

The addition bias in Dutch and Spanish phonological speech errors: The role of structural context

Robert J. Hartsuiker

*Nijmegen Institute for Cognition and Information, University of
Nijmegen, The Netherlands*

In speaking, both content (words, segments) and structure (syntactic trees, metrical structures) need to be produced. There is controversy about which aspects of metrical structure are generated during phonological encoding. In particular, do speakers retrieve a unit that represents CV structure? The present study shows that the incorporation of units for syllable CV structures in a connectionist model of phonological encoding enables us to explain empirical patterns of speech errors. The model accounts for the finding of a bias towards additions of segments. First, corpus analyses in Dutch and Spanish showed an addition bias in both languages. A second corpus analysis showed that in apparently noncontextual deletion errors the resulting CV structure is often identical to that of syllables in the immediate context. Furthermore, simulations are reported with two connectionist models, one that incorporates representations for CV structures and one that does not. Only the first model correctly simulated the empirically obtained addition bias, as well as the general pattern of substitution errors. Further simulations showed that the model predicts effects of the structural context. A final simulation tested a second explanatory mechanism for the addition bias: feedback from segments to syllables.

Requests for reprints should be addressed to Robert J. Hartsuiker, University of Edinburgh, Department of Psychology, 7 George Square, Edinburgh, EH8 9JZ, UK., Email: Rob.Hartsuiker@ed.ac.uk

This research is supported by grant 575-21-004 from the Dutch organization for Scientific Research (NWO). I am grateful to Sieb Nooteboom for providing me with the Dutch corpus, Gabriella Vigliocco for providing me with the Spanish corpus. Àngels Colomé for correcting the Spanish syllabifications, Thomas Berg for providing me with German data, but more importantly for stimulating discussion, to Nuria Sebastián for providing me with a Spanish lexical database, and to Ardi Roelofs, Gary Dell, Herman Kolk, Judy Kroll, and two anonymous reviewers for very helpful comments on previous versions of this article.

INTRODUCTION

When producing speech, we generate both content and structure. In grammatical encoding, a syntactic tree is constructed and the lexical items to be associated with that tree are retrieved. Evidence for the independent representation of syntactic structure comes from syntactic priming studies (Bock, 1986, 1989; Bock & Loebell, 1990; Branigan, Pickering, & Cleland, 2000; Hartsuiker & Kolk, 1998; Hartsuiker, Kolk, & Huiskamp, 1999; Hartsuiker & Westenberg, 2000; Pickering & Branigan, 1998) that show the syntactic form of a sentence can be primed by a preceding sentence, irrespective of lexical content. In phonological encoding too, there is a separation between content (the phonological segments of a word) and structure (the metrical structure of a word). A theory of phonological encoding, proposed by Levelt and Wheeldon (1994), explicitly assumes that the retrieval of segments and metrical structures are two independent processes. After retrieval, segments are associated with the structures.

The assumption of a separation between structure and content is relatively uncontroversial (see however, Dell, Juliano, & Govindjee, 1993, for an alternative model). A more controversial issue is which aspects of the word are represented at the structural level. In particular, an important question is whether there is an abstract representation for the sequence of consonants and vowels that make up the word (the word's CV structure) or not. The evidence, to be discussed shortly, is mixed. A number of studies reported evidence in favour of the hypothesis that CV structure is a relevant unit of speech production (Costa & Sebastián-Gallés, 1998; Meijer, 1996; Sevald, Dell, & Cole, 1995) whereas other studies did not find evidence for the retrieval of CV structure in speech production (Meijer, 1994; Roelofs & Meyer, 1998). Computational models of phonological encoding also differ in whether they contain explicit representations for CV structure (Dell, 1988) or not (Dell, 1986; Dell et al., 1993; Roelofs, 1997).

In this article, I present evidence that supports the CV structure hypothesis. The evidence comes both from the analysis of speech error corpora and from simulations with two connectionist models, one of which does and one of which does not contain explicit representations for CV structure. The article is organised as follows. I will first briefly review the empirical evidence with respect to the status of CV structures. Then, I will describe two existing connectionist models of speech production, that differ with respect to whether CV structure is explicitly represented or not. I will argue that a specific type of speech error, "wordshape errors" constitute an empirical benchmark for these models. The predictions are subsequently tested in analyses of speech error corpora, and I test whether the models are able to simulate the obtained data patterns.

What is the empirical status of CV structure? To begin with, Stemberger (1990) investigated an English corpus of phonological speech errors. A number of his findings support the hypothesis that CV structure plays a role in production. First, substitution errors involving initial consonants were significantly more frequent when the source and target words had the same CV structure. Second, errors leading to the formation of a consonant cluster were more frequent when the source word also contained such a cluster (i.e., there was similarity in the CV structure of the words).

Further evidence was provided by Sevald et al. (1995). These authors showed that it is easier to repeat nonwords, when a monosyllabic item and the first syllable of a consecutive disyllabic item shared CV structure. This effect occurred independently from whether, in addition to CV structure, segmental information also was shared.

Furthermore, there are phonological priming studies in which Dutch/English bilinguals translated English words into Dutch target words (Meijer, 1994, 1996; cf. LaHeij, De Bruyn, Elens, Hartsuiker, Helaha, & Van Schelven, 1990). These studies show that the production of a word with a given CV structure is faster when it is accompanied by a (Dutch) prime word that has the same CV structure than when the prime word has a different CV structure, even when there was no overlap in the actual segments. However, this pattern was not found in every experiment of this type (Meijer, 1994).

Costa and Sebastián-Gallés (1998) conducted four experiments in Spanish with two paradigms: picture-word interference (cf. Schriefers, Meyer, & Levelt, 1990) and a paradigm in which a list of words was read aloud and subsequently a picture was named. In both sets of experiments, small but consistent priming effects were found for shared CV structure. Furthermore, one of the experiments revealed that shared structure of the first syllable was sufficient to generate this effect.

