STRATEGIES FOR EDITING OUT SPEECH ERRORS IN INNER SPEECH

Sieb Nooteboom & Hugo Quené

Utrecht institute of Linguistics OTS, Utrecht University
Sieb.Nooteboom@let.uu.nl

ABSTRACT

In a classical SLIP task spoonerisms are elicited with either a lexical or a nonlexical outcome. If the frequency of a particular class of responses is affected by the lexicality of the expected spoonerisms, this indicates that many such responses have replaced elicited spoonerisms in inner speech. This is shown in early interrupted speech errors and in completed speech errors that deviate from the elicited spoonerisms.

Keywords: Speech errors, lexical bias, feedback, self-monitoring, inner speech.

1. INTRODUCTION

In a classical SLIP (Spoonerisms of Laboratory-Induced Predisposition) task [1], consonant exchanges are elicited that have either a lexical outcome, as in BARN DOOR > DARN BORE, or a nonlexical outcome, as in BAD GAME > GAD BAME. (Recently, it has been made plausible that such 'consonant exchanges' really are full or partial exchanges of 'articulatory gestures' [4]. In this paper we stick to the notion 'spoonerisms' without committing us to either 'phonemes' or 'articulatory gestures' as the units involved). Spoonerisms with a lexical outcome were shown to be more frequent than those with a nonlexical outcome [1]. This so-called 'lexical bias' effect was explained from pre-articulatory editing of inner speech: Nonwords are supposed to be more frequently detected and repaired before speech is initiated than real words. This explanation is supported in [7, 8]. Others have explained lexical bias in phonological speech errors from feedback of activation between phonemes and words in the mental production of speech [2, 11]. Such feedback increases the activation of lexical but not of nonlexical errors, because the latter have no lexical representations. The two mechanisms do not exclude each other, and there is increasing evidence that the relative frequencies of real-word errors and nonword errors are affected both by feedback and by self-monitoring [5, 6].

An important source of evidence stems from errors in a SLIP task that are not identical with the elicited spoonerisms, for example 'early interrupted' speech errors, such as G..BAD GAME. If such interruptions are more frequent in the condition with expected lexical outcomes (i.e., positive lexical bias), this might reflect an underlying effect of feedback on the frequency of covert spoonerisms. If interruptions are less frequent in the condition with expected lexical outcomes (i.e., negative lexical bias), this might reflect that self-monitoring detects nonlexical errors more frequently than lexical ones. Note that interrupted errors are detected in inner speech, not overt speech, because speech fragments before interruption are generally shorter than humanly possible reaction times and offset-to-repair times are often close to 0 ms.

If in a SLIP task the frequency of speech errors, that are neither completed spoonerisms nor interrupted spoonerisms, is also significantly affected by the lexicality of elicited spoonerisms, then, we argue, that this would suggest that self-monitoring inner speech has replaced elicited spoonerisms with other speech errors. When these other errors are *more* frequent in the condition with expected lexical outcomes, this would possibly reflect an underlying effect of feedback on the frequency of the covert spoonerisms. When these other errors are *less* frequent in the condition with expected lexical outcomes, however, this would possibly reflect that self-monitoring more frequently detects and replaces nonlexical than lexical spoonerisms. Note that the standard argument in [1, 7, 8] would be that self-monitoring replaces the spoonerisms with the correct targets. Obviously, this can not be observed in the error counts.

Below we describe two experiments employing the SLIP technique for eliciting spoonerisms [1]. In the first experiment the participants are under considerable time pressure, and have the explicit order to correct any speech error as quickly as possible. In the second experiment, conditions are much more relaxed, and there is no instruction to correct the speech errors. This difference was introduced because published error patterns vary enormously from experiment to experiment, apparently as a function of relative time pressure.

