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

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.

It does not generate interactional or movement errors. Shattuck-Hufnagel (1987) observed that the word-onset effect is strongest for interactional errors (although for monosyllables still 62% of all noninteractional errors are word-initial; for polysyllables this is only 28%). The word onset effect is simulated in the model by Dell et al. (1993) because of the sequential nature of its operation: Segments are retrieved in the serial order of the canonical word form, and contextual uncertainty is maximal in the onset segment, and then gradually decreases towards the final segment of the word form. During a postlexical process of serial ordering of segments, segments of different syllables and words must be present simultaneously, as shown most clearly by typical sound exchanges such as *bood geer* for *good beer*, where obviously the /b/ is activated when the /g/ of *good* is to be pronounced. Such errors are not accounted for by the matching distributed processing model proposed by Dell et al. (1993).

There is an interesting difference between interactional errors and noninteractional errors. As discussed above, the first are most easily (although perhaps not necessarily) explained by a postlexical mechanism of serial ordering of activated sound segments. However, the latter can have two different origins. As mentioned earlier, noninteractional errors can result from problems with the initial lexical retrieval or activation of the sound segments making up the word form. However, they can also result from problems during postlexical serial ordering of sound segments, for example because sound segments are activated by retrieval of word forms that compete with word forms actually to be uttered. Such competing word forms and their component segments remain hidden for the listener and for the student of speech errors, but they can affect the postlexical process of serial ordering of segments. We can call such cases "covert interactional errors". Assuming these two different origins of noninteractional speech errors may help to explain seemingly contradictory effects reported for these errors. There is evidence that word initial segments are not less but more resistant against retrieval errors than segments in other positions, as seen from the resistance of word initial segments against errors in retrieval attempts during the tip of the tongue state (Brown & McNeill, 1966; Burke et al., 1991). This is confirmed for aphasic patients by Schwartz, Wilshire, Gagnon, and Polansky (2004), who found in a picture naming task with many aphasic patients that at least "remote" nonword errors (errors that are phonologically dissimilar with the target segment and lead to nonwords), show a decline in frequency over serial position within the word, i.e. the opposite of the word initialness effect in interactional speech errors. But according to Shattuck-Hufnagel (1987) noninteractional segmental speech errors occur not less but more frequently in word onsets than in other positions, although this word onset effect is less for noninteractional than for interactional errors. The latter difference can be explained by our assumption that noninteractional errors have two different sources, one that is suffering from the word onset effect and one that is suffering from the opposite effect. It is noteworthy that Schwartz et al. (2004), who reported a reverse word onset effect for aphasic patients, found no evidence for a dual origin (lexical retrieval and postlexical processing) of nonword sound errors in 457 nonword errors made by 18 subjects with fluent aphasia. This may indicate that most errors made by patients suffering from aphasia are made during word form retrieval and not during subsequent postlexical processing or ordering. Below, we will concentrate on interactional errors and we will assume that these are caused during postlexical processing, i.e. during the serial ordering of segments. We also propose that the so-called word onset effect in segmental speech errors does not stem from the way the serial ordering of speech segments is organized, but rather from the phonotactic structure of the language. This proposal is motivated by the fact that the word onset effect appears to be non-existent in Spanish (Berg, 1991; Del Viso, Igoa, & García-Albea, 1991; García-Albea, Del Viso, & Igoa, 1989). The process of serial ordering of activated sound segments presumably does not depend on the language spoken.

The present study examines whether the word onset effect found by Shattuck-Hufnagel (1987) may in a certain sense be an artifact. We argue that the reported effect is caused by using the relative frequencies of segmental positions in running speech (word onset as compared to other positions) as a predictor of relative frequencies of speech errors for these positions. We propose that a more reasonable predictor is provided by the number of phonotactic opportunities which segments in each position in the word have to interact with other segments in the immediate context. If this alternative predictor would indeed explain the so-called word onset effect, then there would be no need for special mechanisms in the process of mental preparation of speech to account for the word onset effect. Below, we will examine the evidence for the word onset effect in interaction errors more closely, by means of an independent analysis of interactional segmental ordering errors (i.e. substitution errors: exchanges, anticipations and perseverations) , for which a source of the error can be found within the immediate environment, in the Utrecht corpus of errors of speech (Schelvis, 1985). If the relative frequencies of interaction speech errors can indeed be predicted from the numbers of opportunities for interactions each segment position has, this would vindicate models in which all segment positions are treated equally (e.g. Dell, 1986) , and would rule out models with special provisions for word onset consonants (e.g. Shattuck-Hufnagel, 1992)

As mentioned before, the main idea underlying this investigation is that, where relative frequency of word onset consonants as compared to other consonants is not a useful predictor of relative frequency of word onset consonantal speech errors as compared to consonantal errors in other positions, these relative frequencies can best be predicted from the number of opportunities segments in different positions have to interact with other segments. Consider the following utterance:

more beer for Kerr

We count four word onset consonants and four other consonants. From the relative frequency of word onsets and other consonantal positions one would predict an equal number of word onset errors and other errors. Of course, this would be far from a reasonable prediction for this particular word sequence. If we count the number of opportunities for each consonant to be substituted by another consonant, assuming that word onset consonants and word final consonants do not interact, we see that in this phrase word onset consonants provide $4 \times 3 = 12$ and word final consonants zero opportunities of being substituted by interaction with another consonant in the immediate environment. This is a somewhat extreme example, but the point is that each constraint on speech errors that is not equally distributed over word onset and other positions may be a confound, and hence may invalidate the prediction of relative number of speech errors from the relative frequencies of word onset consonants and other consonants. If word final consonants of neighboring words are more often identical than word onset consonants, this would upset a valid prediction of numbers of consonantal errors from relative frequencies of word onsets and other consonantal positions: word onset consonants would be much more often involved in consonantal substitutions than predicted from their relative frequency. For this reason we propose that it makes more sense to predict relative frequencies of segmental speech errors in different positions in words from an estimate of the number of opportunities which speech segments in each position in the word have to interact with other segments in the immediate environment, taking into account the phonotactic constraints on segmental speech errors. If we then still find many more word onset errors than predicted, then we can confirm the special role of word onset consonants in speech preparation.

Preliminaries

We want to examine whether the relative frequencies of segmental speech errors can be predicted from the relative number of opportunities segments in each position have to interact with segments in the immediate environment. Before we do this, we will in these preliminaries first examine whether the predominance of word onsets as reported by Shattuck-Hufnagel (1987) also occurs in Dutch. Then we will define what for our purpose a word is, and thus what word initial, word medial and word final means. Next we will discuss a number of phonotactic constraints on speech errors that might create problems for comparing our predictions with actual frequencies, leading to a reasoned exclusion of a number of types of speech errors from our analysis. Finally, before we will report the results of our analysis, we will describe how we constructed a corpus of spontaneous speech that reflects as closely as possible the statistical properties of spontaneous speech as used in the speech errors studied.

