Phonological Encoding and Monitoring in Normal and Pathological Speech

Edited by
Robert J. Hartsuiker, Roelien Bastiaanse,
Albert Postma and Frank Wijnen

Psychology Press
Taylor & Francis Group
HOVE AND NEW YORK

First published 2005
by Psychology Press
27 Church Road, Hove, East Sussex BN3 2FA

Simultaneously published in the USA and Canada
by Psychology Press
270 Madison Avenue, New York NY 10016

Psychology Press is part of the Taylor & Francis Group

Typeset in Times by RefineCatch Limited, Bungay, Suffolk
Printed and bound in Great Britain by TJ International Ltd,
Padstow, Cornwall
Cover design by Anú Design

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
Hartsuiker, Robert J., 1968–
Phonological encoding and monitoring in normal and pathological speech / Robert J. Hartsuiker . . . [et al.].
p. cm.
Includes bibliographical references and index.
ISBN 1-84169-262-X
1. Speech disorders. 2. Linguistics. I. Title.

RC423.H345 2005
616.85′5–dc22 2004012990

ISBN 1-84169-262-X

10 Listening to oneself: Monitoring speech production

Sieb G. Nooteboom

Abstract

According to Levelt (1989) and Levelt, Roelofs, and Meyer (1999) (a) self-monitoring of speech production employs the speech comprehension system, (b) on the phonological level the speech comprehension system has no information about the lemmas and forms chosen in production, and (c) lexical bias in speech errors stems from the same perception-based monitoring that is responsible for detection and overt correction of speech errors. It is predicted from these theoretical considerations that phonological errors accidentally leading to real words should be treated by the monitor as lexical errors, because the monitor has no way of knowing that they are not. It is also predicted that self-corrections of overt speech errors are also sensitive to lexicality of the errors. These predictions are tested against a corpus of speech errors and their corrections in Dutch. It is shown that the monitor treats phonological errors leading to real words in all respects as other phonological, and not as lexical errors and that no criterion is applied of the form "is this a real word?" It is also shown that, whereas there is considerable lexical bias in spontaneous speech errors and this effect is sensitive to phonetic similarity, self-corrections of overt speech errors are not sensitive to lexical status or phonetic similarity. It is argued here that the monitor has access to the intended word forms and that lexical bias and self-corrections of overt speech errors are not caused by the same perception-based self-monitoring system. Possibly fast and hidden self-monitoring of inner speech differs from slower and overt self-monitoring of overt speech.

Introduction: Levelt's model of speech production and self-monitoring

We all make errors when we speak. When I intend to say "good beer" it may come out as "bood beer" or even as "bood gear"; or when I want to say "put the bread on the table" I may inadvertently turn it into "put the table on the table" or into "put the table on the bread". Let us call errors like "bood beer" or "bood gear", where phonemes are misplaced, phonological errors, and ones like "table on the table" or "table on the bread", where meaningful items show up in the wrong positions, lexical errors. Lexical errors supposedly arise during grammatical encoding, phonological ones during phonological encoding (Levelt, 1989). Errors as given in our examples are syntagmatic speech errors, involving two elements in the intended utterance, a source and a target, the source being the intended position of an element, the target being the position where it ends up. So in the intended utterance "bread on the table", underlying the error "table on the table", "table" is the source and "bread" the target. Speakers also make paradigmatic speech errors, involving only a single intruding element, but here I will only be concerned with syntagmatic speech errors (cf. Fromkin, 1973).

The fact that we know that speech errors exist implies that we can detect them. And we not only detect errors in the speech of others, but also in our own speech. In the collection used for the current study, roughly 50 per cent of all speech errors were detected and corrected by the speakers (an earlier analysis of Meringer's, 1908, corpus suggested somewhat higher values; Nooteboom, 1980). Apparently, part of a speaker's mind is paying attention to the speech being produced by another part of the same mind, keeping an ear out for inadvertent errors that may be in need of correction. Let us call this part of the speaking mind the "monitor", and its function "self-monitoring" (Levelt, 1983, 1989). The general question I am focussing on here is: "How is self-monitoring of speech organized, and what information does it operate on?" The question is not new. A firm stand on this issue, based on extensive empirical evidence, has been for example taken by Levelt (1989), and by Levelt et al. (1999). The reason to take their theory as a starting point is that it is the most constrained, most parsimonious, theory of speech production available. In many ways it predicts what it should and does not predict what it should not. Alternative theories will be mentioned in the discussion section.

For the present purposes the following properties of the spreading-activation theory proposed by Levelt and his associates are relevant: (1) Speech production is strictly serial and feedforward only, implying that there is no cascading activation and no immediate feedback from the level of phonological encoding to the level of grammatical encoding; (2) self-monitoring employs the speech comprehension system, also used in listening to the speech of others; (3) the speech being produced reaches the comprehension system via two different routes, the inner route feeding a covert form of not-yet-articulated speech into the speech-comprehension system, and the auditory route feeding overt speech into the ears of the speaker/listener; (4) on the phonological level there is no specific information on intended phonological forms leaking to the speech comprehension system. The monitor must make do with a general criterion of the form "is this a real word?" instead of a criterion such as "is this the word I wanted to say?"; (5) lexical bias in speech errors is caused by the same perception-based self-monitoring system that is responsible for the detection and correction of overt speech errors.

This theory leads to some predictions that can be tested by looking at properties of speech errors in spontaneous speech and their corrections. The following predictions are up for testing:

- The monitor treats phonological errors that lead to real words, such as "gear" for "beer", as lexical errors.
- If spontaneous phonological speech errors show lexical bias, as has been suggested by Dell (1986), then one should also find a lexical bias effect in self-corrections of overt speech errors.

