Different sources of lexical bias and overt self-corrections

Sieb G. Nooteboom

UiL OTS
Utrecht University, Utrecht
Sieb.Nooteboom@let.uu.nl

Abstract

In this paper it is argued, on the basis of a quantitative analysis of spontaneous speech errors and their corrections in Dutch, that the mechanism leading to lexical bias in speech errors cannot be same as that leading to overt self-corrections. Although spontaneous speech errors show a strong lexical bias, overt self-corrections do not. Lexical bias strongly increases with dissimilarity between target phoneme and source phoneme No such effect is found in overt self-corrections. Several possible sources of these differences are discussed.

1. Introduction

Baars, Motley and MacKay (1975) elicited spoonerisms by having subjects read aloud a target like *darn bore* preceded by bias items in which at least the first phoneme in this case was a *b*, triggering the spoonerism *barn door*. They observed that the error rate for cases such as *darn bore*, triggering lexically viable outcomes, was higher than the error rate for cases like *dart board*, triggering non-word outcomes. This lexical bias was not found when the context contained non-words only. The authors explained this result by positing an output-editing mechanism suppressing non-words that arise from speech errors in inner speech. Levelt, Roelofs and Meyer (1999) recently supported this original explanation by Baars et al. (1975) and suggested that the pre-articulatory editing leading to lexical bias is a form of covert self-correction of internal speech by the self-monitoring system that is also responsible for overt detection and correction of speech errors. A different approach has been suggested by Dell & Reich (1980), and Dell (1986), who proposed that lexical bias is caused by "phoneme-to-word" feedback during production processes, and therefore obviously not by the same mechanism that is responsible for the overt detection of speech errors. Postma (2000) recently suggested that there is production-based monitoring during speech production. If there is, such monitoring can also be responsible for lexical bias.

If overt speech errors and overt self-corrections are affected by the same perception-based monitoring system, one expects the distributions of both to reflect the same kinds of bias. If, on the other hand, overt speech errors are only affected by production-based mechanisms and overt self-corrections are affected by perception-based monitoring, one would not necessarily expect these distributions to be similar. Of course, before we can study possible similarities between these distributions it should first be assessed whether spontaneous speech errors do show an effect of lexical status. Note that this effect is compatible with both a production-based origin, as suggested by Dell & Reich (1980), Dell (1986) and Postma (2000), and with a perception-based origin (Levelt, 1989; Levelt et al. 1999). Below I will examine four predictions

1) The first prediction is that there is a lexical bias effect not only in the laboratory task used by Baars and his associates, but also in spontaneous speech errors. Garrett (1976) did not find much evidence for lexical bias in his MIT corpus of spontaneous errors. Dell and Reich (1981) report a considerable lexicality effect for another corpus. Here a new attempt will be described, on the basis of speech errors drawn from two collections of spontaneous speech errors in Dutch, and a new proposal for a null hypothesis.

2) If there is lexical bias in spontaneous speech errors and it is caused by perception-based self-monitoring, then one would expect to find a similar lexical bias in the overt corrections of spontaneous errors, meaning that spontaneous non-word errors are significantly more often corrected than real-word errors.

3) As self-monitoring is self-perception, and smaller differences are less easily perceived than larger differences, one may predict that speech errors differing minimally from the intended forms will be more often go unnoticed than speech errors differing greatly from the intended forms. One should therefore expect suppression of non-words to be less likely when two phonemes involved in a substitution differ in only one feature than when they differ in more features.

4) Obviously, if we do find an effect of phonetic similarity in lexical bias, and suppose that lexical bias is caused by the same mechanism that is responsible for overt detection of speech errors, then we expect to find the same sensitivity to phonetic similarity in the distribution of overt self-corrections. Notably we expect that errors involving a single-feature difference with the intended form are less likely to be corrected than errors involving more features. Sensitivity to phonetic similarity is compatible both with a production-based and with a perception-based mechanism. Therefore the interesting case would be if one were to find such sensitivity in the one and not in the other. This would be evidence for different mechanisms.

2. Method

2.1. Two collections of speech errors

The above predictions have been tested against errors drawn from two different collections of spontaneous speech errors in Dutch, for predictions 1 and 3, and from one of these collections for prediction 2 and 4. The reason for the difference is that one of the available collections alone provided too few relevant cases for testing predictions 1 and 3, and the oldest of the two collections could not be used for testing prediction 2 and 4, because in that collection corrections were not reliably noted down.

- The oldest collection 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 out-numbering other types, such as lexical syntagmatic errors, blends, and intrusion errors. The collection was never put into a digital data base 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 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 2,500 errors of various types, of which more than 1,100 are phonological syntagmatic errors and some 185 lexical syntagmatic errors. The collection was put into a digital data base, currently accessible with Microsoft Access.