On the other hand, Roelofs and Meyer (1998) conducted experiments with the implicit priming paradigm, developed by Meyer (1990). In this type of experiment, word pairs are first associated during a learning phase. Then, in the test phase, only the first word in the pair is presented and the participant has to produce the second word (the "target" word). Crucially, the word pairs are divided into sets that are either homogeneous or heterogeneous. The target words in homogeneous sets have form overlap; targets in heterogeneous sets do not. The so-called "preparation effect" is the difference between the average reaction times in the two sets (each word occurs in both sets). The experiments revealed an effect of shared word onsets, when the words had the same number of syllables and the same stress pattern. However, this effect occurred independently of whether the words in each set had the same CV structure.

In sum, the evidence for the existence of CV structure representations in speech production is mixed. It is not always found, and studies with different paradigms yielded different results. On the basis of some of these results Levelt, Roelofs, and Meyer (1999) admit that “the CV structure of words is in some sense psychologically real” (p. 22). However, these authors note that it is not clear at present in what way this psychological reality comes about. Before starting the empirical part of this paper, let us review the status of CV structures in two classical models of phonological encoding (Dell, 1986, 1988).

Models of phonological encoding

An influential computational model of speech production was proposed by Dell (1986). This is a connectionist model, which was constructed with the explicit purpose of fitting speech error data, naturalistic as well as experimental. Subsequent papers provided additional data that supported the model (Dell, 1988) and extended the scope of the model to disorders of lexical access in fluent aphasia (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Martin, Dell, Saffran, & Schwartz, 1994).

The Dell (1986) model belongs to the class of “interactive activation models” (e.g., McClelland and Rumelhart, 1981).¹ The model consists of a number of layers of nodes, each of which represents a single linguistic unit (e.g., word, syllable, coda, segment, etc.). Each layer is directly connected with the layer directly above it and directly below it. All connections are bi-directional. A schematic drawing of the model is provided in Figure 1. How is a word produced in this model? If a word is set to be the “current word” (i.e., this is done by the computer program controlling the simulation), that word is given a “jolt” of activation. At the same time, the next word is given a smaller amount of “anticipatory” activation. Activation spreads from the word layer to other layers below it for a specified number of time steps. Spreading in the network leads to the activation of syllable nodes, which in turn activate relevant onsets, codas, and vowels. In addition to activation spreading from the “current” word unit, there is also “noise” in the model, both random noise and activation that emanates from previous words and from upcoming words. After a number of time steps, the segments are selected. The selected segments are then set to an activation level of 0 for a single time step, in order to prevent excessive perseveration in upcoming syllables.

¹ That is to say, the model has localist representations, separate layers of nodes representing linguistic levels, and bidirectional connections. The model does not belong to a more restricted class of “IA” models in the sense of Jacobs and Grainger (1994), who propose a nested style of modelling in which the pattern of connectivity and parameter settings are set to be as similar as possible to the original model.

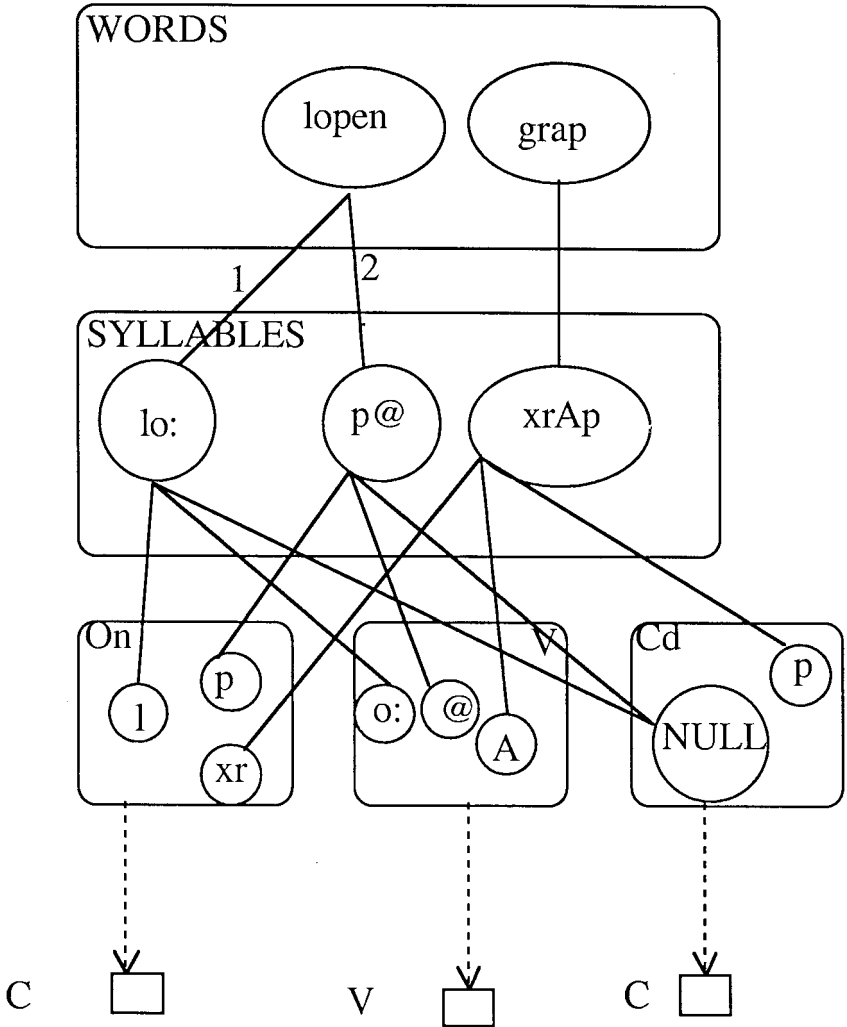


Figure 1. Simplified overview of the Dell (1986) model of phonological encoding. For simplicity's sake, nodes for phonological features and for rhymes (consisting of a nucleus and a coda) are left out.