Before testing the first prediction, it should be assessed that so-called real-word phonological errors are indeed caused during phonological and not during grammatical encoding. This question will be dealt with first. Also, it will appear below that there may be a problem in testing the first prediction, caused by the fact that many overtly corrected anticipations, such as "Yew . . . New York", may not be anticipations at all, but rather halfway-corrected transpositions. If so, there is no way of telling whether the error triggering the monitor was the real word "Yew" or the non-word "Nork" (cf. Cutler, 1982; Nooteboom, 1980). The question is whether or not this observation potentially invalidates the interpretation of a comparison between correction frequencies of phonological non-word errors, phonological real-word errors and lexical errors. It will be shown that it does. To circumvent this problem, a separate analysis will be made in which non-word and real-word phonological errors are limited to perseverations, such as "good gear" instead of "good beer", because there no part of the error can hide in inner speech. With respect to the prediction concerning lexical bias, it should be noted that reports on the existence of lexical bias in spontaneous speech errors differ. Garrett (1976) did not find evidence for lexical bias, Dell (1986) did, but Del Viso, Igoa, and Garcia-Albea (1991) did not for Spanish, although using a measure for lexical bias that is very similar to Dell's. So before studying lexical bias in self-corrections of overt speech errors, it should be assessed that there really is lexical bias in spontaneous speech errors. As will be seen, there is ample evidence for lexical bias in Dutch spontaneous speech errors. Therefore it makes sense to ask whether or not there is lexical bias in self-corrections in overt speech errors, as predicted from Levelt's theory. The reader will see that there is not. A related question is whether lexical bias is sensitive to phonetic distance between target and error phoneme, as predicted from perception-based monitoring but also from production-based theories, and if so whether the same is true for the probability of self-corrections of overt speech errors. Finally, there is the question whether the structure of the current data rather stems from a collector's bias than from the mechanisms underlying the production and perception of speech.

The following questions will now be dealt with in succession:

- Are alleged real-word phonological errors actually made during phonological or grammatical encoding?
- Does the fact that alleged corrected anticipations might sometimes have been halfway-corrected transpositions hinder the interpretation of comparisons between correction frequencies for non-word and real-word errors?
- Does the monitor treat phonological errors that lead to real words, such as "gear" for "beer", as lexical or as phonological errors?
- Do spontaneous phonological speech errors show lexical bias?
- Do self-corrections of overt speech errors show lexical bias?
- Are lexical bias and probability of self-corrections of overt speech errors equally sensitive to phonetic distance between target and error?
- Do the current data suffer from a collector's bias invalidating otherwise plausible conclusions?

Several possible explanations of the current findings will be discussed in the final section of this chapter.

The corpus

To answer the above questions two different collections of spontaneous speech errors in Dutch were used, the first collection only being used in studying lexical bias, because for these speech errors no overt self-corrections were available.

The oldest collection (AC/SN corpus) is basically the same as the one described by Nooteboom (1969). The errors were collected and noted down in orthography during several years of collecting by two people, the late Anthony Cohen and myself. Unfortunately, corrections were not systematically noted down. Collection of errors continued some time after 1969, and in its present form the collection contains some 1000 speech errors of various types, phonological syntagmatic errors outnumbering other types, such as lexical syntagmatic errors, blends, and intrusion errors. The collection was never put into a digital database and is only available in typed form, each error on a separate card. Selection of particular types of errors for the present purpose was done by hand.

The second collection (Utrecht corpus) stems from efforts of staff members of the Phonetics Department of Utrecht University, who, on the initiative of Anthony Cohen, from 1977 to 1982 orthographically noted down all speech errors heard in their environment, with their corrections, if any (cf. Schelvis, 1985). The collection contains some 2500 errors of various types, of which more than 1100 are phonological syntagmatic errors and some 185 lexical syntagmatic errors. The collection was put into a digital database, currently accessible with Microsoft Access.

Are alleged real-word phonological errors actually made during phonological or grammatical encoding?

Before making any comparisons between non-word phonological errors, real-word phonological errors and lexical errors, we have to make sure that in production alleged real-word phonological errors really arise at the level of phonological encoding and not at the level of grammatical encoding. In Table 10.1 we see confusion matrices for source and target of phonological non-word errors, phonological real-word errors and lexical errors.

These data show that in lexical errors an open-class word is never replaced by a closed-class word and a closed-class word never by an open-class word. In fact, closer analysis shows that syntactic word class is nearly always preserved (cf. Nooteboom, 1969). This is quite different for non-word phonological errors where the distribution of word-class preservation and violation is entirely predictable from relative frequencies and chance. So how do our alleged phonological real-word errors behave? Obviously they behave like non-word phonological errors, not like lexical errors. So we can be reassured that in the bulk of such errors lexical status is purely accidental. Now we are in a better position to ask whether the monitor treats real-word phonological errors as lexical errors, as predicted by Levelt et al., or rather as phonological errors. But first there is this problem with corrected anticipations perhaps being misclassified transpositions.

Corrected anticipations or halfway-corrected transpositions?

It has been observed that relatively many corrected anticipations in collections of speech errors, such as: "Yew. . . . New York", may be misclassified halfway-corrected transpositions (Cutler, 1982; Nooteboom, 1980). If we assume that speech errors can be detected in inner speech before becoming overt, in all these cases the monitor has not one but two opportunities to detect an error, and for all we know the second, hidden, part of the transposition may have been a non-word, as in the current example. This state of affairs potentially upsets any statistical differences we find in a comparison

Table 10.1 Three confusion matrices for source and target being, or belonging to, a closed- versus an open-class word, separately for phonological non-word errors, phonological real-word errors, and lexical errors

	Phonological non-word errors		*Phonological real-word errors*		*Lexical errors*	
Source	Open class	Closed class	Open class	Closed class	Open class	Closed class
Target						
Open class	303	55	169	30	135	0
Closed class	58	21	25	10	0	24