We have not limited the segmental positions to word onset on the one hand and other consonantal positions on the other hand. We think that the special role of word onset consonantal positions can be tested more precisely when we distinguish between word onsets, word internal consonantal positions, word final consonantal positions and vowel positions. One would expect that, if word onsets are special, word onset errors would be much more frequent than predicted from the relative number of opportunities for speech errors, and that in word medial, word final and vowel positions segmental speech errors would be less frequent than predicted. We have chosen to limit our analysis to a subset of speech errors from our corpus, in counting observed speech errors, in counting the frequencies of segmental positions involved (as was done by Shattuck-Hufnagel, 1987), and in counting the number of phonotactically allowed opportunities for speech errors to occur. This subset was established by excluding cases that suffer from phonotactic constraints that are obviously or potentially unbalanced over the positions to be compared, as we will explain below.

Comparing American English and Dutch

Before we endeavor to examine the possible cause of the word onset effect by analyzing consonantal speech errors in our corpus of speech errors in spontaneous Dutch (Schelvis, 1985), we have to know whether the situation in Dutch is similar to what is reported for American English. This is important because, as mentioned in the introduction, the effect may be language dependent. So, similar to what Shattuck-Hufnagel did, we counted consonantal errors of all types in the corpus of speech errors, distinguishing between word onset errors and errors in other positions. We used the same corpus of misspoken utterances (after repairing the speech errors, see below) to estimate the relative frequencies of consonants in word onsets relative to consonants in other positions. Words were defined as the units that would have been separated by blanks in writing. This definition of words for the current purpose will be justified below. As was done by Shattuck-Hufnagel, all consonants in word onset consonant clusters were counted as word onset consonants. We found that 50% (322 vs 317) of consonantal errors were word onset errors, whereas only 39% (2356 vs 5250) of consonants in the corpus are found in word onset positions, including the consonants in consonant clusters. The difference is considerable and highly significant ($\text{chi}^2 = 32.2$; df = 1; $p < 0.0001$). Yet, the word onset effect in consonantal speech errors in Dutch does not seem to be equally massive as the effect reported for American English by Shattuck-Hufnagel (1987). Possibly this is related to a difference in the definition of words. It is not clear to us how words were defined in the counts reported by Shattuck-Hufnagel (1987). However, the situation in the two languages is sufficiently similar to warrant an attempt to search for possible causes of the effect by further analyzing segmental errors in our Dutch corpus.

Defining words and positions within words

Before we will discuss phonotactic constraints on segmental speech errors that potentially may upset our predictions of relative frequencies of speech errors in different positions, we must first define word onsets and word endings. The suggestion by Meringer (1895) that both word onsets and root onsets have greater “weight” than other positions, requires the notion “word” to be interpreted. It obviously refers here to those lexical units that are activated during speech planning and then given form by a sequence of phoneme-like speech segments, in which process the very first segment might play a special role. According to linguistic analysis the smallest such units would be morphemes. If we would follow Meringer and Mayer, a Dutch word form such as *gewerkt*, being the past participle of the verb root *werk* ('to work') formed by adding *ge-* and *-t* to the root morpheme *werk*, would contain two heavy-weight onsets, the *g* being a word onset and the *w* being a root onset. Shattuck-Hufnagel (1987) does not discuss this problem explicitly, but from her examples we see that an English word such as *repaid* has only one onset consonant, and that is the word onset *r*. The root onset *p* is not considered an onset consonant. We agree with this analysis, since it is reasonable to argue that a frequent word like *repaid* is highly lexicalized and readied for speaking as a single unit. This is possibly so for virtually all words encountered in spontaneous speech. If indeed the degree of lexicalization affects whether or not sub-word units such as root morphemes are separately activated, it becomes difficult to know in each specific case what the relevant lexical units are, because it is not always clear to what extent a particular complex word form is lexicalized. However, in spontaneous speech in Dutch most words are monomorphemic. Therefore, very likely for our purpose it would not make much difference whether we define the lexical units involved as words or as word roots or content morphemes. To see whether this is indeed so, we counted in our collection of speech errors all substitutions of single consonants by single consonants, and counted the number of cases where position of the error and position of the source of the error in the word (initial, medial, final), and position of the error and position of the source of the error in the (non-inflectional) morpheme (initial, medial and final), and position with respect to word stress (collapsing primary and secondary word stress) coincided. An example of a substitution in which error and source have the same position with respect to word, morpheme and stress is the following substitution (in exchanges we only count the first substitution, for reasons to be explained later):

geld en tijd ('money and time') > *teld en gijd*

An example of a substitution in which error and source of the error have nothing in common is the following substitution :

polynoom ('polynomial') > *polymoon*

As it happens, in 66% of 218 substitutions of single consonants by single consonants error and source have the same position with respect to word, morpheme and stress because these positions coincide. In 72% error and source have the same position with respect to word and root morpheme, but not always with respect to stress. If we exclude substitutions where error and source are within the same word, because then for obvious reasons the coinciding of all three positions is impossible, we find that in 76% of 172 substitutions positions of error and source with respect to word, root morpheme and stress, coincide, and in 79% of these substitutions positions with respect to word and morpheme but not always with respect to stress, coincide. This harmony between words and morphemes has the advantage that, if onset consonants are special, this at least statistically will follow from our analysis if we simply define words as the units that in standard writing would have been separated by blanks. This is what we have done. This strategy also has the advantage that it is easily and exactly reproducible by other researchers. Only consonants in onsets of these words are counted as word initial consonants, and only consonants in word endings are counted as word final consonants. All consonants not being initial and not being final are considered to be medial. All vowels (except schwas, see below) are simply counted as vowels, and are not further classified as to position within the word.

Limiting the speech errors to be considered

Our proposal that the frequencies of segmental speech errors in different positions may be predicted from the number of opportunities each segment has to be substituted by another segment in the immediate environment brings along its own problems in testing. Not all constraints on segmental speech errors are all-or-none and not all constraints are well known. Take the following example taken from our corpus:

Nijmegen weer ('Nijmegen again') > *Nijwegen weer.*

In assessing the number of possible substitutions in this string of two words, should we consider this substitution as one of the opportunities or not? As we will see below, we do not, because, although such substitutions of medial by initial word consonants (and vice versa) apparently do occur, they are quite rare. Therefore it seems warranted not to use them in our predictions. We will now briefly discuss those factors that we know affect the probability of speech errors in ways that might upset our predictions:

1. Type of speech error: Substitution, addition and omission.

Examples of these:

Substitution:	*zou kunnen kijken* ('could have a look') > *zou kunnen zijken*
Addition:	*bel de blauwe* ('ring the blue') > *blel de blauwe*
Omission:	*drukte door* ('pushed through') > *dukte door*

All types of segments suffer from substitutions, but generally only consonants, not vowels, suffer from addition and omission. Thus including additions and omissions would make it difficult to compare consonants and vowels. It should also be noted that the probabilities for additions of consonants to consonants are severely constrained by the consonant clusters that are allowed by the phonotactics of the language. It is a priori unlikely that such constraints on clusters are completely comparable between different positions within the word or morpheme, because the clusters are different for different positions. For these two reasons we have limited our analysis to substitutions, excluding additions and omissions. In the class of substitutions we have collapsed anticipations, perseverations and exchanges. Examples of these three forms of substitutions are:

