WORD-ONSETS AND STRESS PATTERNS: SPEECH ERRORS IN A TONGUE-TWISTER EXPERIMENT

Sieb Nooteboom & Hugo Quené

Utrecht institute of Linguistics OTS, Utrecht University
Utrecht, Netherlands

ABSTRACT

Consonants in word onsets are more often than other consonants involved in interactional speech errors [8] [9]. This has been explained from the process of speech preparation [9] or from the higher degree of activation of initial versus other consonants [2], or from phonotactic constraints on speech errors [6]. Here we report a tongue-twister experiment showing (a) that words in each other's immediate context produce more interactional errors if the words share their stress patterns than if they do not, (b) a considerable and highly significant word-onset effect that cannot be explained from phonotactic constraints on speech errors. The latter effect is explained as a frequency effect.

Keywords: Word onset, stress pattern, speech error, interaction, substitution.

1. INTRODUCTION

It has been reported that in American English word onset consonants are more often involved in interactional speech errors (i.e. speech errors with an obvious source in the immediate context) than consonants in other positions and than one would expect from chance [8] [9]. [9] explained this word onset effect from a model of serial ordering of segments in which word initial segments were treated differently from other segments. [2] suggested that the effect might be caused by a higher degree of activation of initial than of other segments. [6] demonstrated that the relative frequencies of interactional substitutions in a corpus of spontaneous errors in Dutch can be fully accounted for by the phonotactic structure of the language: there are simply more initial consonants than consonants in other positions available for interaction in the immediate context. However, a considerable word-onset effect was found in tongue twister experiments [9] [5], where the effect could not easily be explained from the phonotactics of the immediate context. Thus there is a potential conflict between error frequencies in Dutch spontaneous speech [6] and in tongue-twister experiments in Dutch [5] and American English [9]. In order to investigate this conflict, we have conducted a tongue twister experiment, similar to experiment 2 described by [9]. However, we changed the experiment slightly, in order to prevent elicited interactions between initial and medial consonants. This was done because recent findings have shown that initial, medial and final consonants interact almost only with initial, medial and final consonants respectively [6]. There seems to be considerable resistance against interactions between consonants in different positions. Otherwise the four conditions were copied from [9]: condition B in which the potentially interacting consonants shared <u>both</u> word onset position and pre-stress position, condition W in which they shared <u>word-onset</u> position and not pre-stress position, condition S in which they shared <u>pre-stress</u> but not word onset position and condition N in which they shared <u>neither</u> position. These four conditions in experiment 2 reported by [9] are exemplified, together with the numbers of elicited interactional errors in that experiment 2 in Table 1. Targeted interacting consonants are in italics.

Table 1. Examples of stimuli in 4 conditions, with numbers of elicited targeted errors, in exp. 2 of [9]. For the sake of clarity the consonants targeted for interaction are printed in italics.

condition	stimulus	nr errors
B	*p*ack *f*ussy *f*ossil *p*ig	253
W	*p*ad *f*orsake *f*oresee *p*ot	132
S	*p*in su*ff*ice su*ff*use *p*et	75
N	*p*od so*f*a su*ff*er *p*eg	26

We assume that in conditions S and N, the number of interactions is artificially constrained by forcing interaction between two consonants in different positions [6]. Therefore we have in the experiment to be described below replaced the stimulus pattern by four disyllabic words in all conditions, so that the interacting pairs of consonants may be in the same position in the word (see Method). We predict that, if frequencies of interactional errors are indeed predominantly determined by phonotactic opportunities [6], in our new set-up conditions B and W will have equal error frequencies, because in these conditions the numbers of opportunities for interaction are kept equal.

2. METHOD

The method used was basically the same as the one used in [9]: each participant was asked to read each four-word sequence that appeared on a screen aloud three times. Then the word sequence disappeared and the participant had to speak the same sequence from memory, also three times.

2.1. Stimuli

Stimuli consisted of 24 quartets of Dutch tongue twister word sequences, each quartet having one stimulus for each of four conditions. Examples are given in Table 2.

Table 2. Example of a quartet of stimuli. Consonants targeted for interaction are in italics and stressed vowels indicated with an accent sign. This was not done in the actual stimuli.

condition	stimulus
B	*w*áter *r*ápper *r*óeper *w*állen
W	*w*óeker *r*appórt *r*apíer *w*íkkel
S	be*w*íjs pa*r*íjs po*r*éus ju*w*éel
N	la*w*áai píe*r*en pá*r*el ge*w*ín

There were also 24 corresponding quartets having four stimuli of the form used by [9]. The above stimulus B had as its counterpart "*w*ok *r*ápper *r*óeper *w*al". Results for these corresponding stimuli will not be reported here, but because of this we had two stimulus lists with 12 stimulus quartets of each form. So each list had a total of 24 quartets (equalling 96 stimuli) of which 12 (equalling 48 stimuli) were of the form reported on here.

2.2. Participants

There were 28 participants, 20 females and 8 males, all students of 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.