Table 10.2 Numbers of corrected and uncorrected speech errors, separately for perseverations, anticipations, and transpositions

	Perseverations	*Anticipations*	*Transpositions*	*Total*
Corrected	103	*442(?)*	*42(?)*	587
Not corrected	153	238	175	566
Total	256	*680(?)*	*217(?)*	1153

Note: There is a strong interaction between error class and correction frequency ($chi^2 = 153$; $df = 2$; $p < 0.001$). Cursive numbers are suspected not to correspond to what happened in inner speech. Utrecht corpus only.

between lexical and phonological real-word errors. That this is a serious threat may be shown by the following estimates of the relative numbers of anticipations and transpositions in inner speech. Let us assume that the probability of detecting an error in internal speech is not different for anticipations and perseverations (to the extent that this assumption is incorrect the following calculations will be inaccurate; but if the underlying reasoning is basically sound, they will at least provide a plausible rough estimate). We know the number of uncorrected perseverations, the total number of perseverations, and the number of uncorrected anticipations (Table 10.2). From the numbers in Table 10.2, using an equation with one unknown, one can easily calculate what the total number of anticipations, and therefore also the number of corrected anticipations, would have been, without the influx of halfway-corrected transpositions. The equation runs as follows:

103 corrected perseverations : 153 not corrected perseverations,
= ? corrected anticipations : 238 not corrected anticipations

The estimate number of corrected anticipations would then be:

$$(103 \times 238) : 153 = 160$$

The total number of anticipations would be 160 + 238 = 398. The estimate number of misclassified halfway-corrected transpositions is 442 – 160 = 282. Note that this brings the total number of transpositions in internal speech to 282 + 42 + 175 = 499 instead of 217, making transpositions by far the most frequent class of speech errors (Table 10.3). These estimates are further confirmed in the following way: The probability of remaining uncorrected is 0.6 for both perseverations and anticipations. A transposition contains an anticipation plus a perseveration. The probability of remaining uncorrected should therefore be $0.6 \times 0.6 = 0.36$. The new estimate of the fraction of transpositions remaining uncorrected equals:

$$1 - (42 + 282) : 499 = 0.35$$

Table 10.3 Numbers of corrected and uncorrected speech errors in inner speech, separately for perseverations, anticipations, and transpositions

	Perseverations	*Anticipations*	*Transpositions*	*Total*
Corrected	103	*160*	*324*	587
Not corrected	153	238	175	566
Total	256	*398*	*499*	1153

Note: Cursive numbers are estimated (see text). Utrecht corpus only.

Apparently, still provided that our assumption that the probability of being detected in internal speech is the same for perseverations and anticipations was correct, both parts of the error contribute equally and independently to the probability of remaining uncorrected.

From these calculations, it is at least plausible that a great many corrected anticipations in our corpus originated as halfway-corrected transpositions in inner speech. Of course we have no way of knowing which are and which are not. In all such cases in which the error is phonological we do not know whether the error triggering the monitor was a real word or a non-word. We therefore should treat any comparison between numbers of correction for real-word and non-word anticipations with caution.

Does the monitor treat phonological errors that lead to real words as lexical or phonological?

We know that lexical errors and phonological errors are treated differently by the monitor: Both the distribution of the number of words a speaker goes on speaking before stopping to correct a speech error and the distribution of the number of words a speaker retraces in his correction is different for lexical and phonological errors (Nooteboom, 1980). Our corpus of speech errors noted down with their corrections, makes it possible to compare the distributions of the number of words spoken before the speaker stops for correction, and the number of words included in the correction, between different classes of speech errors. If Levelt et al. are right in assuming that the monitor has no way of knowing whether a particular error was made during grammatical or during phonological encoding, these distributions should be different for non-word and real-word phonological errors, and the same for real-word phonological and lexical errors.

Contrary to this prediction, Figure 10.1. suggests that the distribution of the numbers of words spoken before stopping is very similar for non-word and real-word phonological errors and rather different for real-word phonological errors and lexical errors.

To test these predictions statistically, for the moment neglecting the threat stemming from corrected anticipations being halfway-corrected transpositions, the numbers underlying Figure 10.1 were collapsed into a 2 × 2 matrix

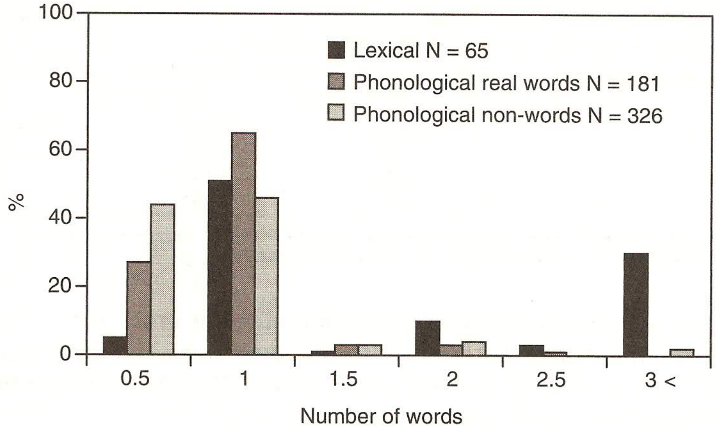

Figure 10.1 Percentage of speech errors as a function of the number of words spoken before stopping for correcting a speech error, plotted separately for lexical errors, phonological errors leading to non-words, and phonological errors accidentally leading to real words.