Anticipation:	*pit los zit* ('wick is loose') > *pis…pit los zit*
Perseveration:	*dus ja* ('thus yes') > *dus da*
Exchange:	*tuin en keuken* ('garden and kitchen') > *kuin en teuken*

Of course in fact an exchange consists of two substitutions, an anticipatory and a perseveratory substitution. However, we agree with Shattuck-Hufnagel (1983), that these two substitutions are not independent. The second appears to be an automatic result of the first. We have shown elsewhere that in internal speech, before self-monitoring for speech errors has applied, exchanges are considerably more frequent than anticipations and perseverations together. This is clear evidence that the two parts of a sound exchange are not independent (Nooteboom & Quené, 2013). With respect to exchanges, Shattuck-Hufnagel (1987) has observed that the relative frequency of word onset errors is greater for exchanges than for anticipations and perseverations. Also she mentions that in exchanges there is less doubt about the interaction between the two segments involved than in anticipations and perseverations. Thus she excluded anticipations and perseverations from further assessment of the word onset effect, limiting her further analysis to exchanges and incomplete errors such as *bos pakken* ('bundle take') > *bok…bos pakken*. This selection was not possible here, simply because our corpus of speech errors in spontaneous Dutch is somewhat too small to exclude so many cases. But we will later examine whether in our corpus word onset consonants are relatively more often substituted in exchanges than in anticipations and perseverations.

2. Single consonants versus clusters.

Whole consonant clusters are much less frequently substituted by other clusters or by single consonants than single consonants are. The frequency of cluster substitutions probably depends on cluster properties in unknown ways. Also, cluster types differ for different segmental positions, potentially causing imbalance over different positions. If our corpus consisted of many thousands of segmental errors instead of a few hundred we could probably assess the relative weight of factors involved in substitutions of clusters and single consonants. As it is, we cannot. This would bring quite some uncertainty to our predictions of relative frequencies of segmental errors in different positions. Also clusters basically are limited to consonants (we treat diphthongs as single vowels, because in speech errors they behave as single vowels; cf. Nooteboom, 1973). This would make comparisons between consonants and vowels more difficult. For these reasons we exclude substitutions in which whole clusters are replaced both in counting speech errors and in establishing predictions, and count only substitutions involving *single* consonants. For defining what single consonants are, a particular problem is created by the fact that in words beginning with a vowel the empty word onsets quite regularly interact with initial single consonants, examples being *eet heel* > *heet heel*, *meerdere eer* > *eerdere eer*, *eerder willen* > *weerder willen, aap uit de mouw* > *maap..aap uit de mouw* and *elfde hokje* > *helfde hokje*. On the other hand, in words ending with a vowel, word final empty positions never interact with word final single consonants. For this reason, we have counted each word initial empty position, preceding a vowel, as a word initial consonant but we have not done so for word final empty positions.

3. Single consonants within clusters versus single consonants not within clusters.

Clusters severely constrain the number of opportunities for substitutions. For example the /s/ in a word initial cluster like /stʁ/ is never substituted with another consonant; the /t/ in the same cluster can only be replaced with /p/ or /k/, whereas the /ʁ/ again is never substituted by another consonant. These constraints may easily upset the balance between different positions. Also constraints within clusters inhibit a fair comparison between consonantal errors and vowel errors. Therefore we exclude substitutions of consonants within word-internal clusters, both in the predictions and in the counting of substitutions. We do not only exclude consonants belonging to morpheme-internal clusters, for example the *p* and *l* in *diploma*, but also consonants belonging to clusters containing a morpheme boundary, as the *r* and *b* in *deurbel* ('doorbell'). The reason is the following. We use the number of opportunities of segments to be substituted by other segments in their immediate environment as predictor of frequencies of segmental substitutions. Now if words and content morphemes are both lexical units of which the onsets are potentially special, we should count both as word onset consonants. If only words and not content morphemes are lexical units with potentially special onsets, we should count word internal morpheme initial consonants as word medial consonants. As cases in which word initial consonants and morpheme initial consonants do not coincide are relatively scarce, we have no way to find out (without doing experiments). But it appears that if we want to include these consonants in clusters that contain a morpheme boundary in making predictions of relative frequencies of substitutions, we run into difficulties because we do not know what the phonotactic constraints on such potential substitutions are. And again, our corpus is too limited to find out. Such uncertainties upset our predictions. Therefore we have *excluded all word internal clusters*, limiting our analysis to substitutions of single vowels and single word initial, word medial and word final consonants by single vowels and single word initial, medial and final consonants. This also implies that all word medial consonants in our analysis are intervocalic.

4. Position in syllable, morpheme and word.

In assessing the opportunities each segment has to interact with other segments in the immediate context, we have to determine what the constraints on such interactions are. For vowels, other than schwa, this is relatively simple. In Dutch short vowels are basically excluded as word endings, and diphthongs can not fill a position immediately before /ʁ/. Long vowels and diphthongs are excluded before /ʁ/. So potential substitutions leading to violations of these constraints are not counted. For single consonants (we have already excluded consonants within word-internal clusters) the situation is somewhat more complex. It has often been observed that substitutions of consonants virtually always are limited to consonants occupying the same position with respect to the vowel: Prevocalic and postvocalic consonants do not interact with each other. Consonantal interaction seems to be limited to interactions between consonants that occupy the same position within the syllable (Del Viso et al., 1991; Dell, 1986, 1988; García-Albea et al., 1989; Nooteboom, 1973; Poulisse, 1999 and others). However, Shattuck-Hufnagel (1983, 1987, 1992, 2011) has claimed repeatedly that hard evidence for the syllable as a framework for serial ordering of segments is nonexistent. It seems possible to account for all known facts without resorting to the syllable, notably by employing constraints imposed by words and morphemes. Of course, word initial single consonants are necessarily prevocalic and word-final single consonants are necessarily postvocalic. We know that word-initial and word-final consonants hardly ever interact, but of course this can be accounted for without resorting to the syllable. The question is how other consonants, and particularly intervocalic consonants, behave. Some phonologists plead for syllabification rules in speech production which would make virtually all intervocalic consonants prevocalic (For English see Kahn, 1976; for Dutch see Booij, 1995). This would mean that intervocalic consonants would preferably interact with clearly prevocalic (word- or morpheme-initial) consonants, and rarely with clearly postvocalic consonants. They would not necessarily have a preference to interact with other intervocalic consonants above word-initial consonants, because both clearly prevocalic and intervocalic consonants would occupy the same position within the syllable. In order to see how such constraints work out in segmental substitutions we have assessed the error and source positions of all substitutions of consonants that are not part of clusters. Note again that all medial consonants are intervocalic. The results of this count are given in Fig. 1.

