Timing and selective attention in self-monitoring for speech errors.

Sieb Nooteboom and Hugo Quené

Utrecht institute of Linguistics OTS, Utrecht University

Paper to be submitted to the Journal of Memory and Language

Running title: Timing and attention in self-monitoring

Corresponding author:

Name: Sieb Nooteboom

Address: Cor Ruyslaan 20

3584 gd UTRECHT

The Netherlands

Telephone: +31651212060

Email: s.g.nooteboom@uu.nl

Abstract

Some predictions are made on detecting and repairing speech errors, mainly from a computational model of self-monitoring (Hartsuiker & Kolk, 2001). These predictions were tested in two four-word tongue twister experiments eliciting such errors and their repairs, in word initial and medial position. Findings are: (1) The distributions of error-to-cutoff times, although truncated close to 0 ms, are nearly complete for both positions, implying that against prediction interruption takes more time than speech initiation. This also implies that so-called "prepairs" are rare. (2) The distributions of cutoff-to-repair times are censored at 0 ms, but cutoff-to-repair times are longer for medial than initial consonants, showing that against prediction repairing takes more time for medial than for initial errors. (3) Detection rate is much lower for medial than for initial consonants and decreases with position of the misspoken word in the tongue twister sequence. This probably reflects predicted variation in selective attention.

Key words: Speech errors, self-monitoring, timing, attention, tongue twisters

1. **Introduction** *The main questions*

The observed frequency with which segmental speech errors occur potentially differs for different positions in words and utterances. For example, it has been reported that segmental speech errors are more frequent in initial position in the word than in other positions. (Dell, 1986, 1988; Nooteboom & Quené, 2015a; Shattuck-Hufnagel, 1983, 1987, 1992; Wilshire, 1998). Also it has been reported that the frequency of segmental speech errors increases from earlier to later within intonational units (Choe & Redford, 2012). Typically, it is attempted to explain such effects from the organisation of the process of speech preparation. For example, the so-called word-onset effect was explained by Shattuck-Hufnagel by assuming that word-onset segments are treated differently than other segments in the process of serial ordering of segments. Dell (1986) assumed that word initial segments are activated more strongly than other segments during serial ordering. Nooteboom and Quené (2015a) argued that in spontaneous Dutch relative frequencies of segmental errors can be predicted rather precisely from the differences between the numbers of opportunities for interaction in different positions. However, such explanations do not take into account the potential effects of self-monitoring and repair on the observed frequencies of segmental speech errors. This is the main focus of the current paper. We will explore how the processes of self-monitoring and repair can affect the observed frequencies of interactional segmental speech errors in different positions in the word (lexical form) and in the utterance (or intonational unit). Our two main questions we will atttempt to answer in this paper are the following: (1) How are observed frequencies of segmental speech errors affected by the timing of various processes of error detection and repair during self-monitoring? (2) How are observed frequencies of segmental speech errors affected by variations in selective attention?

1.2. Timing in self-monitoring

To elaborate our first question we take our starting point in a computational implementation of the dual perceptual loop theory of Levelt, Roelofs and Meyer (1999), as reported by Hartsuiker and Kolk (2001, from now on to be referred to as H&K), with some additional assumptions described by Nooteboom and Quené (2017). This also means that the experiments to be described potentially provide an interesting test of some major aspects of this computational model. H&K provide a table with basic durations of time intervals in their model of phonological encoding and self-monitoring. This table is reproduced here as table 1.1.

Table 1.1. Basic durations of each time interval in speech generation and self-monitoring according to H&K. In the "Stage" column we added some terms between brackets for the sake of clarification. б stands for syllable, ω for lexical item c.q. lexical form.

Stage	Symbol	Duration (ms)	Per unit
Phonological encoding	T_{phon}	110	б
Selection (of action plan)	T_{sel}	100	ω
Command (to execute action plan)	T_{com}	100	б
Audition (in case of external monitoring)	T_{sel}	50	ω
Parsing (either internal or perceived lexical form)	T_{pars}	100	ω
Comparing (encoded or perceived form with correct target)	T_{comp}	50	ω
Interrupting (execution of action plan or overt speech)	T_{int}	150	ω
Restart planning (of a repair)	$T_{restart}$	50	ω

Note (by H&K). "Restart planning" is a parameter that represents the duration of repeated execution of selection processes before phonological encoding minus the time benefit from priming the to-be-selected units.

We add the following reflections and assumptions about the process of speech preparation, mainly taken from Nooteboom and Quené (2017):

Phonological encoding: This stage follows after activation of a lexical item or morpheme that provides a stress pattern and slots for segments and also provides the segments for filling these slots during phonological encoding (Levelt et al. 1999). The buffer for phonological encoding can contain more than one lexical item. Phonological encoding turns lexical items into pronounceable lexical forms. At this stage segments for similar positions can interact (within a lexical form syllable initial, medial or final consonants with syllable initial, medial or final consonants respectively; between lexical forms word initial, medial or final consonants and vowels with word initial, medial or final consonants and vowels respectively in comparable syllables; cf Nooteboom & Quené, 2015a). Interaction can lead to full replacement of a segment by a competing segment, but also to an articulatory blend of the two competing segments (cf. Goldstein et al., 2007; McMillan & Corley, 2010).

Selection of an action plan: We assume that selection of an action plan only occurs after phonological encoding of a lexical form is completed. This implies that in comparing timing aspects of self-monitoring for speech errors, the duration of phonological encoding is not involved. Let us call the moment phonological encoding of a lexical unit is completed and an action plan is to be selected T_1.

Command to execute action plan: Whereas, according to H&K, the selection of an action plan takes 100 ms per lexical form, the command to execute an action plan takes 100 ms per syllable. This implies that, for example, articulation is started 200 ms after T_1 for a one-syllable word, but 300 ms after T_1 for a two-syllable word.

Audition (in case of external monitoring): The dual perceptual loop theory by Levelt et al (1999) states that overt speech is monitored via audition. H&K assume that self-monitoring needs at least 50 ms of perceived speech to start parsing the overt speech in search of errors. Nooteboom and Quené (2017) have demonstrated that self-monitoring overt speech does not depend on audition. The implication is that after articulation has started parsing can be applied to somatosensory and / or proprioceptive feedback from the articulators (Lackner, 1974; Hickok, 2012; Pickering and Garrod, 2013). However, there is no reason to assume that this would change the timing involved.

Parsing (either internal or perceived form): The assumption here is that the time involved in parsing is independent of word length. This seems to imply that parsing is applied to a window of either internal speech or external speech, the window corresponding to 100 ms of speech, and the result is fed into comparing the error form with a correct target form.

Comparing (internal or perceived form with correct target): Here the assumption by H&K seems to be that comparing an error form with a correct target form takes 50 ms independent of word length. To us this seems unlikely. In self-monitoring overt speech the comparison, if done in something like real time, has to follow the perceived speech as it is realized in time. This obviously takes longer for long words than for short words. This also means that errors later in a word are detected later than errors early in the word. For self-monitoring internal speech we observe that parsing and comparing together can be interpreted as scanning a lexical form in internal speech for speech errors. This seems very similar to detecting a particular speech segment in an internal representation of a lexical form. Wheeldon and Levelt (1995) found in a phoneme detection task applied to internal speech, by having their subjects inspect unspoken Dutch translations of English words, that scanning internal speech costs time, the time involved roughly corresponding to speaking time. Later segments were detected later than earlier segments. This obviously has consequences for self-monitoring. Whereas according to Hartsuiker and Kolk (2001) parsing and comparing together cost 150 ms after T_1, our assumption that parsing and comparing together can be seen as time-consuming scanning the lexical form for errors, implies that these 150 ms may be correct for initial segments, but if so, later segments would be detected later with respect to T_1, the difference roughly corresponding to speaking time. This means that, both in self-monitoring internal speech and in self-monitoring overt speech, errors against later segments are detected later than errors against earlier segments, the difference between earlier and later segments being roughly the same for internal and overt speech. The reader may observe that we can calculate from the H&K implementation that the time gap between error detection in internal and in overt speech is in the order of 350 ms. Nooteboom and Quené (2017) have found that in their experiments this time gap was in the order of 500 ms. If our assumptions about self-monitoring being scanning for errors are correct, this time gap of 500 ms only holds when we compare error detection in similar positions in the word.

Interrupting (execution of action plan or overt speech): For a two-syllable word, speech initiation after T_1, i.e. after phonological encoding of a lexical form is completed, would cost 300 ms (100 ms per lexical form for selection of an action plan and 100 ms per syllable for the command to execute the action plan). Interestingly, error detection plus interrupting also would cost 300 ms for a two-syllable word, viz. 150 ms for parsing plus comparing and 150 ms for interruption. This means that in self-monitoring internal speech for a two-syllable word the moment of speech initiation and the moment of interrupting are supposed to coincide. It is natural to assume that there is some statistical noise in all processes involved (such noise is also modeled in the H&K implementation). If so, the moment of interruption will sometimes be earlier and sometimes be later than the moment of speech initiation. If the moment of speech initiation falls before the moment of interruption, the speech error detected in internal speech will not be spoken. The error remains covert, hidden from observation. If this happens frequently, it can drastically change the frequency of observed segmental speech errors. The assumption that the moment of speech initiation and the moment of interruption on average coincide for two-syllable words leads to the following prediction:

Prediction 1: Error-to-cutoff times have a distribution that is truncated close to 0 ms, and has a (virtual) peak at 0 ms. (Note that error-to-cutoff times of precisely 0 ms or less reflect unobservable errors. They do not figure in a distribution of observable errors).

In case of word initial consonants, early interruption will cause that no part of the error word form is spoken. In case of, for example, a word medial consonant, the fragment of the word being spoken before interruption does not contain an error, as in *ba..uhh..baker*, where the internal error may have been *baper*. So here also, the error remains unobservable. The reader may note that if after error detection in internal speech the moment of speech initiation and the moment of interruption would coincide for word initial segments, from all assumptions made so far one would predict that these moments would also coincide in the same way for non-initial segments. This is so, because the moment of interruption is supposed to fall 150 ms after the moment of error detection. So, if the moment of error detection is a certain amount of time later for later segments, the moment of interruption will be later by the same amount of time. This leads to our second prediction:

Prediction 2: Error-to-cutoff times, i.e. time intervals between the onset of the error segment and the moment of interruption, are independent of the position of the error segment in the word.

Restart planning (of a repair): Hartsuiker and Kolk (2001) assume that after an error is detected, a repair is planned by restarting the whole process for selecting a lexical item before phonological encoding and then re-running phonological encoding, selecting an action plan and executing that action plan. They assign only 50 ms for this whole process and make no difference between repairing internally and externally detected errors. Nooteboom and Quené (2017) have found that, whereas a repair is often available very fast after error detection in internal speech, planning a repair for an externally detected error is extremely time consuming, taking some 900 ms on average in their experiments, with a minimum of some

400 ms and a maximum of some 1800 ms. They have explained the enormous difference in error-to-repair times between internally and externally detected errors by assuming that the correct target is available in competition with the error form after internal error detection, but after external error detection it has been de-activated during the time gap of 500 ms between internal and external error detection. In that view, 50 ms or even less for re-activating a correct target form that never was fully de-activated is realistic after internal but not after external error detection. We propose that phonological encoding of a lexical unit immediately after it has started is updated one or more times during further processing (this is a possible interpretation of what it would mean that the correct target form is sustained from the lexical level, cf. Nooteboom & Quené, 2017). This means that a second copy of the lexical form readied for being spoken is available for comparison with the first instant of the lexical form. In case a speech error is made during encoding of the first lexical form, the error can be detected by comparing the two versions. The idea that the correct target form is available and in competition with the error form immediately after phonological encoding is supported by the frequent cases of articulatory blends in segmental speech errors between correct and erroneous segment, as reported by Frisch & Wright (2002). Goldrick & Blumstein, (2006), Goldstein et al. (2007), McMillan and Corley (2010) and Mowrey and MacKay (1990). Of course, both the very short time reserved for re-activating the correct target introduced by Hartsuiker and Kolk, and our assumption that the correct target form is made available by phonological encoding for a second time, nearly simultaneously with the error form, explain that often a repair is available at the moment of speech interruption, even in cases with very short error-to-cutoff times, as has been demonstrated by Blackmer and Mitton (1991; see also Hartsuiker & Kolk, 2001 and Nooteboom & Quené, 2017).

It may be observed that if indeed repairs are often available at or before the moment of interruption, this implies that the distribution of cutoff-to-repair times is censored at 0 ms: The part of the density distribution of time intervals between the moment the speaking process is interrupted and the moment a repair comes available that falls before the error is spoken (i.e. with such time intervals less than 0 ms), is turned into cases with cutoff-to-repair times of precisely 0 ms. (The reader may have observed that our assumption about the distribution of cutoff-to-repair times, the distribution being censored, is different from our earlier assumption about the distribution of error-to-cutoff times. That distribution was supposed to be truncated). This leads to our third prediction:

Prediction 3. The distribution of cutoff-to-repair times is censored at 0 ms, cutoff-to-repair times of precisely 0 ms being overrepresented.

Because both the moment of interruption and the moment the process of repairing is started are triggered by the moment of error detection, we expect no difference in cutoff-to-repair times between initial and later segmental errors:

Prediction 4: Cutoff-to-repair times are independent of position of the error in the word.

1.3. Selective attention in self-monitoring

The above predictions follow from the computational model of Levelts perceptual loop theory of self-monitoring proposed by H&K, augmented by some assumptions proposed by Nooteboom and Quené (2017). However, this account of self-monitoring does not include effects of selective attention. Self-monitoring, as a semi-conscious process, is supposed to need some measure of selective attention (cf. Levelt, 1989: p. 463ff; also see Hartsuiker, Kolk & Martensen, 2005). Because other processes involved in speech preparation and speaking also need attention, the amount of attention paid to self-monitoring either internal or overt speech, may vary from moment to moment. That variation of the amount of selective attention paid to self-monitoring may have considerable effects, is probable given that overall rate of segmental errors that are detected and repaired by the speaker is in the order of 50%

(Nooteboom, 1980). In speaking a multi-word utterance, during speech preparation of a particular lexical form, going from word beginning to word end, more and more attention is needed for preparing the next lexical form. If that is indeed the case, one may expect that the amount of selective attention paid to self-monitoring decreases from word beginning to word end. This immediately leads to the following prediction:

Prediction 5: Within internal lexical forms, rate of error detection by self-monitoring decreases from earlier to later.