Table 10.4 Numbers of speech errors as a function of number of words spoken before stopping to correct a speech error, separately for non-word phonological errors, real-word phonological errors, and lexical errors

n	*1 or less*	*More than 1*
Phonological non-word	294	32
Phonological real word	163	18
Lexical	36	29

Note: Phonological non-word errors do not differ significantly from phonological real-word errors ($chi^2 = 0.00217$; $df = 1$; $p > 0.95$); real-word phonological errors differ significantly from lexical errors ($chi^2 = 37$; $df = 1$; $p < 0.0001$). Utrecht corpus only.

in order to avoid extremely small expected values, while keeping the relevant differences. The collapsed matrix is shown in Table 10.4. Phonological real-word errors differ significantly from lexical but not from phonological non-word errors. This suggests that the monitor treats the phonological real-word errors as phonological ones.

Figure 10.2 presents a similar comparison for the number of words repeated in the correction. The corresponding collapsed matrix of the underlying numbers is given as Table 10.5. Again, these data suggest that the monitor treats phonological real-word errors as phonological and not as lexical ones.

A great proportion of the data in Tables 10.4 and 10.5 concern corrected anticipations. As discussed in the previous paragraph, we should treat these data with some caution. In Tables 10.6 and 10.7 data are presented limited to phonological non-word and real-word perseverations to be compared with

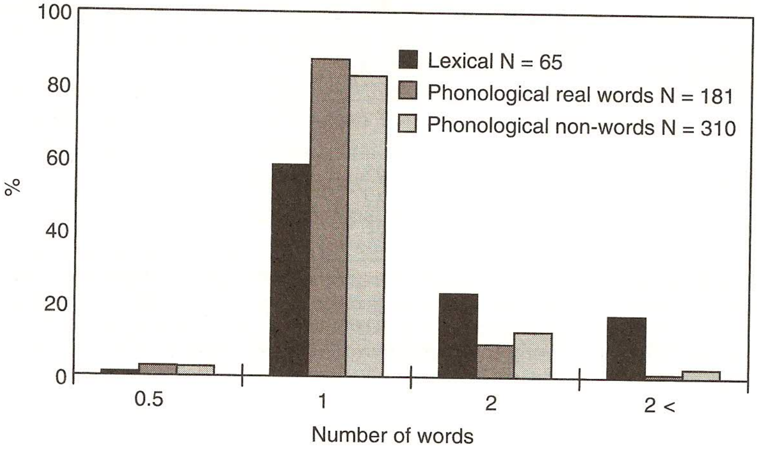

Figure 10.2 Percentage of speech errors as a function of the number of words spoken in the correction, plotted separately for lexical errors, phonological errors leading to non-words, and phonological errors accidentally leading to real words.

Table 10.5 Numbers of speech errors as a function of number of words repeated in the correction, separately for non-word phonological errors, real-word phonological errors and lexical errors

n	*1 or less*	*More than 1*
Phonological non-word	264	46
Phonological real word	161	20
Lexical	39	26

Note: Less than 1 indicates that the speaker not even went back to the beginning of the word containing the error. This occurred only in compounds. Phonological non-word errors do not differ significantly from phonological real-word errors ($chi^2 = 1.41$; $df = 1$; $p > 0.1$); real-word phonological errors differ significantly from lexical errors ($chi^2 = 26$; $df = 1$; $p < 0.0001$). Utrecht corpus only.

Table 10.6 Numbers of speech errors as a function of number of words spoken before stopping for correction, separately for non-word phonological perseverations, real-word phonological perseverations, and lexical errors

n	*Less than 1*	*1*	*More than 1*
Phonological non-word	24	25	0
Phonological real word	15	11	0
Lexical	4	32	29

Note: Less than 1 indicates that the speaker did not complete the word containing the error. Phonological non-word perseverations do not differ significantly from phonological real-word perseverations ($chi^2 = 0.52$; $df = 1$; $p > 0.3$); real-word phonological perseverations differ significantly from lexical errors ($chi^2 = 35$; $df = 2$; $p < 0.0001$). Utrecht corpus only.

Table 10.7 Numbers of speech errors as a function of number of words repeated in the correction, separately for non-word phonological perseverations, real-word phonological perseverations, and lexical errors

n	*1 or less*	*More than 1*
Phonological non-word	45	7
Phonological real word	32	6
Lexical	38	27

Note: Less than 1 indicates that the speaker not even went back to the beginning of the word containing the error. Phonological non-word perseverations do not differ significantly from phonological real-word perseverations ($chi^2 = 0.096$; $df = 1$; $p > 0.9$); real-word phonological perseverations differ significantly from lexical errors ($chi^2 = 7.3$; $df = 1$; $p < 0.01$). Utrecht corpus only.

lexical errors. In a perseveration, no part of the error triggering the monitor can hide in inner speech. Although the data are rather sparse, we find again a significant difference between phonological real-word errors and lexical errors but not between phonological non-word and real-word errors. A concern might be that with real-word errors sometimes syntax is violated, potentially providing an extra cue to the monitor. However, over those phonological real-word anticipations for which it could be assessed whether or not syntax was violated by the error, probability of correction appeared to be equal for errors with violated and with intact syntax ($N = 150$; $chi^2 = 0.465$; $df = 1$; $p > 0.3$). It seems safe to conclude that the monitor treats phonological real-word errors as phonological and not as lexical errors.

Do spontaneous phonological speech errors show lexical bias?