Obviously, initial single consonants are never substituted by single final consonants, and single final consonants are hardly ever substituted by single initial consonants. This was well known (Dell, 1986; Fromkin, 1971; Nooteboom, 1973; Shattuck-Hufnagel, 1987). Single intervocalic consonants (such as the /m/ in *diploma*) are rarely substituted by single initial or final consonants. They are mostly substituted by other single intervocalic consonants. Apparently, word-internal single medial, intervocalic, consonants form a separate class in themselves, and substitutions of these intervocalic consonants are largely constrained to other intervocalic consonants. These data throw doubt on the syllable as a framework for the serial ordering of segments. In counting the opportunities specific segments have to be substituted by other segments in their immediate environment, we have only counted substitutions of initial by initial, medial by medial, and final by final consonants, neglecting the relatively rare other combinations. One may observe that the constraint by these positions also solves the problem created by the fact that voiced plosives and voiced fricatives are excluded from word endings, and that /h/ is limited to word beginnings and /ŋ/ to word endings.

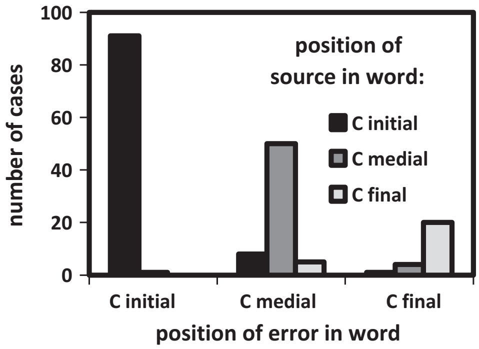

Fig. 1. Numbers of substitutions of single word initial, medial and final consonants by single word initial, medial and final consonants. $N = 180$. Data from speech errors in spontaneous Dutch.

5. Phonetic similarity.

The probability that one segment replaces another increases with increasing phonetic similarity between the two segments (among others Dell, 1986; Nooteboom, 1973, 2005; Nooteboom & Quené, 2008). This is a strong effect and makes that most segmental substitutions are between phonetically similar segments. This factor is difficult to control for in an analysis of segmental errors made in spontaneous speech. We assume here that the effect of this factor is on average similar for different segmental positions.

6. Stress and accent.

As the speech errors in our corpus were noted down in writing, we do not know which words were pitch-accented and which were not. We therefore did not consider pitch accent as a factor. One might argue that all full vowels in Dutch carry some degree of stress. The only fully unstressed vowel is the schwa. Interestingly, our corpus of 999 segmental speech errors contains not a single case of schwa interacting with another vowel. This is remarkable because the schwa is the most frequent phoneme in Dutch (token frequency: 10.97%; cf. Luyckx, Kloots, Coussé, & Gillis, 2007). We interpret this as an effect of stress: complete absence of stress apparently shields the schwa from being vulnerable to substitutional interaction with other vowels. But this also means that counting the schwa as a vowel in our assessment of the opportunities each vowel has of being replaced by another vowel might seriously bias our prediction. We solved this by simply not considering the schwa to be a speech segment that can be involved in segmental substitutions. We did not further distinguish between degrees of stress. A special case that seems to be stress-related is that not only schwas themselves but also word-final consonants following schwas in our corpus are never, as single segments, involved in speech errors: The final consonants of words like *jager* (/ja:χəʁ/), *matig* (/ma:təχ/) or *lelijk* (/le:lək/) seem to be shielded from interaction with other segments. As such consonants are relatively frequent in Dutch, we have left out those consonants both from the predictions and from counting substitutions. Shattuck-Hufnagel (1987, 1992) demonstrated that pre-stress consonants are more vulnerable to interactions with other consonants than consonants that are not in pre-stress position, but otherwise comparable. If so, this should be included in our prediction. However, we begin in not doing this.

7. Word class.

Segmental speech errors are more frequent in content words than in function words (cf. Nooteboom, 1973). We have assumed that this factor has basically similar effects for the positions to be compared. There is one exception. Segments belonging to articles (in Dutch *een* (/ən/), *de* (/də/) and *het* (/ət/), all three with the vowel schwa) are, with one exception, in our corpus never, as single segments, involved in speech errors, although articles constitute 14% of all word tokens in spontaneous Dutch speech (cf. Nooteboom, 1973). The exception is the following:

gehemelte en de huig ('palate and the velum') > *gehemelte en ge huig*

That articles seem to be shielded from substitutional interactions would potentially upset our predictions, if only because the relative frequencies of initial and final consonants, contributing to our predictions, differ within this set of very frequent words. Again we solved this by *leaving out articles* both from the predictions and from the counting of speech errors.

Constructing a corpus of spontaneous speech

After all the above exclusions were applied, there remained a set of 415 cases involving substitutions of vowels and word initial, medial and final single consonants not being part of word internal clusters. For making two different predictions of the distribution of substitutions over the four positions within words, we constructed a matching corpus of spontaneous speech in the following way. We selected all segmental speech errors with an identifiable source and target within the utterance in the entire corpus of speech errors in spontaneous Dutch. We then removed all speech errors which are in another language, mostly in English, some in French. This gave a set of 971 segmental speech errors. We then removed all speech errors that we found too ambiguous as to whether they were segmental or lexical after which there remained a set of 632 speech errors. For this set each speech error was repaired into the intended correct utterance. After that we removed all words before and after the sequence of words containing both source and target of the segmental error (see Nooteboom, 2011, for further details on this corpus of word sequences). Next we removed the definite and indefinite articles from the word strings. The remaining strings of words varied in length from 1 to 9 words. However, the number of opportunities for segmental substitutions within word sequences increases enormously with greater length of the strings. Because speech errors in longer strings of words are extremely rare, this may bias our predictions disproportionately. We therefore limited the length of strings of words to five words by removing all words after the first five of the string concerned. Here follows an example:

1. Speech error: *mijn voeder...eh mijn moeder had het verhaal* ('my stother...uh my mother had the story')
2. Repaired speech error: *mijn moeder had het verhaal*
3. Removing words before and after the string containing error and source: *moeder had het verhaal*
4. Removing article: *moeder had verhaal*
5. (Removing all words after the first five; not applicable here).

For the current purpose this manner of constructing a matching corpus of spontaneous speech has the advantage that counting the number of opportunities that segments have to be substituted by other segments is limited to word sequences that reflect the statistical properties of the often brief immediate context of the spontaneous speech errors in the corpus.

Counting and results

Applying all exclusions of speech errors as discussed above, we counted the numbers of substitutions separately for vowels, and for word initial, medial and final single consonants. We also counted the numbers of all segments in the corpus, separately for each position (this is the predictor used by Shattuck-Hufnagel, 1987) and the number of all phonotactically allowed interactional substitutions of single segments for each of the four positions. An example of how we counted the number of segments for each position is given for the following word string:

heterdaad betrappen ('catch redhanded')

The string has two words, both starting with a single consonant. Therefore there are two word initial single consonants. Note that if we had defined content morphemes as the relevant lexical units, we would have very likely counted the initial /d/ of *-daad* as an onset segment. But we did not. There are also two medial single consonants, the /t/ in *heterdaad* and the /p/ in *betrappen* (*pp* corresponds with only a single /p/ in Dutch). There is only a

Table 1
Each column gives numbers referring to a particular segmental position. The first row of numbers gives actual numbers of single segment interactional substitutions in the corpus, for the position concerned. The second row gives the total numbers of segments in the corpus for each position. The third row gives the total numbers of opportunities for interactional substitution in the corpus for each segmental position. Between brackets are the proportions per segmental position.

