How do speakers resist distraction? Evidence from a taboo picture-word interference task

Elisah Dhooge and Robert J. Hartsuiker

Department of Experimental Psychology, Ghent University

Running head: response exclusion

Word Count: 2478

Address for correspondence:

Elisah Dhooge

Department of Experimental Psychology

Ghent University

Henri Dunantlaan 2

9000 Ghent

Belgium

e-mail: Elisah.Dhooge@ugent.be

Abstract

Even in the context of irrelevant stimuli, word production is still a highly accurate and fluent process. But how do speakers prevent themselves from naming the wrong things? One possibility is that an attentional system inhibits task-irrelevant representations. Alternatively, a verbal self-monitoring system might check speech for accuracy and remove errors stemming from irrelevant information. Because self-monitoring is sensitive to social appropriateness, taboo errors should be intercepted more than neutral errors. To prevent embarrassment, speakers might also speak slower in the context of taboo distractors. Consistent with the self-monitoring account, two picture naming experiments, which varied in their emphasis on speed (Experiment 1) or accuracy (Experiment 2), found fewer incorrect naming of an irrelevant distractor word but longer picture naming latencies for pictures with taboo distractors. When intrusions of irrelevant words are highly undesirable, speakers do not just inhibit these words: the language production system adjusts itself to the context.

Introduction

One of the most intriguing features of word production is its effortlessness. For example, Levelt (1989) estimated that speakers produce about 150 words a minute with an accuracy of one error each thousand. This effortlessness becomes even more impressive given that speaking is not an isolated process but takes place in a social and perceptual context. As a result, speakers are continuously bombarded with many stimuli that could distract them. But apparently, the speech production system has mechanisms that handle distracting information well. For example, people usually do not have problems speaking with the television or radio on and even though large billboards on the side of the street may distract them, interference with speech production is rare (although it can happen, see Harley, 1984). The present study examines how people resist the temptation to name distracting words.

To do so, we used the picture-word interference task (i.e., PWI) in which participants see a picture with a superimposed word (i.e., the distractor). Their task is to name the picture and ignore the word. Previous research showed that participants can easily resist the temptation to name the distractor, indexed by low error rates (e.g., Dhooge & Hartsuiker, 2010). However, distracting information does influence the time course of word production. For example, compared to an unrelated distractor, a distractor from the same semantic category as the picture (e.g., DOG – cat) will slow down naming latencies whereas a phonologically related distractor (e.g., DOG – doll) will generally speed up picture naming latencies (e.g., Schriefers, Meyer, & Levelt, 1990).

Even though many studies investigated how and up to which level the distractor is processed (e.g., Levelt, Roelofs, & Meyer, 1999; Roelofs, 1997), it is unclear how distracting

information is *blocked out* of the production process instead of being spoken out loud.

According to the response exclusion hypothesis (Dhooge & Hartsuiker, 2010; Finkbeiner & Caramazza, 2006; Janssen, Schirm, Mahon & Caramazza, 2008; Mahon, Costa, Peterson, Vargas & Caramazza, 2007; Miozzo & Caramazza, 2003), speakers automatically formulate a covert verbal response when confronted with a word. This entails that in the PWI task, the name of the distractor always needs to be excluded from an output buffer before naming the picture.

Alternatively, the WEAVER++ model of word production (e.g., Roelofs, 1997; 2003) assumes that irrelevant information is blocked out of the production system by an attentional modulation. To name the picture, distractor information is filtered out at an early stage, resulting in less interference and increased ability to select the correct response. Importantly, both views assume that distracting information is processed and excluded from the system. However, what mechanism enables speakers to detect and remove the erroneous response from the speech production stream?

The goal of this paper is to assess whether this is accomplished by an established mechanism in speech production, the verbal self-monitor (e.g., Hartsuiker & Kolk, 2001; Levelt, 1989; Levelt et al., 1999). The monitor is a mechanism that allows speakers to attend their speech and check it for accuracy, allowing them to intercept errors and repair them. Importantly, it allows speakers to inspect internal speech. Thus, errors can be intercepted and corrected even before they are pronounced (e.g., Dell & Repka, 1992; Oomen, Postma, & Kolk, 2001). The monitor can be easily integrated into response exclusion accounts, because the monitor's function is to detect and block out unintended verbal responses.