Lexical bias here is taken to mean that, in case of a phonological speech error, the probability that the error leads to a real word is greater, and the probability that the error leads to a non-word is less than chance. Lexical bias has been shown for experimentally elicited speech errors, where chance level could be experimentally controlled (Baars & Motley, 1974; Baars, Motley, & Mackay, 1975). The problem with spontaneous speech errors, of course, is to determine chance. Garrett (1976) attempted to solve this problem by sampling word pairs from published interviews and exchanging their initial sounds. He found that 33 per cent of these "pseudo-errors" created words. This was not conspicuously different from real-word phonological speech errors, so he concluded that there was no lexical bias in spontaneous speech errors. One may note, however, that Garrett did not distinguish between monosyllables and polysyllables. Obviously, exchanging a phoneme in a polysyllabic word hardly ever creates a real word. This may have obscured an effect of lexical bias. Dell and Reich (1981) used a more elaborate technique to estimate chance level, involving "random" pairing of words from the error corpus in two lists of word forms, exchanging of the paired words' initial sounds, and determining how often words are thereby created, normalizing for the frequency of each initial phoneme in each list. They found a significant lexical bias in anticipations, perseverations and transpositions. In the latter, involving two errors ("Yew Nork" for "New York") lexical bias was stronger in the first ("Yew") than in the second ("Nork") error. Del Viso et al. (1991), using a method very similar to Dell's, found no evidence for lexical bias in Spanish spontaneous speech errors. Note, however, that Dell's method is not very straightforward. The greater number of longish words in Spanish as compared with English may have obscured an effect of lexical bias.

In the current study I followed a different approach for assessing lexical bias, restricting myself to single-phoneme substitutions in monosyllables, i.e., errors where a single phoneme in a monosyllable is replaced with another single phoneme, in this way capitalizing on the fact that replacing a phoneme much more often creates a real word in a monosyllable than in a polysyllable. I did not, however, as Garrett (1976) and Dell and Reich (1981) did, restrict myself to initial phonemes, but took all single-phoneme substitutions in monosyllables into account. The two collections of Dutch speech errors together gave 311 such errors, 218 of which were real-word errors and 93 non-word errors. Although these numbers suggest a lexical bias, this may be an illusion, because it is unknown what chance would have given. It is reasonable to assume that a major factor in determining the probability of the lexical status of a phoneme substitution error is provided by the phonotactic alternatives. If, for example, the *p* of *pin*, is replaced by a *b*, the phonotactically possible errors are *bin, chin, din, fin, gin, kin, lin, sin, shin, tin, thin* (with voiceless *th*), *win, yin, *guin, *hin, *min, *nin, *rin, *zin, *zhin, *thin* (with voiced *th*). In this case there are 21 phonotactic alternatives, of which 13 are real words and 8 are nonsense words.

Of course, if all phonotactic alternatives are real words (which sometimes happens), the probability that the error produces a real word is 1; and if all alternatives are nonsense words (which also happens) the probability of a real word error is zero. In the case of *pin* turning into *bin*, the chance level for a real-word error would have been 13/21 = 0.62. I have assessed the average proportions of real-word phonotactic alternatives for all 311 single-phoneme substitutions in monosyllables (not only initial phonemes), taking only into account the phonotactically possible single phonemes in that position.

The average proportions of real-word and non-word alternatives in this particular set of monosyllables are both 0.5. The expected numbers of real-word and non-word speech errors therefore are both 311/2 = 155.5, whereas the actual numbers are 218 and 93 (Table 10.8). There is a strong interaction between error categories and expected values based on average proportions of phonotactic real-word and non-word alternatives. Evidently there is a strong lexical bias in spontaneous speech errors.

Table 10.8 Observed numbers of real words and non-words in single-phoneme substitutions in monosyllables only and numbers expected on the basis of the average proportions of real-word and non-word alternatives

	Observed values	*Expected values*
Real words	218	155.5
Non-words	93	155.5

Note: $chi^2 = 26$; $df = 1$; $p < 0.0001$. AC/SN corpus plus Utrecht corpus.

Do self-corrections of overt speech errors show lexical bias?

As we have seen, spontaneous speech errors show a strong lexical bias. If self-monitoring were responsible for lexical bias, by applying a lexicality test, as has been suggested by Levelt et al. (1999), then one would expect the same lexicality test to affect overt self-monitoring. This should lead to non-word errors being more often detected and corrected than real word errors. Indeed, if Levelt et al. were correct in their suggestion that monitoring one's own speech for errors is very much like monitoring someone else's speech for errors, listening for deviant sound form, deviant syntax, and deviant meaning, real-word errors cannot be detected in self-monitoring on the level of phonology. By definition real-word errors would pass any lexicality test, and therefore could only be detected as if they were lexical errors causing deviant syntax or deviant meaning. If, among other criteria, a lexicality test is applied by self-monitoring for phonological errors, we may expect the correction frequency to be higher for non-word errors than for real-word errors. Table 10.9 gives the relevant breakdown for all 315 single-phoneme substitutions in the Utrecht corpus and Table 10.10 gives the relevant breakdown of all 1111 phonological speech errors in this collection.

Obviously, there is no evidence of non-word errors being more frequently corrected than real-word errors. The data in Table 10.10 show that, if we consider all phonological errors instead of single-phoneme substitutions only, the probabilities for correction of real-word and non-word errors are exactly equal. It thus seems very unlikely that a lexicality test is applied in self-monitoring for overt speech errors during spontaneous speech production.

Table 10.9 Numbers of corrected and uncorrected single-phoneme substitutions in monosyllables and polysyllables together, separately for real-word errors and non-word errors

	Real words	*Non-words*
Corrected	99	69
Uncorrected	98	49

Note: $chi^2 = 2$; $df = 1$; $p > 0.1$. Utrecht corpus only.

Table 10.10 Numbers of corrected and uncorrected phonological errors in monosyllables and polysyllables together, separately for real-word errors and non-word errors

	Real words	*Non-words*
Corrected	218	341
Uncorrected	210	342

Note: $chi^2 = 0.117$; df = 1; $p > 0.7$. Utrecht corpus only.

Are lexical bias and probability of self-corrections of overt speech errors equally sensitive to phonetic distance between target and error?