	Vowel	Initial cons.	Medial cons.	Final cons.	Grand total
Nr. segment subst.	235 (0.57)	92 (0.22)	63 (0.15)	25 (0.06)	415 (1)
Nr. segments in corpus	2056 (0.43)	1241 (0.26)	739 (0.16)	713 (0.15)	4749 (1)
Nr. of opportunities for subst.	4929 (0.61)	1670 (0.20)	966 (0.12)	580 (0.07)	8145 (1)

single final consonant, the final /t/ of *heterdaad* (spelled as *d*). The *n* of *betrappen* is not pronounced in standard Dutch. If it were pronounced it would not have been counted because it would have been a word final consonant preceded by schwa, and those we have excluded. Finally, there are three full vowels, the /e:/ in the first syllable and the /a:/ in the third syllable of *heterdaad*, and the /ɑ/ in the second syllable of *betrappen*. The other three vowels are all schwas, and those we had excluded from contributing to the predictions. In this way we treated all 632 word strings, and summed the results for each of the four positions, and also assessed the grand total. We then divided each of the four sums by the grand total in order to get a prediction of the relative number of segmental substitutions in each of the four segment positions. The resulting numbers and the calculated relative frequencies are given in the second row of numbers in Table 1, where the first row of numbers gives the observed numbers and relative frequencies of substitutions of single segments for each of the four positions.

As mentioned, we also counted for each vowel and each single word initial, medial and final consonant all phonotactically allowed opportunities for being substituted by another segment within the same word string. Let us use the same example:

heterdaad betrappen

The three full vowels generate together $3 \times 2 = 6$ potential vowel substitutions. There are two potential word initial substitutions, /h/ by /b/ and /b/ by /h/. There are also two potential word medial substitutions, the first /t/ by /p/ and vice versa. There are no potential word final substitutions, because the last /t/ of "heterdaad" is the only word final consonant. For this word sequence the total number of substitutions is 10, and the relative number of substitutions as fractions of 1 are 0.6 for vowels, 0.2 for initial consonants, 0.2 for medial consonants and 0 for final consonants. We did this for the whole corpus of word sequences, assessing the grand total and dividing the sum for each of the four positions by the grand total of all possible substitutions in order to get a prediction of the relative numbers of substitutions. The resulting numbers and fractions are given in the bottom row in Table 1.

Now we have two separate predictions for the relative frequencies of segmental substitutions for four different positions, one prediction (a) from the relative frequencies of those positions in spontaneous speech, and the other (b) from the relative number of potential substitutions in the immediate context for each of these positions. By multiplying the total number of segmental speech errors, 415, with each of the proportions given in the second and third row of numbers in Table 1, we obtain expected frequencies that can be compared with the observed numbers of errors. We hypothesize that the observed frequencies of errors are better predicted by prediction (b) than by prediction (a), taking into account all we know about the phonotactic constraints on segmental speech errors.

In Fig. 2 we can compare the actual numbers found with the two sets of predictions stemming from the relative frequencies of the four positions (a) and stemming from the number of opportunities for interaction for each position (b).

The observed distribution differs significantly from prediction a on a chi^2 test (chi^2 = 41.9; df = 3; $p < 0.0001$). Clearly, vowel substitutions greatly outnumber substitutions in each of the other three positions. This may be an effect of excluding many consonantal substitutions. Also there are more vowel substitutions than predicted from the frequencies of positions for vowels. Interestingly, there are not more but fewer substitutions of initial consonants than predicted from the relative frequencies of the four positions. This can be compared with the very different numbers we obtained earlier, when counting consonant errors of all kinds (substitutions, additions, omissions of segments and in all contexts, including consonant clusters). There we found that 50% of all consonant errors were onset errors, whereas only 39% of all consonants were onset consonants. Obviously, excluding additions and omissions, and excluding consonants within clusters leads to very different results, doing away with the unexplained

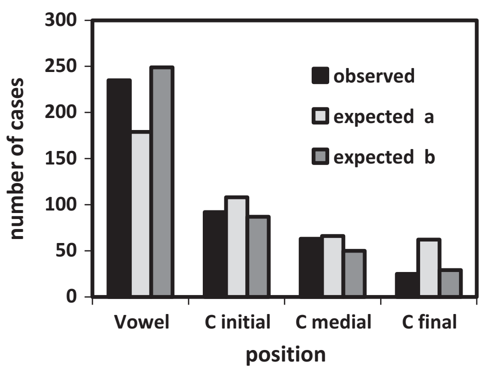

Fig. 2. Observed numbers of substitutions of single segments (black), numbers predicted from the frequencies of the four positions in spontaneous speech: expected a (light grey), and numbers predicted from the numbers of opportunities for interaction for each position; expected b (dark grey). $N = 415$.

predominance of word onset consonants in segmental speech errors. This can be accounted for by assuming that constraints on additions and omissions and constraints on segmental errors in consonant clusters are very different for the three consonantal positions.

In Fig. 2 we can also compare the actual numbers of substitutions of vowels and single initial, medial and final consonants with those predicted from the relative numbers of potential phonotactically allowed substitutions within the immediate context (b). Here, the two distributions are very similar. Our attempt to exclude dubious cases from the predictions has worked well, despite the uncertainties in the relative probabilities of the possible substitutions. We conclude that these uncertainties, such as the neglected effect of phonetic similarity and the neglected relative probability of interactions between word initial, medial and final consonantal segments, and the neglected effect of pre-stress position for consonants must be evenly distributed over the different positions. There is indeed no significant difference between the two distributions on a chi^2 test ($chi^2 = 6.2$; df = 3; $p = .1042$). Apparently, the frequency of substitutions of single segments in different positions in words can be accurately predicted from the number of probabilities specific segments have to be substituted by segments in the same position in words within the immediate context.

As mentioned earlier, Shattuck-Hufnagel found in her analysis of segmental speech errors in spontaneous American English evidence that word onset errors are relatively more frequent in exchanges plus interruptions (most interruptions stem from halfway repaired exchanges in inner speech, cf. Shattuck-Hufnagel, 1983 and Nooteboom & Quené, 2013) than in anticipations plus perseverations. To see whether this would also be true for our corpus of segmental errors in spontaneous speech we counted the substitutions of single segments in the different positions within the word separately for exchanges plus interruptions and for anticipations plus perseverations. The result of this count is given in Table 2. The two distributions do not differ significantly on a chi^2 test ($chi^2 = 2.6$; df = 3; $p = .4567$). Fig. 3 gives the observed numbers of substitutions for exchanges plus interruptions and the numbers as expected on the basis of the numbers of segments in each position (a) and on the basis of the numbers of opportunities for substitutions to occur in each position (b). Fig. 3 thus is comparable to Fig. 2, but now limited to exchanges plus interrupted speech errors.

