Journal of Memory and Language 95 (2017) 19–35

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Self-monitoring for speech errors: Two-stage detection and repair with and without auditory feedback

Sieb G. Nooteboom *, Hugo Quené

Utrecht University, Utrecht Institute of Linguistics OTS, Trans 10, 3512 JK Utrecht, The Netherlands

ARTICLE INFO

Article history:
Received 21 March 2016
revision received 22 December 2016
Available online 16 February 2017

Keywords:
Speech errors
Self-monitoring
Repairs
Audition
Somatosensory

ABSTRACT

Two experiments are reported, eliciting segmental speech errors and self-repairs. Error frequencies, detection frequencies, error-to-cutoff times and cutoff-to-repair times were assessed with and without auditory feedback, for errors against four types of segmental oppositions. Main hypotheses are (a) prearticulatory and postarticulatory detection of errors is reflected in a bimodal distribution of error-to-cutoff times; (b) after postarticulatory error detection repairs need to be planned in a time-consuming way, but not after prearticulatory detection; (c) postarticulatory error detection depends on auditory feedback. Results confirm hypotheses (a) and (b) but not (c). Internal and external detection are temporally separated by some 500 ms on average, fast and slow repairs by some 700 ms. Error detection does not depend on audition. This seems self-evident for prearticulatory but not for postarticulatory error detection. Theoretical implications of these findings are discussed.

Introduction

The main questions

This paper is about self-monitoring for speech errors during speech production. We know that speakers often detect their own speech errors, because in spontaneous speech more than 50% of all speech errors against sound forms are repaired by the speaker (cf. Levelt, 1983; Levelt, 1989; Nooteboom, 1980; Nooteboom, 2005a). Also other types of speech errors are often repaired. This paper asks whether we can classify observed repairs into speech errors detected by self-monitoring before and after speech initiation, and if so, how we can distinguish between these two classes of repaired speech errors; whether there are two different processes for repairing a speech error, one leading to very fast and one leading to slow repairs; and to what extent the detection of speech errors by self-monitoring depends on auditory feedback.

Typical examples of repaired speech errors, taken from Blackmer and Mitton (1991), are the following:

"if Quebec can have a ba/ a Bill 101"

"behownd her/ behind her own closed doors"

The "/" in both cases indicates speech interruption, often followed by a silent interval. These two examples differ in an interesting way: In the first example the speech fragment containing the error "ba/ " is very short, and in many such cases fragments like these are shorter than a humanly possible reaction time. As pointed out by Levelt (1983) and Levelt (1989), who gave the example "v/ horizontal" in which the "v" is supposed to be the first speech sound of the word "vertical", this demonstrates that speech errors can be detected before speech initiation. However, the number of speech sounds spoken before interruption is not necessarily proof that the error was detected before speech initiation. We will call cases in which the error form is not fully spoken 'interruptions'.

In the second example we see that speech was only interrupted after both the word containing the error, "behownd" and the following word "her" were spoken. It is generally assumed that in such cases the speech error was detected by the speaker after speech initiation, via auditory perception of her or his own speech (Cf. Hartsuiker, Corley, & Martensen, 2005; Hartsuiker & Kolk, 2001; Levelt, 1989; Levelt, Roelofs, & Meyer, 1999; Hartsuiker, Kolk & Martensen, 2005; Nooteboom & Quené, 2008 and many others). Of course, a priori it is imaginable that also in this and similar cases the speech error was detected before speech initiation and the speaker just waited before interrupting the utterance, for example in order to gain time for planning a repair (cf. Seyfeddinipur, Kita, & Indefrey, 2008). We call such cases as "behownd her/ behind her own closed doors", in which the error form

* Corresponding author at: Cor Ruyslaan 20, 3584 GD Utrecht, The Netherlands.
E-mail address: S.G.Nooteboom@uu.nl (S.G. Nooteboom).

http://dx.doi.org/10.1016/j.jml.2017.01.007
0749-596X/

is fully spoken, 'completed', supposedly but not necessarily reflecting detection of speech errors in overt speech. Although we will initially distinguish between 'interrupted' and 'completed' spoken error forms, this initial classification will have to be replaced by another classification of repaired errors as probably detected in internal or in overt speech.

In this paper we focus on interactional segmental speech errors. Interactional errors are errors following from interaction between two different units in the speech program. Examples of interactional segmental errors include exchanges such as *Yew Nork* for *New York*, anticipations such as *Yew York* for *New York* and perseverations such as *New Nork* for *New York*. There are also segmental errors in which speech sounds are added or omitted under the influence of other speech sounds in the context. We will not consider additions and omissions, because we focus on experimentally elicited errors and we have not elicited additions and omissions. In this paper we will also not consider lexical, syntactic, semantic or appropriateness errors. There is no a priori reason to suppose that our results will also be valid for these other error categories. They probably are not, because temporal constraints on detecting and repairing segmental errors on the one hand and lexical, syntactic or semantic errors on the other hand appear to be rather different (cf. Nooteboom, 2005a). If indeed speakers can detect segmental speech errors both before and after speech initiation, this raises the question whether and how we can observationally distinguish between these two classes of repaired speech errors. This is the first question we will attempt to answer.