In the Dell (1986) model, selection of the segments in each syllable is performed with respect to a strict syllable template containing three slots: initial consonant, vowel, and final consonant. There are three sets of segments nodes, each corresponding to one of these slots. In each set, the segment node with the highest activation level is selected and placed in the appropriate slot. In principle, such a syllable template only allows for the

production of syllables with the structure “CVC”. However, in order for the model to be able to produce different kinds of CV structures, there are so-called “null-nodes” and “cluster-nodes”. To illustrate the model, let us suppose that the word to be uttered is the Dutch word grap /xrAp/ (English translation: joke). If all goes well, the highest activated segment nodes at the time of selection will be the cluster node /xr/ for the onset slot, an /A/ for the vowel slot, and a /p/ for the final-consonant slot. A speech error would occur when the wrong node is activated in any of the three segment pools. For instance, if the segment representing /o:/ is selected instead of /A/, the substitution error /xro:p/ will result. The model also predicts that sometimes consonants are added or deleted. Suppose the null node were accidentally selected instead of /p/ as the final consonant. Then there would be a deletion, /xrA/. If the target word were lopen /lo: p@/ (to walk) and if /p/ were selected as the final consonant of the second syllable, then there would be an addition /lo: p@p/.

A problem with the Dell (1986) model was the inability to account for the pattern of “wordshape errors” in corpora of speech errors. Wordshape errors are phonological speech errors, in which a segment is either added or deleted. Examples of errors observed in speech error corpora of English (Stemberger, 1990), Spanish (Del Viso, Igoa, & García-Albea, 1987), and Dutch (Schelvis, 1985) are provided for additions in (1) and for deletions in (2):

- (1a) The same as the hit FRATE—hit rate for low frequency items
- (1b) Pero un globo GROSA (Intended: rosa)
(But a pink globe)
- (1c) LEEUWEBLEKKEN (Intended: leeuwebekken)
(snapdragons)
- (2a) So I assume it’s probably TILL cheaper (Intended: still)
- (2b) ¿quieres unas patatas FRÍAS—unas patatas fritas?
(Do you want some french fries?)
- (2c) ACHTIG procent (Intended: tachtig procent).
(Eighty percent)

These should be distinguished from so-called substitution errors in which a segment is replaced by a different segment, leaving the wordshape intact. Substitution errors can be divided into anticipations (3), perseverations (4), and exchanges (5) (Examples from Shattuck-Hufnagel & Klatt, 1980.)

- (3) Twenty-pive percent (Intended: twenty-five percent)
- (4) You can tell Ten (Intended: you can tell Ken)
- (5) waple malnut (Intended: maple walnut)

In corpora of spontaneous speech errors in Dutch (Nooteboom, 1969) and American English (Shattuck-Hufnagel, 1983) anticipations outnumbered

perseverations, and exchanges occurred much more infrequently than either anticipations or perseverations. The Dell (1986) model accounted quite well for this pattern of substitution errors. However, the model did not correctly predict the pattern of wordshape errors. Nooteboom (1969) reported that in the corpus of Dutch errors, there were far more additions ($N = 173$) than deletions of single consonants ($N = 21$). Dell's (1986) model predicted the opposite pattern: In the simulation reported in that paper there was a higher percentage of deletions (19% of all errors) than additions (between 5% and 8%, depending on a model parameter).

Why does the model tend to delete instead of add segments? Dell (1986) explained this tendency as follows. The null nodes are very frequent "sounds" and thus there are many connections to them. The more connections a node has, the more input it will receive from other nodes and thus the more active it will become. This results in an asymmetry: Null nodes will more often replace a consonant (creating a deletion) than a consonant will replace a null node (creating an addition). Likewise, cluster nodes are relatively infrequent "segments." Therefore, the probability that a single segment will replace a cluster (creating a deletion) is larger than the probability of the opposite pattern (creating an addition). This results in a bias towards deletion.

In 1988, Dell provided a verbal description of a model that might solve the problem with wordshape errors. In this new model, selection of segments is no longer organised by way of a strict CVC-template and there are no longer any cluster nodes or null nodes. Instead, a new set of nodes is created that represent CV structures. Dell referred to these nodes as "wordshape headers". In the remainder of this paper I will refer to them as "syllabic structure nodes".² An illustration of the new model is presented in Figure 2.

In order to illustrate the new model, let us take the example of the word "lopen" again. Producing this word requires the successive activation of the syllables /lo:/ and /p@/. Both syllable nodes are connected to the syllabic structure node "CV". When it is time for selection, the syllabic structure node with the highest activation will be selected. This syllabic structure node determines the segment pools from which a segment will be selected. As in the previous model, the segment with the highest activation level in each pool is selected. How do errors come about in this model? A substitution error occurs when the wrong segment node is the most active one at selection time. Wordshape errors occur when the wrong syllabic structure node wins the competition. Suppose that it is time to select the second syllable /p@/ of /lo: p@/ and the node with structure "CVC" wins

² Dell (1988) only presented a description of a model for monosyllabic words, so that word shape and syllable shape were identical.

