Lexical Bias in Phonological Speech Errors: Phoneme-to-Word Feedback or Output Editing?

Sieb G. Nooteboom[†]

† Utrecht institute of Linguistics OTS, Utrecht University, The Netherlands
E-mail: Sieb.Nooteboom@let.uu.nl

ABSTRACT

This paper attempts to answer two questions: (a) What is the cause of lexical bias in phonological speech errors? (b) Is there phoneme-to-word feedback in speech production? An experiment is reported adapted from Baars, Motley, and MacKay [1], eliciting word-word and nonword-nonword spoonerisms. Three modifications were made: (1) There was ample time for overt corrections. (2) Phonetic distance between targets and errors was used as an independent variable. (3) Response times were measured. Results on the incidence of spoonerisms convincingly show that lexical bias is caused by nonwords being more often rejected and repaired than real words in inner speech by self-monitoring. Results on response latencies clearly demonstrate phoneme-to-word feedback in speech production. It is argued here that phoneme-to-word feedback may be an unavoidable side-effect of our own perception of inner speech for the purpose of self-monitoring.

1. INTRODUCTION

The main focus of the present paper is on the question whether or not there is phoneme-to-word feedback in speech production. Dell [2] argues there is, Levelt, Roelofs and Meyer [3] argue there is not. One of Dell's arguments in favour of feedback is the existence of lexical bias in phonological speech errors, i.e. the phenomenon that a real-word outcome of a speech error is more likely than a nonword error, given equal a priori probabilities. Levelt c.s., allowing no feedback, explain lexical bias by supposing that nonword errors have a higher probability than real-word errors of being rejected and repaired in inner speech by self-monitoring. Although lexical bias has played a major role in the discussion so far, it may be observed that there are two separate questions here. One question is: What is the cause of lexical bias? Another question is: Is there phoneme-to-word feedback in speech production? In order to answer these questions, an experiment has been done adapted from Baars, Motley and MacKay [1]. These ran an experiment eliciting spoonerisms by phonological priming, comparing the number of nonword errors with the number of real-word errors given equal a priori probabilities. They convincingly showed that real-word errors are more frequent than nonword errors and explained this by assuming that nonword errors stand a higher chance than real-word errors of being rejected and repaired by some form of output editing of inner speech, an explanation still adhered to by Levelt et al. [3]. A self-monitoring account is supported by lexical feedback being sensitive to contextual and situational information and to social appropriateness [3]. Evidence against a self-monitoring and in favour of a feedback account has recently been put forward by Humphreys [4]

In the current experiment some modifications were made to the design by Baars et al.. One modification was meant to get data on overt error rejection and correction. The reason is that if indeed lexical bias stems from self-monitoring of inner speech by the same monitor that is used for listening to other-produced speech, one expects nonwords also to be more frequently rejected from overt speech than real-words. This prediction is up for testing here. A second modification is that phonetic distance between error and target was used as experimental independent variable. The reason is that arguably according to a feedback account of lexical bias, phonetic distance should mainly act on the number of real-word errors produced, whereas according to a self-monitoring account phonetic distance is likely to act on the number of nonwords produced (see discussion). Thus this experimental variable can help us to discriminate between the two hypothetical explanations of lexical bias. Thirdly, it was decided to measure response latencies, not so much in the erroneous responses, but more importantly in the great many error-free and fluent responses. The reason for this is that a spreading activation model like WEAVER++ [1, 5] predicts faster responses with higher activation. Adding a feedback loop à la Dell [2] to WEAVER++ would yield higher activation in case real-word errors are primed for than in case non-word errors are primed for. Thus response times for the former would be shorter than those for the latter. In case there is no feedback, one expects no difference because the system has no way to know anything about a speech error that has never been made. If there is feedback, one also expects the shortening effect of it to get stronger as the phonetic distance between error and target gets smaller, because the effect of feedback would increase with decreasing distance.

 ISBN 1-876346-48-5

2. METHOD

The method used was basically the same as the one applied by Baars et al. [1] with some minor modifications, as explained above.