It has been shown, particularly with interrupted error forms, that frequently but not always, very short error-to-cutoff times are followed by very short cutoff-to-repair times, even of 0 ms (Blackmer & Mitton, 1991). Blackmer and Mitton concluded that in such cases a repair is available at the moment of speech interruption. This suggests that possibly there are two classes of repairs, distinguished by the moment the repair comes available to the speaker. If indeed this is the case, one may ask where this difference comes from. Thus the second question we will focus on is whether and how we can distinguish between fast and slow repairs, and if so, where this difference comes from. It seems reasonable to assume that there is an immediate connection with the detection of speech errors in internal versus overt speech. Later in this introduction we will explain why we think that after internal error detection it is often not necessary to plan a repair, whereas after external error detection often no repair is available, and a repair has to be planned in a time-consuming way.

Evidence in favour of the distinction between self-monitoring internal and self-monitoring overt speech is formed by demonstrations that the detection rate of speech errors by self-monitoring is affected negatively by loud masking noise (e.g. Lackner & Tuller, 1979; Oomen, Postma, & Kolk, 2001; Postma & Kolk, 1992; Postma & Noordanus, 1996). Because speech errors can be detected both before and after speech initiation, one would not expect that the error detection rate would drop to zero in the absence of auditory feedback. Effects of noise masking would be limited to error detection in overt speech, at least in as far as we assume that error detection in overt speech depends on hearing one's own voice, as is proposed by Levelt (1989) and Levelt et al. (1999). But some researchers believe that errors can be detected after speech initiation on the basis of somatosensory and / or proprioceptive feedback from the articulators (Hickok, 2012; Lackner, 1974; Pickering & Garrod, 2013). So far, the question to what extent error detection by self-monitoring overt speech depends on audition, remains unanswered. This is because we do not know which repaired speech errors are detected in internal and which in overt speech. If our attempt to distinguish between these two classes of repaired speech errors is successful, we can find out to what extent self-monitoring of overt speech depends on audition. This is the third main question we will try to answer in this paper.

The three main questions that we focus on in this paper are:

(1) Are speech errors detected by self-monitoring both before and after speech initiation, and if so, how can we distinguish between these two classes of detected speech errors?
(2) Can it be that there are two different processes for repairing a speech error, one leading to very fast and one leading to slow repairs?
(3) To what extent does the detection of speech errors by self-monitoring depend on auditory feedback?

Detection of speech errors before and after speech initiation

Theories of self-monitoring for speech errors are often classified as perception-based and production-based. For our purposes we consider so-called forward-modeling accounts of self-monitoring as a third category. The most influential theory of self-monitoring for speech errors is the perceptual loop theory proposed by Levelt (1989) and Levelt et al. (1999). In this theory both error detection in internal speech and error detection in overt speech employ the same speech comprehension system that is also employed in listening to other-produced speech. Internal speech is fed into the speech comprehension system directly, not following the route via articulation, acoustics and audition. It is assumed that errors made during the mental generation of speech, for example errors in phonological encoding, can be detected and repaired before speech initiation, leading to so-called covert repairs or 'prepairs' (cf. Postma and Kolk, 1992; Schlenk, Huber, & Wilmes, 1987). In this paper we will not consider covert repairs because we have no relevant observational evidence. Nevertheless errors detected in internal speech lend themselves to investigation because they are often articulated, leading to so-called early interruptions as in "if Quebec can have a ba/ a Bill 101". According to the perceptual loop theory errors can also be detected in overt speech, via audition and speech comprehension. For both the internal and external loop, the output of the speech comprehension system is fed into a centrally located monitor by which errors can be detected and repair planning initiated. Repair planning is supposed to start at speech interruption. It should be pointed out here that, although it may be convincingly argued that very short error-to-cutoff times necessarily correspond to speech errors detected in internal speech (simply because error-to-cutoff time is shorter than a humanly possible reaction time), this is not necessarily so for all interrupted error forms. The distinction 'interrupted' versus 'completed' does not necessarily correspond to the distinction 'internally' versus 'externally' detected. In this paper it will be attempted to find a way to tell at least statistically which repaired speech errors were detected in internal speech and which were detected in overt speech.

In production-based theories of self-monitoring it is assumed that there is some mechanism or mechanisms within the mental process of speech generation by which errors are detected. Examples are provided by Laver (1980; see also Schlenk et al., 1987), assuming special purpose editors within the speech generation system, and MacKay (1987), proposing that, because a speech error in some sense is a relatively new structure, it will cause prolonged activation of some node in the neural network generating speech; this prolonged activation will increase awareness and thereby lead to error detection. A different mechanism for error detection is proposed by Nozari, Dell, and Schwartz (2011). These authors made a model of error detection by conflict between simultaneously activated and competing units during speech coding. Interestingly, these production-based monitors are all directed at error detection in internal speech, before speech initiation. It is generally assumed