If lexical bias results from editing out of non-words by self-monitoring, one would expect that errors differing from the correct form in only a single distinctive feature would be missed more often than errors differing in more features. The reason is that self-monitoring is supposed to depend on self-perception (Levelt et al., 1999), and it is reasonable to expect that in perception smaller differences are more likely to go unnoticed than larger differences. As lexical bias is supposed to be the effect of suppressing non-words, one expects lexical bias to increase with dissimilarity between the two phonemes involved. To test this prediction I divided the 311 single-phoneme substitution errors in monosyllables into three classes, viz errors involving 1 feature, errors involving 2 features, and errors involving 3 or more features. For consonants I used as features manner of articulation, place of articulation, and voice. For vowels features were degree of openness, degree of frontness, length, roundedness, and monophthong versus diphthong. Table 10.11 gives the numbers of real-word and non-word errors for the three types of single-phoneme substitutions in monosyllables, in the AC/SN corpus and Utrecht corpus together.

These results clearly suggest that lexical bias is sensitive to phonetic (dis)similarity, as predicted both from a perception-based theory of pre-articulatory editing, but also from "phoneme-to-word" feedback (Dell & Reich, 1980; Dell, 1986; Stemberger, 1985). If self-corrections are also sensitive to phonetic (dis)similarity this would favor the hypothesis that both effects stem from the same mechanism. If they are not, this would suggest

Table 10.11 Numbers of real-word errors and non-word errors in monosyllables only, separately for errors involving 1, 2, or 3 or more features

	1 feature	*2 features*	*3 features*
Real words	95	96	27
Non-words	59	29	5

Note: $chi^2 = 7.29$; df = 2; $p < 0.01$. AC/SN corpus plus Utrecht corpus.

Table 10.12 Numbers of corrected and uncorrected single-phoneme substitutions in monosyllables and polysyllables together, separately for errors involving 1, 2, or 3 features

	1 feature	*2 features*	*3 features*
Corrected	94	85	15
Uncorrected	60	65	19

Note: $chi^2 = 3.3$; $df = 2$; $p > 0.1$; n.s. Utrecht corpus only.

different mechanisms for lexical bias and self-detection of overt errors. Table 10.12 gives the relevant data taken from the Utrecht corpus. Obviously, there is little evidence that self-corrections are sensitive to phonetic (dis)-similarity, although one would predict such an effect from perception-based monitoring. This finding is corroborated by experimental data reported by Postma and Kolk (1992), to be further discussed in the following section. Self-correction of overt speech errors differs in this respect from whatever mechanism is responsible for lexical bias in speech errors.

Do the current data suffer from a collector's bias invalidating otherwise plausible solutions?

Perhaps the current data suffer from a collector's bias, invalidating the otherwise plausible conclusions (cf. Cutler, 1982). Of course, here the two possible sources of such a bias are phonetic similarity and lexical status. It seems unlikely, however, that such biases hold equally for corrected and uncorrected speech errors. The reason is that correction presents a very clear clue to the collector, easily overriding any more subtle difference due to phonetic similarity or lexical status. Thus, if there is a collector's bias due to phonetic similarity or to lexical bias, there should be an interaction between corrected versus uncorrected and lexical status combined with phonetic similarity. The data in Table 10.13 strongly suggest that there is no such interaction. This makes it implausible that the absence of effects of lexical status and phonetic similarity in correction frequencies is due to a collector's bias.

Table 10.13 Numbers of corrected and uncorrected single-phoneme substitutions in monosyllables and polysyllables together, separately for errors involving 1, 2 or more features, and for real-word errors and non-word errors

	1 feature, real word	*1 feature, non-word*	*2/3 features, real word*	*2/3 features, non-word*
Corrected	52	41	47	28
Uncorrected	52	26	53	23

Note: $chi^2 = 3.6$; $df = 3$; $p > 0.3$, n.s. Utrecht corpus only.

That the sensitivity of lexical bias and the insensitivity of self-detection of speech errors to phonetic similarity do not stem from a collector's bias is supported by experimental data provided by Lackner and Tuller (1979), and by Postma and Kolk (1992). Lackner and Tuller had speakers recite strings of nonsense syllables of CV structure, both with and without auditory masking of their own overt speech by noise. Subjects were instructed to press a telegraph key when they detected an error in their speech. Speakers made many errors with a difference of a single feature between error and target, but hardly any with more than a single feature. Apparently such multifeature errors were suppressed more often. This replicates the sensitivity of lexical bias to phonetic distance, assuming that the repertoire of nonsense syllables to be recited form a temporary lexicon in such an experiment. Because of the lack of multifeature errors, no useful comparisons could be made in terms of detection frequencies. This is different in the experimental data reported by Postma and Kolk. They replicated the Lackner and Tuller experiment, this time with both CV and VC syllables, and with normal speakers and stutterers. They also found many single-feature errors and hardly any multi-feature errors in the CV syllables. Surprisingly, in the VC syllables there were relatively many multifeature errors. Whatever the cause of this, detection frequencies showed hardly any effect of phonetic distance, precisely as in the current data on self-corrections of spontaneous speech errors. It seems safe to conclude that the current findings cannot be explained away by a collector's bias.

Discussion

In this chapter I have set out to test two predictions derived from the theory of speech production and self-monitoring proposed by Levelt et al. (1999):

- The monitor treats phonological errors that lead to real words, such as "gear" for "beer", as lexical errors.
- If spontaneous phonological speech errors show lexical bias, as has been suggested by Dell (1986), then the same lexical bias should be found in self-corrections of overt speech errors.

Both predictions have been falsified: Real-word phonological errors are clearly treated by the monitor as phonological, not as lexical errors. And although spontaneous speech errors in Dutch show a clear lexical bias, the probability of self-correction of overt speech errors does not show a trace of lexical bias.