Stimuli

Priming word pairs consisted of pairs of monosyllabic Dutch words, visually presented in clear capital print on a computer screen and intended to be read silently. Before each test stimulus there were 3, 4, 5, 6 or 7 priming word pairs, chosen to prime a spoonerism, as in the sequence "GIVE BOOK, GO BACK, GET BOOT" preceding the test stimulus "BAD GOOF". In total there were 144 priming word pairs preceding test stimuli, and 144 non-priming word pairs preceding filler stimuli. The initial consonants of priming word pairs and test word pairs were chosen from the set /f, s, X, v, z, b, d, p, t, k/. There were 18 test stimuli primed for nonword-nonword spoonerisms, as "BAD GOOF" giving "GAD BOOF", and 18 test stimuli primed for word-word spoonerisms as "BARN DOOR" giving "DARN BORE". Each set of 18 was divided in 3 groups of 6 stimuli with equal phonetic distance between initial consonants, viz. 1 , 2 or 3 distinctive features. For example, /f/ vs /s/ differ in 1 feature, /f/ vs. /p/ differ in 2 features, and /f/ vs. /z/ differ in 3 features. There were 36 filler stimuli preceded by 144 non-priming word pairs and not controlled for expected outcomes of spoonerisms, class of initial consonants, or phonetic distance between target and potential error. In all other respects they were similar to the test stimuli. After each test and each filler stimulus word pair the subject saw on the screen a prompt SPREEK UIT (= "SPEAK"). After that the subject saw a second prompt CORRECTIE (= "CORRECTION"). In addition to the set of test and filler stimuli described so far there was a set of 7 stimuli with a variable number, on the average 4, of non-priming preceding word pairs to be used as practice for the subjects, and of course also followed by two prompts each. The total number of visually presented priming word pairs (144), non-priming word pairs (144 + 28 = 172), practice stimuli (7), test stimuli (36), filler stimuli (36) and prompts (144 + 14) was 553.

Subjects

There were 50 subjects, 17 male and 33 female, all of them naive as to the purpose of the experiment. They were staff members and students of Utrecht University, all with standard Dutch as their mother tongue and with no known history of speech or hearing pathology. Subjects varied in age from 17 to 56.

Procedure

Each subject was tested individually in a sound proof booth. The timing of visual presentation on a computer screen was computer controlled. The order in which test and filler stimuli, along with their priming or non-priming preceding word pairs were presented was randomized and different for each subject. Each (non-)priming word pair, each SPEAK-prompt and each CORRECTION-prompt was visible during 900 ms and was followed by 100 ms with a blank screen. The subject was instructed, on seeing the "SPREEK UIT" (=SPEAK) prompt to speak aloud the last word pair presented before this prompt. The subject was instructed to correct the spoken word pair in case of error. It was not necessary to wait for the "CORRECTION" prompt. The purpose of the latter was only to provide each subject with plenty of time for correction in case an error was made. All speech of each subject was recorded, and digitally stored on one of two tracks of DAT. On the other track of the DAT two tones of 1000 Hz and 50 ms duration were recorded with each test or filler stimulus, one starting at the onset of the visual presentation of the "SPEAK" stimulus, the other starting at the onset of the presentation of the "CORRECTION" prompt. These signals were helpful for orientation in the visual oscillographic analysis of the speech signals, and the first of these was indispensable in measuring response times.

Collecting the data

Reactions to all remaining test and filler stimulus presentations were transcribed either in orthography, or, where necessary, in phonetic transcription by two phonetically trained transcribers, viz. the present author and one of his students, using a computer program for the visual oscillographic display and auditory playback of audio signals. Transcriptions differed in less than 2% of all utterances and in less than 10% of all utterances containing an error. Response times for all correct and incorrect responses, to both fillers and test stimuli were measured by hand in the two-channel oscillographic display from the onset of the 50 ms tone coinciding with the onset of the presentation of the visual "SPEAK" prompt to the onset of the spoken response.

3. RESULTS

Incidence of spoonerisms

All in all there were 381 errors of many different kinds made in the two conditions primed for spoonerisms. Of these 381 errors most had no relation with the experimental conditions. The only errors that will concern us here are 56 completed spoonerisms (3%) and 67 aborted spoonerisms, i.e. spoonerisms of the expected kind but stopped before completion. The completed spoonerisms show a clear lexical bias: 37 real-word errors versus 19 nonword errors ($p < 0.01$ in a binomial test). Lexical bias appears to depend on phonetic distance, as shown in Table 1. The effect of phonetic distance is significant for the number of nonword errors, rapidly decreasing with distance, but not for the number of real-word errors, that increases somewhat with phonetic distance. This is a first indication that lexical bias is caused by self-monitoring rather than by feedback. The question why effects of phonetic distance and lexical status on self-monitoring are not independent will be taken up in the discussion. If indeed lexical bias is caused by nonwords standing a higher chance than real words to be rejected and repaired in inner speech, one may ask whether we can find evidence for this in overt rejections and/or repairs. Of all the cases where expected spoonerisms are stopped before completion, most were stopped in the first syllable, many even after the first consonant.