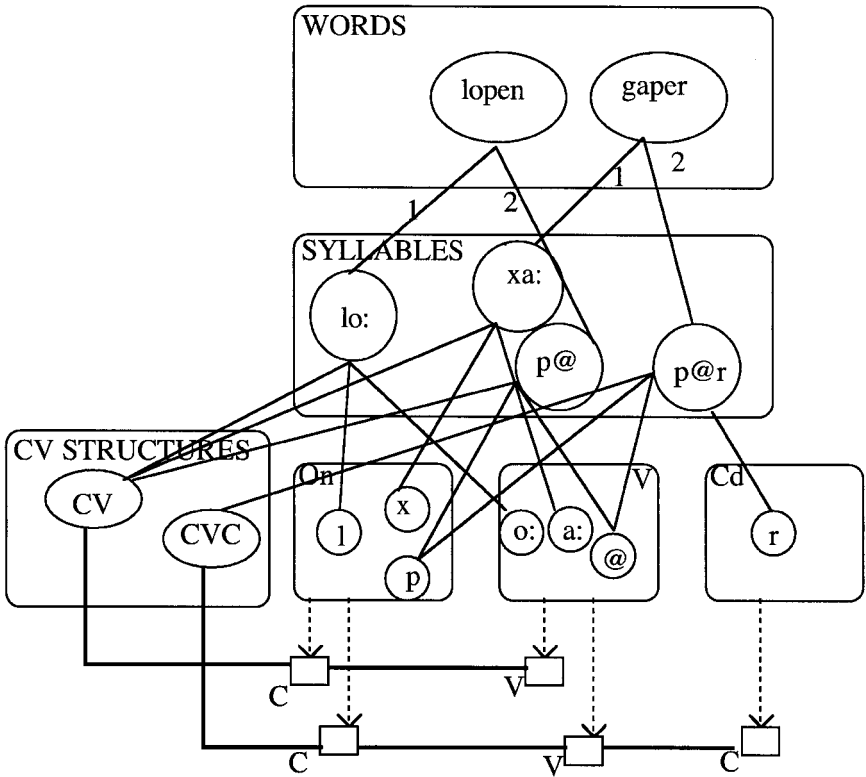


Figure 2. The Dell (1988) model in the version I implemented. As opposed to Dell (1988), I do not assume inhibitory connections within the structure layer. Furthermore, in this version of the model, selection of the segments runs in parallel for each position.

the competition. Then a final consonant, possibly /p/ will be added, resulting in the addition error /lo: p@p/. If the target word is /xA p@r/ and the syllable structure node CV is accidentally selected for the second syllable, there would be a deletion error, i.e., /xA p@/.

The present paper presents an elaboration of the Dell model, based on the verbal description in Dell (1988). I will test whether the introduction of nodes that represent a syllables' CV structure indeed solves the problem with wordshape errors.

The effect of context

We have seen that there is evidence for retrieval of CV structures in phonological encoding, and that an influential model of phonological encoding (Dell, 1986) that does not incorporate CV structures runs into

problems with wordshape errors. Before we turn to simulations that test whether the Dell (1988) model solves these problems, it is important to discuss a number of issues with respect to the empirical data, especially with respect to the meaning of "context".

There is a complication concerning the distribution of additions and deletions. According to Stemberger (1990), who restricted his analysis to wordshape errors with consonant clusters, there is an addition bias in English, especially towards adding a second consonant to create a consonant cluster (see also Stemberger & Treiman, 1986). Importantly, in Stemberger's (1990) English data, there was only an addition bias in contextual errors. For noncontextual errors, there was a bias towards deletion errors.

However, what are contextual errors? Usually, the distinction between contextual and noncontextual errors is based on the presence or absence of identical segments in the target word and a "source" word (a word occurring in close vicinity to the target, that is suspected to be the source of the error). This is an unproblematic distinction when we consider phonological substitution errors. However, what happens when a wordshape error is made? In the Dell (1988) model, a precondition for any wordshape error is the misselection of a syllable structure node. But why would a syllable structure node be misselected? I propose that a determinant of such misselections is the presence of a similar CV structure in the context. The reason for these context effects is that the corresponding structure nodes will be activated, either because of anticipatory activation, or because they have been selected earlier, and their activation has not yet decayed towards resting level. The implication is that many so-called non-contextual errors are in fact contextual: they come about because of other CV structures in the immediate context.

Given this redefinition of contextuality, it is no longer obvious that there is an addition bias for contextual errors and a deletion bias for non-contextual errors. A new analysis is warranted and will be presented shortly. This is especially so, because misselections of syllabic templates have different consequences in the case of additions as opposed to deletions. If the error is an addition, an extra segment will be added. The probability is quite high that the added segment appeared somewhere in the near context. Take example (6), which is actual model output:

- (6) hoog-hoek (Intended: oog-hoek)
(corner of one's eye)

In (6), the target syllable has the structure VC. The intruding syllable structure has the pattern CVC, which appears in the immediate context as the second syllable. The added onset segment happens to be /h/. This is not coincidental, since it is also the onset of the second syllable and has

therefore been primed. If this error were observed in spontaneous speech, it would be classified as an anticipatory addition error. Now consider example (7), which was also generated by the model:

(7) pa-li ri-te (Intended: pa-ling ri-te)³
 (eel) (rite)

In (7), the syllabic structure node CVC (the second syllable of the first word) is replaced by the syllabic node CV, which appears both in the preceding and in the subsequent context. However, since there is no segment present in the context that also appears in the error, this error would be classified as non-contextual in speech error corpora.

This then shows that the opposition contextual/non-contextual wordshape errors is fictitious: Both types of error may have a contextual source, but this is only revealed in the addition errors in which a segment is copied from the source. This is not to say that *all* non-contextual errors are in fact contextual: Wordshapes and segments may certainly be misselected without being triggered by the context (i.e., as a result of random noise). The claim here is that some errors that are classified as non-contextual are really contextual errors in disguise. They are triggered by the syllabic structure context.

Crosslinguistic differences

In the previous section, we proposed the hypothesis that the CV structure of neighbouring syllables influences the pattern of wordshape errors. An implication of that hypothesis is that the distribution of CV structures in the language is a determinant of that pattern. This is because syllable types that are frequent in a certain language, will occur often in the context. Thus, in a language with more CVC syllables than CV syllables (e.g., Dutch), one would expect relatively many addition errors, since a CV syllable will often be surrounded by CVC syllables. In a language with more CV syllables than CVC syllables (e.g., Spanish), one would expect relatively few addition errors.