The observed frequencies in Fig. 3 differ significantly from those expected from prediction a, ($chi^2 = 198.8$; df = 3; $p < .0.0002$), but do not differ significantly from those expected from prediction b ($chi^2 = 3.8$; df = 3; $p = .288$), as was the case in Fig. 2 for the whole set of substitutions. These results provide no evidence that in our

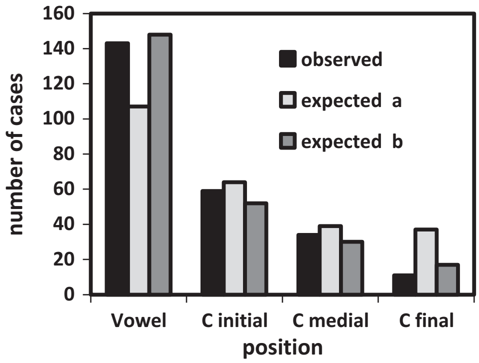

Fig. 3. Observed numbers of exchanges plus interrupted errors of single segments (black), numbers predicted from the frequencies of the four positions in spontaneous speech: expected a (light grey), and numbers predicted from the numbers of opportunities for interaction for each position; expected b (dark grey). $N = 247$.

corpus errors in word onsets would be more frequent in exchanges plus interruptions than in anticipations plus perseverations. Also these results are in good agreement with our proposal that frequencies of speech errors can be accurately predicted from the numbers of opportunities the segments in each position within the word have to interact with other segments in the immediate context.

In our analysis so far we have defined initial, medial and final consonants only with respect to the whole word, as separated by blanks in standard writing, as a unit. This was warranted because there is in our material a high degree of coincidence between word boundaries and morpheme boundaries, as discussed above (see Section 'Defining words and positions within words'). If we would have used root morphemes or content morphemes as units we would have found very similar results. But this presentation of our data in terms of word initial, medial and final consonants may suggest to the reader that all word medial consonants are equal, and that their position in a content morpheme has no effect. However, this is not necessarily the case. Of course, if in a segmental error the error and its source are both in word initial position, then error and source are of necessity also both in morpheme initial position and if they are both in word final position then they are of necessity also both in morpheme final position. For examining how morpheme boundaries affect the serial ordering of segments, we have to focus on those speech errors in which error and source of the error are not both word initial and not both word final. We can include the cases where error and source are both word medial because word medial consonants can be morpheme initial, medial or final. There are in the set of 415 errors represented earlier in Table 1, 68 errors where in principle

Table 2
Numbers of substitutions of single segments by single segments in four different positions in the word, separately for exchanges plus interruptions and for anticipations and perseverations. $N = 413$. The distributions do not differ significantly ($chi^2 = 2.604$; df = 3; $p = 0.4567$).

	Vowel	Word initial	Word medial	Word final
Exchanges + interruptions	142	60	34	12
Anticipations + perseverations	93	32	29	11

positions of error and source can differ with respect to word and morpheme. Some examples of such errors, with word internal morpheme boundaries indicated with "-":

(1) *ber-en-kuil* ('bear pit') > *ber-en-ruil*
(2) *Nijmegen weer* ('Nijmegen again') > *Nijwe…Nijmegen weer*
(3) *seminarie* ('seminary') > *serinarie*
(4) *sokk-en te stoppen* ('to mend socks') > *sopp-en te stokk-en*
(5) *bos pakk-en* ('bundle take') > *bok…bos pakk-en*

In (1) the morpheme initial intervocalic *k* (note that the *n* is not pronounced here) is substituted with the morpheme final *r*.

In (2) the morpheme medial *m* is substituted with the morpheme initial *w*.

In (3) the morpheme medial *m* is substituted with the morpheme medial *n*.

in (4) the morpheme final *k* is substituted with the morpheme final *p* (note that pp and kk stand for single /p/ and /k/ respectively).

In (5) the morpheme final s is substituted with the morpheme final k.

Fig. 4 gives the number of substitutions of morpheme initial, morpheme medial and morpheme final single consonants by other morpheme initial, medial and final consonants, limited to those substitutions in which the two interacting consonants are not both word initial, and thus forced to be morpheme initial, and are not both word final, and thus forced to be morpheme final. Obviously, substitutions are constrained as much by content morpheme boundaries as by word boundaries.

That within this set of 68 segmental errors morpheme initial substitutions are relatively rare and morpheme final substitutions are relatively frequent can be easily explained by the fact that word forms with a prefix are relatively rare and word forms with a suffix are relatively frequent in Dutch. Speech errors such as *bos pakken* ('bundle take') > *bok…bos pakken* are much more to be expected than speech errors such as *benut je het beste*('exploit you the best') > *benut je het neste*, once we have excluded all cases where word and morpheme boundaries coincide.

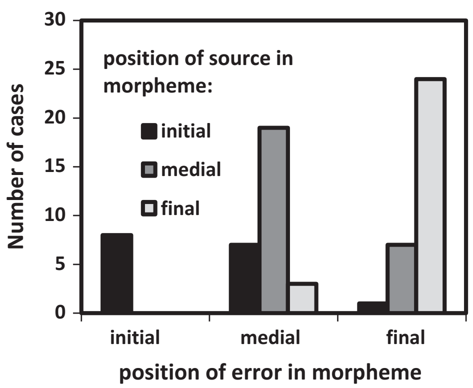

Fig. 4. Numbers of substitutions of morpheme initial, medial and final consonants in the set of segment substitutions in which word initial and morpheme initial, and word final and morpheme final positions of error and source are not forced to coincide ($N = 68$).

In sum, the results of this investigation provide no evidence that consonants in word onsets are treated differently from segments in other positions.

Discussion

We have seen above that the relative frequencies of Dutch interactional segmental speech errors in various positions in word forms can be predicted rather precisely from the numbers of phonotactically allowed opportunities segments in each position have for being substituted by other segments in the immediate environment. We have also seen that intervocalic consonants are mostly substituted by other intervocalic consonants, not by word initial or word final consonants. Frequencies of segmental substitutions are not only constrained by positions in the word, but also by segmental positions in the morpheme. Finally, we found that interactional substitutions of speech sounds are equally frequent in exchanges and interrupted speech errors as in anticipations and perseverations.