In multi-word utterances, for example four-word tongue twisters as often used in experiments eliciting segmental speech errors, often more than a single error is made during a single utterance. Therefore, on average more and more attention is absorbed by dealing with these errors, going from beginning to end of such utterances. This implies that less and less attention is available for each particular following error. Therefore one would expect that the probability of error detection will decrease from beginning to end in such an utterance:

Prediction 6: Within spoken utterances (not longer than a single intonational unit), rate of error detection decreases from earlier to later.

2. Experiment 1

Experiment 1 was originally not designed to test the above predictions. It was designed to investigate the rate of single segment interactional speech errors as a function of segment position in the word and segment position relative to stress (cf. Shattuck-Hufnagel, 1992). However, it appeared to us that the results of such an experiment are not easy to interpret when the effects of self-monitoring on observable error frequencies are unknown. As it happens, the design of the experiment is suitable to investigate effects of the time course of self-monitoring and variations of selective attention available for self-monitoring on observed error rates. This we set out to do.

2.1. Method of Experiment 1

The experiment consists of two parts. One part is a replication of Experiment 2 reported in Shattuck-Hufnagel (1992), but this time in Dutch. The other part is a modification of Shattuck-Hufnagel's experiment by employing two-syllable Dutch words only. The basic idea of the entire experiment is to elicit segmental errors by having participants rapidly and repeatedly speak word sequences that have properties of tongue twisters, and compare error frequencies between conditions that differ in what properties the potentially interacting consonants share or not share, in particular their word onset position and their pre-stress position. In one half of the experiment all sequences of four words had an initial and final monosyllabic word and two intermediate disyllabic words, replicating Shattuck-Hufnagel's (1992) Experiment 2. We will refer to these stimuli as the "1+2+2+1" stimuli. In the other half of the experiment all sequences of four words had disyllabic words only. We will refer to these stimuli as the "2+2+2+2" stimuli. This set-up makes it possible to compare a situation in which consonants sharing pre-stress position were in different positions within the word, viz. word initial and word medial, with a situation in which the consonants sharing pre-stress position were in the same position in the word, which was either word initial or word medial. It also makes it possible to compare segmental interactions between competing words similar or dissimilar in stress pattern, and words in four different positions in the utterance.

Stimuli

A basic unit in constructing the stimuli for the experiment was a quartet of stimuli for the four sharing conditions B, W, S, N, as exemplified for Dutch in the following two quartets, one for the "1+2+2+1" stimuli, and one for the "2+2+2+2" stimuli. In both quartets the potentially interacting consonants are **w** and **r**. The third consonant used by Shattuck-Hufnagel for eliciting unexpected errors, we do not indicate here because this gets fairly confusing in the "2+2+2+2" stimuli. For the sake of clarity the two potentially interacting consonants are given in bold face here, and the stressed vowels are marked as in "á". This was of course not done in the actual visual stimuli. The meaning of the four conditions B, W, S, and N is as follows:

Type B: The two consonants share both word onset position and pre-stress position.

Type W: The two consonants share word onset position but not pre-stress position.

Type S: The two consonants share pre-stress position but not word onset position.

Type N: The two consonants share neither position.

Here follows an example of a set of four stimuli, one for each of the conditions B, W, S, N, for 1+2+2+1 and the 2+2+2+2 stimuli separately:

Table 2.1. Examples of two corresponding sets of four stimuli. “**1+2+2+1**” = “a one syllable word + a two syllable word + a two-syllable word +a one-syllable word”, “**2+2+2+2**”= “four two syllable words”. **B** = all targeted consonants are both word initial and followed by stressed vowel; **W** = targeted consonants word initial, not followed by stressed vowel. **S** = targeted consonants not word initial, followed by stressed vowel; N = targeted consonants not word initial, not followed by stressed vowel.

condition	**stimulus type**	
	1+2+2+1	**2+2+2+2**
B	**w**ok **r**ápper **r**óeper **w**al	**w**áter **r**ápper **r**óeper **w**állen
W	**w**ad **r**appórt **r**apíer **w**ol	**w**óeker **r**appórt **r**apíer **w**íkkel
S	**w**in pa**r**íjs po**r**éus **w**el	be**w**íjs pa**r**íjs po**r**éus ju**wé**el
N	**w**it píe**r**en pá**r**el **w**as	la**w**áai píe**r**en pá**r**el ge**w**ín

As exemplified here, stimulus word pairs of the "2+2+2+2" type were derived from those of the "1+2+2+1" type. We have decided that such related quartets should not be presented to the same participant because this might be confusing. Therefore we created two lists of stimuli each with 12 quartets of the "1+2+2+1" type and 12 quartets of the "2+2+2+2" type, in such a way that for each quartet of the "1+2+2+1" type the corresponding quartet of the "2+2+2+2" type was in the other list and vice versa. Thus each list had 24 quartets and therefore 96 sequences of four words, 12 containing 1+2+2+1 stimuli and 12 containing 2+2+2+2 stimuli. The pairs of potentially interacting consonants were: 1: w/ʁ; 2: w/ʁ; 3: n/m; 4: n/m; 5: b/v; 6: v/b; 7: p/k; 8: k/p; 9: l/ʁ; 10:l/ʁ; 11: j/l; 12: j/l for one half of each list and 1: d/j; 2: z/d; 3: k/χ; 4: χ/k; 5: t/d; 6: t/d; 7: p/t; 8: d/z; 9: ʃ/s; 10: s/ʃ; 11 :v/z; 12: z/v for the other half of each list. The complete lists of stimulus word pairs, organized in quartets, are given in the Appendix A.

Participants

There were 28 participants, 20 females and 8 males, all students at Utrecht University. Their age ranged from 18 to 26. Data from one participant (female, even-numbered) were lost due to technical malfunction. The analysis reported below is based on the remaining 27 participants.

Procedure

Participants were tested individually, in a sound-treated booth, seated in front of a PC screen. The session started with an instruction appearing on the screen. This instruction, translated into English, ran as follows:

"Dear participant,

Thank you for participating in this experiment. Shortly you will see a sequence of four words on the screen. Read these four words aloud as fast as you can. You are to repeat the whole sequence of four words three times. Then you should push the blue button. As a result the words will disappear from the screen. You are to speak the sequence of four words once again, this time from memory. Repeat the sequence once again three times. Thereafter, push the blue button once again. The next sequence of four words will appear on the screen. In total there will be 96 sequences of four words. On the screen, bottom right, you can see how far you have come in the experiment. We start with a set of 10 practice items. Push the blue button to start the experiment".

There were 10 practice items specifically constructed for the purpose. In the test phase, the 96 word sequences were presented in random order to each odd-numbered participant. Each even-numbered participant got the same order of presentation as the immediately preceding odd-numbered participant, but then from List 2 instead of list 1 (cf. Appendix A). All speech produced by each participant in the test phase was recorded with a Sennheiser ME 50 microphone, and digitally stored on disk with a sampling frequency of 48,000 Hz. For each participant two separate audio files were recorded for each stimulus sequence of four words, one recording for the phase in which the words were visible on screen and one for the phase where the words were invisible and they had to be spoken from memory. Thus for each participant 192 audio files were created.

Scoring

All speech of each audio file of each participant was transcribed by the first author, with the help of an audiovisual display in PRAAT (Boersma and Weenink, 2009) in normal orthography or in phonetic transcription when necessary. For each response we recorded the number of the participant, the number of the trial, the stimulus identity, whether the stimulus was of the "1+2+2+1" type or of the "2+2+2+2" type, the condition B, W, S, or N, the list number, the two consonants for which interaction was expected, and whether the stimulus word sequence was visible or invisible. Also we categorized speech errors as "targeted" or not "targeted". "Targeted" were those single consonant substitutions that the condition was intended to elicit, "not targeted" were other single consonant substitutions, involving initial or medial consonants. As valid responses we counted (a) fluent and correct responses, (b) completed exchanges involving initial or medial consonants, (c) interrupted speech errors against single initial or medial consonants, (d) anticipations involving single initial or medial consonants, and (e) perseverations involving single initial or medial consonants. All other error types and errors involving other segmental positions were coded as invalid.

If a response contained more than a single valid speech error, these speech errors were categorized separately. When a participant repeated a stimulus word sequence more than 3 times either in the visible phase or in the invisible phase, the response utterances beyond the third response utterance were discarded. When a participant produced less than three response utterances either in the visible or in the invisible phase, the lacking utterances were counted as omissions, thereby becoming invalid responses. We also coded as invalid all those responses that did deviate from the intended stress pattern or from the intended segment pronunciation, because these responses did not accord with the experimental variables. An example is the spoken response "bot vázal vizier bit" to the stimulus "bot vazal vizier bit", where the participant erroneously stressed the first syllable of "vazal". Another example is the spoken response "geval naatsie neuzen gevang" to the stimulus "geval nazi neuzen gevang", where the participant employed the unintended pronunciation [ʦ] instead of [z] in "nazi".

Unfortunately, there appeared to be quite some hysteresis in the responses in the sense that when a participant made a particular speech error in response to a stimulus, quite often that speech error was repeated unchanged during the six responses to that stimulus. This, of course, violated the required independence of the successive errors made in response to that stimulus. For this reason we regarded as invalid all repetitions by the same speaker of a specific speech error to a certain stimulus. The analysis took each four-word sequence as a stimulus and considered all valid single consonant substitutions in initial or medial consonant position to this stimulus together as one "super response", separately for the visible phase and the invisible phase (but see below). The numbers of these valid interactional single consonant substitutions in initial or medial position were counted for each "super response". This formed a main dependent variable. Each substitutional error was coded as to the word in which it occurred from word 1 to word 4. Complete exchanges were in this respect coded as to the word in which the anticipatory part of the error occurred. Also each valid consonant substitution was coded for the positions, either initial or medial, of the two interacting consonants in the word form. This was done in 4 categories as follows:

1. i<i: Word initial consonant substituted by word initial consonant.
2. i<m: Word initial consonant substituted by word medial consonant.
3. m<i: Word medial consonant substituted by word initial consonant.
4. m<m: Word medial consonant substituted by word medial consonant.

Due to the strict constraints on the stimuli, the phonotactic opportunities for targeted errors in initial and medial position were exactly equal, at least in the "2+2+2+2" stimuli.

Each valid speech error was coded as "unrepaired" or "repaired". For all repaired valid speech errors we measured word onset-to-cutoff times (from the onset of the word containing the error to the moment of interruption) and error-to-cutoff times (from the onset of the error segment to the moment of interruption. Of course word onset-to-cutoff times and error-to-cutoff times are identical for errors in initial consonants. All repaired errors were classified as "interrupted" when the word containing the error was not completed or "not interrupted" when the error word was completed. We also measured cutoff-to-repair times in all valid errors. Error-to-repair time is defined as the sum of error-to-cutoff and cutoff-to-repair time.

2.2. Results of Experiment 1

Table 2.2 gives a first breakdown of the responses we obtained in this experiment.

Table 2.2. Numbers of responses obtained in Experiment 1. Invalid errors comprise all errors other than non-repeated single segment interactional substitutions in initial and medial position. “1+2+2+1” stands for a sequence of one syllable+ two syllables + two syllables + one syllable. “2+2+2+2” stands for a sequence of four two-syllable words.

	1+2+2+1	2+2+2+2	total
fluent correct	6244	5436	11680
valid errors	315	509	824
invalid errors	1509	2441	3950
total	8068*	8386*	16454

*Note: If only a single error could have been made per spoken tongue twister, these numbers would have been 27 x 48 x 6= 7776. The surplus stems from multiple errors per response utterance.

In our further analysis we mainly focus on the category of valid errors. From Table 2.2 it is clear that the 2+2+2+2 stimuli are much more successful than the 1+2+2+1 stimuli in eliciting interactional segmental errors. This was to be expected, because, given that initial and medial segments rarely interact with each other, segments in the 1+2+2+1 stimuli simply have fewer opportunities for interaction than segments in the 2+2+2+2 stimuli (for the strong effect of shared position on error frequencies see Nooteboom & Quené, 2015a). Also the lack of prosodic similarity between one-syllable and two-syllable words may be involved (Nooteboom & Quené, 2015b).

Table 2.3 gives a further breakdown of the valid errors by the positions of the intended error being elicited by the tongue twister and of the realized valid error. For example, in the stimulus “nut lamel limiet nies” the initial consonant position is targeted for interaction, and the elicited error (if any) is therefore in initial position. This applies to all 1+2+2+1 stimuli. The intended position of the error corresponds with the condition: B and W conditions are meant to elicit errors in word initial position, and the S and N positions to elicit errors in medial positions (cf Table 2.1 above).

Table 2.3. Numbers of valid errors, broken down by numbers of syllables in the stimulus, by position of the intended error (rows) and by position of the realized error (columns).

		position of realized error	
		initial	medial
1+2+2+1			
	initial (B, W)	201	8
	medial (S, N)	94	12
2+2+2+2			
	initial (B ,W)	215	70
	medial (S, N)	142	82

Table 2.3. shows that for the 1+2+2+1 stimuli, by far the most errors are indeed elicited in initial position, in accordance with the position targeted for interaction. Note that in the 1+2+2+1 stimuli medial consonants cannot be successfully targeted for interaction because they do not share this position in the four words. This is because initial consonants tend to interact with initial consonants and medial consonants with medial consonants (Nooteboom & Quené, 2015a). In this sense, in the 2+2+2+2 stimuli both positions are equal. Yet, we see that also in the 2+2+2+2 stimuli eliciting errors in initial position is more successful than eliciting errors in medial position. This was investigated by means of a Generalized Linear Mixed Model (GLMM, Quené & Van den Bergh, 2008) on the 2+2+2+2 stimuli only, with intended position as the only fixed predictor. Participants and item sets (matching stimuli) were included as random intercepts, and intended position was included as a random slope at the participant level. Results confirm the lower prevalence of realized errors when elicited in

medial position (beta=–0.2883, Z=–2.729, p=.0064) as compared to initial position (baseline, beta=–2.414). Interestingly, targeting errors in medial position yields most errors in the non-targeted, i.e. initial position (142/224=63%), similar to the pattern observed for errors elicited in initial position (215/285=75%): By far the most of the valid errors were made in initial position, irrespective of which position was targeted for interaction.