In this paper, I compare the pattern of wordshape errors in Dutch and Spanish. There are some important differences in the syllabic structures of these languages. First, in Spanish, clusters can contain up to two consonants. Word-final syllables have a maximum of one consonant in the coda (Navarro, 1966). In Dutch, the onset can contain up to three

³ Notice that the resulting error violates the phonological constraint that a minimal rhyme in Dutch consists of two skeletal positions, either VV or VC. Since the present model does not distinguish between short and long vowels, this constraint is not incorporated in the model.

consonants, but if there are three consonants, the first one must be an /s/. The rhyme in Dutch contains between two and three segments. However, the Dutch syllabic template also allows for a so-called “appendix”, with maximally three extra consonants at the end of a word (Booij, 1995). An example of a word with an appendix is herfst (autumn) which has the structure CVCCCC.

Second, in both languages syllabification is governed by the so-called “maximal onset constraint” (Booij, 1995; Van der Hulst, 1984). That entails that as many consonants as possible are assigned to the onset position. In Dutch, there is also the “minimal rhyme constraint”. The latter constraint entails that a rhyme needs to contain at least two positions (a short vowel and a consonant, or a long vowel). The word aspect (id.), in which the first vowel /A/ is short, will be syllabified as as-pect in order to comply with the minimal rhyme constraint, even though /sp/ is a legal onset cluster in Dutch.

Third, as suggested above, the frequency of different syllable types differs between Spanish and Dutch. Navarro (1966) reported the frequency of different syllable types in a text corpus. Many more syllables had structure CV (58.5%) than CVC (27.4%). I also computed the frequencies of syllable types in 3582 mono- and disyllabic words, in a lexical database of Spanish (Sebastián-Gallés, Martí, Carreiras, & Cuetos, 2000), excluding syllables with consonant clusters, and found similar results: CV was much more frequent (53.9%) than CVC (38.9%). In an analysis of all mono- and disyllabic words in the Dutch CELEX corpus (Baayen, Piepenbrock, & Van Rijn, 1993), excluding syllables with consonant clusters, 62.4% of the syllables had structure CVC, and 30.6% had structure CV.

Feedback

A final issue concerns feedback in the model. Models of phonological encoding differ with respect to whether feedback from “lower” levels to “higher” levels is assumed or not. The WEAVER model, proposed by Roelofs (1997) does not have feedback connections. All versions of the Dell (1986, 1988) model however, do postulate that activation spreads both top-down and bottom-up.

One of the goals of this paper is to propose that feedback is a conceivable mechanism for explaining the tendency to add rather than delete segments. The reason is that longer words and syllables receive more feedback, because they are connected to more segments. They thus tend to replace shorter words and syllables. If this is true, the model should generate more addition errors when the amount of feedback is increased. This hypothesis is tested in Simulation 5.

Plan of this paper

In the remainder of this paper, I will first report a corpus analysis of Dutch and Spanish speech errors, in order to establish the distribution of addition and deletion errors in those languages, given the redefinition of context supplied above. A second corpus analysis concerns the effect of preceding or following context of syllable structures on the distribution of wordshape errors. Subsequently, a number of simulation studies will be reported. The first simulation, testing Dutch, replicated Dell's (1986) finding that the model with null nodes and cluster nodes predicts the wrong pattern of wordshape errors. More importantly, it showed that the 1988 version of the model adequately accounts for the data. The second simulation showed that the model also correctly predicts the distribution of substitution errors. The third simulation correctly predicted an addition bias in Spanish. The fourth simulation explored the effect of syllabic structure context. Finally, the fifth simulation tested the effects of another possible determinant of wordshape errors: feedback from segments, syllable structures, and syllables.

CORPUS ANALYSIS 1: THE ADDITION BIAS

There is one paper that reports the distribution of wordshape errors in Dutch (Nooiteboom, 1969) and one paper that reports such errors in Spanish (Del Viso, Igoa, & García-Albea, 1991). Unfortunately, the data listed in those reports are not complete for our purposes. In the report on Dutch errors, only *contextual* wordshape errors were listed (either anticipations, perseverations, or exchanges). As would be predicted by Stemberger, there were many more addition than deletion errors for Dutch contextual errors. On the other hand, Del Viso et al. only reported the (so-called) *non-contextual* wordshape errors (1991, p. 167). These authors showed, again unsurprising in light of Stemberger's findings, that deletions were much more frequent than additions in Spanish non-contextual errors. In the present analysis, I re-analysed the Del Viso corpus and I also analysed a Dutch corpus, the Utrecht corpus of speech errors (Schelvis, 1985). In both corpora, all addition and deletion errors at the segmental level were counted, irrespective of (apparent) contextuality.

Method

In order to establish the overall rates of additions and deletions, I analysed two corpora of speech errors. The first analysis concerned the Utrecht corpus of Dutch speech errors (Schelvis, 1985), a corpus of about 2500 spontaneous speech errors which have been collected by staff members of the institute of phonetics in Utrecht, The Netherlands. The errors were

collected in the period from 1979–1985. Errors were written down, including self-corrections if any and details about the speaker and the context, if relevant.

The second analysis concerned the Del Viso et al. (1987) corpus. This corpus consists of about 3600 spontaneous speech errors in Spanish. These errors were also noted down by observers, who further paid attention to such aspects of the error as self-corrections and the linguistic context.