2.2. Assessing 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 nonsense word is less than chance. The problem here, 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% percent 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.

In the current study I followed a different approach, 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 optimally 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 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 taken these proportions of real-word and non-word phonotactic alternatives for all individual errors included in the analysis, averaged these proportions, and I have taken the average as null hypothesis for the proportions of real-word errors and non-word errors in the subset of the collection at hand. The proportions of real- and non-word phonotactic alternatives seem to offer a workable null hypothesis for assessing any other effects.

Of course, with overt self-corrections there are fewer methodological problems in assessing lexical bias. If, among other criteria, a lexicality test is applied by self-monitoring, we may expect the correction frequency to be higher for non-word errors than for real-word errors.

3. Testing four predictions

3.1. Lexical bias in spontaneous speech errors

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 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. There is a strong interaction between error categories and expected values based on average proportions of phonotactic real-word and non-word alternatives (chi^2=50; df=2; p<0.0001). Evidently there is a strong lexical bias in spontaneous speech errors, as predicted.

3.2. Lexical bias in self-corrections of overt speech errors

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, then one would expect the same lexicality test to affect overt self-monitoring, as has been suggested by Levelt et al. (1999). This should lead to non-word errors being more often detected and corrected than real word errors. Indeed, if Levelt were correct in his 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. Elsewhere (Nooteboom, submitted) I have shown that phonological real-word errors are treated by the monitor as phonological errors, not as lexical ones. The distributions of numbers of words speakers move on before stopping for correction differ significantly between phonological errors and lexical errors, but are the same for phonological real-word and phonological non-word errors. The same is true for the distributions of numbers of words included in the correction. These findings confirm evidence from Shattuck-Hufnagel and Cutler (1999), who demonstrated that lexical errors tend to be corrected with a pitch accent on the corrected word, whereas both phonological real-word errors and phonological non-word errors do not.

Clearly, phonological real-word errors are detected on the level of phonological, not of lexical processing. 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 1 gives the relevant breakdown for the 315 single-phoneme substitutions, and Table 2 gives the relevant breakdown of all 1,111 phonological speech errors in the collection.

Table 1. Numbers of corrected and uncorrected single-phoneme substitutions, separately for real-word errors and non-word errors. (chi^2=1.95; df=2; p>0.3).

	Real words	Non-words
Corrected	99	69
Uncorrected	98	49

Obviously, there is no evidence of non-word errors being more frequently corrected than real-word errors. If there is any tendency in Table 1, it goes the wrong way. The data in Table 2 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 2. Numbers of corrected and uncorrected phonological errors, separately for real-word errors and non-word errors (chi^2=0.117; df=2; p>0.5)

	Real words	Non-words
Corrected	218	341
Uncorrected	210	342

3.3. Lexical bias and phonetic dissimilarity

If lexical bias results from editing out of non-words by self-monitoring, one would expect errors differing from the correct form in only a single distinctive feature 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 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 3 gives the numbers of real-word and non-word errors for the three classes.

Table 3. Numbers of real words and non-word errors, separately for errors involving 1, 2, or 3 or more features (chi^2 = 11.31; df=4; p<0.05).

	1 Feat.	2 Feat.	3 Feat.
Real words	95	96	27
Non-words	59	29	5

These results clearly suggest that lexical bias is sensitive to phonetic similarity, as predicted not only from a perception-based theory of pre-articulatory editing, but also from "phoneme-to-word" feedback (Dell & Reich, 1980; Stemberger 1985; Dell 1986).

3.4. Self-corrections and phonetic similarity

If self-corrections are sensitive to phonetic similarity, as lexical bias is, this would favour the hypothesis that both effects stem from the same mechanism. If they are not, this would suggest different mechanisms. Table 4 gives the relevant data.

Table 4. Numbers of corrected and not corrected single-phoneme substitutions, separately for errors involving 1 feature, 2 features of 3 features (chi^2=3.995; df=4; p>0.05; n.s.)

	1 Feat	2 Feats	3 Feats
Corrected	94	85	15
Not corrected	60	65	19

Obviously, there is little evidence that self-corrections are sensitive to phonetic similarity, although one would predict such an effect from perception-based monitoring.

4. A collector's bias?

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 differences 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 5 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 5. Numbers of Corrected and Uncorrected Single–phoneme Substitutions, Separately for Errors Involving 1, 2 or More Features, and for Real-word Errors and Non-word Errors (chi^2=3.18; df=6; p>0.7).