Table 2.4. Numbers of unrepaired and repaired single segmental interactional substitutions separately for 1+2+2+1 and 2+2+2+2 stimuli and for word initial and medial positions.

	1+2+2+1			2+2+2+2		
	initial	medial	total	initial	medial	total
unrepaired	82	9	91	73	59	132
repaired	213	11	224	284	93	377
tot	295	20	315	357	152	509

For testing our predictions 1 and 2, we focus on the error-to-cutoff times of the repaired errors. One positive outlier value exceeding 1000 ms was discarded (1 out of 601 observations). The error-to-cutoff times were analyzed by means of tobit regressions for censored data in R (Tobin, 1958; Kleiber & Zeileis; 2008, R Core Team, 2017). Although the error-to-cutoff times were all positive and above zero, tobit regression was used here because the same technique was used in the analysis of cutoff-to-repair times, reported below. Separate models were fit for 1+2+2+1 and for 2+2+2+2 stimuli, using lognormal distributions. Figure 2.1 shows the observed (histograms) and fitted (curves) distributions of lognormal error-to-cutoff times, as well as bootstrapped 95% confidence intervals (Efron & Tibshirami, 1993) for the means of the fitted distributions.

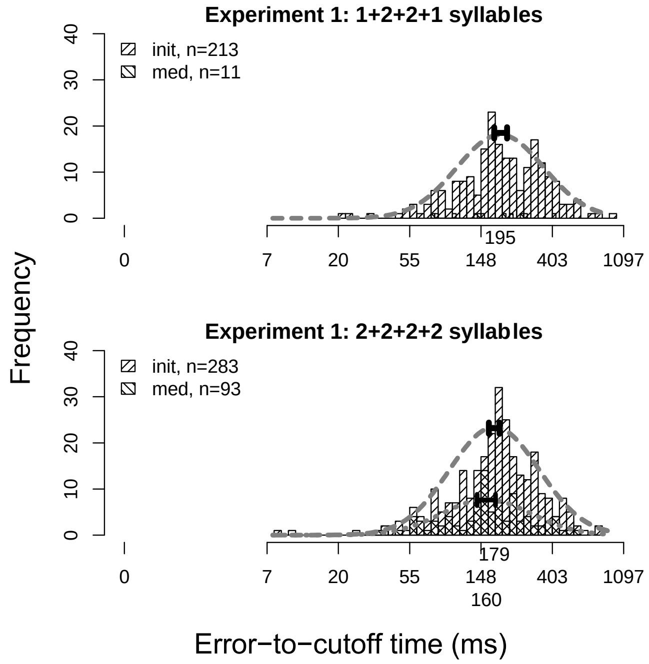

Figure 2.1: Histograms of observed error-to-cutoff times, broken down by syllable structure (upper panel 1+2+2+1 syllables, lower panel 2+2+2+2 syllables) and by position of the error, with lognormal density distributions fitted by a tobit regression model (initial errors: dashed, medial errors: dotted). The horizontal error bars near the peak of a distribution indicate the bootstrapped 95% confidence interval of the location of that peak (over 500 replications).

Our first prediction was that error-to-cutoff times would be truncated close to 0 ms with a (virtual) peak at 0 ms. The histograms in Figure 2.1 indicate that the error-to-cutoff times indeed are, if at all, truncated close to 0 ms (truncation at precisely 0 ms would have been impossible because error-to-cutoff times of 0 ms would render the errors unobservable), but also that the lognormal distributions are nearly complete. This implies that very often the moment of speech initiation and the average moment of interruption do not coincide. Note that when the error-to-cutoff time is close to 0 ms, the error must have been detected in internal speech (cf. Blackmer & Mitton, 1999). Errors detected in overt speech are relatively few and have error-to-cutoff times of many hundreds of ms (Nooteboom & Quené, 2017). That the peak of the distribution of error-to-cutoff times lies far above 0 ms but below 200 ms, can only mean that interruption after internal error detection often comes much later than speech initiation. The difference is in the order of 190 ms for the 1+2+2+1 stimuli and close to 180 ms for the 2+2+2+2 stimuli.

Our second prediction was that error-to-cutoff times are independent of position in the word. For the 1+2+2+1 stimuli, there are too few observations to test this prediction. For the 2+2+2+2 stimuli (Figure 2.1, lower panel), the tobit regression analysis did not yield a significant effect of consonant position (beta=–0.111 on lognormal scale, Z=–1.479, p=.139), conform our prediction. This suggests that, as expected, both the moment of internal error detection and the moment of interruption are shifted from early to later when we compare medial with initial consonant errors. This shift is in the order of 150 ms.

For testing our predictions 3 and 4, we focus on the *cutoff-to-repair times* of the repaired errors only. Outlier values exceeding 1000 ms were discarded (10 out of 601 observations). The cutoff-to-repair times were again analyzed by means of tobit regression for censored data (Tobin, 1958; Kleiber & Zeileis, 2008), separately for 1+2+2+1 and 2+2+2+2 stimuli, using lognormal distribution. Thus the cutoff-to-repair times scored as 0 ms are still included in the model, and contribute to the resulting estimated lognormal distribution. Figure 2.2 shows the observed and fitted lognormal distributions, as well as the bootstrapped 95% confidence intervals for the cell means of the fitted distributions.

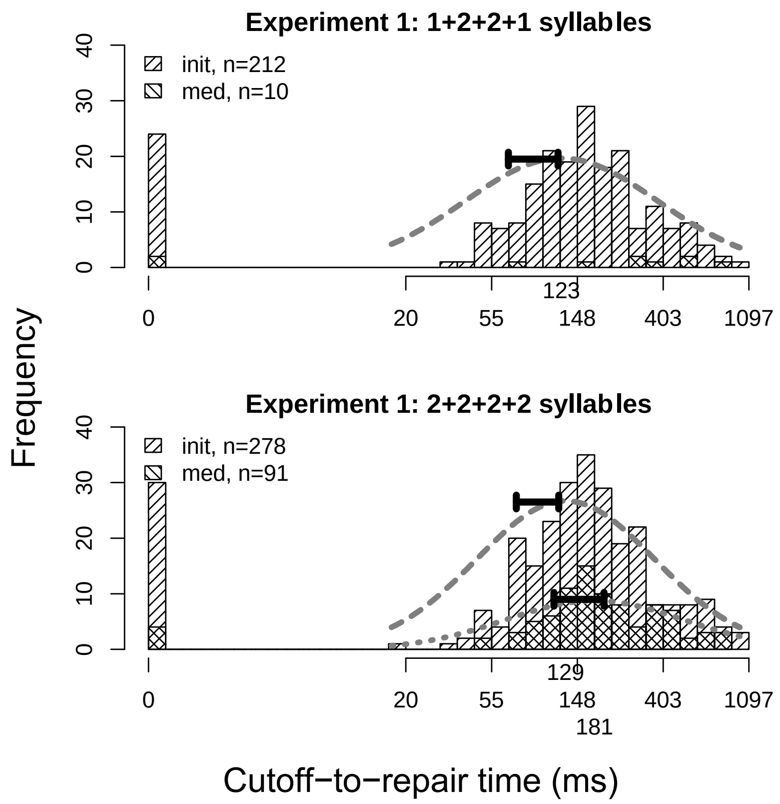

Figure 2.2. Histograms of observed cutoff-to-repair times, broken down by syllable structure (upper panel 1+2+2+1 syllables, lower panel 2+2+2+2 syllables) and by position of the error, with lognormal density distributions fitted by a tobit regression model (initial errors: dashed, medial errors: dotted, initial and medial errors combined: dashed+dotted). The horizontal error bars near the peak of a distribution indicate the bootstrapped 95% confidence interval of the location of that peak (over 500 replications).

Our third prediction was that the distribution of cutoff-to-repair times is censored at 0 ms, and cutoff-to-repair times of 0 ms are overrepresented (note that the clear separation of 0 ms from the rest of the distribution in Fig. 2.2 is an artifact of the lognormal distribution; see below for further discussion). Obviously, the three distributions with enough observations in Figure 2.2 are indeed censored at 0 ms, suggesting that quite a number of cases have an observed cutoff-to-repair time of 0 ms. These potentially would have had a negative value if only we could have assessed the actual moment that the repair came available to the mind of the speaker.

Our fourth prediction was that cutoff-to-repair times are independent of the position of the error in the word. For the 1+2+2+1 stimuli, there are too few repaired errors to test this prediction. For the 2+2+2+2 stimuli (Figure 2.2, lower panel), the tobit regression analysis yielded a significant effect of consonant position (beta=+0.338 on log scale, Z=+2.69, p=.0072), however, this position effect was only weakly supported by bootstrap validations (over 500 replications) of the model, as illustrated by the overlapping bootstrapped 95% confidence intervals of cell means in Figure 2.2 (lower panel).

A tobit regression analysis, as well as similar methods for censored or truncated data, assumes that all observations are generated by a single process, here fit by a lognormal distribution. This assumption is questionable here , however, for two reasons. From an empirical perspective, the estimated lognormal distributions fit somewhat poorly to the observed cutoff-to-repair times (see Figure 2.2); this fit did hardly improve when we attempted to fit the data to gaussian or weibull rather than lognormal distributions. For example, in Fig.2.2 (lower panel), the higher number of censored (0 ms) observations for initial consonants exert a stronger leftward pull on the centre of the estimated lognormal distribution, relative to the final consonants, even though the nonzero distributions seem to overlap at first glance. Secondly, from a theoretical perspective, there is no reason for us to assume that a single process, corresponding to a unimodal lognormal distribution, has generated the observed cutoff-to-repair times. Very likely the distribution captures both internally and externally detected repaired errors, which have very different temporal properties (cf. Nooteboom & Quené, 2017). For these reasons, we also inspected the odds of a repair being immediate (i.e., having a censored cutoff-to-repair time of 0 ms) vs non-immediate (cutoff-to-repair time longer than 0 ms), by means of a Generalized Linear Mixed Model (GLMM; Quené & Van den Bergh, 2008; Bates, Maechler, Bolker & Walker, 2016; R Core Team, 2017; participants and item sets were used as random intercepts). Models including Stimulus Type and Position as fixed predictors did not provide a better fit than the intercept-only model (which had beta=–2.5566, Z=–8.489, p<.0001), in spite of the different log odds for initial (-2.58) and medial (-3.22) consonants. Thus, neither consonant position nor stimulus type did affect the odds of a repair being immediate, if the random variation between participants and between item sets in these odds was taken into account. This may well have been due to a power problem, as the number of repaired errors per participant and per item have been too low to assess effects of interest.

Because our prediction is about the detection of errors in internal speech, we might also focus on cutoff-to-repair times of which we can be reasonably certain that they correspond to internally detected errors. We might therefore limit the observations to repaired errors with an error-to-cutoff time lower than 350 ms (n=327 responses; note that this selection is bases on short error-to-cutoff times, not on cutoff-to-repair times). We re-ran the tobit regression analyses on these selected lower-censored responses again separately for the two stimulus types, and again with position in the word (initial vs medial) as a fixed factor. For the 1+2+2+1 stimuli, there are again too few repaired errors for modeling. For the 2+2+2+2 stimuli, the tobit regression analysis on these selected responses yielded a significant effect of consonant position (beta=+0.388 on lognormal scale, Z=+2.891, p=.0038). If real, this effect implies that, against prediction, repairing speech errors in medial position is slower than repairing speech errors in initial position.

For testing our predictions 5 and 6, relating to possible effects of variation in selective attention, we focus on the *odds of detection* (detection rate) of the valid errors. Detected errors were coded as hits, and undetected errors as misses; these binomial responses were analyzed by means of a single mixed-effects Generalized Linear Model (GLMM; Quené & Van den Bergh, 2008), with position (initial vs medial), word number in stimulus (1 to 4) and stimulus structure (1+2+2+1 v s 2+2+2+2) as three fixed predictors. Participants and item sets (matching stimuli) were included as random intercepts, and position and word number were also included as random slopes at the participant level. The log odds of detection are summarized in Figure 2.3, broken down by the three fixed predictors in the GLMM.