The analysis of the Dutch corpus excluded all errors in written language, all slips of the ear, and all errors in which part of the utterance was in a different language (usually English). Only phonological speech errors were considered. The addition and deletion errors were not scored as such in the original classification scheme. All phonological anticipations, perseverations, and exchanges were reclassified as either substitution errors or wordshape errors. Word or phrasal blends were excluded, as well as deletions/additions of vowels, syllables, or morphemes. Errors in a so-called “dustbin”-category, which contained “non-contextual” additions and deletions, were also reclassified.

The analysis of the Del Viso corpus considered all anticipations, perseverations, anticipation/perseverations (where the error could be either—or both), and “incomplete movement errors” involving consonants. Errors involving units larger than the consonant or cluster were excluded, because those errors fall outside the scope of the models tested here (see General Discussion). Furthermore, I analysed the categories of deletions and additions of consonants (non-contextual errors), which are explicitly incorporated in the corpus.

Results and discussion

Table 1 lists the number of additions and deletions in Dutch and Spanish. As the table shows, there is a bias towards addition errors, both in Dutch and in Spanish.⁴ If we compare the two corpora directly, the addition bias is stronger in Dutch (85% additions) than in Spanish (72%). Indeed, a

TABLE 1

Total numbers of addition and deletion errors in Dutch and Spanish, proportions in parentheses

	<i>Addition</i>	<i>Deletion</i>	<i>Total</i>
Dutch	236 (.85)	43 (.15)	279
Spanish	164 (.72)	65 (.28)	229

⁴ We also know of German error data (Berg, personal communication), in which there is an addition bias (for contextual errors).

chi-square test revealed that the distribution of additions and deletions in Dutch and Spanish is significantly different, ($\chi^2(1) = 12.64; p < .05$).

CORPUS ANALYSIS 2: CONTEXT OF SYLLABLE STRUCTURES

If it is true that a representation of the syllable CV structure becomes active during speech production, one expects that the presence or absence of competing CV structures is a determinant of wordshape errors. Because CV structures for upcoming syllables receive anticipatory priming, and because recently produced CV structures are also primed, such structures might occasionally replace the intended structure. I suggest then, that many wordshape errors are contextual in nature, but that not only segmental context plays a role, but also the context of neighbouring CV structures. Since this variable has not been taken into account into the classification of speech errors, many speech errors may be incorrectly labelled “non-contextual”.

Importantly, such misclassification is particularly likely for (non-contextual) deletion errors. As mentioned in the introduction, if an addition error has a structural source, the added segment will probably also appear in the context, thus resulting in the error being classified as contextual. This does not apply to deletion errors, in which no contextually active segment appears in the error.

An example of a deletion in which the intruding CV structure is in the near context is given in (8) taken from the Del Viso et al. (1987) corpus:

- (8) Yo iba a echar una cata (Intended: carta)
I went to write a “cata” (Intended: letter; “cata” is a nonword)

In (8) the phrase u-na car-ta, with pattern V-CV CVC-CV is replaced with u-na ca-ta, with pattern V-CV CV-CV. This may be both an anticipation and a perseveration of CV structure.⁵

A counterexample is (9), from the same corpus.

- (9) estando en la segunda plata (Intended: planta)
being on the second “plata” (Intended: floor; “plata” means silver)

In (9), the syllabic structure CCVC (plan) is replaced with CCV (pla), even though CCV does not appear in the immediate context.

In sum, the (structure) context hypothesis predicts that in most of the deletion errors *classified* as non-contextual, but not in the addition errors

⁵ Example (8) would be classified as contextual in Stemberger’s (1990) coding scheme, since the missing final consonant /r/ was presented earlier.

so classified, there will be syllables with the same structure as the intruding syllable in close vicinity to the target syllable.

Method

I took from the Utrecht corpus and from the Del Viso corpus all non-contextual deletion and addition errors at the single consonant level. There were 20 deletions and 8 additions in the Dutch corpus. In the Spanish corpus, there were 44 deletions and 13 additions. The corpus listed each phrase that contained the error. I reconstructed each phrase to the intended phrase, and divided it into syllables. The syllabification was done by the author, a native speaker of Dutch with training in Spanish. The Spanish syllabification was checked and corrected by a native speaker of Spanish.

The resulting syllable strings were then analysed for the occurrence of syllables with a similar structure as the intruding one. I will call such syllables “context syllables”. The analysis was done with three different window sizes: One syllable before or after the target syllable (10); At most 2 syllables before or after the target syllable (11); At most 5 syllables before or after the error syllable (12). In the examples of items falling into each window size (10–12), the error syllable is underlined, and the context syllable(s) are set in boldface. All examples are taken from the Spanish corpus.

- (10) (= 1) yo quie ro un **ca** le **da** rio (Intended: ca len da rio)
(I want a calendar)
- (11) (≤ 2) gru **pos** de pác ti **cas** de tres (Intended: prác ti cas)
(exercise groups of three)
- (12) (≤ 5) po é mi co po **e** ma (Intended: po lé mi co po e ma)
(controversial poem)

Results and Discussion

In Table 2, I have listed the percentage of cases in which there was a context syllable for each window size (1, 2 or less, 5 or less).

Consistent with the context hypothesis, there were many cases in which the immediately following or preceding syllable shared CV structure with the intruding syllable. This effect becomes more pronounced when we consider a larger window of syllables. As predicted, this effect is restricted to the deletion errors

In order to compare the findings with a chance baseline, we determined the distribution of syllable types in the complete, corrected phrases in Spanish (in Dutch, there were too few observations to warrant such an analysis). If the results are due to chance rather than to a context effect,

TABLE 2

Percentages of the errors in which context syllables with the same CV structure as the error occur within 1, 2, or 5 syllables from the target syllable for non-contextual additions and deletions in the Dutch and Spanish corpora. Final column reports the total number of errors.

