Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Memory and Language 54 (2006) 515–540
Journal of Memory and Language
www.elsevier.com/locate/jml

Halting in single word production: A test of the perceptual loop theory of speech monitoring ☆

L. Robert Slevc *, Victor S. Ferreira

Department of Psychology, University of California, San Diego, La Jolla, CA 92093-0109, USA

Received 20 February 2005; revision received 26 October 2005
Available online 24 January 2006

Abstract

The *perceptual loop theory* of speech monitoring (Levelt, 1983) claims that inner and overt speech are monitored by the comprehension system, which detects errors by comparing the comprehension of formulated utterances to originally intended utterances. To test the perceptual loop monitor, speakers named pictures and sometimes attempted to halt speech in response to auditory (Experiments 1 and 3) or visual (Experiments 2, 4, and 5) words that differed from the picture name. These *stop-signal* words were varied in terms of their semantic or phonological similarity to the intended word. The ability to halt word production was sensitive to phonological similarity and, in Experiment 5, to emotional valence, but not to semantic similarity. These results suggest that the perceptual loop detects errors by making comparisons at a level where phonological knowledge is represented. These data also imply that dialogue, back channeling, and other areas where speech production is affected by simultaneous comprehension may operate based on phonological comparisons.

Keywords: Self monitoring; Perceptual loop theory; Speech errors

One of the more striking features of language production is its efficiency and accuracy. Levelt (1989) estimates that we produce about 150 words per minute, but make only one lexical error per 1000 words. This is especially impressive considering the complexities of word production. When speakers produce words, they start with an idea they wish to communicate, then must retrieve both lexical and phonological information, and finally program a set of motor movements that can then be comprehended by listeners. Despite these complexities, speech production seems relatively effortless and error-free.

One reason that errors are relatively infrequent may be that speakers comprehend their own speech to inspect it for errors, thereby allowing them to inhibit and repair erroneous utterances and speak relatively fluently. It is not unusual for speakers to stop and correct themselves

☆ Portions of this work were presented at the Architectures and Mechanisms for Language Processing conference in September, 2002. This research was supported by National Institute of Health Grant R01 MH-64733. We thank Wind Cowles, Gary Dell, Zenzi Griffin, and the members of the psycholinguistics group at UCSD for helpful discussions, Carla Firato, Cassie Gipson, Ana Molina, and Mary Tibbs for assistance collecting and transcribing data, and Jeff Bowers, Rob Hartsuiker, Sieb Nooteboom, and an anonymous reviewer for their comments on an earlier version of this paper.
* Corresponding author.
E-mail address: slevc@psy.ucsd.edu (L.R. Slevc).

0749-596X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2005.11.002

when they make an error, sometimes even before the error is externally apparent. This idea has been formalized as the *speech monitor*.

A number of findings in the speech-error record have been used as evidence for an inner speech monitor, such as the fact that phonological speech errors are more likely to result in real words than non-words (the *lexical bias effect*). Baars, Motley, and MacKay (1975) provided experimental evidence of this effect by using a procedure to elicit Spoonerisms (exchanges of the initial sounds of a pair of words), and found that exchanges that formed other words (e.g., "darn bore" from "barn door") were more likely than exchanges that formed non-words (e.g., "dorn bef" from "born deaf"). Motley, Camden, and Baars (1982) used the same procedure to show that exchanges that resulted in taboo words (e.g., making a Spoonerism out of "hit shed") were made less often than exchanges that did not result in taboo words (the *taboo-words effect*). Baars and colleagues proposed that these effects could result from use of an inner monitor or editor that is sensitive to both lexical status and social appropriateness, and thus would be more likely to detect and prevent articulation of an error resulting in a non-word or a taboo word than one that results in a real word or a more appropriate word. Further supporting this account, subjects showed elevated galvanic skin responses on trials where they avoided making taboo-word errors (relative to cases where they avoided making errors that would not result in taboo words), suggesting that the taboo errors were, in fact, generated internally even when they were not overtly produced.

The speech monitor has been proposed to be sensitive to more than just lexical status and social appropriateness—it has been claimed to be sensitive to a wide variety of errors, including conceptual errors, syntactic errors, lexical errors, phonemic errors, prosodic errors, morphemic errors, errors in appropriateness of speech, and errors relating to social context (Levelt, 1989; Postma, 2000). In fact, the inner-speech monitor has been proposed to be sensitive to nearly everything to which listeners are sensitive, leading to the relatively natural theory that the monitor uses the comprehension system to listen to both inner and outer speech (Hartsuiker & Kolk, 2001; Levelt, 1983, 1989; Postma, 2000). This proposal is known as the *perceptual loop theory* of speech monitoring, and claims that a speaker's prearticulatory output (or phonetic plan) is processed by the language comprehension system, which allows the speaker to compare the comprehension of what he or she is about to say to what he or she originally intended to express. Speakers are also hypothesized to listen to their own overt speech, giving them another chance to catch errors through the same mechanism (and although it is too late to prevent errors at that point, they can still be corrected). Fig. 1 depicts the perceptual loop theory, as proposed by Levelt (1983, 1989).

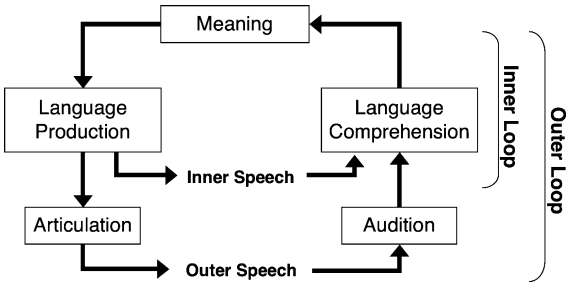

Fig. 1. The perceptual-loop theory of self-monitoring (Levelt, 1983, 1989).

A number of results support the claim that speakers monitor both their inner and outer speech via self-comprehension. Speakers detect similar kinds of errors in their silent speech and in their overt speech (Dell & Repka, 1992), suggesting that the same mechanism is used for both internal and external error detection. Speakers' error rates are similar when producing silent, noise-masked, or mouthed speech (Postma & Noordanus, 1996), suggesting that this monitoring does not occur at the motor level. And not only can speakers detect many of their own errors when their overt speech is noise masked, but they actually do so more quickly (on average) than when they can hear their own speech (Lackner & Tuller, 1979), suggesting that monitoring of inner speech is faster than monitoring of external speech. Speakers also show evidence of capacity restrictions on error detection and correction (Oomen & Postma, 2002; Postma, 1997), suggesting that speech monitoring is a centrally regulated (or controlled) process, which is consistent with the perceptual-loop theory (Postma, 2000). Finally, there is at least some evidence of a link between disordered speech comprehension and monitoring deficits (e.g., Marshall & Tompkins, 1982), which fits with the idea that monitoring is carried out through the comprehension system.

There is, however, some evidence that is inconsistent with the idea that monitoring relies on the perception of self-produced speech. One type of discrepant evidence comes from language-impaired populations. If monitoring relies on the comprehension system, one would expect patients with neurological impairments in comprehension to have difficulty monitoring, and patients with impairments in monitoring (such as jargon aphasics, who often are unaware of their speech errors despite suffering from severe anomia) to have corresponding problems in comprehension. However, a number of studies have found dissociations in these processes in neurologically damaged patients (e.g., Schlenk, Huber, & Wilmes, 1987; McNamara, Obler, Au, Durso, & Albert, 1992; Maher, Rothi, & Heilman, 1994). Hartsuiker and Kolk (2001) point out that these findings are not necessarily evidence against the perceptual loop