2.3. 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. There were 10 practice items specifically constructed for that 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. 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 there were created two separate audio files for each sequence of four words, one for the phase in which the words were visible on screen and one for the phase were the words were invisible and they had to be spoken from memory. Thus for each participant 192 audio files were created.

2.4. Scoring

Each audio file was transcribed using Praat [1] by the first author and each targeted and not targeted interactional substitution of a single consonant by a single consonant from the immediate context was coded separately per participant, per condition and per stimulus, as to the type of speech error (exchange, anticipation, perseveration, interruption). Unfortunately, there were strong effects of hysteresis in the sense that if a participant made a particular speech error on a stimulus, he or she very often repeated that same error one or more times during the six responses to that stimulus. Therefore we have for each stimulus only kept the first error of a particular type. Multiple errors of different types in response to the same stimulus were not attributed to hysteresis and were therefore not excluded.

3. RESULTS

First, we were interested in confirming interaction patterns observed in interactional speech errors in spontaneous Dutch, where word initial, medial and final consonants were found to be mostly substituted by initial, medial and final consonants respectively. [6]. Figure 1 gives the numbers of single consonant (both targeted and not targeted) substitutions in our tongue-twister experiment collapsed over the four conditions, separately for position of substitution and source position.

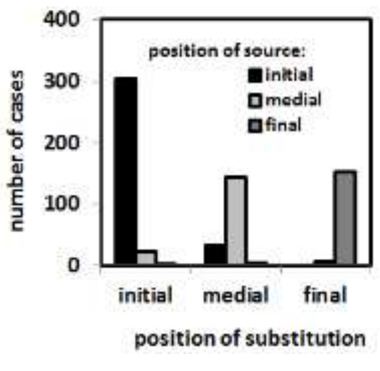

Figure 1. Numbers of single consonant substitutions, collapsed over the four conditions and separately for position of substitution and source position (N = 665).

Clearly, consonants were hardly ever substituted by consonants in a different position. Also, the few cases that do cross positions obviously have very little to tell about a potential word-onset effect. Therefore we have focused for the further analysis on the substitutions that do not cross positions. Table 3 gives the numbers of single consonant substitutions per condition and per position in the word.

Table 3. Numbers of single consonant substitutions per condition and per position in the word. The numbers of interactional errors in targeted positions are in bold face. Numbers for visible and invisible stimuli are collapsed here.

condition	initial	medial	final	total
B	**127**	57	73	257
W	**51**	12	13	76
S	83	**58**	28	169
N	41	**15**	39	95
total	302	142	153	597

It should be noted that, given the structure of the stimuli, initial and medial consonants have in all conditions the *same* numbers of opportunities to interact with a consonant in the same position in the word in the immediate context. The higher numbers of substitutions in initial position as compared to medial position, in all four conditions, seem to suggest a substantial word-onset effect. There is more uncertainty in this respect for final consonants, because these were less strictly controlled. Nevertheless we have included the numbers for final consonants, because these add to the picture how vulnerable the stimuli in each condition are to speech errors. The differences between conditions are, of course, enormous. This can evidently not be explained from the success of the elicitation procedure, targeting specific consonant positions. In conditions S and N, the targeted substitutions are in medial position, but yet the numbers of substitutions in initial position are much and significantly higher. Obviously, we have to look for another factor or other factors potentially explaining the considerable differences between conditions.

As it happens, in conditions B and S all four words share the same stress pattern, in condition B the pattern stressed-unstressed (Sw) and in condition S the pattern unstressed-stressed (wS). In conditions W and N this is not the case: In condition W the order of stress patterns is Sw-wS-wS-Sw, in condition N this is reversed. We have submitted the data of the experiment to a mixed-effects logistic regression model (GLMM; [7]) with as random intercepts participants and sets of matching stimuli, and as fixed factors condition (with N as baseline) and visible versus invisible stimulus (with invisible as baseline). The condition factor was coded as three contrasts, viz. targeted word onset versus targeted medial position (B and W versus S and N), shared versus not shared stress pattern (B and S versus W and N), and Sw versus wS in 2nd syllable (B and N versus W and S); this factor was also added as a random slope between sets of stimuli Analysis results are presented in Table 4.

Table 4. Estimated coefficients of a mixed logistic regression model (see text). Significant fixed effects ($p < .05$) are in boldface.