2. EXPERIMENT 1

This experiment used the standard method for eliciting spoonerisms: Successive word pairs like DOVE BALL, DEER BACK, DARK BONE, BARN DOOR, are to be read silently. On a prompt, the last word pair seen, in this example BARN DOOR, has to be spoken aloud. The experiment investigated relative frequencies of completed spoonerisms, interrupted spoonerisms, speech errors sharing the initial consonant with the elicited spoonerisms, and unrelated speech errors, as a function of expected outcome. Participants were under considerable time pressure. Under time pressure we expected a negative lexical bias in the interruptions and possibly in other speech errors.

2.1. Method

2.1.1. Stimulus material

The priming and test word pairs all consisted of Dutch CVC-words. The test word pairs with nonlexical outcomes were derived from those with lexical outcomes by only changing the coda of each word. The precursor priming word pairs all had the reverse initial consonants as compared to the following test word pair. The last word pair priming for a spoonerism always had the same vowels as the test word pair. Each test and each base-line stimulus was preceded by five word pairs, the last three of which were, for the test stimuli, priming an exchange of the initial consonants. To these test and base-line stimuli were added 46 filler stimuli, 4 of which with 4 preceding pairs (no one primed for spoonerisms), 4 with 3 preceding word pairs (no one primed for spoonerisms), 12 with 2 preceding word pairs (6 of which primed for spoonerisms by both preceding word pairs), 8 with 1 preceding word pairs (4 primed for spoonerisms), and 18 with 0 preceding word pairs. The participants could never anticipate when a response had to be given, so that they had to pay full attention also to the first word pair seen after a response.

There were 2 stimulus lists, being complementary in the sense that the 18 word pairs that were primed for spoonerisms in the one list were identical to the 18 word pairs providing the base-line condition in the other list, and vice versa,

Independent variables were expected outcome, lexical vs nonlexical, and phonetic similarity of the to-be-spoonerized consonants, similar (1 feature) vs dissimilar (more than 1 feature). Dependent variables were error rates of the different categories of speech errors.

2.1.2. Participants

There were 102 participants, most of them students and employees of the faculty of humanities in Utrecht University with no known or self-reported hearing or speech deficit.

2.1.3. Procedure

Each participant 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 base-line stimuli, along with their priming or non-priming preceding word pairs, were presented was randomized and different for each participant. The order of the stimuli for each even-numbered participant was basically the same as the one for the immediately preceding odd-numbered participant, except that, by exchanging the stimulus lists (see 2.1.1.), base-line and test stimuli were interchanged. Fifty-one participants were, after the practice word pairs, presented with list 1 immediately followed by list 2, the 51 other participants were presented with list 2 immediately followed by list 1. After each test / base-line word pair a “?????”-prompt, meant to elicit pronunciation of the last word pair seen, was visible during 900 ms and followed by a loud buzz sound and blank screen, both of 100 ms duration. The participants were urgently instructed to speak the last word pair seen before this buzz sound. The buzz sound was immediately followed the Dutch word for “correction”, visible during 900 ms again followed by 100 ms with a blank screen, after which the first word pair of the following test cycle became visible. The participants were instructed to correct themselves immediately whenever they made an error. All speech of each participant was recorded and digitally stored.

2.1.4. Collecting the data
Responses to all test and stimulus presentations were transcribed by the first author, and classified as correct responses, completed spoonerisms, interrupted spoonerisms, so-called 'competing errors' sharing the initial consonant with elicited spoonerisms, and other, unrelated, speech errors. For the sake of brevity and convenience, results will be discussed for the 2 experiments together.

3. EXPERIMENT 2

Experiment 2 was set up to investigate whether under more relaxed conditions than those in Experiment 1, error patterns would be different.

3.1. Method

The method used was basically the same as the one applied in Experiment 1, with some modifications. There were no filler word pairs. This ensured that the participants could relax during the first two precursor word pairs, because these were never followed by a prompt to speak. There was no buzz sound before which a response had to be given and no urge for correction. None of the 102 participants had taken part in Experiment 1.