The first finding corroborates a finding by Shattuck-Hufnagel and Cutler (1999), who showed that lexical errors tend to be corrected with a pitch accent on the corrected item, whereas both non-word and real-word phonological errors do not. This suggests that the monitor has access to the intended phonological form. Instead of asking, "is this a real word?", it appears to ask, "is this the word I intended to say?" One may note that this may be related to self-monitoring being a relatively slow, conscious, or at least a semi-conscious process (Levelt, 1989). Hartsuiker and Kolk (2001) estimate that the sum of auditory input processing, parsing and comparing is minimally 200 ms, and add another 200 ms for interrupting (cf. Levelt, 1989). Blackmer and Mitton (1991) have provided evidence suggesting that error detection and correction can take place before message interruption, such that speaking and reconstructing the intended message are incremental processes. Both accounts of temporal aspects of self-monitoring do not conflict with the suggestion that overt self-monitoring to some extent may depend on time-consuming conscious processing. Fodor (1983) and Baars (1997) both suggested that consciousness provides access to otherwise hidden subconscious information. In other words, slow and conscious processes are not modular (although they may suffer from limited resources), whereas fast and subconscious processes are often modular. If we take these ideas seriously, self-correction of overt speech errors may have access to intended phonological word forms through (semi-) conscious processing, and therefore has no need for a general criterion of the form "is this a real word?" In this way the current data may be reconciled with the idea that speech production and perception (but not monitoring) are to a large extent modular.

The second main finding of the current study is that lexical bias and self-correction of overt speech errors differ in some important respects, suggesting that they do not stem from the same underlying mechanism. As we have seen, self-correction of overt speech errors is a relatively slow, semi-conscious process. Lexical bias must be due to a very fast process that does not interrupt the stream of speech and that never seems to reach consciousness. Self-correction of overt errors does not seem to be sensitive to lexical status, the mechanism responsible for lexical bias obviously is. The latter mechanism is also sensitive to phonetic similarity between target and error, as lexical bias significantly increases with phonetic similarity. In contrast the probability of self-correction of overt speech errors appears to be independent of phonetic similarity. It may be noted that the sensitivity of lexical bias to phonetic similarity in itself is not an argument in favor of either perception-based self-monitoring or a production-based mechanism as the source of the effect, as on the face of it the phonetic-similarity effect is compatible with both explanations. However, the finding that lexical bias is and self-correction of overt speech errors is not sensitive to phonetic similarity suggests that there are two different mechanisms involved.

Because self-correction of overt speech errors obviously is perception based it may be unexpected that there is no effect of phonetic similarity, not only in our data on spontaneous speech but also in the experimental data provided by Postma and Kolk (1992). It seems reasonable to expect that small differences would be more easily not heard than greater differences. Possibly the absence of a similarity effect is related to the assumption that the monitor compares intended with perceived form, instead of checking whether the perceived form is or is not part of the lexicon. One may also note that the absence of a lexicality effect in self-corrections of overt speech errors contradicts a prediction from Mackay's Node Structure Theory (1992). This theory predicts that non-word errors are more easily detected than real-word errors because there is a level where non-word errors are novel combinations, and real-word errors are not.

The properties of the mechanism causing lexical bias in spontaneous speech errors seem to be different from those of self-correction of overt speech errors: Lexical bias is caused by a mechanism that is fast and unconscious, is sensitive to the lexicality of the error and sensitive to phonetic distance between error and intended form. Self-correction of overt errors is time-consuming, and is not sensitive to the lexical status of the error and phonetic distance between error and target. There are several possible explanations for this difference.

One is that lexical bias is caused by a feedback mechanism as suggested by Dell (1986). Dell and Reich (1980) describe the proposed mechanism as follows: "An activated set of phonemes that corresponds to a word is continually reinforced by reverberation with a single word node because most of the activation from the phonemes converges and sums up at that node." Of course, an erroneous set of phonemes would either "reverberate" with the wrong word node, explaining the lexical bias, or with no word at all, explaining the "suppression of non-word outcomes" (Dell & Reich, 1980). A set of phonemes differing minimally from the activated word node would still reverberate" considerably with it, but as the difference increases, "reverberation" would diminish. That lexical bias decreases with phonetic similarity between intended and erroneous form, thus is entirely in tune with the feedback model. Of course, a feedback model is less parsimonious than a strictly serial feedforward-only model. However, it has been shown computationally that lexical bias is consistent with an architecture of speech production in which the interactivity introduced by cascading activation and phoneme-to-word feedback is severely restricted, and thereby seriality to a large extent preserved (Rapp & Goldrick, 2000). An argument against the feedback model, as pointed out by Levelt et al. (1999), is that it does not easily explain that lexical bias is sensitive to contextual and situational information and to social appropriateness (Motley, 1980; Motley, Camden, & Baars, 1982). One may note, however, that an architecture of speech production that is not strictly modular but rather has restricted interactivity, with some "leakage" of information from one module to another, as suggested by Rapp and Goldrick (2000), would more easily allow for such effects.

Alternatively, the current data can be explained by a fast automatic production-based monitor that is completely separate from the perception-based monitor responsible for self-corrections of overt speech errors. Such a production-based monitor has been suggested by Nickels and Howard (1995) and Postma (2000). This would more easily account for the fact that lexical bias is sensitive to contextual and situational information, and social appropriateness (Motley, 1980; Motley et al., 1982), something one would intuitively rather expect from a monitoring system than from a fully automatized speech production system.

A third possibility is that lexical bias is caused by output editing of inner speech, as suggested by Levelt et al. (1999). This would, of course, also account for output editing being sensitive to contextual and situational information and social appropriateness. We then have to assume that output editing of inner speech differs in its properties from output editing of overt speech errors, notably in being fast, unconscious and sensitive to a general criterion of lexicality and to phonetic distance between error and intended form. Perhaps the properties of output editing by the self-monitoring system change as a function of the time the system is allowed to do its job. Fast and hidden editing of unspoken errors remains unconscious, and has to depend on general criteria, slow editing of errors already spoken may become conscious, and may have access to more detailed information about the intended form.