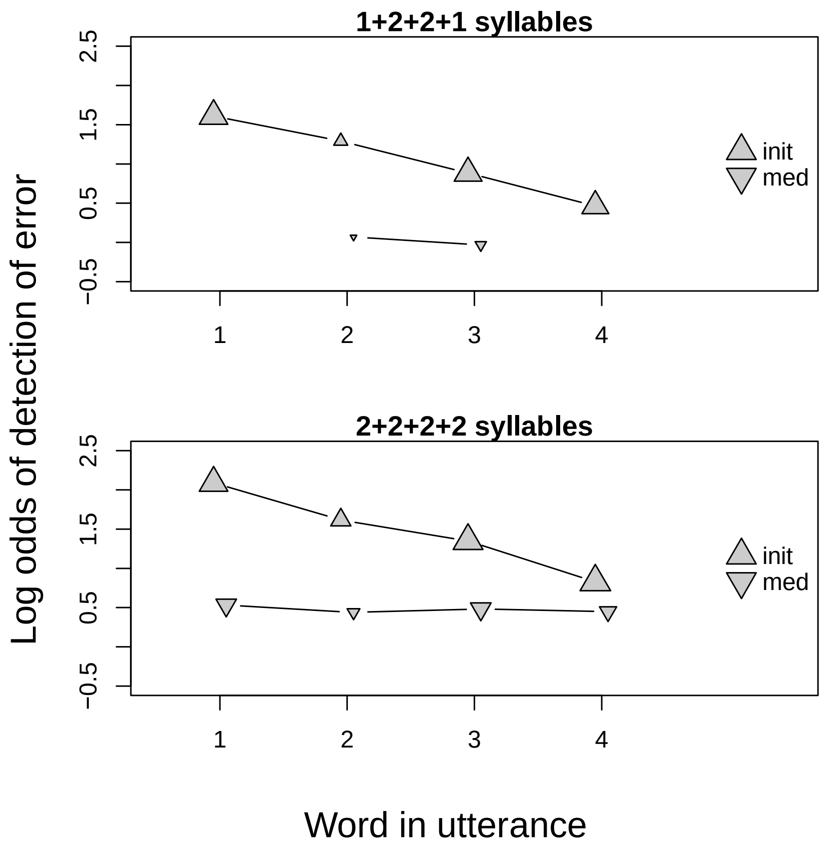

Figure 2.3: Estimated log odds of detection of valid errors, broken down by syllable structure (upper panel 1+2+2+1 syllables, lower panel 2+2+2+2 syllables), by position of the error (initial: upward triangles, medial: downward triangles), and by word number in the response utterance (1 to 4, along horizontal axis). Symbol sizes correspond with the numbers of detectable valid errors in each cell.

Our fifth prediction was that within spoken lexical forms, rate of error detection is higher for earlier segments (here in word-initial position) than for later segments (here in word-medial position). This predicted difference is indeed clearly visible in both panels of Figure 2.3, and is confirmed by the main effect of position in the GLMM (beta=–1.398, Z=–3.56, p=.0004). Speech errors in word-initial position have a much higher probability to be detected in self-monitoring than speech errors in word-medial position. This points at a difference between initial and medial consonants in amount of selective attention available for self-monitoring.

The sixth prediction was that within multi-word utterances (not longer than a single intonational unit), the odds of error detection decrease from earlier to later words in the utterance. This predicted effect of word number is also visible in both panels of Figure 2.3, and is also confirmed by the significant main effect of word number (beta=–0.3747, Z=–4.160, p<.0001). Although the patterns in Figure 2.3 suggest an interaction effect between position and word number, this was only marginally significant in the GLMM (beta=+0.320, Z=–1.794, p=.0728), most likely due to the low numbers of valid, detectable errors for word-medial consonants. Other main effects and interactions in the GLMM were not significant. The significant effect of word number suggests that amount of selective attention available for self-monitoring decreases from earlier to later within utterances. (The reader may also note that the total numbers of errors per word position, coded in the symbol sizes in Figure 2.3, does not increase from earlier to later as one would expect from the results reported by Choe & Redford, 2012).

2.3. Discussion of Experiment 1

In the current experiment we have set out to test some predictions of effects of timing and selective attention in self-monitoring on rates of detection of segmental speech errors in different positions We will briefly discuss the results in terms of our six predictions:

Prediction 1: Error to-cutoff times have a distribution that is truncated close to 0 ms and has a (virtual) peak at 0 ms. This prediction stems from the property of the H&K model that speech initiation of the error form and interruption of the speech process are simultaneous and take an equal amount of time. What we have found is that the distribution of error-to-cutoff times indeed is truncated at 0 ms, as predicted, but not close to the peak of the distribution. The distributions appear to be nearly complete. This cannot be attributed to the effect of externally detected errors, because we know from Nooteboom and Quené (2017) that the great majority of repaired errors reflect internally detected errors. Also, the results remains the same when we limit the analysis to error-to-cutoff times smaller than 350 ms. Apparently, in this experiment the processes of self-monitoring plus interruption take somewhat more time than the processes of speech preparation after completion of phonological encoding. But our result clearly confirms that the processes of speech preparation and the processes of error detection and interruption occur in parallel. In this aspect it supports an important aspect of the computational model by Hartsuiker and Kolk (2001). Our result also suggests that the number of cases in which the internal speech error remains invisible because the process of speaking is interrupted before the error is spoken, is very limited. The corresponding bias in observed rate of segmental errors is quite small.

Prediction 2: Error-to-cutoff times, i.e. time intervals between the onset of the error segment and the moment of interruption, are independent of the position of the error segment in the word. This prediction was made from our assumption that parsing an internal lexical form and comparing it with the intended form, may be looked at as scanning the internal form for errors. We indeed found no significant difference in error-to-cutoff times between initial and medial consonant errors. This should be interpreted against the background of the difference in word onset-to-interruption time between initial and medial consonant errors, that is found to be 155 ms. If internal error detection of initial and medial consonant errors were (nearly) simultaneous, we would have found that error-to-cutoff times would have been some 155 ms shorter for medial than for initial consonant errors. The absence of such a difference confirms that internal scanning for errors happens in something close to speaking time. This also implies that the observed error frequency is distorted by interruption falling before speech initiation in the same way for different positions in the word. Results also indicate that this distortion is very small.

Prediction 3: The distribution of cutoff-to-repair times is censored at 0 ms. We derived this prediction from the assumption that after internal error detection often a repair is available before the moment of interruption. In our view this is so, because the correct target form which is going to serve as a repair, often is active simultaneously with and competing with the error form, and not yet de-activated at the moment of early interruption after internal error detection. Inspecting the distributions of cutoff-to-repair times separately for the 1+2+2+1 and the 2+2+2+2 conditions and for initial and medial positions has shown that these distributions are censored at 0 ms. This suggests that the columns containing cases of 0 ms hide quite a number of cases where the actual moment a repair came available occurred sometime before the moment of interruption. Of course, this effect does not distort the observed error frequencies, because in all those cases the spoken error was observed. But it does confirm that after internal error detection often repairs are available before interruption. The distributions also show that cutoff-to-repair times may be very long, even in the order of 1000 ms (of course this implies that error-to-repair times, not reported here, are even much longer, cf. Nooteboom & Quené, 2017). As demonstrated by Nooteboom and Quené (2017), these long cutoff-to-repair times mostly stem from externally detected errors, where generating a repair costs much more time and effort because at the time of external detection the correct target has been de-activated.

Prediction 4: Cutoff-to-repair times are independent of position of the error in the word. This prediction was made because after internal error detection both interruption and repairing are started at the moment of error detection, and we had no a priori reason to assume that the time needed to repair would differ between the two positions. We found, against prediction, that cutoff-to-repair times were (marginally) significantly longer in medial than in initial positions. If confirmed, this would suggest that repairing medial consonant errors takes more time than repairing initial consonant errors. We will come back to this in reporting experiment 2.

Prediction 5: Within spoken lexical forms, rate of error detection decreases from earlier to later.

This was predicted from our assumption that amount of selective attention available for self-monitoring decreases during the scanning of a lexical form for errors, because more and more attention will be needed for preparing articulation of the next lexical form. We found a strong and highly significant difference between initial (c. 80%) versus medial (c. 60%) position in the rate of error detection. We explain this from a rapidly diminishing amount of selective attention available for self-monitoring during scanning of a lexical form for errors.

Prediction 6: Within spoken utterances (not longer than a single intonational unit), rate of error detection decreases from earlier to later. This was also predicted from variation in amount of selective attention available for self-monitoring. Particularly in the current task with tongue twisters meant to elicit segmental speech errors, very often more than a single error is made in one utterance. This supposedly decreases the selective attention for each speech error as we proceed from early to late in the utterance. Also this predicted effect was found to be significant, with percentages of repaired speech errors decreasing from 86% to 66% in initial position and from 63% to 53% for medial position. The data obtained in this experiment suggest that there may be an unpredicted interaction between the within-word and within-utterance effect of selective attention. However, this may have been the result of lack of statistical power. We will come back to this in describing experiment 2.

In sum, current findings confirm and refine a number of aspects of the Hartsuiker and Kolk computational implementation of the dual loop theory proposed by Levelt et al. (1999), and they confirm predicted effects of selective attention available for error detection in self-monitoring. Although the results basically confirm the predictions made, they are not always very robust. Below we will present experiment 2, which is limited to 2+2+2+2 stimuli, in order that we obtain more data at least for some of our predictions.

3. Experiment 2.

3.1. Method of Experiment 2

Experiment 1 has shown that the 1+2+2+1 stimuli, derived from the original Experiment 2 in Shattuck-Hufnagel (1992), are not very efficient in eliciting segmental interactions. In our Experiment 2 we will refrain from using these 1+2+2+1 stimuli. Instead we have opted for a setup that makes it possible to investigate the contribution of targeting specific consonant positions for interaction, by repeating or not repeating consonants in these positions, as in (2a), where interaction between /w/ and /r/ is elicited in initial position and (2b), where no such interaction is elicited by consonant repetition in initial position:

(2a) water rapper roeper wallen

(2b) water roeper lommer bikkel

We will refer to these two groups of stimuli as “eliciting” versus “not eliciting”. Note that these terms apply to a specific segmental position. We reserve the terms “targeted” versus “not targeted” for distinguishing between the specific position in which interaction is or is not elicited, such as the initial position in the above example, and other positions. In the examples above the initial position is “targeted” for interaction, but interaction is only elicited in the (2a), not in (2b).

Stimuli

Whereas in Experiment 1 we had created quartets of stimuli targeting the same consonants for interaction, we succeeded not always in doing this for Experiment 2, due to limitations in the Dutch vocabulary. Here follows an example of two sets of four stimuli, one stimulus for each of the conditions B, W, S, N, for eliciting and not eliciting stimuli separately:

Table 3.1. Examples of two corresponding sets of four stimuli. “Eliciting” = interaction provoked by consonant repetition in the targeted position; “not eliciting” = no interaction provoked by consonant repetition in the targeted position. **B** = all targeted consonants both word initial and followed by stressed vowel; **W** = word initial, not followed by stressed vowel. **S** = not word initial, followed by stressed vowel; N = not word initial, not followed by stressed vowel.

condition	eliciting	not eliciting
B	**w**áter **r**ápper **r**óeper **w**állen	**w**áter **r**óeper **l**ómmer **b**íkkel
W	**w**óeker **r**appórt **r**apíer **w**íkkel	**w**ijzer **p**arijs **d**orien **g**ozer
S	be**w**íjs pa**r**íjs po**r**éus ju**wé**el	be**w**íjs pa**r**íjs lo**k**aal ge**n**oot
N	la**w**áai píe**r**en pá**r**el ge**w**ín	la**w**áai píe**r**en bá**kk**en ge**s**óp

We again created two lists of stimuli each with 12 quartets of the "eliciting" type and 12 quartets of the "not eliciting" type, in such a way that for each quartet of the "eliciting" type the corresponding quartet of the "not eliciting" type was in the other list and vice versa. Thus each list had 24 quartets and therefore 96 sequences of four words. The complete lists of stimulus word pairs, organized in quartets, are given in the Appendix A.

Participants

There were 30 participants, 25 females and 5 males, all students at Utrecht University. Their age ranged from 18 to 53, with an average of 24.8. All participants reported having no hearing, speech or vision problems. They were paid for their participation.

Procedure

The procedure was the same as in Experiment 1.

Scoring

Scoring was the same as in Experiment 1, except that the labels "1+2+2+1" and "2+2+2+2" were replaced by the labels "eliciting" and "not eliciting".

3.2. Results of Experiment 2.

Table 3.2 gives first a breakdown of the responses we obtained in this experiment.

Table 3.2. Numbers of responses obtained in Experiment 1. Invalid errors comprise all errors other than non-repeated single segment interactional substitutions in initial and medial position.

	eliciting	not eliciting	total
fluent correct	6139	6752	12891
valid errors	509	467	976
invalid errors	2720	1985	4705
total	9368*	9204*	18572

*Note: If only a single error could have been made per response utterance, these numbers would have been 30 x 48 x 6= 8640. The surplus stems from multiple errors per response utterance.

In our further analysis we mainly focus on the category of valid errors.

Table 3.3 gives a further breakdown of the valid errors as to whether the error was "targeted" or not. By "targeted" we mean that the error was elicited in the "eliciting" condition (but not in the "not eliciting" condition) by the structure of the tongue twister. In a stimulus such as "vader bellen builen veter", in the eliciting condition the initial position is targeted for interaction by consonant repetition in that position. In the corresponding stimulus "vader bellen kommer pooier" in the "non eliciting" condition, although no interaction is elicited, we still refer to the initial position as "targeted", for reasons of comparison. This enables us to assess the effect of eliciting versus not eliciting interaction in a specific position under otherwise comparable conditions.

Table 3.3. Numbers of valid errors, broken down by eliciting versus not eliciting, by position of the intended error (rows) and by position of the realized error (columns).

		position of realized error	
		initial	medial
eliciting			
	initial (B, W)	206	76
	medial (S, N)	120	107
not eliciting			
	initial (B ,W)	149	70
	medial (S, N)	152	96

The effect of the position of elicitation on the number of errors in the two positions of realized errors together was investigated by a GLMM with the intended position as the only fixed predictor, with participants and item sets (matching stimuli) as random intercepts, and with intended position as random slope at the participant level. Results indicate that the overall error rates are approximately equal for errors in initial position (6.6%) and in medial position (6.2%) (beta=–0.076, Z=–1.06. p=0.289). Clearly, most of the valid errors were made in initial position, as shown in Table 3.3, irrespective of the position in which the error was elicited.

In Table 3.4. we provide a breakdown of all valid single interaction errors as initial versus medial and not repaired versus repaired.

Table 3.4. Valid single interactional substitution separately for initial and medial position and separately for not repaired and repaired errors

	initial	medial	total
unrepaired	211	206	417
repaired	416	142	558
total	627	348	975

For testing our predictions 1 and 2, we focus again, as we did in Experiment 1, on the error-to-cutoff times of the repaired errors. No observations exceeded the outlier criterion value of 1000 ms (all 558 valid observations remaining). Error-to-cutoff times were again analyzed by means of tobit regressions for censored data in R (Tobin, 1958; Kleiber & Zeileis, 2008; R Core Team, 2017), using a lognormal distribution.