4. RESULTS OF BOTH EXPERIMENTS

Data of both experiments were analyzed with multinomial logistic regression followed by 250 two-stage bootstrap replications with replacement over 18 pairs of test word pairs, each pair consisting of the word pair with expected lexical outcome and the derived pair with expected nonlexical outcome. Differences between cells were evaluated by means of sign tests of the estimated means, using Bonferroni adjustment for multiple comparisons. Relevant results are given in Figure 1 for Experiment 1, and Figure 2 for Experiment 2.

In Figure 1, we see a significant positive lexical bias in the completed spoonerisms in both the similar ($p<.001$) and the dissimilar ($p<.001$) condition. The interruptions show a significant positive lexical bias in the similar ($p<.001$) and a significant negative lexical bias in the dissimilar condition ($p<.001$). The positive lexical bias in the similar condition for interruptions cannot be explained from self-monitoring: self-repairs following interrupted errors are overt, not covert. We interpret this effect as following from a feedback mechanism as proposed in [2, 11]. The negative lexical bias in the dissimilar condition cannot be explained from feedback. We interpret this as evidence that, at least under time pressure, nonlexical spoonerisms in inner speech are more often detected than lexical ones.

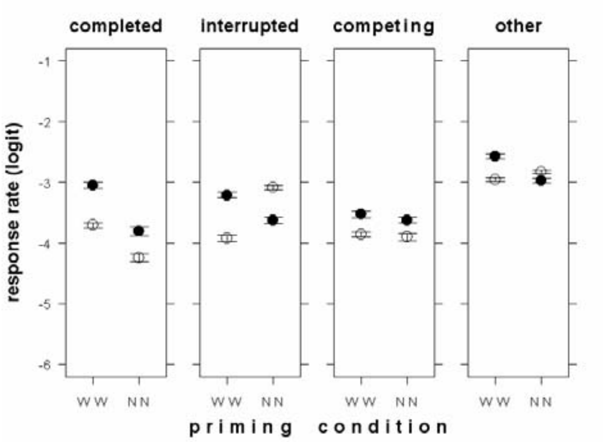

Figure 1: Observed logits of estimated error rates in Experiment 1, broken down by priming condition (WW=lexical; NN=nonlexical), by phonetic similarity or dissimilarity (filled and open symbols, respectively), and by response category (between panels). Error bars correspond to 95% confidence intervals of the bootstrapped logistic-regression coefficients (over 250 replications).

In line with [7, 8], we suppose that self-monitoring inner speech employs speech perception mechanisms that are also used in the perception of other-produced speech. These include a word recognizer. When no fitting word is found, an error is detected. However, if the error in inner speech is similar to the target, frequently the target may be recognized (incorrectly). This explains the absence of a negative lexical bias in the similar condition for the interruptions.

In the 'competing errors' lexical bias is absent for both similar ($p=.254$), and dissimilar consonants ($p=.259$). Note that this can either mean that the frequency of these errors is not affected by lexicality of the elicited spoonerisms, or that an effect of feedback is cancelled out by an effect of self-monitoring. The other, unrelated errors in the similar condition show a significant positive lexical bias ($p<.001$). In the dissimilar condition there is a significant negative lexical bias ($p<.001$). Apparently, the frequency of these 'other speech errors' is sensitive to the lexicality of the primed-for spoonerisms, again showing effects of both feedback (similar) and self-monitoring (dissimilar).

Figure 2 shows the relevant data obtained in Experiment 2, with little time pressure. Error rates are overall considerably lower than in Experiment

1, leading to more uncertainty in the data. The positive lexical bias in 'completed spoonerisms' is not significant for similar consonants (p=0.046, n.s. after Bonferroni correction) but is significant for dissimilar spoonerisms: p<.001).

