Journal of Memory and Language 78 (2015) 33–46

Contents lists available at ScienceDirect

Journal of Memory and Language

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

Word onsets and speech errors. Explaining relative frequencies of segmental substitutions

Sieb Nooteboom *, Hugo Quené

Utrecht Institute of Linguistics, UiL OTS, Utrecht University, the Netherlands

ARTICLE INFO

Article history:
Received 10 February 2014
revision received 30 September 2014
Available online 1 November 2014

Keywords:
Word onset
Speech error
Interaction
Substitution
Phonotactic constraint

ABSTRACT

Consonants in word onsets are, in English and Dutch, more frequently misspoken than consonants in other positions, and also more frequently than expected from the relative numbers of onset consonants and other consonants. We argue here that relative numbers of segments in specific positions in the word is not a valid predictor of relative frequencies of segmental speech errors. A more valid predictor would be the relative number of phonotactically allowed opportunities segments in different positions have to be involved in interactional speech errors. Analysis of segmental speech errors in spontaneous Dutch shows that relative frequencies of interactional substitutions of single segments in vowel positions, and word initial, medial and final consonant positions, may indeed be predicted rather precisely from the allowed opportunities for segments in different positions to be involved in interactional speech errors, and that there is no additional 'word onset' effect in these speech errors.

Introduction

Consonants in word onsets are more often involved in segmental speech errors than consonants in other positions. This was already observed by Meringer and Mayer, who, on the basis of their study of segmental errors of speech, ascribed more "weight" to word onsets and root onsets, and also to vowels, than to other positions (Meringer & Mayer, 1895, p. 162: "Die höchstwertigen Laute sind also der Anlaut der Wurzelsilbe und der Wortanlaut und der oder die betonten Vokale"; in English: "So the sounds with most weight are the onset of the root syllable, the word onset and the stressed vowel or vowels", cf. also Levelt, 2013, p. 159). Shattuck-Hufnagel (1983, 1987, 1992) independently noted the predominance of word onset consonants in segmental speech errors, and also the vulnerability to speech errors of consonants in pre-stress position. She demonstrated the word onset effect by counting segmental speech errors in an extensive corpus of speech errors made in spontaneous speech in American English, and also in experiments eliciting speech errors by having speakers speak aloud tongue twisters rapidly and repeatedly. A typical result for errors in spontaneous speech is that in the MIT corpus of speech errors 66% of 1520 consonantal errors occur in word onsets whereas in a corpus of running speech (Carterette & Jones, 1974) only 33% of consonants happen to be in word onsets (Shattuck-Hufnagel, 1987), i.e. word onset consonants are overrepresented in speech errors. This is most clearly so for so-called interactional errors or movement errors, i.e. errors that have an obvious source in the immediate context. For noninteractional errors the effect seems to be less clear. If Shattuck-Hufnagel limited the count to completed exchanges (because these are most clearly interactional errors), even 91% of clearly interactional single consonant speech errors were in word onset position. We will refer to this presumed predominance of word onset consonants in speech errors as the word onset effect or word initialness effect.

* Corresponding author at: Cor Ruyslaan 20, 3584 gd Utrecht, Netherlands.

E-mail address: S.G.Nooteboom@uu.nl (S. Nooteboom).

http://dx.doi.org/10.1016/j.jml.2014.10.001
0749-596X/

In a series of tongue twister experiments Shattuck-Hufnagel (1992) confirmed the special position of word onset consonants. Results of these studies also show that pre-stress consonants are confused somewhat more than consonants not sharing their position with respect to lexical stress. Shattuck-Hufnagel (1992) proposes a model of speech preparation based on the "scan-copy" model proposed in Shattuck-Hufnagel (1987) which has separate nodes for word onset consonants, and which has a separate selection of lexical candidates with their phonemic make-up on the one hand and of "prosodic frames" with marked word-onset slots and stress positions on the other (for the separate roles of sets of segments and prosodic frames also see Dell, 1986; Fromkin, 1971; Garrett, 1975; Levelt, 1989 and Levelt, Roelofs, & Meyer, 1999; but also see Dell, Juliano, & Govindjee, 1993). Shattuck-Hufnagel states that "in scanning for a segment to fill a word-onset slot in the frame, the scan-copy mechanism isolates all word-onset segments in the buffer, which are already represented as separate from the rest of the word, and then scans across the candidate segments in this set". Obviously, in this view the evidence for the special position of word onset consonants as obtained by studying segmental errors of speech has become a major factor in modeling the process of serial ordering of speech segments during speech preparation.