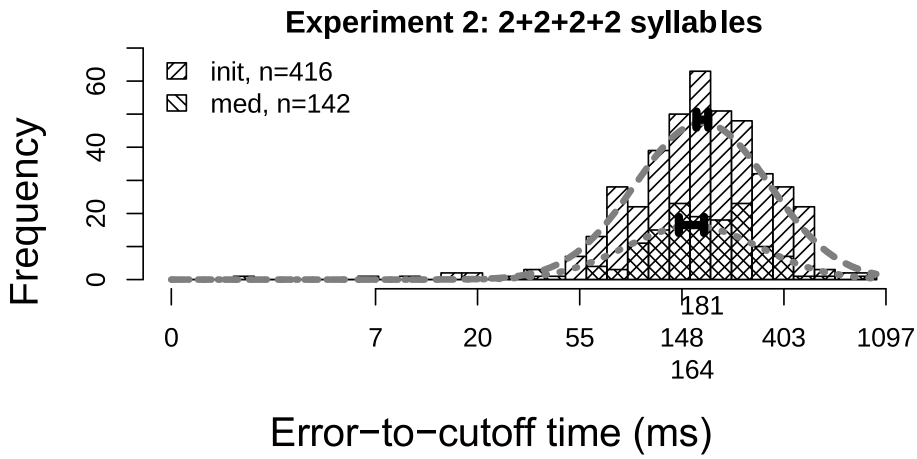

Figure 3.1. Histograms of observed error-to-cutoff times, broken down by position of the error, with lognormal density distributions fitted by a tobit regression model (initial errors: dashed, medial errors: dotted). The horizontal error bars near the peaks of the distributions indicate the bootstrapped 95% confidence interval of the location of that peak (over 500 replications).

Figure 3.1 shows the observed (histograms) and fitted (curves) distributions of lognormal error-to-cutoff times, as well as bootstrapped 95% confidence intervals (Efron & Tibshirami,

1993) for the mean of the fitted distribution. The histograms in Figure 3.1 indicate that the error-to-cutoff times are again nearly complete: If the distributions are truncated at all, they are so only in the lower end tails of the distributions. This implies that the average moment of speech initiation and the average moment of interruption do not coincide. Obviously, interruption after internal error detection is, on average, later than speech initiation. The difference is in the order of 180 ms for initial and in the order of 160 ms for medial consonant errors. This basically confirms what we found in Experiment 1.

Our second prediction was that error-to-cutoff times are independent of position in the word. The tobit regression analysis did not yield a significant effect of consonant position (beta=–0.098 on log scale, Z=–1.541, p=.123), conform our prediction. This suggests again that interruption requires the same amount of time after initial and after medial errors: later error detection in medial than in initial position (in absolute time) is compensated by later interruption in medial than in initial position (in absolute time).

For testing our predictions 3 and 4, we focus on the *cutoff-to-repair times* of the repaired errors only. Outlier values exceeding 1000 ms were discarded (20 out of 558 observations). The cutoff-to-repair times were again analyzed by means of tobit regression for censored data (Tobin, 1958; Kleiber & Zeileis, 2008; R Core Team, 2017), separately for initial and medial position, using lognormal distribution. Thus the cutoff-to-repair times scored as 0 ms are still included in the model, and contribute to the resulting estimated lognormal distribution. Figure 3.2 shows the observed and fitted lognormal distributions, as well as the bootstrapped 95% confidence intervals for the means of the fitted distributions.

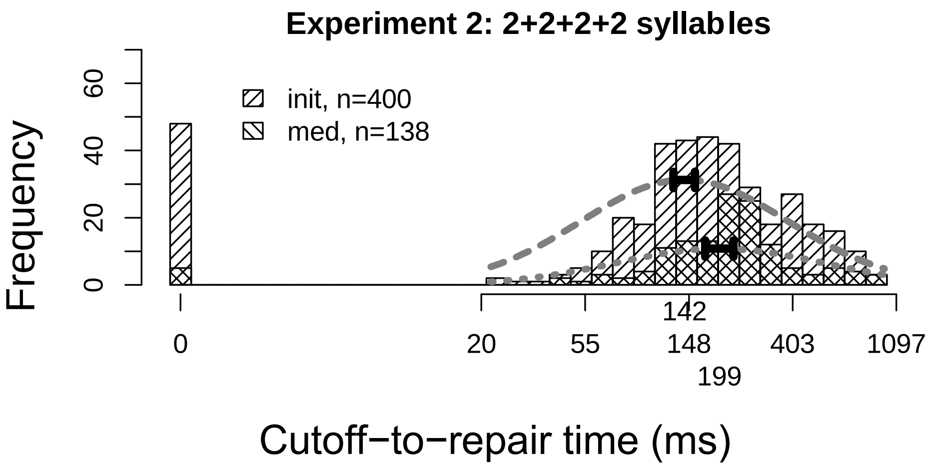

Figure 3.2. Histograms of observed cutoff-to-repair times, broken down by position of the error, with lognormal density distributions fitted by a tobit regression model (initial errors: dashed, medial errors: dotted. The horizontal error bars near the peak of a distribution indicate the bootstrapped 95% confidence interval of the location of that peak (over 500 replications).

Our third prediction was that the distribution of cutoff-to-repair times is censored at 0 ms. Obviously, the distributions in Figure 3.2 are indeed censored at 0 ms, which suggests once more that a number of cases showing immediate repairs correspond to negative time intervals between moment of interruption and the moment a repair comes available to the mind of the speaker. Of course, the censored distribution deviates from lognormal because of the overrepresentation of intervals of 0 ms. A second reason to expect a distribution deviating from lognormal is that the distribution reflects not only internally but also externally detected errors, the two underlying distributions separated by some 500 ms (cf. Nooteboom & Quené, 2017). We find indeed quite a number of cases with relatively long cutoff-to-repair times. If we omit two extremely long cutoff-to-repair times longer than 3.5 seconds, we find that the distribution runs from 0 to 1548 ms, with 72 cases longer than 500 ms.

Our fourth prediction was that cutoff-to-repair times are independent of the position of the error in the word. As was the case in Experiment 1, the tobit regression analysis yielded a significant effect of error position (beta=+0.336 on log scale, Z=+3.414, p=.0006). In this experiment, this position effect was indeed supported by the bootstrapped 95% confidence intervals of the peaks of the distributions (initial 128...157 ms, medial 174…228 ms). Again, we see that the tobit-modeled distributions fit somewhat poorly to the observed distributions of log-transformed error-to-cutoff times. For these reasons, we again inspected the odds of a repair being immediate (i.e., having a censored cutoff-to-repair time of 0 ms) vs non-immediate (cutoff-to-repair time longer than 0 ms), by means of a Generalized Linear Mixed Model (GLMM; Quené & Van den Bergh, 2008; Bates, Maechler, Bolker & Walker, 2016; R Core Team, 2017; participants and item sets were used as random intercepts). A GLMM including position as fixed predictor performed significantly better than the intercept-only model [Likelihood Ratio Test, chi2(1)=8.9, p=.0028; Pinheiro & Bates, 2000]. For initial consonant errors, the odds of a repair being immediate were 0.091 (or 8%), whereas for medial consonants the odds were significantly lower at 0.024 (or 2%) [beta=–1.32, Z=–2.62, p=.009].

As in Experiment 1, we re-ran the tobit regression analyses on the repaired errors with error-to-cutoff times lower than 350 ms (*n*=470 responses) again with position in the word (initial vs medial) as a fixed factor. And again, the tobit regression analysis with lower and upper censoring yielded a significant effect of consonant position (beta=+0.367 on lognormal scale, Z=+3.321, p=.0009). This confirms the preliminary finding of Experiment 1 that, against prediction, repairing speech errors in medial position is slower than repairing speech errors in initial position. The exact difference in cutoff-to-repair times between initial and medial consonant errors is difficult to assess, given the non-gaussian distributions, but Figure 3.2 suggests this difference to be in the order of 50 ms.

For testing our predictions 5 and 6, relating to possible effects of variation in selective attention, we focus again on the *odds of detection* (detection rate) of the valid errors, modeled by a single mixed-effects Generalized Linear Model (GLMM; Quené & Van den Bergh, 2008), with position (initial vs medial), word number in stimulus (1 to 4), and elicitation status (true: eliciting or false: not eliciting) as three fixed predictors. Participants and matching item sets (of matching stimuli) were included as random intercepts. Models including elicitation status or including this main effect plus its interactions did not perform better than models without these terms, according to Likelihood Ratio Tests [$\chi^2(1)=0.1422$ and $\chi^2(3)=0.1077$, respectively, both n.s.], so these terms were dropped from the GLMM. The log odds of detection are summarized in Figure 3.3, broken down by the two remaining fixed predictors in the GLMM.

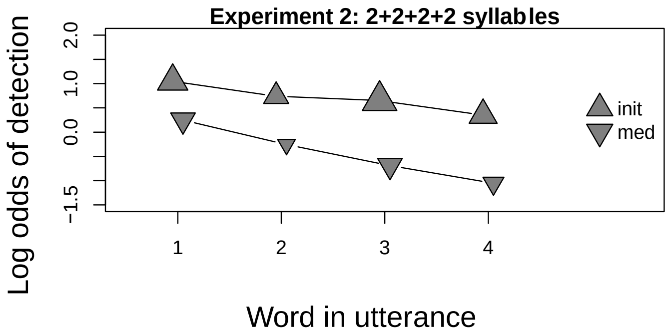

Figure 3.3: Estimated log odds of detection of valid errors, broken down by position of the error (initial: upward triangles, medial: downward triangles), and by word number in the response utterance (1 to 4, along horizontal axis). Symbol sizes correspond with the numbers of detectable valid errors in each cell.

Our fifth prediction was that within spoken lexical forms, rate of error detection is higher for earlier segments (here in word-initial position) than for later segments (here in word-medial position). This predicted difference is indeed clearly visible in both panels of Figure 2.3, and it is confirmed by the main effect of position in the GLMM (beta=–0.954, Z=–3.78, p=.0002). As in Experiment 1, speech errors in word-initial position have a much higher probability to be detected in self-monitoring than speech errors in word-medial position. Note that also the total numbers of detectable errors, as coded in the symbol sizes, are systematically lower in medial than in initial position.

The sixth prediction was that within multi-word utterances (not longer than a single intonational unit), the odds of error detection decrease from earlier to later words. This predicted effect of word number is also visible in Figure 3.3, and is also confirmed by the significant main effect of word number (beta=–0.237, Z=–2.83, p=.0046). The interaction effect between position and word number was not significant in the GLMM (beta=–0.205, Z=–1.517, p=.1293). The significant effect of word number, for initial and medial consonant errors, on the odds of detection confirms that the amount of selective attention available for error detection decreases from earlier to later words within utterances. (Note that the total numbers of detectable errors as coded in Figure 3.3 in the symbol sizes do not increase from early to late as one would expect from the results reported by Choe & Redford, 2012. There rather seems to be an alternating pattern).

3.3. Discussion of Experiment 2.

In Experiment 2 we have set out to see whether some results obtained in Experiment would stand further testing. We will shortly discuss the results of Experiment 2 in terms of our six predictions. Before we do that, we wish to point out that our data also show some unpredicted and unexpected results: In Experiment 1 the data suggested that eliciting versus not eliciting interaction between two consonantal segments by repetition of a consonant in a specific position has a rather strong effect in initial position, but not in medial position. This unexpected finding was confirmed in Experiment 2, in a much more convincing test, because now the “eliciting” and “not eliciting” stimuli were in all other respects comparable. We also found that the error rate is much higher in initial than in medial position. We will argue below that both effects are related to speakers’ selective attention.

Prediction 1: Error to-cutoff times, i.e. time intervals between onset of the error segment and the moment of interruption, have a distribution that is truncated close to 0 ms and the distribution has a (virtual) peak at 0 ms. What we had found in Experiment 1 is that the distribution of error-to-cutoff times indeed is truncated at 0 ms, as predicted, but not close to the peak of the distribution. The distributions appeared to be nearly complete. This finding is confirmed in Experiment 2. This implies that, although the processes of speech preparation and the processes of error detection and interruption occur in parallel, in the sense that they overlap in time, on average interruption is much slower that speech initiation. Of course, if speech initiation and interruption would have been exactly equally fast after phonological encoding is completed, the distribution of error-to-cutoff times would have its then unobservable peak at 0 ms. As it is, the peaks of the distributions are at about 181 ms for initial and 164 for medial position in Experiment 2, comfortably close to the peak values of 179 and 160 ms for initial and medial position in the 2+2+2+2 stimuli in Experiment 1.

Prediction 2: Error-to-cutoff times are independent of the position of the error segment in the word. As in Experiment 1, we again found no significant difference in error-to-cutoff times between initial and medial consonant errors. Because the shift in speaking time from initial to medial segments (of about 150 ms) is not reflected in error-to-cutoff times, this suggests that internal error detection comes later for medial than for initial segments, thus providing evidence that internal scanning for errors happens in something close to speaking time. This also implies that the observed error frequency is distorted by interruption falling before speech initiation in the same way for different positions in the word. However, this distortion appears to be very small, given that generally interruption of the speaking process follows, not precedes, speech initiation.

Prediction 3: The distribution of cutoff-to-repair times is censored at 0 ms. Inspecting the distributions of cutoff-to-repair times separately for the “eliciting” and the “non eliciting” conditions and for initial and medial positions has shown that these distributions indeed are censored at 0 ms. This suggests that, as in Experiment 1, the columns containing cases of 0 ms hide quite a number of cases where the actual moment a repair came available fell a varying amount of time before the moment of interruption. Of course, this effect does not distort the observed error frequencies, because in all those cases the spoken error was observed. This result demonstrates that after internal error detection often repairs are available before interruption. We also find confirmed that cutoff-to-repair times may be very long, even in the order of 1000 ms.