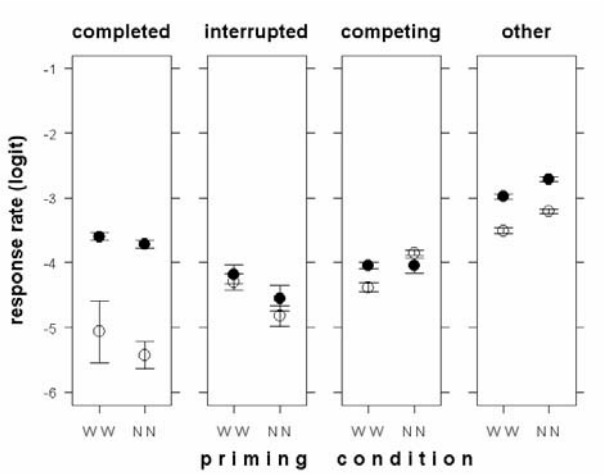

Figure 2. Results of Experiment 2, as for Figure 1.

In Experiment 2 we find a significant positive lexical bias in the interruptions for both similar (p<.001) and dissimilar (p<.001) condition, interpreted by us as an effect of feedback. The absence of a negative lexical bias in the dissimilar condition confirms that the variability of data patterns found in the literature, may be due to differences in time pressure [10].

In Figure 2 we find no positive or negative lexical bias in the competing errors for similar consonants (p=0.9), but, in contrast with Experiment 1, a considerable and significant negative lexical bias for dissimilar consonants (p<.001). Apparently, the frequency with which spoonerisms in inner speech are replaced with competing speech errors, is higher for nonlexical than for lexical ones, at least in the dissimilar condition.

As for the other, unrelated speech errors, there is a significant negative lexical bias in both the similar and the dissimilar condition, supposedly showing the effect of a lexicality criterion in self-monitoring inner speech. That in this case the effect is not much smaller in the similar than in the dissimilar condition, as it was found to be in the interruptions in Experiment 1, is probably because here there is much more time available. Response times (not reported here) for interruptions were found to be extremely short.

5. DISCUSSION

We interpret the patterns in the data described in the preceding section as demonstrating that the frequency of overt speech errors is affected by both feedback of activation, as proposed in [2, 11], and self-monitoring inner speech, employing a criterion of lexicality, as proposed in [7, 8]. Time pressure does not only increase the number of speech errors (cf. [3]), but it also affects the focus of self-monitoring, leading to interruptions (with time pressure) vs. to replacements of spoonerisms (without time pressure). This supports the proposal [5,6] that there are two sources of lexical bias: feedback and self-monitoring. A new finding is that elicited spoonerisms may not only be repaired, completed, or interrupted, but also may be replaced in inner speech with either form-related or unrelated other speech errors. Further details and discussion can be found in [10].

6. REFERENCES

[1] Baars, Motley, M.T. & MacKay, D.G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior, 14,* 382-391.

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

[3] Ganushchak, L.Y. & Schiller, N.O. (2006). Efects of time pressure on verbal self-monitoring: An ERP study. *Brain Research 1125,* 104-115.

[4] Goldstein, L., Pouplier,M., Chen, L., Saltzmann, E. Byrd, D. (2007). Dynamic action units slip in speech production errors. *Cognition 103,* 386-412.

[5] Hartsuiker, R.J. (2006). Are speech error patterns affected by a monitoring bias? *Language and Cognitive Processes, 21* (7–8), 856–891.

[6] Hartsuiker, R., Corley, M. and Martensen, H. (2005). The lexical bias effect is modulated by context, but the standard monitoring account doesn't fly: Related Beply to Baars, Motley, and MacKay (1975). *Journal of Memory and Language, 52*, 58–70.

[7] Levelt, W.J.M. (1989). *Speaking. From intention to articulation*. Cambridge Massachusetts: The MIT Press.

[8] 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.

[9] Nooteboom, S.G. (2005). Lexical bias revisited: Detecting, rejecting and repairing speech errors in inner speech. *Speech Communication, 47* (1-2), 43-58.

[10] Nooteboom, S.G. & Quené, H. (in press). Self-monitoring and feedback: A new attempt to find the main cause of lexical bias in phonological speech errors. *Journal of Memory and Language.*

[11] 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 143-186). London: Erlbaum.