The monitor seems to be sensitive to several criteria, such as social appropriateness. Although taboo words are quite common in daily life (with estimates ranging from .5% to .7%,

see Jay, 2009, for a discussion), the appropriateness criterion is relevant for the speaker because speech errors resulting in taboo words might be highly offensive to the listener and highly embarrassing for the speaker. Indeed, Motley, Camden and Baars (1981, 1982) found evidence for an appropriateness criterion using the so-called SLIP task. In this task, participants read word pairs of which sometimes one had to be read aloud (i.e., target pair). The word pairs preceding a target pair bias them into making spoonerisms on the target trials. Motley et al. found that taboo spoonerisms (e.g., tool kits -> cool tits) occurred less often than neutral errors. Furthermore, target trials that were read correctly but could have yielded a taboo error were accompanied by a higher galvanic skin response and longer naming latencies. Similarly, Severens Janssen, Kühn, Brass, and Hartsuiker (in press) showed that such trials elicited an EEG-effect; however, there was a floor effect in their error data. The SLIP task findings suggest that the taboo sequence was internally formulated, detected, and corrected, resulting in a slower response. This last observation is mirrored in a taboo Stroop task (e.g., Siegrist, 1995; MacKay, Shafto, Raylor, Marian, Abrams & Dyer, 2004). In the Stroop task, participants need to name the ink color of a word. If the word is socially inappropriate, it takes longer to name the color compared to when the word is neutral. Thus, speech becomes more careful.

Summarizing, up to now, no study has addressed the mechanism by which distracting information is filtered out of the speech production system. To investigate whether speakers handle distracting information by means of the monitor, we presented to be named pictures with taboo and neutral words. Experiment 1 elicited errors (i.e., naming the distractor instead of the picture) whereas in Experiment 2 naming latencies were of importance. We distinguish three possible scenarios. First, distracting information might be inhibited by an early, attentional system (e.g., In WEAVER++ this attentional mechanism is implemented as a production rule

which reduces the activation of the node representing the distractor word when it is the task to name the picture, Roelofs, 2003, p. 101). Assume that this system is sensitive to whether the to-be-inhibited item is socially appropriate or not. It is then possible that taboo words are inhibited more strongly than neutral words. This predicts that taboo words intrude less frequently in production (fewer errors) than neutral words and that they cause less interference in correct naming (faster reaction times). Second, the attentional system might inhibit taboo words *less* strongly than neutral words (for example because such words intrinsically capture attention and are difficult to ignore). This predicts more errors and slower reaction times with taboo distractors. Third, it is possible that word distractors are planned for production, but then need to be ruled out by the monitor that is sensitive to social appropriateness. This predicts that taboo words intrude less frequently than neutral words. Additionally, assuming that language production is adaptive to the situation at hand, speech might slow down when one risks making a taboo error. Thus, only a monitoring account predicts fewer distractor intrusions and slower naming times with taboo distractors. To test these hypotheses, Experiment 1 used instructions that emphasized speed, a manipulation that reliably elicits distractor intrusions (cf. Starreveld & La Heij, 1999). Experiment 2 used standard instructions, so as to evaluate naming latencies.

Experiment 1: Speed emphasis

Method

Participants. Twenty participants took part in the experiment. All reported normal or corrected to normal vision, and were native speakers of Dutch.

Design. Naming (2, named picture vs. named the distractor) was the dependent variable. Level of ‘tabooness’ was the within-subjects and within-items variable with two levels: taboo or

neutral. Note that due to technical limitations imposed by the voice key, the additional recording of naming latencies unfortunately proved to be impossible.

Materials. A pretest was used to select taboo words and neutral words. Sixty-one taboo words and 191 control words were presented to 33 further participants. One participant was excluded because Dutch was not his native language. Participants were asked to rate how taboo they found the words, on a scale from 1 to 7, with 1 being ‘not taboo at all’ and 7 ‘very taboo’. This resulted in the selection of 20 taboo words (taboo score range 3.94 – 5.16) and 20 control words (taboo score range 1.00 – 1.88). Twenty black and white pictures were selected from the Severens, Van Lommel, Ratinckx, and Hartsuiker (2005) database. Each picture was paired with a taboo word and a neutral word that was semantically and phonologically unrelated to the picture. Taboo and control words were matched perfectly on the number of letters and phonemes. Furthermore, there were no significant differences in log frequency, number of syllables, number of neighbors or bigram frequency (smallest p-value .16). There was a significant difference on tabooeness, $t(19) = 41.15$, $p < .001$. Details on the selected stimuli are presented in Table 1.

Insert Table 1 about here

Pictures were 300 x 300 pixels large. Distractors were presented in a Times New Roman 26 point font, in black capital letters. A plus sign served as fixation point.