Prediction 4: Cutoff-to-repair times are independent of position of the error in the word. In Experiment 2 we find confirmed that cutoff-to-repair times are significantly longer in medial than in initial positions. This suggests that repairing medial consonant errors takes more time than repairing initial consonant errors. The difference is in the order of 50 ms.

Prediction 5: Within spoken lexical forms, rate of error detection decreases from earlier to later. If so, we expect that rate of error detection is much lower in medial than in initial position. This we found to be so in Experiment 1 and it is strongly confirmed in Experiment 2, for all four word positions in the tongue twisters. Simultaneously, we find that error frequency is much lower in medial than in initial position and, as we have seen earlier, that the effect of consonant repetition on error frequency is much lower in medial than in initial position. Possibly these three findings are related. We propose that they reflect variation in the amount of selective attention.

Prediction 6: Within spoken utterances (not longer than a single intonational unit), rate of error detection decreases from earlier to later. This predicted effect was found to be significant in Experiment 1, with percentages of repaired speech errors decreasing from 86% to 66% in initial position and from 63% to 53% for medial position. In Experiment 2 we find also this effect to be highly significant, with percentages of repaired speech errors decreasing from 73% to 59% for word initial positions and from 56% to 26% for word medial positions. The impression in Experiment 1 that there might be an interaction between the effects of position in the word and position in the utterance was not confirmed in Experiment 2. The two effects seem to be independent of each other. (The total numbers of detectable errors seem to follow a high-low-high-low zigzag pattern both for initial and medial consonants).

In sum, current findings again confirm and refine a number of aspects of the Hartsuiker and Kolk computational implementation of the dual loop theory proposed by Levelt et al. (1999), and also confirm proposed effects of predictable variation in the amount of selective attention available for error detection in self-monitoring. We will further discuss implications of these findings in the general discussion.

4. General discussion

We have attempted to answer in this paper two main questions, viz 1) How are observed frequencies of segmental speech errors affected by the timing of various processes of error detection and repair during self-monitoring? (2) How are observed frequencies of self-monitoring affected by variations in selective attention? Below we will discuss these two questions in that order. In answering the first question we will find occasion to discuss some aspects of the computational model of self-monitoring proposed by Hartsuiker and Kolk (2001). In answering the second question we will also attempt to explain some unexpected results in our experiments.

4.1. Effects of timing in self-monitoring

From the computational model of self-monitoring proposed by Hartsuiker and Kolk (2001), which is largely a model of timing of various processes in self-monitoring, one would predict that not all internally generated segmental errors become observable, overt, errors of speech. This is so because the model predicts that after internal error detection the moment of speech initiation of an error form and the moment of interrupting the speech process tend to coincide. All cases where interruption would occur before speech initiation remain unobservable. The current investigation is among other things an attempt to find out how frequencies of speech errors and their repairs are affected by the processes involved in self-monitoring. For practical reasons we have limited this investigation to interactional consonantal speech errors. There is no a priori reason to believe that our results would also have validity for other types of speech errors. The H&K model predicts that, at least for two-syllable words, both the moment of speech initiation and the moment of interrupting the speech process occur some 300 ms after phonological encoding of the lexical form has become completed. This would inevatibly lead to a distribution of error-to-cutoff times that is truncated close to 0 ms (error-to-cutoff times of precisely 0 ms would reflect unobservable errors). However, this is not what we find. The distributions of error-to-cutoff times are complete or nearly complete, both for initial and for medial consonant errors, in both experiments. If truncated at all, they are truncated in the lower tail of the distribution. If there are cases obscured because interrupting the process of speaking occurred before the initiation of overt speech, these are relatively few. This means that the set of speech errors for which an error-to-cutoff time can be obtained roughly corresponds to the set of all internally detected speech errors. This finding suggests that so-called “prepairs” (Schlenk, Huber, & Wilmes, 1987), i.e. cases where internally detected speech errors are not only detected but also repaired before speech initiation making both error and repair unobservable (cf. Levelt, 1989; p. 466; Levelt et al., 1999), are very rare. Note that there is nothing in our data to refute the assumption by Hartsuiker and Kolk that the command to initiate speech and the command to interrupt speech are executed in parallel. However, the H&K model could be improved on by assuming that after error detection, interrupting takes some 170 ms longer than initiating speech. The older assumption that interrupting only starts after speech is initiated (cf. Levelt, 1989), is refuted by the relatively many cases with error-to-cutoff times that are so short that there is not enough time for execution of an interruption command. In this sense the H&K model is corroborated by our data. (In our Experiment 2, for example, we find 89 error-to-cutoff times shorter that 100 ms).

The H&K model does not seem to make a difference between speech errors in different positions in the word. One way to interpret this is that both the timing of speech initiation and the timing of interruption are started at the moment of error detection. Of course, once speech is initiated speech errors later in the word are realized later than speech errors earlier in the word: a medial error comes later than an initial error, in our experiments by some 155 ms. Interestingly, the distributions of error-to-cutoff times in initial and medial position give information on the timing of error detection in internal speech: if there is no difference between these two distributions relative to 0 ms, then this implies that medial errors are detected in internal speech later than initial errors, in fact roughly the same amount of time as medial errors are spoken later than initial errors in the overt error form. Indeed no significant difference was found between initial and medial errors in error-to-cutoff times, which means that the detection of segmental errors can be compared to scanning the encoded lexical form in internal speech from early to later in something resembling real time. This ties in with a demonstration by Wheeldon and Levelt (1995) who reported a phoneme detection experiment with unspoken Dutch words, silently translated from English. They found that scanning a word internally for a particular segment is time-consuming and takes place in something close to real time.

We have assumed that after internal error detection a repair is rapidly available, because the correct target lexical form remains being activated from the lexical level whereas the competing error form is not. After external error detection, in overt speech, the correct target form would have been de-activated during the 500 ms delay separating internal from external error detection (Nooteboom & Quené, 2017). Therefore, after internal error detection, there would be relatively many cases in which the cutoff-to-repair time would be 0 ms, reflecting all cases in which a repair has come available to the speaker's mind at or before the moment of interruption. This led to our prediction that the distribution of cutoff-to-repair times would be censored at 0 ms. This is indeed what we found in both experiments. Of course, there are also a number of cases in which the error was detected externally. In these experiments we had no reliable way to separate between internally and externally detected errors, but on the basis of results obtained by Nooteboom and Quené (2017) it is reasonable to assume that there should be quite a number of cases with relatively long cutoff-to-repair times reflecting externally detected errors. This is what we found in both experiments. We have found both an overrepresentation of the number of cutoff-to-repair times of 0 ms and of cutoff-to-repair times longer than 1000 ms, which we interpret as confirmation of a result obtained by Nooteboom and Quené (2017), viz. that repairing after internal error detection is rapid whereas repairing after external error detection is slow, and this in turn suggests that there are two distinct processes of repairing a segmental speech error, differing in the time it takes to make a repair available. In this respect the H&K computational model can be improved on by incorporating these two different processes of repair.

A priori we saw no reason to expect that the cutoff-to-repair times would depend on position in the word. However, we actually found in both experiments that the relative number of cutoff-to-repair times of 0 ms (reflecting cases in which a repair was available at or before the moment of interruption) was significantly less in medial than in initial positions. This means that on average repairing a segmental speech error takes more time in medial than in initial position. Of course, as we will discuss below, we had predicted and we found that rate of repair is significantly lower in medial than in initial position. This prediction was made from our assumption that selective attention available for self-monitoring decreases from early to late within lexical forms. It seems not unnatural to assume that less selective attention not only leads to a decrease in rate of error detection but also to slower error detection. This brings us to our next subsection of this general discussion.

4.2. Effects of selective attention in self-monitoring

We have assumed that selective attention for self-monitoring internal speech decreases both within lexical forms and within utterances from early to late during speech preparation. From this we predicted that rate of internal error detection would decrease from early to late in both lexical forms and utterances. This is what we found. At the outset of this investigation we had, on the basis of the H&K computational model of self-monitoring, reason to suppose that a considerable percentage of internally detected errors remain unobservable because the speaking process would be interrupted before speech would be initiated. If that indeed would have been so, the effects of position in word and utterance discussed here would probably have interacted with the causes that make internally detected errors unobservable. However, given that, as discussed in the previous subsection, the distributions of error-to-cutoff times for both word initial and word medial consonant errors were found to be nearly complete, relatively few internally detected errors remain unobservable. Therefore the rather strong effects of position in word and in utterance discussed here, do hardly affect the total numbers of observable errors. Of course, studies of repair rates should take into account the current strong effects of position (The reader may also have noted that, whereas Choe & Redford, 2012, found that the number of segmental errors increases from early to late within intonational units, this was not replicated in the current experiments: See the symbol sizes in Figures 2.3 and 3.3. Instead we see an alternation between more and less errors over the four word positions probably resulting from the peculiar structure of our tongue twisters).

We also found some other, not predicted, patterns in our data that could be explained by variations in selective attention. To begin with, error rate is significantly and much higher in word initial than in word medial consonants. Nooteboom and Quené (2015a) made the same observation in speech errors made in spontaneous Dutch, but they could explain this difference from the varying numbers of opportunities for segmental interaction between initial and medial positions. This explanation does not work for the current experiments, however, simply because the number of opportunities for segmental interaction was kept constant for the two positions compared. In the current experiments we apparently find a significant and rather strong word onset effect, i.e. relatively more segmental errors in initial than in medial position, that we did not expect. A second unexpected finding is that eliciting interactions by repeating consonants in certain positions in the word strongly increases the number of interactional errors in initial position but has no effect whatsoever in medial position. A third unexpected finding was that in both experiments cutoff-to-repair times are significantly longer for medial consonant errors than for initial consonant errors. We propose that these various differences between initial and medial consonants reflect variations in selective attention both during phonological encoding and during self-monitoring. Selective attention both for phonological encoding and for scanning encoded lexical forms for speech errors would strongly decrease from beginning to end, attention being focused mainly on initial segments. This explanation also implies that selective attention affects speed of processing: less attention for medial consonants than for initial consonants makes that repairing medial consonants is slower than repairing initial consonants. That variations in selective attention may affect speed of processing has been known for a long time: In 1879 Obersteiner, quoted by Guilford and Ewart (1940), stated that “retardation of the reaction stands in inverse proportion to the intensity of attention”. More recently Bates and Stough (1997) demonstrated that both attention and IQ affect reaction time in visual perception, and Tünnermann, Petersen, and Scharlau (2015) showed that increased selective attention increases speed of visual processing.

Conclusion

We have set out to investigate to what extent timing and selective attention in self-monitoring affect numbers of observable segmental errors in speech, using four-word tongue twisters in two experiments. As to timing, we started from a computational implementation of Levelt’s perceptual loop theory of self-monitoring proposed by Hartsuiker and Kolk (2001). Our results confirm that after internal error detection, the execution of the command to initiate speech and the execution of the command to interrupt speech at least partly run parallel. However, against prediction, the results also suggest that interrupting speech is much slower than initiating speech. Therefore, relatively few internally detected speech errors remain hidden from observation. We further found that scanning internal speech for errors is time-consuming, and that the time involved is roughly equivalent with speaking time. Our results also imply that after internal error detection often a repair is available at the moment of interruption and that repairing word initial errors is faster than repairing word medial errors. This is possibly an effect of variation in selective attention. Variation in selective attention is probably also responsible for a considerable and highly significant decrease in rate of error detection both within words and utterances.

Author Note

The authors are grateful to the Utrecht institute of Linguistics OTS for providing the technical facilities that enabled us to do the two experiments. In particular we are grateful to the technical staff of the institute. Alex Manus helped us with setting up and running both experiments, and Dr. Iris Mulders was extremely helpful in finding enough speakers. The raw data of the two experiments in .xlsx files can be found through the following web page: http://www.let.uu.nl/~Sieb.Nooteboom/personal/Experimentaldata.htm.

References

Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.<doi:10.18637/jss.v067.i01>.

Bates, T. & Stough, C. (1997). Processing speed, attention, and intelligence: Effects of spatial attention on decision time in high and low *IQ* subjects. *Science Direct, 23 (5)*, 861-868.

Blackmer, E. R., & Mitton, J. L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition*, 39, 173–194.

Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer (Version 5.1.05) [Computer program]. <http://www.praat.org/>.

Choe, W.K. & Redford, M.A. (2012). The distribution of speech errors in multi-word prosodic units. *Laboratory Phonology, 3 (1)*, 5-26.

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

Dell, G.S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, *27*, 124-142.

Efron, B., & Tibshirami, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

Frisch, S. A., & Wright, R. (2002). The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue. *Journal of Phonetics*, 30, 139–162.

Goldrick, M., & Blumstein, S. E. (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes*, 21, 649–683.

Goldstein, L., Pouplier, M., Chen, L., Saltzman, E., & Byrd, D. (2007). Dynamic action units slip in speech production errors. *Cognition*, 103, 386–412.

Guilford, T.P. & Ewart, E. (1940). Reaction-Time during distraction as an indicator of attention value. *American Journal of Psychology, 53,* 554-563.

Hartsuiker, R.J. & Kolk, H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychol;ogy*, *42*, 113-157.

Hartsuiker, R. J. Kolk, H.J. & Martensen, H. (2005). Division of labor between internal and external speech monitoring. In R. Hartsuiker, Y. Bastiaanse, A. Postma, & F. Wijnen (Eds.), *Phonological Encoding and Monitoring in Normal and Pathological Speech* (pp. 187–205). Hove: Psychology Press.

Hickok, G. (2012). Computational neuroanatomy of speech production. Nature Reviews Neuroscience, 13, 135–145.Lackner, J. R. (1974). Speech production: Evidence for corollary discharge stabilization of perceptual mechanisms. *Perceptual and Motor Skills*, 39, 899–902.

Kleiber, C. & Zeileis, A. (2008). *Applied Econometrics with R.* Springer-Verlag, New York. ISBN 978-0-387-77316-2.

Lackner, J. R. (1974). Speech production: Evidence for corollary discharge stabilization of perceptual mechanisms. *Perceptual and Motor Skill*s, 39, 899–902.

