-980.

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review, 13*(3), 361-377.

Garnsey, S. M., & Dell, G. S. (1984). Some neurolinguistic implications of prearticulatory editing in production. *Brain and Language, 23*(1), 64-73.

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

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*(3), 558-565.

Goldrick, M. (2004). Phonological features and phonotactic constraints in speech production. *Journal of Memory and Language, 51*(4), 586-603.

Goldrick, M., & Rapp, B. (2007). Lexical and post-lexical phonological representations in spoken production. *Cognition, 102*(2), 219-260.

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*(4), 949-953.

Hartsuiker, R. J. (2006). Are speech error patterns affected by a monitoring bias? *Language and Cognitive Processes, 21*, 856-891(36).

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 et al. (1975). *Journal of Memory and Language, 52*(1), 58-70.

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

Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology, 25*(4), 463-492.

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

Lakoff, G. (1987). Women, fire, and dangerous things : What categories reveal about the mind. Chicago: University of Chicago Press.

Levelt, W. J. M. (1989). *Speaking : From intention to articulation*. Cambridge, Mass.: 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*(01), 1.

Levelt, W. M. (1983). Monitoring and self-repair in speech. *Cognition, 14*(1), 41-104.

Liberman, A. M., Delattre, P., & Cooper, F. S. (1952). The role of selected stimulus-variables in the perception of the unvoiced stop consonants. *The American Journal of Psychology, 65*(4), 497-516.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21*(1), 1-36.

MacKay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia, 8*(3), 323-350.

MacKay, D. G. (1981). The problem of rehearsal or mental practice. *Journal of Motor Behavior, 13*(4), 274-285.

MacKay, D. G. (1992). Constraints on theories of inner speech. In D. Reisberg (Ed.), *Auditory imagery* (pp. 121-149). Hillsdale, N.J.: L. Erlbaum Associates.

Macken, W. J., & Jones, D. M. (1995). Functional characteristics of the inner voice and the inner ear: Single or double agency? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(2), 436-448.

McGuire, P. K., Silbersweig, D. A., Murray, R. M., David, A. S., Frackowiak, R. S., & Frith, C. D. (1996). Functional anatomy of inner speech and auditory verbal imagery. *Psychological Medicine, 26*(1), 29-38.

Nooteboom, S., & Quené, H. (2008). Self-monitoring and feedback: A new attempt to find the main cause of lexical bias in phonological speech errors. *Journal of Memory and Language, 58*(3), 837-861.

Nozari, N., & Dell, G. S. (2009). More on lexical bias: How efficient can a "lexical editor" be? *Journal of Memory and Language, 60*(2), 291-307.

Oppenheim, G. M., & Dell, G. S. (2008). Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition, 106*(1), 528-537.

Palmer, E. D., Rosen, H. J., Ojemann, J. G., Buckner, R. L., Kelley, W. M., & Petersen, S. E. (2001). An event-related fMRI study of overt and covert word stem completion. *NeuroImage, 14*(1 Pt 1), 182-193.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature, 362*(6418), 342-345.

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.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews.Neuroscience, 6*(7), 576-582.

Reisberg, D., Smith, J. D., Baxter, D. A., & Sonenshine, M. (1989). "Enacted" auditory images are ambiguous; "pure" auditory images are not. *The Quarterly Journal of Experimental Psychology.A, Human Experimental Psychology, 41*(3), 619-641.

Shergill, S. S., Bullmore, E. T., Brammer, M. J., Williams, S. C., Murray, R. M., & McGuire, P. K. (2001). A functional study of auditory verbal imagery. *Psychological Medicine, 31*(2), 241-253.

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

Shuster, L. I., & Lemieux, S. K. (2005). An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. *Brain and Language, 93*(1), 20-31.

Smith, J. D., Reisberg, D., & Wilson, M. (1992). Subvocalization and auditory imagery: Interactions between the inner ear and the inner voice. In D. Reisberg (Ed.), *Auditory imagery* (pp. 95-119). Hillsdale, N.J.: L. Erlbaum Associates.

Stefan, F. (1996). *Similarity and frequency in phonology* . Ph. D. dissertation, Northwestern University.

Stemberger, J. P. (1991). Apparent anti-frequency effects in language production: The addition bias and phonological underspecification. *Journal of Memory and Language, 20*(2), 161-185.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America, 94*(26), 14792-14797.

Vallar, G., & Cappa, S. F. (1987). Articulation and verbal short-term memory: Evidence from anarthria. *Cognitive Neuropsychology, 4*(1), 55.

Vygotskiĭ, L. S., Hanfmann, E., & Vakar, G. (1965). *Thought and language* [Myshlenie i rech′.]. Cambridge, Mass.: M.I.T. Press.

Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language, 34*(3), 311-334.

Wilson, M. (2001). The case for sensorimotor coding in working memory. *Psychonomic Bulletin & Review, 8*(1), 44-57.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625-636.

Yetkin, F. Z., Hammeke, T. A., Swanson, S. J., Morris, G. L., Mueller, W. M., McAuliffe, T. L., et al. (1995). A comparison of functional MR activation patterns during silent and audible language tasks. *American Journal of Neuroradiology, 16*(5), 1087-1092.