Table 1. Effect of Phonetic Distance for Completed Spoonerisms, Separately for Word-word (p=0.59 in a Binomial Test), and for Nonword-nonword errors (p < 0.01).

	word-word	nonword-nonword
1 feature	10	12
2 or 3 features	27	7

Obviously the speakers in these cases did not have time to react to their own overt speech, but rather reacted to their inner speech (cf. [3]). So if indeed nonwords have a higher probability to be rejected in inner speech than real words, one expects this effect also to show up in the overt rejections. Therefore there should be an interaction between completed versus aborted spoonerisms and lexical status. Table 2 suggests this interaction (although the effect of lexical status is not significant for aborted errors) thereby supplying further support for the hypothesis that lexical bias is caused by self-monitoring rather than by feedback.

Table 2. The Effect of Lexical Status on the Frequency of Expected Spoonerisms that are Completed or Aborted before Completion.

	completed	aborted
word-word	37	28
nonword-nonword	19	39

If this hypothesis is correct, one also expects an effect of phonetic distance on the probability of nonwords to be rejected, and not necessarily on real words. This is tested in Table 3.

Table 3. Effect of Phonetic Distance on the Number of Completed and Aborted Nonword-nonword Spoonerisms

	completed	aborted
1 feature	12	11
2 features	6	12
3 features	1	16

These data suggest an interaction between lexical status and completed versus aborted, as one would predict from a self-monitoring account of lexical bias. Although the effect of phonetic distance on the number of aborted errors is in itself not significant, under the assumption that there are no intrasubject correlations, the interaction is found to be significant in a Chi square test (Chi^2=9.51; df =2; p<0.01).

Response times

Fig. 1 shows average response times for all correct and fluent filler stimuli, not primed for spoonerisms, and test stimuli, primed for spoonerisms.

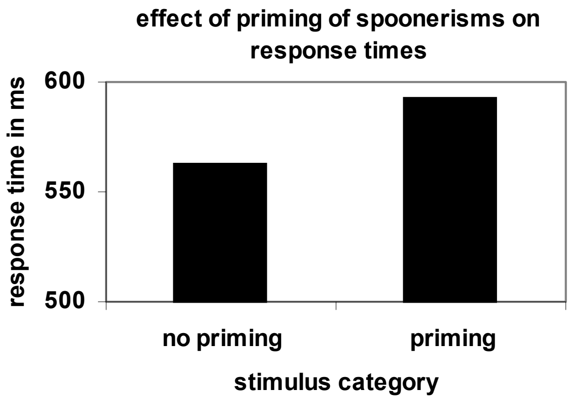

Fig. 1. Mean response times in error-free and fluent productions as a function of presence or absence of phonological priming of spoonerisms.

Means (and standard errors) are 563 (3.9), and 593 (4.4) ms respectively. The priming effect is found to be highly significant in a repeated measures ANOVA over subjects [test stimuli versus fillers: F(1,49)=20.5, p<.001; Subjects: F(49,2732)=21.1, p<.001]. This ANOVA used a univariate. A multivariate design was impossible here because of the great differences between subjects in number of observations per cell, caused by enormous differences in error rate and rate of hesitations between subjects. These results strongly suggest that extra activation by heavy phonological priming of competing phonemes delays onset of production, the delay being in the order of 30 ms.

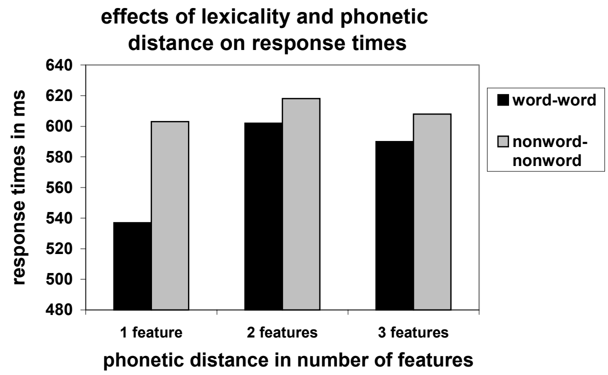

Fig. 2. Mean response times in error-free and fluent productions as a function of lexicality of expected but not realized spoonerisms.