As pointed out in the introduction, the word onset effect appears to be non-existent in Spanish (Berg, 1991; Del Viso et al., 1991; García-Albea et al., 1989). Against the background of the current findings we can interpret this as indicating that in Spanish segments in different positions do not differ much in the number of opportunities they have for interaction with other segments in the immediate environment. Whether or not this is so, is still to be investigated. Some of our findings are different from what Shattuck-Hufnagel (1987) found for American English. She found that interactional segmental speech errors are much more frequent in word initial position than in other positions. She explained this from a model of speech preparation that had special nodes for word onset consonants. From the current results one might expect that if, for American English, one would, as we have done for Dutch, use as predictor the number of opportunities segments in each position have for interaction with other segments in the immediate environment, then the difference between word initial segments and segments in other positions (the word onset effect) perhaps would be accounted for. However, the reader may recall that we have defined words as the units that in standard writing would have been separated by blanks. Of course, this makes what we consider to be words dependent on a spelling convention. The convention that determines which units are and which are not separated by blanks differs between Dutch and English. This may be a possible cause of differences between what Shattuck-Hufnagel (1987) reported for American English and what we have found for Dutch. Whether this is the case or not is still to be investigated, but the effect of the different spelling conventions on the notation of segmental errors of speech in spontaneous speech may turn out to be quite limited. This is so because the convention only differs with respect to not very frequent compounds. These are supposedly rare in spontaneous speech. For frequent compounds like 'cupboard, 'daredevil, headlight', 'homework', 'horsefly' 'houseboat', 'jawbone', 'landlady', 'landlord', 'railway', 'rainbow' etc. the convention is the same for English and Dutch. For this moment we do not have a plausible explanation for the quantative differences between Shattuck-Hufnagel (1987) and our results for Dutch.

Shattuck-Hufnagel (1987) also found that word initial interactional speech errors are much more frequent for exchanges plus interrupted errors than for anticipations and perseverations. For Dutch we found no such difference. For American English this effect also should be assessed anew, using our newly proposed predictor. Of course, component segments of word initial clusters may be particularly prone to be exchanged with each other, as in *broek blijft* ('trousers remain') > *bloek brijft*. If so, we have in our current investigation missed this effect because we have excluded clusters from consideration. If the effect would be found to be real, then it may be attributed to greater contextual similarity in these clusters.

The main finding in this study is that the relative frequencies of interactional segmental speech errors in various positions in word forms can be predicted rather precisely from the numbers of phonotactically allowed opportunities segments in each position have for being substituted by other segments in the immediate environment. If this finding holds under further examination, then this poses certain constraints on theories of serial ordering of sound segments during speech preparation. Notably, the introduction of separate nodes and separate activation for word onset consonants, as proposed by Shattuck-Hufnagel (1992) would be superfluous. Also the suggestion by Dell (1986) that word onset consonants are more highly and more quickly activated than consonants in other positions would be superfluous. The property of Dell's model (1986, 1988) that segments in all positions are treated in the same way would be supported. We propose that Dell's 1986 model would generate the word onset effect if the model would restrict interactional speech errors to those pairs of segments that share the same position in the word and would be set up in such a way that it would take as input sequences of segments as they occur in spontaneous connected speech with all its normal statistical properties. This reflects our conviction that the proportion of errors in each position in the word that the model would generate is directly tied to the number of opportunities there are for errors in each position to arise, provided that the model does not allow interaction between segments in different positions in the word. Note that our results have nothing to say about the model by Dell et al. (1993). In our view that model only simulates the lexical retrieval or activation of sound segments for each word form, not the postlexical serial ordering of the activated segments. The same is true for the conclusion by Schwartz et al. (2004) with respect to problems in retrieval of sound segments caused by aphasia. As most of such errors are non-interactional, they probably have little to do with the process of serial ordering of activated sound segments.

The fact that frequencies of interactional substitutions in different positions in word forms can be rather precisely predicted from the relative numbers of opportunities there are for such interactional substitutions, suggests that virtually all interactional substitutions are generated during a single process in which sound segments are serially ordered. This confirms that it makes sense to distinguish between two processes, a lexical process of activating segments of word forms and a postlexical process of serially ordering those segments. It has been shown that during word form retrieval sound segments within syllables are activated from early to late (Meyer, 1991). One may wonder why, if segments of canonical word forms are activated sequentially, these same segments have to be ordered serially again after their activation. Levelt et al. (1999) have suggested that such serial (re)ordering is necessary because of a process of resyllabification and prosodification: The word and syllable final /t/ of *escort* will be pronounced as a syllable initial /t/ in *escorting* /ɛs-kɔʁ-tɪŋ/. The phrase *escort us* will turn into a prosodified phonological word /ɛs-kɔʁ-tœs/. Possibly this argument can be extended to other cases where the phonetic and / or phonological properties of segments are changed under the influence of the immediate environment. Particularly in rapid fluent Dutch one may find many cases similar to *ik haal hem op*/ɪk+ha:l+hɛm+ɔp/('I fetch him') being spoken as ['ka:-lə-mɔp], where the final /k/, the final /l/ and the final /m/ are turned into initial consonants, and where the segments /I/, second /h/ and /ɛ/ are completely deleted. We propose that such processing for articulation requires a serially ordered string of pre-activated segments in a buffer corresponding to some form of inner speech. It is at this stage that we may envision that exchanges, anticipations and perseverations of segments can occur.

Our finding that word initial, intervocalic and word final consonants are virtually only substituted by members of their own class, suggests on the one hand that in their coding these consonants at that stage already have properties related to how they would be produced with respect to the surrounding vowels, and on the other hand that these properties have not yet been changed by resyllabification. If they were, we would find many cases in which word final consonants interacted with word initial consonants. The special position of intervocalic consonants, not interacting with word initial or word final consonants, also suggests that at the level of processing where interactional speech errors are generated there are no clear syllables. Said differently, in line with suggestions by Shattuck-Hufnagel (1987, 1992, 2011), at the level of processing where interactional speech errors are generated, no syllabification has been applied obligatorily of the kind suggested by some phonologists (see Booij, 1995; Kahn, 1976) and by the psycholinguists Levelt et al. (1999).

Shattuck-Hufnagel (1992) found experimentally that the word onset effect in errors elicited in tongue twisters like *from the leap of the note to the nap of the lute* is greater than the word onset effect in errors elicited in word sequences *like leap note nap lute*. This is as we would expect from our proposal that interactional error frequencies are a function of the numbers of opportunities for interaction. If we exclude interactions involving segments belonging to articles, as we have done for Dutch, each of the two l's of *leap* and *lute* still has six opportunities for interaction, viz. with the initial segments of 'from, of, note, to, nap, of' (interaction involving the seemingly empty position in words starting with a vowel is in fact quite frequent in speech errors). Similarly, each of the two n's of *note* and *nap* has six opportunities for interaction. All in all, for the initial sounds of the four words *leap, note, nap, lute* there are 24 opportunities for interaction. In the word sequence leap, note, nap, lute we count eight opportunities for interaction. This fully accounts for the proposed difference. The same is true of an observation by Dell, Reed, Adams, and Meyer (2000) who found very little tendency for a word onset effect in nonsense tongue twisters such as *hef seng nem keg*. In such sequences all segments have equal numbers of opportunities for interaction. Therefore, both observations are in agreement with the present proposal that the apparent word onset effect is a function of the number of opportunities segments in different positions have to interact with segments in other positions.