	1 Feat.; Real word	1 Feat.; Non-word	2/3 Feat.; Real word	2/3 Feat.; Non-word
Corr.	52	41	47	28
Not corr.	52	26	53	23

5. Discussion

What have we found? Phonological speech errors show lexical bias in the sense that in real words there are more and in non-words there are fewer such errors than expected on a chance basis. The size of the lexical bias in phonological speech errors decreases with phonetic similarity. Both effects seem compatible with the perceptual-loop theory of self-monitoring, suggested by Levelt et al. (1999). But contrary to expectation, the correction frequency of phonological speech errors is not influenced by either lexical status, or phonetic similarity. Note that these are not the only differences between lexical bias in speech errors and self-correction of speech errors. There are at least two other differences. One is a difference in speed, the other a difference in degree of consciousness.

The difference in speed is obvious: Lexical bias must be due to a mechanism operating before the error is made overt. Overt detection and correction of a speech error often, although not always, happens after the error has become overt. Another, possibly related, difference is in degree of consciousness. Speakers are often, although not always, conscious of having made a speech error, and then in many cases stop for correction. Note that a speech error that becomes sufficiently conscious to make the speaker stop for correction, has not necessarily become overt. As pointed out by Levelt (1989), stopping after an error sometimes occurs after only the first phoneme of the mispronounced word has been produced, suggesting that the stopping must have been initiated before the error had become overt. These cases can be explained by the less time-consuming inner-loop monitoring, i.e. by monitoring of inner speech via the speech comprehension system. Detecting self-produced speech errors in one's inner speech often reaches consciousness. This contrasts with the process leading to lexical bias in phonological speech errors. If lexical bias results from editing out speech errors leading to non-words, this editing process seems to be entirely subconscious. The apparent differences between lexical bias and overt self-correction strongly suggest that they are not both effects of perception-based self-monitoring. Lexical bias must have another origin. One candidate is the mechanism of "phoneme-to-word" feedback (Dell, 1986). Another is a process of production-based monitoring (Postma, 2000).

A final question is why perception-based overt detection of speech errors is not sensitive to lexical status and phonetic similarity, as one would expect from a perception-based mechanism. This is unclear. Whether speech perception really is sensitive enough to lexical status and phonetic similarity is still to be verified in perception experiments. If it turns out that perception of speech errors made by others is sufficiently sensitive to lexical status and phonetic similarity, the absence of these effects in detecting self-produced errors suggests that the perception-based self-monitoring system has immediate access to the intended form. Binary comparison of intended and produced form is an easy task, which is likely not to be sensitive to lexical status and phonetic similarity.

6. Conclusion

Lexical bias in phonological speech errors and overt detection of self-produced phonological speech errors are not products of the same mechanism.

7. References

[1] Baars, B.J. & Motley, M.T., "Spoonerisms: Experimental elicitation of human speech errors: methods, implications, and work in progress", *Journal Supplement Abstract Service, Catalog of selected documents in psychology*, (1974, fall).

[2] Baars, B.J., Motley, M.T., & McKay D., "Output editing for lexical status from artificially elicited slips of the tongue", *Journal of verbal learning and verbal behavior, 14*, 382-391, 1975.

[3] Cutler, A., "The reliability of speech error data", In: A. Cutler (ed.) *Slips of the tongue and language production*, Mouton, Amsterdam, 1982

[4] Dell, G.S., "A spreading-activation theory of retrieval in sentence production", *Psychological Review*, 93, 283-321, 1986.

[5] Dell, G.S. and Reich, P.A., "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* New York: Academic Press, 1980, pp. 273-286.

[6] Garrett, M. F., Syntactic processes in sentence production. In: R. J. Wales and E. Walker (eds.), *New approaches to language mechanisms* ,Amsterdam: North-Holland Publishing Company, 1976, pp. 231-256.

[7] Levelt, W.J.M. *Speaking. From intention to articulation*, MIT Press, 1983.

[8] Levelt, W.J.M., Roelofs, A. and Meyer, A.S., "A theory of lexical access in speech production", *Behavioral and Brain Sciences*, 22, 1-75, 1999.

[9] Postma, A., "Detection of errors during speech production: a review of speech monitoring models", *Cognition*, 77, 97-131, 2000.

[10] Schelvis, M., "The collection, categorisation, storage and retrieval of spontaneous speech error material at the Institute of Phonetics", *PRIPU* No 10, Utrecht University, 1985, pp. 3-14.

[11] Shattuck-Hufnagel, S. and Cutler, A. (1999). "The prosody of speech error corrections revisited", [CD-ROM]. *Proceedings of the XIVth International Congress of Phonetic Sciences, San Francisco, August 1-6*, San Francisco, Regents of the University of California, 1999, pp. 1483-1486.