Fig. 2 shows average response times for correct and fluent test stimuli, separately for expected word-word and expected nonword-nonword errors, and for phonetic distances between initial phonemes of targets and expected errors of 1, 2 and 3 distinctive features. These data were analyzed in a univariate analysis of variance with subjects as repeated measures; main factors were word-word versus nonword-nonword expected spoonerisms, and number of features difference between onset consonants of the expected spoonerisms. Both main factors and their interaction are significant [word-word versus

ISBN 1-876346-48-5 © 2003 UAB

nonword-nonword: F(1,49)=18.0, p<.001; number of features: F(2,98)=15.5, p<.001; Interaction: F(2,98)=3.3, p=.043]. The random main effect of subjects was significant [F(49,1107)=12.7, p<.001] but its interactions with the fixed effects were not. These effects require a model of speech production with phoneme-to-word feedback. In the absence of phoneme-to-word feedback the production system can have no access to the lexical status of an expected spoonerism, that has not been actually made. The extra activation caused by phoneme-to-word feedback apparently speeds up the firing of the target units. This interpretation is supported by the interaction between response time and lexical status, an interaction predicted from a model with phoneme-to-word feedback.

4. DISCUSSION

The current data on response latencies can be accounted for by a spreading-activation model of speech production like WEAVER++ [5] with a) an added phoneme-to-word feedback loop and b) a self-monitoring component using the speech perception system taking both inner and overt speech as an input and employing a criterion of lexicality. In WEAVER++ a unit (word, phoneme, or syllable) exceeds its firing threshold earlier with stronger activation.

Phonological priming decreases activation for the target phoneme by diverting activation to the expected error phoneme. This explains the effect of phonological priming in response times.

Phoneme-to-word feedback increases activation of the target phonemes because they get activation from both the target words and the expected error words that share most phonemes with the target words. This is only the case for expected real-word errors, not for expected nonword errors, because nonwords have no representation in the mental lexicon. This increase in activation explains the effect of lexical status on response latencies in error-free and fluent responses. Note that this effect is considerable, in the order of 70 ms, for expected spoonerisms with only a single feature difference between target and expected error.

Although the data on response latencies demonstrate phoneme-to-word feedback, the data on the incidence of spoonerisms convincingly show that lexical bias is caused not by feedback but rather by self-monitoring being more effective for nonwords than for real words. Of course, logically it is possible that there are two causes of lexical bias. However, as argued elsewhere [6], the current data provide good quantitative reasons to believe that the number of spoonerisms that are covertly repaired in inner speech roughly equals the number of spoonerisms that have become overt. This is a very low number. The great bulk of error-free and fluent responses can not have been influenced by covert rejection and repairing. Output editing can never explain an effect of lexical status in the order of 70 ms, as found for a phonetic distance of a single feature, in error-free cases.

The data on the rejection rate of spoonerisms show an effect of phonetic distance between error and target for nonwords, and not for real words. A priori one would have expected effects of lexical status and phonetic distance to be independent of each other. The difference can be explained by the word-recognition part of self-monitoring: An erroneous real word is most likely to be recognized as itself, also when the difference is small. An erroneous nonword is likely to be recognized as the closest real word, which in the current experiment is the correct target word, at least if the phonetic distance is small. Increasing phonetic distance will rapidly lead to a higher rejection rate, causing the effect of phonetic distance to interact with lexical status.

Finally, it should be mentioned that the main reason for Levelt et al. [3] to reject phoneme-to-word feedback is that they see no function for it. This argument fails if we assume that phoneme-to-word feedback is not caused by a direct backward link in speech production, but rather by immediate connections between perception and production. If so, feedback would be an unavoidable side-effect of self-monitoring.

5. CONCLUSIONS

Lexical bias in phonological speech errors is caused by output editing. There is phoneme-to-word feedback in speech production, most likely as an unavoidable side-effect of self-monitoring.

REFERENCES

[1] Baars, B.J., Motley, M.T. & MacKay D., "Output editing for lexical status from artificially elicited slips of the tongue", *Journal of Verbal Learning and Verbal behavior*, **14**, pp. 382-391, 1975.

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

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

[4] Humphreys, K., "Lexical bias in speech errors", unpublished doctoral dissertation, University of Illinois at Urbana Champaign, 2002.

[5] Roelofs, A., "Storage and computation in spoken word production", in: S. Nooteboom, F. Weerman & F. Wijnen [Eds.] Storage and Computation in the Language Faculty, pp. 183-216, Dordrecht: Kluwer Academic Publishers, 2002.

[6] Nooteboom, S.G., "Does perceiving your own inner speech cause phoneme-to-word feedback in speech production:?", submitted for publication, 2003.

ISBN 1-876346-48-5 © 2003 UAB