The current investigation suffers from some limitations. First of all, Cutler (1982) has warned that speech error data collected by listening to spontaneous speech and writing down the speech errors heard, may suffer from listening biases and thus are not necessarily very reliable. We are not aware of any systematic investigation of the effect of such biases on the statistical properties of collections of speech errors. Our corpus of speech errors was collected by quite a number of young researchers asked to note down every speech error they heard in their daily life, together with its repair if any and together with its context. There is no guarantee whatsoever against serious biases in detecting and writing down speech errors in our corpus. However, if the distribution of interactional substitutions over segmental positions in word forms would have been seriously distorted by a listening bias, then our main result, viz. that the distribution of interactional substitutions matches the distribution of opportunities for interactional errors, would be the result of chance. This, although not impossible, is unlikely. Secondly, we have, for good reasons, in our investigation excluded a great many cases, thus reducing a corpus of 860 segmental speech errors to only 415. Perhaps in doing this we have obscured interesting differences in frequencies of speech errors between segmental positions in word forms. In particular, our approach and the limited size of our corpus made it impossible to explore the relative relevance or irrelevance to relative frequencies of segmental speech errors of word-onset position, root position, and pre-stress position (cf. Shattuck-Hufnagel, 1992). We intend to explore this later in an experiment eliciting speech errors, where such factors can be manipulated systematically.

Conclusions

The main result from this study of segmental substitutions in spontaneous speech is that the relative frequencies of these errors in different positions in the word can be predicted effectively and rather precisely from the numbers of opportunities which segments in different positions have to interact with other segments in the intermediate environment. There is no 'word onset' effect other than that may be attributed to the phonotactic structure of the language. Our findings also suggest that segmental interactions are not controlled by syllable position but are controlled by initial, medial and final position in both words and morphemes. If word onsets are special in attracting relatively many speech errors, this is because they provide more phonotactically allowed opportunities for segmental interactional substitution errors than other positions. Our results can help in constraining future models for the serial ordering of segments during speech preparation.

Authors note

The authors are grateful to Gary Dell and Rob Hartsuiker; the text of this paper has profited considerably from their thoughtful comments. Several files in which readers may verify our counts of speech errors, counts of numbers of segments in different positions, and counts of numbers of opportunities for interactions in different positions are currently available online in the form of excel documents at [http://www.hum.uu.nl/medewerkers/s.g.nooteboom/Experimentaldata.htm].

References

Berg, T. (1991). Phonological processing in a syllable-timed language with pre-final stress: Evidence from Spanish speech error data. *Language and Cognitive Processes, 6*, 265–310.

Booij, G. (1995). *The phonology of Dutch*. Oxford: Clarendon Press.

Brown, R., & McNeil, D. (1966). *Journal of Verbal Learning and Verbal Behavior, 5*(4), 325–337.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language, 30*, 542–579.

Carterette, E. C., & Jones, M. H. (1974). *Informal speech: Alphabetic and Phonemic texts with statistical analyses and tables*. Berkeley, CA: University of California Press.

Cutler, A. (1982). The reliability of speech errors. *Linguistics, 19*, 561–582.

Del Viso, S., Igoa, J. M., & García-Albea, J. E. (1991). On the autonomy of phonological encoding: Evidence from slips of the tongue in Spanish. *Journal of Psycholinguistic Research, 20*, 161–185.

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.

Dell, G. S., Juliano, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science, 17*, 149–195.

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.

Fromkin, V. A. (1971). The nonanomalous nature of anonmalous utterances. *Language, 47*, 27–52.

García-Albea, J. E., Del Viso, S., & Igoa, J. (1989). Movement errors and levels of processing in sentence production. *Journal of Psycholinguistic Research, 18*(1), 145–161.

Garrett, F. (1975). The analysis of sentence production. In G. Bower (Ed.), *The psychology of learning and motivation* (pp. 133–177). New York: Academic Press.

Hartsuiker, R. (2002). The addition bias in Dutch and Spanish: The role of structural context. *Language and Cognitive Processes, 17*, 61–96.

Kahn, D. (1976). *Syllable-based generalizations In English Phonology*. Unpublished Doctoral dissertation. MIT, Cambridge.

Levelt, W. J. M. (1989). *Speaking. From intention to articulation*. Cambridge, Massachusetts: The MIT Press.

Levelt, W. J. M. (2013). *A history of psycholinguistics. The pre-Chomskyan era*. Oxford: Oxford University 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.

Luyckx, K., Kloots, H., Coussé, E. & Gillis, S. (2007). Klankfrequenties in het Nederlands. In: D. Sandra et al. (Eds.) *Tussen taal, spelling en onderwijs*. Essays bij het emeritaat van Frans Daems (pp. 141–154). Gent: Academia Press.

Meringer, R. & Mayer, C. (1895). Versprechen und Verlesen. Eine Psychologisch-Linguistische Studie. New edition by A. Cutler & D. Fay (1978). *Amsterdam Studies in the theory and history of linguistic science* (Vol. 2). Amsterdam: John Benjamins b.v.

Meyer, A. S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language, 30*, 69–89.

Nooteboom, S. G. (2005). Lexical bias revisited: Detecting, rejecting and repairing speech errors in inner speech. *Speech Communication, 47*, 43–58.

Nooteboom (2011). Self-monitoring for speech errors in novel phrases and phrasal lexical items. *Yearbook of Phraseology, 2*, 1–16.

Nooteboom, S. G. (1973). The tongue slips into patterns. In V. A. Fromkin (Ed.), *Speech errors as linguistic evidence* (pp. 144–156). The Hague: Mouton.

Nooteboom, S. G., & Quené, H. (2008). Self-monitoring and feedback: A new attempt to find the main cause of lexical bias in phonological speech errors. *Journal of Memory and Language, 58*, 837–861.

Nooteboom, S. G., & Quené, H. (2013). Heft lemisphere: Exchanges predominate in segmental speech errors. *Journal of Memory and Language, 68*, 26–38.

Nozari, N., & Dell, G. S. (2012). Feature migration in time: Reflection of selective attention on speech errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 1084–1090.

Poulisse, N. (1999). *Slips of the tongue: Speech errors in first and second language production*. Amsterdam & Philadelphia: John Benjamins.

Roelofs, A. (2000). WEAVER++ and other computational models of lemma retrieval and word-form encoding. In L. Wheeldon (Ed.), *Aspects of language production* (pp. 71–114). Sussex, UK: Psychology Press.

Rumelhart, D.E. & McClelland, J.L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations*. Cambridge, MA: MIT Press.

Schelvis, M. (1985). *The collection, categorisation, storage and retrieval of spontaneous speech error material at the Institute of Phonetics* (Vol. 10, pp. 3–14). Utrecht: PRIPU.

Schwartz Wilshire Gagnon & Polansky (2004). Origins of nonword phonological errors in aphasic picture naming. *Cognitive Neuropsychology, 21*, 159–186.

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

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* (pp. 251–289).

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