Levelt, W.J.M. (1989). *Speaking. From Intention to Articulation*. Harvard, MA: MIT Press

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.

McMillan, C. T., & Corley, M. (2010). Cascading influences on the production of speech: Evidence from articulation. *Cognition*, 117, 243–260.

Mowrey, R., & MacKay, I. (1990). Phonological primitives: Electromyographic speech error evidence. *Journal of the Acoustical Society of America*, 88, 1299–1312.

Nooteboom, S.G. (1980) Speaking and unspeaking: detection and correction of phonological and lexical errors of speech. In: V.A. Fromkin (ed.) *Errors in linguistic Performance: Slips of the Tongue, Ear, Pen, and Hand,* Academic Press, 87-96.

Nooteboom, S.G. & Quené, H. (2015a). Word onsets and speech errors. Explaining relative frequencies of segmental substitutions. *Journal of Memory and Language, 78,* 33-46.

Nooteboom, S.G. & Quené, H. (2015b). Word onsets and stress patterns: speech errors in a tongue-twister experiment. *Proceedings of the 18th Congress of Phonetic Sciences, Glasgow 10-14 August 2105*. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/proceedings.html.

Nooteboom, S.G. & Quené, H. (2017). Self-monitoring for speech errors: Two-stage detection and repair with and without auditory feedback. *Journal of Memory and Language, 95,* 19-35.

Obersteiner, H. (1979). Experimental researches in attention. *Brain, 1*, 439-453.

Quené, H., & Van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, 59, 413–425.

R Development Core Team (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available:<http://www.R-project.org>.

Schlenk, K., Huber, W., & Wilmes, K. (1987). ‘‘Prepairs” and repairs: Different monitoring functions in aphasic language production. *Brain and Language*, 30, 226–244.

Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In: P. MacNeilage (Ed.) *The Production of Speech* (pp. 109-136). New York: Springer.

Shattuck-Hufnagel, S. (1987). The role of word-onset consonants in speech production planning: New evidence from speech error patterns. In: E. Keller & M. Gopnik (Eds.) *Motor and sensory processes of language* (pp. 17-52). Hillsdale: Lawrence Erlbaum Associates.

Shattuck-Hufnagel, S. (1992). The role of word structure in segmental serial ordering. *Cognition*, *42*, 213-259.

Tobin, J., (1958) Estimation of relationships for limited dependent variables. *Econometrica 26,* 24-36.

Tünnermann, J., Petersen, J.A. & Scharlau, I (2015). Does attention speed up processing? Decreases and increases of processing rates in visual prior entry. *Journal of Vision, 15 (3).* doi:10.1167/15.3.1.

Wheeldon, L.R. & Levelt, W.J.M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language 34*, 311-334.

Wilshire, C. E. (1998). Serial order in phonological encoding: an exploration of the word onset effect'using laboratory-induced errors. *Cognition, 68,* 143-166.

Captions Figures

Figure 2.1: Histograms of observed error-to-cutoff times, broken down by syllable structure (upper panel 1+2+2+1 syllables, lower panel 2+2+2+2 syllables) and by position of the error, with lognormal density distributions fitted by a tobit regression model (initial errors: dashed, medial errors: dotted). The horizontal error bars near the peak of a distribution indicate the bootstrapped 95% confidence interval of the location of that peak (over 500 replications).

Figure 2.2. Histograms of observed cutoff-to-repair times, broken down by syllable structure (upper panel 1+2+2+1 syllables, lower panel 2+2+2+2 syllables) and by position of the error, with lognormal density distributions fitted by a tobit regression model (initial errors: dashed, medial errors: dotted, initial and medial errors combined: dashed+dotted). The horizontal error bars near the peak of a distribution indicate the bootstrapped 95% confidence interval of the location of that peak (over 500 replications).

Figure 2.3: Estimated log odds of detection of valid errors, broken down by syllable structure (upper panel 1+2+2+1 syllables, lower panel 2+2+2+2 syllables), by position of the error (initial: upward triangles, medial: downward triangles), and by word number in the response utterance (1 to 4, along horizontal axis). Symbol sizes correspond with the numbers of detectable valid errors in each cell.

Figure 3.1. Histograms of observed error-to-cutoff times, broken down by position of the error, with lognormal density distributions fitted by a tobit regression model (initial errors: dashed, medial errors: dotted). The horizontal error bars near the peaks of the distributions indicate the bootstrapped 95% confidence interval of the location of that peak (over 500 replications).

Figure 3.2. Histograms of observed cutoff-to-repair times, broken down by position of the error, with lognormal density distributions fitted by a tobit regression model (initial errors: dashed, medial errors: dotted. The horizontal error bars near the peak of a distribution indicate the bootstrapped 95% confidence interval of the location of that peak (over 500 replications).

Figure 3.3: Estimated log odds of detection of valid errors, broken down by position of the error (initial: upward triangles, medial: downward triangles), and by word number in the response utterance (1 to 4, along horizontal axis). Symbol sizes correspond with the numbers of detectable valid errors in each cell.

APPENDIX A: Stimuli Experiment 1

	List1					List 2			
nr	1vs2 syll	cons	cond	stimulus		2vs2 syll	cons	cond	stimulus
01	1vs2 syll	w/r	B	wok rapper roeper wal		2vs2 syll	w/r	B	water rapper roeper wallen
02	1vs2 syll	w/r	W	wad rapport rapier wol		2vs2 syll	w/r	W	woeker rapport rapier wikkel
03	1vs2 syll	w/r	S	win parijs poreus wel		2vs2 syll	w/r	S	bewijs parijs poreus juweel
04	1vs2 syll	w/r	N	wit pieren parel was		2vs2 syll	w/r	N	lawaai pieren parel gewin
05	1vs2 syll	w/r	B	wig radar ridder weg		2vs2 syll	w/r	B	waggel radar ridder wegen
06	1vs2 syll	w/r	W	won radauw radijs wiel		2vs2 syll	w/r	W	woning radauw radijs wielen
07	1vs2 syll	w/r	S	web direct dorien wak		2vs2 syll	w/r	S	gewoon direct dorien bewaar
08	1vs2 syll	w/r	N	wis dieren duren wil		2vs2 syll	w/r	N	gewis dieren duren beween
09	1vs2 syll	n/m	B	nap molen maken nok		2vs2 syll	n/m	B	neder molen maken nodig
10	1vs2 syll	n/m	W	nog maleis meloen nuf		2vs2 syll	n/m	W	noten maleis meloen nuffig
11	1vs2 syll	n/m	S	nut lamel limiet nies		2vs2 syll	n/m	S	benut lamel limiet genies
12	1vs2 syll	n/m	N	nuk lemmet lommer net		2vs2 syll	n/m	N	geniet lemmet lommer benul
13	1vs2 syll	n/m	B	nep morren mieren nis		2vs2 syll	n/m	B	nepper morren mieren nissan
14	1vs2 syll	n/m	W	nat merijn marien nek		2vs2 syll	n/m	W	nader merijn marien nimmer
15	1vs2 syll	n/m	S	niet remous romein nul		2vs2 syll	n/m	S	teniet remous romein genoot
16	1vs2 syll	n/m	N	neut raming rommel nies		2vs2 syll	n/m	N	genot raming rommel genies
17	1vs2 syll	b/v	B	bak vazen vezel bel		2vs2 syll	b/v	B	bakker vazen vezel boling
18	1vs2 syll	b/v	W	bot vazal vizier bit		2vs2 syll	b/v	W	botter vazal vizier bitter
19	1vs2 syll	b/v	S	bies zovéél zovér boek		2vs2 syll	b/v	S	debut zovéél zovér tabak
20	1vs2 syll	b/v	N	bon zuivel zever bed		2vs2 syll	b/v	N	gebod zuivel zever gebed
21	1vs2 syll	v/b	B	vak bellen balen vet		2vs2 syll	v/b	B	vader bellen builen veter
22	1vs2 syll	v/b	W	voet ballon balein vis		2vs2 syll	v/b	W	voeder ballon balein visser
23	1vs2 syll	v/b	S	vod labiel libel vies		2vs2 syll	v/b	S	gevat labiel libel devies
24	1vs2 syll	v/b	N	vin lebber lobben val		2vs2 syll	v/b	N	ravijn lebber lobben revier
25	1vs2 syll	p/k	B	pet kamer kommer pop		2vs2 syll	p/k	B	pieter kamer kommer pooier
26	1vs2 syll	p/k	W	pof kameel komeet pin		2vs2 syll	p/k	W	poging kameel komeet peiling
27	1vs2 syll	p/k	S	poet mekaar makaak poch		2vs2 syll	p/k	S	tepas mekaar makaak gepoch
28	1vs2 syll	p/k	N	poes makker mocca peen		2vs2 syll	p/k	N	kaping makker mocca keper
29	1vs2 syll	k/p	B	kas pater peter kil		2vs2 syll	k/p	B	kajak pater peter ketter
30	1vs2 syll	k/p	W	kom patat potent kus		2vs2 syll	k/p	W	koter patat potent kussen
31	1vs2 syll	k/p	S	kop tapijt topaas keel		2vs2 syll	k/p	S	tekoop tapijt topaas bekeer
32	1vs2 syll	k/p	N	kor tepel tapir kiep		2vs2 syll	k/p	N	bekom tepel tapir tekijk
33	1vs2 syll	l/r	B	lies rakker rekel lach		2vs2 syll	l/r	B	liever rakker rekel ladder
34	1vs2 syll	l/r	W	log raket rekest los		2vs2 syll	l/r	W	logger raket rekest lover
35	1vs2 syll	l/r	S	lik karos koraal lef		2vs2 syll	l/r	S	gelik karos koraal beleg
36	1vs2 syll	l/r	N	lak kerel karig lam		2vs2 syll	l/r	N	belet kerel karig meloen
37	1vs2 syll	l/r	B	lis raven rover lied		2vs2 syll	l/r	B	lekker raven rover loeder
38	1vs2 syll	l/r	W	lot ravijn rivier lus		2vs2 syll	l/r	W	lokker ravijn rivier leiding
39	1vs2 syll	l/r	S	les varaan viriel loep		2vs2 syll	l/r	S	gelijk varaan viriel beloop
40	1vs2 syll	l/r	N	lol virus varen lor		2vs2 syll	l/r	N	geloof virus varen zeloot
41	1vs2 syll	j/l	B	jas later loting jek		2vs2 syll	j/l	B	jekker later loting jopper
42	1vs2 syll	j/l	W	juf latijn letaal jap		2vs2 syll	j/l	W	juffer latijn letaal jarig
43	1vs2 syll	j/l	S	jacht talent teloor joch		2vs2 syll	j/l	S	gejacht talent teloor gejuich
44	1vs2 syll	j/l	N	joop teler tema juich		2vs2 syll	j/l	N	gejok teler tema gejaagd