<i>Error type</i>	<i>Distance ≤ 1</i>	<i>Distance ≤ 2</i>	<i>Distance ≤ 5</i>	<i>n</i>
Additions Dutch	12.5	25	25	8
Deletions Dutch	50	65	75	20
Additions Spanish	8	15	23	13
Deletions Spanish	61	80	86	44

one would expect the distribution of error outcomes to be similar to the distribution of all syllables in the set. There were 710 syllables in the full set of phrases. We limited the number of categories to “CV”, “CVC”, and “Others”, because there were too few observations for each of the other types. Of these syllables, 52% were CV, 23% were CVC, and 26% had another structure. The distribution of error outcomes was quite different: 49% CV, 9% CVC, and 42% had another structure. A chi-square test showed that these distributions were significantly different $\chi^2(2) = 10.05$; $p < .01$.

SIMULATION 1: THE ADDITION BIAS IN DUTCH

The first simulation tested whether the Dell (1988) model, as opposed to the Dell (1986) model, correctly generates the addition bias in Dutch. Both models were constructed and tested with the same model lexicon and parameter set. Because it is quite conceivable that the composition of the model lexicon with respect to CV structures influences the outcomes, I constructed the model lexicon so that it reflected the distribution of these patterns in the Dutch language as closely as possible. Indeed, pilot simulations confirmed that this distribution influences the results. These simulations suggested that the more often a given CV structure occurs in the lexicon, the more often it will substitute for a different one. Thus, the addition bias becomes stronger if, for instance, there are relatively more syllables with structure CVC than with structure CV.

Two different versions of each model were constructed, one with and one without feedback from the syllable structure to the syllable level. The reason two versions were tested, was that I wanted a model comparison that is as fair as possible. In contrast to the original Dell (1986) model, the current version of that model has a representation for the syllable structure (CVC), but this is the only syllable structure node. It does not have a functional role in distinguishing between different syllable structures, since all syllables have the same structure. This node does, however, send non-

specific feedback to all syllable nodes, increasing the activation level across the board. In order to prevent the possibility that any differences between the models can be ascribed to the fact that one model includes feedback from a syllable structure layer to the syllable layer, and the other model does not, we varied whether that particular kind of feedback was present or absent.

METHOD

Model lexicon. For the simulation in Dutch, 50 words were selected to serve as the model lexicon. This lexicon was constructed so that it reflected the distribution of CV structures of mono- and disyllabic words in the language. In order to determine that distribution, all Dutch word forms of one or two phonological syllables were taken from the CELEX lexical database (Baayen et al., 1993) with the constraint that the words did not contain clusters. Homophonic words were counted as often as listed. Morphologically decomposable word forms (e.g., *bereid* prepared) were included in the corpus, but verb-particle combinations which are orthographically two words (e.g., *geef op*, give up) were excluded. There remained a corpus of 20,984 word forms. We calculated the frequency of the phonological CV structures of these words, treating long vowels as one V.

From those words, 50 words were chosen at random, with the constraint that the distribution of CV structures was similar to the distribution in the corpus. This resulted in the exclusion of some CV structures, because they occurred very infrequently in the corpus (less than 2%, which corresponds to one word in the model lexicon). Care was taken that the words did not contain ambisyllabic consonants. The distributions in corpus and model lexicon are listed in the first two columns of Table 3.

The selected 50 words are listed in Appendix A.

Model structure and parameters. Two models were constructed, one corresponding to the Dell (1986) model and one to the Dell (1988) model. In each model, the word layer contained 50 word nodes, representing the 50 selected words, and there were also nodes for the appropriate syllables, CV structures, and segments (onsets, vowels, and codas). The appropriate connections were made between the nodes in the different layers. In contrast to Dell (1986), I did not implement a layer of phonological feature nodes and there was no separate layer for rhymes (consisting of a nucleus and a coda).

The model parameters were set as follows: p (spreading rate) 0.1, q (decay rate) 0.5, p_{up} (upward spreading rate) 0.05, Jolt of activation to current word 100, Anticipatory priming to upcoming word 50, R (number of time steps before selection) 5, fraction of spreading to current syllable in

TABLE 3

Distribution of CVC-patterns in the Dutch and Spanish corpora and in the model lexicons (percentages)

<i>CVC-pattern</i>	<i>Dutch corpus</i> (<i>N</i> = 20,984)	<i>Dutch model</i> (<i>N</i> = 50)	<i>Spanish corpus</i> (<i>N</i> = 3582)	<i>Spanish model</i> (<i>N</i> = 50)
Monosyllabic words				
CV	0.9	2	1.3	2
VC	0.6	0	0.4	0
V	0.1	0	0.3	0
CVC	11.6	12	3.3	4
Disyllabic words				
CV-CV	9.7	10	27.7	28
CVC-CV	19.4	20	18.3	18
CVC-CVC	30.1	30	13.1	14
CV-CVC	15.9	16	22.7	22
CVC-VC	0.8	0	0	0
CV-VC	0.2	0	1	0
VC-CV	1.2	0	2.5	2
VC-CVC	7.9	8	2.4	2
V-CV	0.5	2	0.1	0
V-CVC	1.2	0	3.2	4

two-syllable word 1.15. With these parameters the model was correct in more than 93% of the segmental positions (including empty ones) when the words were generated in isolation.⁶

The activation rule was linear, and identical to the activation rule used by Dell (1986, 1988), as follows:

$$(13) a_{i,t+1} = \text{Max}[0, ((1 - q)(a_{i,t} + In_{i,t}))]$$

In (13) $a_{i,t}$ represents the activation of node i at time t , q represents the decay factor, and $In_{i,t}$ represents the input i at time t . The function *Max* denotes that the maximum is taken from the computed activation level and 0. Thus, if a negative activation level is computed, it is set to 0.

The input is the weighted sum of the activations at time t of the nodes j connected to node i :

$$(14) In_{i,t} + \sum_1^j w_{ij} a_j$$

the weights w_{ij} , for a given connection from node j to node i equal p (if the connection is top down), and p_{up} if the connection is bottom up (a "feedback connection").