Procedure. Before the start of the experiment, participants were familiarized with the pictures’ names. The experiment started with a practice phase of 24 trials. Next, the experiment proper started, consisting of three blocks. In a block, each picture was presented once with its neutral distractor and once with its taboo distractor. Each trial started with a fixation cross for

500 ms. After 500 ms, the picture and distractor appeared for 350 ms. Trials were randomized with the restrictions that (1) each picture had to be named once before pictures could be repeated, (2) stimuli from the same condition (i.e., taboo vs. neutral) could not appear on more than 3 consecutive trials, (3) for each 10 trials, five pictures had to be presented with a taboo word and five with a neutral word and (4) whether a picture appeared with its taboo or neutral distractor first was counterbalanced across participants.

Written instructions appeared on screen before the practice phase and were repeated before the experimental phase. Participants were informed that taboo words would be presented and that they could withdraw from the experiment at any time. Crucially, participants were asked to name the picture as fast as possible without worrying about errors (i.e., naming the distractor). They were informed that if they were not making errors, they were not responding fast enough. At each break, they were encouraged to speed up even more.

Results and discussion

Ten of the 2400 responses were scored as other than picture or distractor naming and were removed from the dataset. Responses were fitted using a mixed logit model that predicts the logit transformed likelihood of a picture naming response as implemented in the *lme4* library (Bates, 2007) in R (R Development Core Team, 2009). We included a random intercept for subjects and items. The independent variable significantly increased the fit of the model, $\chi^2(1) = 80.20$, $p < .001$. Importantly, the main effect of level of ‘tabooness’ was significant, $\beta_1 = 1.13$, $SE = 0.13$, Wald’s $Z = 8.58$, $p < .001$, Cohen’s $d = 1.59$. Thus, participants were less likely to name the distractor when it was a taboo word.

The findings are consistent with those of Motley et al.'s (1981) SLIP task experiments, using a different paradigm which furthermore elicits many more errors (14.6%) than the SLIP task (sometimes fewer than 1%; e.g., Severens et al., in press). The findings are compatible with one of the attentional accounts (stronger inhibition of taboo words) and the monitoring account. Experiment 2 will distinguish between them.

Insert Table 2 about here

Experiment 2: Accuracy emphasis

Method

Participants. Twenty further participants were tested.

Design, materials and procedure. All was identical to Experiment 1 except the instructions. Now, participants were asked to ignore the distractor and name the picture as fast as possible without sacrificing accuracy.

Results and discussion

All naming latencies exceeding the participant's mean by more than three standard deviations and naming latencies under 300 ms were discarded from analyses (1.54%). Errors (voice key malfunctioning and verbal disfluencies, and incorrect naming of the picture; 0.83%, cf. Glaser & Dungelhoff, 1984) were also removed. There were no distractor naming errors. Linear mixed effects models were fitted to the data, as implemented in the *lme4* library in R with a random intercept for participants and items. Including level of tabooeness significantly increased the fit of the model, $\chi^2(1) = 19.25$, $p < .001$. Importantly, pictures were named slower

when paired with a taboo distractor than when paired with a neutral distractor, $\beta_1 = -36.78$, $SE = 8.37$, $t = -4.40$, $p < .001$, Cohen's $d = .90$ (see Table 2).

These findings are in line with the self-monitoring hypothesis: detection of a taboo error leads to its interception and to a slow-down of speech production.

Discussion

Two experiments investigated whether the monitor is involved in eliminating irrelevant information from the speech production process. Experiment 1 showed fewer taboo than neutral errors, and Experiment 2 found longer naming latencies with taboo than with neutral distractors. These results indicate that dealing with distracting information is not about ignoring the distractor or not, pleading against an early, attentional mechanism dealing with distracting information. They can, however, be interpreted in terms of the monitor. According to this account, speakers want to avoid taboo errors because of their offensive and embarrassing nature. Because of this property, the monitor will catch and correct taboo errors more often than neutral errors. Additionally, to avoid taboo errors, detection of a taboo word makes speakers more careful, resulting in slower responses.

One might ask whether a monitoring account would predict *faster* naming latencies in the taboo condition, because taboo errors can be detected and removed from the buffer faster than neutral errors, leading to earlier buffering of the picture name. However, we tentatively assume that not only the response to the distractor is subject to monitoring; so is the response to the picture name. And if the monitoring process is adaptive, we expect it to be more stringent when it has just excluded a taboo word. Note that from a theoretical point of view it would be odd if the monitor is indeed the mechanism that checks responses in the PWI task, but is shut down

when a second response enters the buffer. Note that we assume a language production system that adjusts itself fully to the context of speech, not just on a macro level (i.e., general slowing of responses) but also on the micro level of a single trial. Therefore, it is evident that the data do not merely represent a speed-accuracy trade off. Subjects are not simply sacrificing accuracy in order to become faster. On the contrary, our data show that the speech production system adjusts itself in a ‘smart’ way to the context of speech, so that produced speech is not at odds with social customs.