It should be noted that the predominance of word onset consonants was not predicted from the original scan-copy model (Shattuck-Hufnagel, 1987). If there would be another explanation for the word onset effect in speech errors, then this model or similar models of the mental preparation of speech would have no need to accommodate any predominance of word onset consonants in speech errors. The spreading-activation theory of lexical retrieval during speech production by Dell (1986) does not account for the word onset effect in segmental speech errors. Neither does the spreading activation theory of lexical access by Levelt et al. (1999) as implemented in the computational model WEAVER++ (Roelofs, 2000). Of course, if the word onset effect is real, then such theories should account for this effect. Dell (1986, 312) indeed acknowledges that "initial sounds of words and syllables tend to slip more than the other parts", presumably because "they are, in general, easy to retrieve—or, to use activation terms, they become highly activated quickly (...). Thus, although the correct initial sound tends to be highly activated, so do the initial sounds of competing syllables from other parts of the utterance. As a result, these highly active competitors often replace the correct sounds". It should be noted that the above explanation of the word onset effect by Dell (1986) only holds for interactional errors. A word initialness effect in noninteractional errors (errors without an apparent source in the immediate environment), as claimed by Shattuck-Hufnagel (1987) would have to be explained in a different manner.

Dell (1988) introduces an alternative model having nodes corresponding to "word-shape headers", which often are similar to syllable templates, for example CVC, CV, VC. The 1988 model was successful in simulating familiarity and similarity effects in phonological speech errors, but the model did not account for the initialness effect, basically because, as in the 1986 model, all segmental positions are treated in the same way. Assuming that every now and then under the influence of syllable templates in the environment the wrong syllable template is chosen, one can explain (as demonstrated by Hartsuiker, 2002), that sound addition errors occur more frequently than sound deletion errors, simply because CVC is the most frequent syllable template. But again, Hartsuiker's version of Dell's 1988 model does not explain the word onset effect.

The various models of serial ordering of sound segments mentioned so far clearly distinguish between structure ("prosodic frames" with slots specified for segments with specific properties) and content (activated segments to be inserted in the appropriate slots). It should be noted that such a set-up makes it possible to distinguish between *retrieving* or *activating* the phonemic segments of a particular word form on the one hand and *ordering* or *misordering* the activated phonemic segments postlexically on the other. We propose, in line with Shattuck-Hufnagel (1983, 1987, 1992) and with Levelt et al. (1999) that it makes good sense to distinguish between a lexical process of segment retrieval and a postlexical process of serial ordering of segments. We specifically propose that noninteractional errors mainly result from lexical retrieval processes. As we will see below, such noninteractional errors support the claim (Brown & McNeill, 1966; Burke, MacKay, Worthley, & Wade, 1991) that word initial segments are more accurately retrieved than other segments. We also propose that interactional errors result from postlexical ordering of segments that have been retrieved for the two or three words that are about to be uttered. It takes attention to keep these segments apart and in their proper position (cf. Nozari & Dell, 2012). When attention fails, retrieved segments sharing the same position in the word may interact. In this interaction there is no preference for a particular position in the word. From this one could expect that sound retrieval errors and interactional sound errors might have different properties. Such different properties are not allowed by the matching distributed processing model, proposed by Dell et al. (1993). In this model, "linguistic structure is not distinguished a priori from linguistic content. Rather, structural or rule-like effects emerge from the storage of many individual linguistic strings. Storage, or learning, takes place by changing connection strengths or weights among units in a network". The model was trained for generating phonological forms of single words, using a "backpropagation algorithm" (Rumelhart & McClelland, 1986), and various different vocabularies, with the specific purpose of investigating whether certain well known effects in segmental speech errors can be simulated without introducing structure and content explicitly (Dell et al., 1993). These effects were to result simply from the structure of the vocabulary. The model of Dell et al. (1993) appeared to correctly simulate for American English the phonotactic regularity effect, the consonant–vowel category effect, the syllable constituent effect, and, most relevant for the present purpose, the word onset or word initialness effect. Clearly, errors in generating word forms were more frequent for the initial position than for other positions. However, the model has a serious limitation in that the only errors it makes are noninteractional errors.