A serendipitous finding of the current study is that the majority of speech errors in inner speech are transpositions or exchanges, contrary to what counting overt speech errors so far suggested. It remains to be seen whether current theories of speech error generation, in as far as they are based on relative frequencies of different types of speech error (cf. Dell, 1986), can easily be retuned in order to accommodate this finding.

The most important conclusions from the current analysis of speech errors and their corrections seem to be the following.

The part of a speaker's mind that watches out for speech errors in order to correct them has access to the intended phonological forms of misspoken words. In this way, contrary to what has been suggested by Levelt et al. (1999), listening for errors in one's own (overt) speech is quite different from listening for speech errors in the speech of other speakers.

Lexical bias in spontaneous speech errors is not caused by the same mechanism that allows for detection and correction of overt speech errors. It may either be caused by an automatic production-based monitor that is quite different from the semi-conscious perception-based monitor that is responsible for self-corrections of overt speech errors (Nickels & Howard, 1995; Postma 2000), or by a phoneme-to-word feedback mechanism, as proposed by Dell (1986) and Dell and Reich (1980), and more recently by Rapp and Goldrick (2000), or by output editing of inner speech as suggested by Levelt (1989) and Levelt et al. (1999). If so, the current results imply that fast and hidden output editing of inner speech, employing a general criterion of lexicality and thereby rejecting nonwords more frequently than real words, is different from output editing of overt speech, comparing the spoken word form with the intended word form. This difference between output editing of inner speech and of overt speech is supported by more recent experimental evidence (Nooteboom, 2003).

Acknowledgements

I am grateful to those who read and commented on earlier versions of this manuscript, and made suggestions for improvements. These are Bernard Baars, Anne Cutler, Gary Dell, Rob Hartsuiker, Pim Levelt, Albert Postma, Ardi Roelofs, Hugo Quené and Frank Wijnen.

References

Baars, B. J. (1997). *In the theatre of consciousness*. New York: Oxford University Press.

Baars, B. J., & Motley, M. T. (1974). Spoonerisms: Experimental elicitation of human speech errors: Methods, implications, and work in progress. *Journal Supplement Abstract Service, Catalog of Selected Documents in Psychology*.

Baars, B. J., Motley, M. T., & McKay D. (1975). Output editing for lexical status from artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *14*, 382–91.

Blackmer E. R., & Mitton J. L. (1991) Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition*, *39*, 173–94.

Cutler, A. (1982). The reliability of speech error data. *Linguistics*, *19*, 561–82.

Del Viso, S., Igoa, J. M., & Garcia-Albea, J. E. (1991). On the autonomy of phonological encoding: Evidence from slips of the tongue in Spanish. *Journal of Psycholinguistic Research*, *20*, 161–85.

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

Dell, G. S., & Reich, P. A. (1980). Toward a unified model of slips of the tongue, In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 273–86). New York: Academic Press.

Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, *20*, 611–29.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: Bradford Books, MIT Press.

Fromkin, V. A. (1973). *Speech errors as linguistic evidence*. The Hague: Mouton.

Garrett, M. F. (1976). Syntactic process in sentence production. In R. J. Walker, & E.C.T. Walker (Eds.), *New approaches to language mechanisms* (pp. 231–56). Amsterdam: North-Holland.

Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology*, *42*, 113–57.

Lackner, J. R., & Tuller, B. H. (1979). Role of efference monitoring in the detection of self-produced speech errors. In: W. E. Cooper, & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 281–94). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

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

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.

Meringer, R. (1908). *Aus dem Leben der Sprache*. Berlin: V. Behr's Verlag.

Motley, M. T. (1980). Verification of "Freudian" slips and semantic prearticulatory editing via laboratory-induced spoonerisms. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 133–48). New York: Academic Press.

Motley, M. T., Camden, C. T., & Baars, B. J. (1982). Covert formulation and editing of anomalies in speech production: Evidence from experimentally elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *21*, 578–94.

Nickels, L., & Howard, D. (1995). Phonological errors in aphasic naming: comprehension, monitoring and lexicality. *Cortex*, *31*, 209–37.

Nooteboom, S. G. (1969). The tongue slips into patterns. In A. Sciarone, A. J. van Essen, & A. A. van Raad (Eds.), *Nomen, Leyden studies in linguistics and phonetics* (pp. 114–32). The Hague: Mouton. Also in V. A. Fromkin (1973) (Ed.) *Speech errors as linguistic evidence*. (pp. 144–56). The Hague: Mouton.

Nooteboom, S. G. (1980). Speaking and unspeaking: Detection and correction of phonological and lexical errors in spontaneous speech. In V.A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 87–95). New York: Academic Press.

Nooteboom, S. G. (2003). Self-monitoring is the main cause of lexical bias in phonological speech errors. In R. Eklund (Ed.), *Proceedings of DiSS'03, Disfluency in Spontaneous Speech Workshop*. 5–8 September 2003, Göteborg University, Sweden. *Gothenburg Papers in Theoretical Linguistics*, *89*, pp. 25–8.

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

Postma, A., & Kolk, H. (1992). Error monitoring in people who stutter: Evidence against auditory feedback defect theories. *Journal of Speech and Hearing Research*, *35*, 1024–1032.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*, 460–99.

Schelvis, M. (1985). The collection, categorisation, storage and retrieval of spontaneous speech error material at the Institute of Phonetics. Utrecht: *PRIPU*, *10*, 3–14.

Shattuck-Hufnagel, S., & Cutler, A. (1999) The prosody of speech error corrections revisited [CD-ROM]. *Proceedings of the International Congress of Phonetic sciences in San Francisco, 1–6 August 1999* (pp. 1483–6). San Francisco: Regents of the University of California.

Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 1, pp. 143–86). Hove, UK: Lawrence Erlbaum Associates Ltd.