Random effects:	*Estim.*	*95% C.I.*		*n*
Participants	0.0882	(0.0476, 0.4060)		27
Stim set (B)	0.5312	(0.5185,1.0078)		24
Stim set (W)	0.6715	(0.3476,1.1268)		24
Stim set (S)	0.5732	(0.3705, 1.0759)		24
Stim set (N)	1.0572	(0.4886, 1.4471)		24
Fixed effects:	*Estim..*	*std. err.*	*Z*	*Prob.*
(Intercept)	-1.5353	0.1495	-10.27	< .0001
initial vs medial	0.2189	0.2327	0.94	.3470
sharing str. pttrn	**1.1485**	**0.1776**	**6.47**	**<.0001**
Sw vs wS	**0.3858**	**0.1786**	**2.16**	**.0307**
invis, vs vis.	**-1.9809**	**0.1039**	**-9.44**	**< .0001**

The factor visible versus invisible gave a highly significant main effect ($p < .0001$) but no interaction with other fixed effects. The main effect shows that there were many more errors in the invisible part of the experiment, probably due to memory problems. The lack of interaction shows that the distributions of error frequencies in the visible and invisible parts did not differ significantly. Interestingly, the contrast between error rates elicited at word onset (333 errors) versus medial position (264 errors) gave no significant result ($p = .3470$). The contrast between conditions shared versus not shared stress pattern gave a highly significant result ($p < .0001$). Most of the variance in the data is explained by this contrast. The contrast Sw versus wS in 2nd syllable also gave a significant result ($p = .0307$). Inspection of the numbers in Table 3 suggests that this significant contrast must be mainly due to the contrast between conditions B and S, reducing this contrast to sharing Sw versus sharing wS.

From Table3 it is evident that the highest number of single consonant substitutions is in word initial position. This suggests a word-onset effect. This effect could not be analysed by the above logistic regression analysis, because the position of the participant's error was not an experimental variable. The differences between error positions were therefore evaluated by means of a post hoc multinomial logistic regression with 1000 two-stage bootstrap replications over matching stimulus sets and over responses, respectively. The results show that within each condition word-initial errors significantly outnumber both errors in medial and errors in final position ($p < .0001$ for each comparison).

4. DISCUSSION

The main results of the current tongue-twister experiment are:

1. Word initial, medial and final consonants nearly always keep to their own position in interactional substitutions.
2. When words in each other's immediate context share their stress pattern (as in conditions B and S), their segments are more often involved in interactional substitutions than segments of words that do not share a stress pattern (as in conditions W and N). This is a large effect.
3. Sharing the pattern Sw is much more effective in generating interactions than sharing the pattern wS.
4. The distribution of valid single consonant substitutions over initial, medial and final positions in the word shows a considerable and highly significant word-onset effect.

The first finding, that consonants rarely interact with consonants in another position in the word, confirms what was reported earlier on interactional speech errors in spontaneous Dutch [6]. This finding implies that at the level where interactional speech errors are generated, there has been no re-syllabification of the type advocated e.g. in [4]. It also casts doubt on the reality of syllables as organizational units in speech preparation. This supports [10], where it is argued that there is no convincing evidence for the role of syllables as units involved in speech preparation. The strong resistance in speech preparation against interactions between consonants in different positions also suggests that the results obtained in experiment 2 reported by [9], as mentioned in our introduction, are indeed at least partly due to a confound between the conditions and the targeted interactions between different positions in the word.

Secondly, the effect of shared stress pattern suggests that when words in the immediate context share their stress pattern, this increases activation of all their segments, by a process similar to priming.

Thirdly, the fact that the pattern Sw is much more effective than the pattern wS possibly may be related to the fact that the first pattern is the canonical stress pattern for disyllables in Germanic languages, and much more frequent in spoken Dutch than the pattern wS. In a collection of speech errors in spontaneous Dutch [6] of all disyllabic words 84 % were found to be Sw.

Fourthly, although there is no significant effect of targeting either initial (conditions B and W) or medial (conditions S and N) consonants for interaction on the distributions of errors, yet there are significantly more interactional substitutions in initial position than in medial and final positions. This word-onset effect cannot be explained from relative frequencies of opportunities for interaction in the context of this experiment, because here these opportunities were equal for initial and medial consonants. The word-onset effect observed in this experiment matches, however, a similar effect observed in interactional speech errors in spontaneous Dutch where that word-onset effect does indeed correspond to relative frequencies of opportunities for interaction in initial, medial and final positions [6]. We propose that expectations created by that quantitative word-onset effect in spontaneous speech are carried over to the experimental situation, boosting activation of word-onset consonants. This would be a frequency effect.

5. REFERENCES

[1] Boersma, P., & Weenink, D. 2014. *Praat: Doing phonetics by computer(Version 5404)* Computer program]. <http://www.praat.org/>.

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

[3] Dell, G.S., Reed, K.D., Adams, D.R. & Meyer, A.S. 2000. Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 1355-1367.

[4] Levelt, W. J. M., Roelofs, A., & Meyer, A. S. 1999. A theory of lexical access in speech production. *Behavioral and Brain Science* 22, 1–75.

[5] Meijer, P.J.A. 1997. What speech errors can tell us about word-form generation: the roles of constraint and opportunity. *Journal of Psycholinguistics* 26, 141-157.

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

[7] 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.

[8] 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*. Hillsdale: Lawrence Erlbaum Associates, 17-52.

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

[10] Shattuck-Hufnagel, S, 2011. The role of the syllable in speech production in American English: A fresh consideration of the evidence. In C. Cairns & E. Raimy (eds.) *Handbook of the Syllable,* Leyden: Brill, 251-289.