45	1vs2 syll	j/l	B	jacht loper liepen jol		2vs2 syll	j/l	B	jager loper liepen jokker
46	1vs2 syll	j/l	W	jas lapel lipoom jek		2vs2 syll	j/l	W	jammer lapel lipoom joker
47	1vs2 syll	j/l	S	jool paleis poliep jucht		2vs2 syll	j/l	S	gejoel paleis poliep bejuicht
48	1vs2 syll	j/l	N	jaar paling peiler juf		2vs2 syll	j/l	N	bejaard paling peiler gejouw
49	2vs2 syll	d/j	B	dolen jura jarig duiker		1vs2 syll	d/j	B	doof jura jarig duit
50	2vs2 syll	d/j	W	duvel jorien jeroen doper		1vs2 syll	d/j	W	dut jorien jeroen doek
51	2vs2 syll	d/j	S	gedoe radijs redoute gedaas		1vs2 syll	d/j	S	dos radijs redoute dar
52	2vs2 syll	d/j	N	gedut redding roedel bedompt		1vs2 syll	d/j	N	dun redding roedel dom
53	2vs2 syll	z/d	B	zaling dame duimel zakker		1vs2 syll	z/d	B	zaal dame duimel zak
54	2vs2 syll	z/d	W	zuiger domein damast zaling		1vs2 syll	z/d	W	zuig domein damast zaal
55	2vs2 syll	z/d	S	bezaan modern madam seizoen		1vs2 syll	z/d	S	zaak modern madam zoen
56	2vs2 syll	z/d	N	gezag moeder modder bezet		1vs2 syll	z/d	N	zag moeder modder zet
57	2vs2 syll	k/g	B	kennis gekkie gele kater		1vs2 syll	k/g	B	ken golem gele kaai
58	2vs2 syll	k/g	W	kotter geluk geloof kapper		1vs2 syll	k/g	W	kot geluk geloof kap
59	2vs2 syll	k/g	S	bekijk legaat legaal bekort		1vs2 syll	k/g	S	kijk legaat legaal kort
60	2vs2 syll	k/g	N	bekoor lachen lichaam bekoel		1vs2 syll	k/g	N	koor lachen lichaam koel
61	2vs2 syll	g/k	B	gaper kale koele gene		1vs2 syll	g/k	B	gas kale koele geen
62	2vs2 syll	g/k	W	gokker kalot Colijn gister		1vs2 syll	g/k	W	gok kalot Colijn gist
63	2vs2 syll	g/k	S	begoot lakei loket begeef		1vs2 syll	g/k	S	gor lakei loket geef
64	2vs2 syll	g/k	N	begaf lakken lekken tegek		1vs2 syll	g/k	N	gaf lokken lekken gek
65	2vs2 syll	t/d	B	tijdig deken duiker topper		1vs2 syll	t/d	B	tijd deken duiker top
66	2vs2 syll	t/d	W	tukker decaan ducaat tinnef		1vs2 syll	t/d	W	tuk decaan ducaat tin
67	2vs2 syll	t/d	S	getij kadet kado beton		1vs2 syll	t/d	S	teil kadet kado ton
68	2vs2 syll	t/d	N	getik kader koddig getal		1vs2 syll	t/d	N	tik kader koddig tal
69	2vs2 syll	t/d	B	tuigen dapper doping togen		1vs2 syll	t/d	B	tuin dapper doping toog
70	2vs2 syll	t/d	W	tering depot depêche toner		1vs2 syll	t/d	W	teer depot depêche toon
71	2vs2 syll	t/d	S	getik pedant pedaal getob		1vs2 syll	t/d	S	teek pedant pedaal top
72	2vs2 syll	t/d	N	getal peddel padden baton		1vs2 syll	t/d	N	taal peddel padden tok
73	2vs2 syll	p/t	B	peter tekkel tikker poker		1vs2 syll	p/t	B	peet tekkel tikker pook
74	2vs2 syll	p/t	W	pieter tekeer tekort pover		1vs2 syll	p/t	W	pet tekeer tekort pof
75	2vs2 syll	p/t	S	gepees katoen katijf bepoot		1vs2 syll	p/t	S	pees katoen katijf poes
76	2vs2 syll	p/t	N	bepakt ketel kater gepeins		1vs2 syll	p/t	N	pad ketel kater peins
77	2vs2 syll	d/z	B	dader zuilen zaling daler		1vs2 syll	d/z	B	daad zuilen zaling dal
78	2vs2 syll	d/z	W	dekking zeloot zolang duiten		1vs2 syll	d/z	W	dek zeloot zolang duit
79	2vs2 syll	d/z	S	gedaan lazuur lysol gedoopt		1vs2 syll	d/z	S	daan lazuur lysol dop
80	2vs2 syll	d/z	N	bedot lozing lezer gedimd		1vs2 syll	d/z	N	dog lozing lezer dim
81	2vs2 syll	sj/s	B	sjalen sufferd saffie sjezen		1vs2 syll	sj/s	B	sjaal sufferd saffie sjees
82	2vs2 syll	sj/s	W	sjieker sofie saffier sjacher		1vs2 syll	sj/s	W	sjiek sofie saffier sjah
83	2vs2 syll	sj/s	S	gesjouw facet fossiel gesjok		1vs2 syll	sj/s	S	sjouw facet fossiel sjaak
84	2vs2 syll	sj/s	N	gesjor fasces facie gesjans		1vs2 syll	sj/s	N	sjoerd fasces facie sjans
85	2vs2 syll	s/sj	B	sijpel sjoemel sjekkie sollen		1vs2 syll	s/sj	B	sip sjoemel sjekkie sof
86	2vs2 syll	s/sj	W	soppen sjamaan chauffeur cello		1vs2 syll	s/sj	W	sop sjamaan chauffeur cel
87	2vs2 syll	s/sj	S	gesip machien michel gesol		1vs2 syll	s/sj	S	sik machien michel sol
88	2vs2 syll	s/sj	N	besef misje muisje gesim		1vs2 syll	s/sj	N	sein misje muisje sim
89	2vs2 syll	v/z	B	vitter zone zanik voeder		1vs2 syll	v/z	B	vit zone zanik vos
90	2vs2 syll	v/z	W	vielen zonee zonaal vodden		1vs2 syll	v/z	W	vief zonee zonaal vod
91	2vs2 syll	v/z	S	schavuit nasaal nazist tevol		1vs2 syll	v/z	S	vuist nasaal nazist vol
92	2vs2 syll	v/z	N	geval nazi neuzen gevang		1vs2 syll	v/z	N	val nazi neuzen vang

93	2vs2 syll	z/v	B	zessen vutter vette zullen		1vs2 syll	z/v	B	zes vutter vette zal
94	2vs2 syll	z/v	W	zeilen votief vitaal zagen		1vs2 syll	z/v	W	zeil votief vitaal zaag
95	2vs2 syll	z/v	S	bazin tevoet gevat gazon		1vs2 syll	z/v	S	zin tevoet gevat zon
96	2vs2 syll	z/v	N	dozijn tover toeval gezang		1vs2 syll	z/v	N	zijn tover toeval zang

APENDIX B: Stimuli Experiment 2

stim nr	not elic	cons	cond	stimulus		elic	cons	cond	stimulus
01	not elic		B	water rapper lommer bikkel		elic	w/r	B	water rapper roeper wallen
02	not elic		W	nuttig lamel poreus niessen		elic	w/r	W	woeker rapport rapier wikkel
03	not elic		S	noteer maleis rapier varaan		elic	w/r	S	bewijs parijs poreus juweel
04	not elic		N	geniet lemming weiland gevit		elic	w/r	N	lawaai pieren parel gewin
05	not elic		B	nepper morren ridder wegen		elic	w/r	B	waggel radar ridder wegen
06	not elic		W	moter remous zover takken		elic	w/r	W	woning radauw radijs wielen
07	not elic		S	papier merijn vizier bedot		elic	w/r	S	gewoon direct dorien bewaar
08	not elic		N	genot raming lodder bezit		elic	w/r	N	gewis dieren duren beween
09	not elic		B	bakker vazen maken nodig		elic	n/m	B	neder molen maken nodig
10	not elic		W	doeken zoveel romein goten		elic	n/m	W	noten maleis meloen nuffig
11	not elic		S	beman vazal radijs gewoel		elic	n/m	S	benut lamel limiet genies
12	not elic		N	geloof virus tafel tekijk		elic	n/m	N	geniet lemmet lommer benul
13	not elic		B	neder molen roeper wallen		elic	n/m	B	nepper morren mieren nissan
14	not elic		W	wijzer parijs dorien gozer		elic	n/m	W	nader merijn marien nimmer
15	not elic		S	rozijn patat rekest manier		elic	n/m	S	teniet remous romein genoot
16	not elic		N	gebod zuiver lommer rivier		elic	n/m	N	genot raming rommel genies
17	not elic		B	vader bellen kommer pooier		elic	b/v	B	bakker vazen vezel boling
18	not elic		W	vatten labiel topaas doelen		elic	b/v	W	botter vazal vizier bitter
19	not elic		S	tomaat kanon boleet gevat		elic	b/v	S	debuut zovéél zovér tabak
20	not elic		N	ravijn lening zeker gebed		elic	b/v	N	gebod zuivel zever gebed
21	not elic		B	pieter kamer vezel bowling		elic	v/b	B	vader bellen builen veter
22	not elic		W	passen mekaar libel buiging		elic	v/b	W	voeder ballon balein visser
23	not elic		S	rabbijn kameel balein tonijn		elic	v/b	S	gevat labiel libel devies
24	not elic		N	tapijt makker wever gebed		elic	v/b	N	ravijn lebber lobben revier
25	not elic		B	koper pater rekel ladder		elic	p/k	B	pieter kamer kommer pooier
26	not elic		W	kochten tapijt makaak pochen		elic	p/k	W	poging kameel komeet peiling
27	not elic		S	kaneel rapport meloen figuur		elic	p/k	S	tepas mekaar makaak gepoch
28	not elic		N	bekom tepel karig malloot		elic	p/k	N	tapijt matter motte kapot
29	not elic		B	liever rakker peter ketter		elic	k/p	B	kajak pater peter ketter
30	not elic		W	likken karos topaas duiker		elic	k/p	W	koter patat potent kussen
31	not elic		S	balein raket potent kozijn		elic	k/p	S	tekoop tapijt topaas bekeer
32	not elic		N	belet kerel mocca tapuit		elic	k/p	N	bekom tepel tapir tekijk
33	not elic		B	lekker raven builen veter		elic	l/r	B	liever rakker rekel ladder
34	not elic		W	lijken varaan koraal zeggen		elic	l/r	W	logger raket rekest lover
35	not elic		S	loket ravijn banaan pedant		elic	l/r	S	gelik karos koraal beleg
36	not elic		N	lawaai pieren bakken gesop		elic	l/r	N	belet kerel karig meloen
37	not elic		B	jekker later doping koppig		elic	l/r	B	lekker raven rover loeder
38	not elic		W	jager talent viriel zoeker		elic	l/r	W	lokker ravijn rivier leiding
39	not elic		S	meneer latijn rivier moreel		elic	l/r	S	gelijk varaan viriel beloop
40	not elic		N	gejok teler varen bezoek		elic	l/r	N	geloof virus varen zeloot

41	not elic		B	jager loper rover moedig		elic	j/l	B	jekker later loting jopper
42	not elic		W	joelen paleis roman moker		elic	j/l	W	juffer latijn letaal jarig
43	not elic		S	kopij kajuit metaal rozijn		elic	j/l	S	gejacht talent teloor gejuich
44	not elic		N	bejaard paling deren gemaal		elic	j/l	N	gejok teler palet gejaagd
45	not elic		B	waggel radar mieren nissan		elic	j/l	B	jager loper liepen jokker
46	not elic		W	duvel jorien jeroen doper		elic	j/l	W	jammer lapel lipoom joker
47	not elic		S	konijn direct limiet fineer		elic	j/l	S	gejoel paleis poliep bejuicht
48	not elic		N	rebel bader mare bewaar		elic	j/l	N	bejaard paling peiler gejouw
49	elic	d/j	B	dolen jura jarig duiker		not elic		B	dolen jura kater gokken
50	elic	t/d	W	tover dozijn defect toeval		not elic		W	garen radijs legaal beker
51	elic	d/b	S	gedoe robijn gebal gedaas		not elic		S	katoen poliep piraat robijn
52	elic	t/d	N	getut redding roedel beton		not elic		N	genot redding mijlen berijk
53	elic	z/d	B	zaling dame duimel zakker		not elic		B	zaling dame gokken kater
54	elic	z/d	W	zuiger domein damast zaling		not elic		W	bazig tevoet gekat vallen
55	elic	z/d	S	bezaan modern madam seizoen		not elic		S	sigaar domein baron ballon
56	elic	z/d	N	gezag moeder modder bezet		not elic		N	gezag moeder roebel konijn
57	elic	k/g	B	kennis gekkie gele kater		not elic		B	kennis gekkie dadel poeder
58	elic	k/g	W	kotter geluk geloof kapper		not elic		W	boter legaat madam deining
59	elic	k/g	S	bekijk legaat legaal bekort		not elic		S	getob geluk damast venijn
60	elic	k/g	N	bekoor lachen lichaam bekoel		not elic		N	bekoor lachen modder bezet
61	elic	g/k	B	gaper kale koele gene		not elic		B	pater kale gokker later
62	elic	g/k	W	gokker kalot Colijn gister		not elic		W	gotisch pineut legaal teken
63	elic	g/k	S	begoot lakei loket begeef		not elic		S	gemaal kalot majoor rekest
64	elic	g/k	N	begaf lakken lekken tegek		not elic		N	begaf lekken modder facet
65	elic	t/d	B	tijdig deken duiker topper		not elic		B	tijdig deken roken bodem
66	elic	t/d	W	tukker decaan ducaat tinnef		not elic		W	geiten kadet pedaal boete
67	elic	t/d	S	getij kadet kado beton		not elic		S	bevel decaan pineut gemok
68	elic	t/d	N	getik kader koddig getal		not elic		N	gelik kapper padden beton
69	elic	t/d	B	tuigen dapper doping togen		not elic		B	tuigen dapper kikker poker
70	elic	t/d	W	tering depot depêche toner		not elic		W	balen pedant latei kippig
71	elic	t/d	S	getik pedant pedaal getob		not elic		S	teleen depôt metaal gemeen
72	elic	t/d	N	getal peddel vodden baton		not elic		N	getal peddel kamer balein
73	elic	p/t	B	peter tekkel tikker poker		not elic		B	peter tekkel duivel rapper
74	elic	p/t	W	pieter tekeer tekort pover		not elic		W	benig katoen lysol poker
75	elic	p/t	S	gepees katoen katijf bepoot		not elic		S	terrein tekeer zolang roman
76	elic	p/t	N	bepakt ketel kater gepeins		not elic		N	bepakt netel lekken piloot
77	elic	d/z	B	dader zuilen zaling daler		not elic		B	vader zuilen doping toegang
78	elic	d/z	W	dekking zeloot zolang duiten		not elic		W	danig lazuur fossiel kapper
79	elic	d/z	S	gedaan lazuur lysol gedoopt		not elic		S	ducaat zeloot gedimd lazuur
80	elic	d/z	N	gewis dieren bezig palet		not elic		N	bedot lozing jager gepeins
81	elic	sj/s	B	sjalen sufferd saffie sjezen		not elic		B	sjalen sufferd tikker pooier
82	elic	sj/s	W	sjieker sofie saffier sjacher		not elic		W	sjouwer facet katijf bijten
83	elic	sj/s	S	gesjouw facet fossiel gesjok		not elic		S	tevoet sofie zolang bemind
84	elic	sj/s	N	gesjor rozig lezen gesjouw		not elic		N	gesjor fakkel muizen begin
85	elic	s/sj	B	sijpel sjoemel sjekkie sollen		not elic		B	sijpel sjoemel zanik voeder
86	elic	s/sj	W	soppen sjamaan chauffeur cello		not elic		W	sippen machien bevel pater
87	elic	s/sj	S	gesip machien michel gesol		not elic		S	makaak sjamaan banaal kapoen
88	elic	s/sj	N	besef misje muisje gesim		not elic		N	besef misje ketting gevang

89	elic	v/z	B	vitter zône zanik voeder		not elic		B	vitter zône koker gading
90	elic	k/p	W	kamer piloot palet koter		not elic	v/z	W	jochie nasaal gebed kapen
91	elic	v/z	S	schavuit nasaal bezoek tevol		not elic	v/z	S	verrot chinees gezag vitaal
92	elic	v/z	N	geval razen neuzen gevang		not elic	v/z	N	geval nozem fakir gesjans
93	elic	z/v	B	zessen vutter vette zullen		not elic	z/v	B	zessen vutter bellen roedel
94	elic	z/v	W	zeilen votief vitaal zagen		not elic	z/v	W	koter tevoet gekat ratten
95	elic	z/v	S	bazin tevoet gevat gazon		not elic	z/v	S	zeloot votief japon kaneel
96	elic	d/z	N	bedot lozing lezer gedimd		not elic	z/v	N	dozijn tover toekan kapel