In conclusion, the reason that speakers resist the temptation to say something irrelevant is not that they simply inhibit irrelevant information. On our account, they rather scrutinize what they are about to say and adapt the speech production system to the situation at hand.

References

- Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). The CELEX lexical database. CD-ROM. Philadelphia, PA.
- Bates, D.M. (2007). *Linear mixed model implementation in lme4*. Unpublished manuscript, University of Wisconsin, Madison.
- Dell, G. S., & Repka, R. J. (1992). Errors in inner speech. In B. J. Baars (Ed.), *Experimental slips and human error: Exploring the architecture of volition*. New York: Plenum Press.
- Dhooge, E., & Hartsuiker, R. J. (2010). The distractor frequency effect in picture-word interference: evidence for response exclusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 878-891.
- Finkbeiner, M., & Caramazza, A. (2006). Now you see it, now you don't: On turning semantic interference into facilitation in a Stroop-like task. *Cortex*, 42, 790–796.
- Glaser, W. R., & Dungelhoff, F. J. (1984). The time course of picture word interference. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 640-654.
- Harley, T.A. (1984). A critique of top-down independent levels models of speech production: Evidence from non-plan-intentional speech production. *Cognitive Science*, 8, 191-219.
- Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology*, 42, 113-157.
- Janssen, N., Schirm, W., Mahon, B. Z., & Caramazza, A. (2008). Semantic interference in a delayed naming task: Evidence for the response exclusion hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 249-256.
- Jay, T. (2009). The utility and ubiquity of taboo words. *Perspectives on Psychological Science*, 4, 153-161.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

- Levelt, W. M. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral & Brain Sciences*, 22, 1-75.
- MacKay, D.G., Shafto, M., Taylor, J.K., Marian, D.E., Abrams, L., & Dyer, J.R. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition*, 32, 474-488.
- Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 503-533.
- Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in picture–word interference paradigm. *Journal of Experimental Psychology: General*, 132, 228–252.
- Motley, M. T., Camden, C. T., & Baars, B. J. (1981). Towards verifying the assumptions of laboratory- induced slips of the tongue: The output-error and editing issues. *Human Communication Research*, 8, 3-15.
- 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, 578-594.
- Oomen, C. C. E., Postma, A., & Kolk, H. H. J. (2001). Prearticulatory and postarticulatory self-monitoring in Broca's aphasia. *Cortex*, 37(5), 627-641.
- R Core Development Team (2009). *R: A language and environment for statistical computing*. Available from the R Foundation for Statistical computing Web site: <http://www.R-project.org>.
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production.

- Cognition*, 64, 249-284.
- Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review*, 110, 88–125.
- Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production – picture word interference studies. *Journal of Memory and Language*, 29, 86-102.
- Severens, E., Janssen, I., Kühn, S., Brass, M., & Hartsuiker, R. J. (in press). When the brain tames the tongue: Covert editing of inappropriate language. *Psychophysiology*.
- Severens, E., Van Lommel, S., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, 119, 159-187.
- Siegrist, M. (1995). Effects of taboo words on color-naming performance on a Stroop test. *Perceptual & Motor Skills*, 81, 1119-1122.
- Starreveld, P. A., & LaHeij, W. (1999). Word substitution errors in a speeded picture-word task. *American Journal of Psychology*, 112, 521-553.

Author note

This research was supported by Research Foundation Flanders Grand FWO08/ASP/070 to Elisabeth Dhooge.

Table 1. Properties of items (standard deviations between parentheses) used in Experiment 1 and Experiment 2. For the stimulus list, see <http://users.ugent.be/~eldhooge/>.

	Taboo	Neutral
Log frequency *	0.54 (0.49)	0.54 (0.49)
Letters	5.55 (1.79)	5.55 (1.79)
Syllables	1.65 (0.67)	1.75 (0.79)
Phonemes	4.70 (1.66)	4.70 (1.72)
Neighbors *	8.55 (6.81)	7.10 (6.04)
Bigram frequency *	46179.80 (32564.11)	46385.35 (30996.01)
Taboo score	4.52 (0.33)	1.20 (0.21)

* based on Celex lexical database (Baayen, Piepenbrock, & Van Rijn, 1993)

Table 2. Percentage of distractor naming responses in Experiment 1; mean naming latencies (standard errors between parentheses) for Experiment 2.

	Experiment 1	Experiment 2
	% Distractor naming	Naming latency (ms)
Neutral	21%	803 (14)
Taboo	9%	841 (21)