⁶ Thus there were < 20 errors per 279 positions (3 positions for each of 7 monosyllabic words and 6 positions for each of 43 disyllabic words). No attempt was made to optimise performance of the model.

The version of the Dell (1986) model without feedback from the syllabic structure layer to the syllable layer is as similar as possible to the original simulation reported in Dell (1986). The version with this feedback can be compared more readily with the 1988 model, because that model also includes feedback from syllabic structures to syllables. In both model versions there was always feedback from segments to syllables and from syllables to words.

Procedure. On each trial, the model produced a word pair. These word pairs were drawn at random from the model lexicon, and the corresponding word nodes were set to be the “current” and “upcoming” word. One thousand word pairs were produced in each of the four cells of the design, obtained by crossing Model (Dell, 1986 or Dell, 1988) and Feedback from syllable structure to syllables (present or absent). I found it desirable to test a large number of word pairs in order to reduce variability in the outcomes. In pilot simulations, a large variability occurred. This variability comes about because of an important random factor: each word pair was randomly drawn from the model lexicon. Because each word can be drawn together with each of the other words, including itself and because the order of the two words may be of importance, there are $50^2 = 2500$ possible word pairs. This implies that there will be considerable variation in the composition of any number of trials.

Once the two words in a pair were set to be current and upcoming, a boost of activation was given to the unit representing the first word and anticipatory priming was given to the unit for the next word. The model then started phonological encoding of each of the two words, using the processing principles and the parameter set reported above. The output (a selected segment for each position in each syllable) was written to a computer file. After a word pair was encoded, all units in the model were set to a resting activation level (0), and the next word pair was drawn from the lexicon. After generation of all the word pairs, the computer program controlling the simulation produced a report that listed the numbers of additions and deletion errors. Shifts, in which a segment was deleted from one syllable, and added to another syllable, counted as one addition and one deletion.

Results and discussion

The number of additions and deletions for each model and value of the feedback parameter is reported in Table 4.

The simulation shows that the Dell (1986) model has a strong tendency towards deletion and that this deletion bias is independent of the existence of feedback from syllable structure nodes to syllable nodes. The Dell

TABLE 4

Comparison between the Dutch versions of the Dell (1986) model and the Dell (1988) model with feedback from the syllable structure layer to the syllable layer ($w_{\text{syll, shape}} = 0.05$) or without such feedback ($w_{\text{syll, shape}} = 0.0$)

	<i>Utrecht corpus</i>	<i>Dell (1986)</i>	<i>Dell (1986)</i>	<i>Dell (1988)</i>	<i>Dell (1988)</i>
$w_{\text{syll, shape}}$		0	0.05	0	0.05
Additions	236 (85%)	47 (9%)	25 (4%)	461 (89%)	530 (93%)
Deletions	43 (15%)	455 (91%)	575 (96%)	56 (11%)	41 (7%)

(1988) model on the other hand predicts an addition bias. The proportion of additions predicted by the model is close to the empirical proportion obtained from the Utrecht corpus. Furthermore, the addition bias is stronger if there is feedback from syllable structure nodes to the syllable nodes.

SIMULATION 2: SUBSTITUTION ERRORS IN DUTCH

The previous simulation revealed that the Dell (1988) model, as opposed to the Dell (1986) version, is able to mimic the empirical pattern of wordshape errors. However, is the 1988 model also able to capture the pattern of substitution errors? To test the model, I ran additional simulations with the Dutch versions of the 1986 and the 1988 model. These data are compared with the results of two corpus analyses: the percentages of anticipations, perseverations, and exchanges of phonemes as reported by Nootboom (1969), and a count of these error types in the Utrecht corpus, which I conducted myself.

Method

The same model lexicon and parameter set was used as in the previous simulation. There was feedback from the syllable structure layer to the syllable layer in both versions of the model ($w_{\text{syll, shape}} = 0.05$). I tested the models with 200 word pairs, randomly drawn from the model's lexicon, and counted the number of anticipations, perseverations, exchanges, and substitutions involving syllables. There were also errors that could be either an anticipation or a perseveration. Half of these were added to the count of anticipations, and the other half to the count of perseverations. This was the same procedure as followed by Dell (1986).

Results and discussion

All additions, deletions, and substitutions involving entire syllables were excluded. I counted the first 100 single-segment substitution errors for each

TABLE 5

Comparison between empirically obtained distributions of anticipations, perseverations and exchanges, and simulated distributions using the Dutch versions of the Dell (1988) and Dell (1986) models

	<i>Corpus data</i>		<i>Simulation results</i>	
	<i>Utrecht corpus</i>	<i>Nooteboom (1969)</i>	<i>Dell (1988)</i>	<i>Dell (1986)</i>
Anticipation	55%	76%	51%	51%
Perseverations	26%	17%	43%	46%
Exchanges	18%	7%	6%	9%

version of the model. Table 5 lists the percentages of anticipations, perseverations, and exchanges generated by each model. Table 5 also reports empirically obtained percentages, obtained from the Nooteboom (1969) corpus and the Utrecht corpus.

Table 5 shows that there is substantial variation between the error patterns obtained in each error corpus. In particular, the number of exchanges in the Utrecht corpus is much larger than in the earlier Nooteboom (1969) corpus. Nevertheless, there is a general trend in the data. In particular, anticipations outnumber perseverations, and perseverations outnumber exchanges. How well do the models capture this pattern? As can be seen in Table 5, the predicted error patterns of the Dell (1986) and the Dell (1988) model are very similar. Each model predicts more anticipations than perseverations, and less exchanges than perseverations.

However, both models predicted a larger proportion of perseverations than is empirically obtained. It should be noted though that these simulations were done without attempting to obtain the best possible fit. In fact, in another simulation with the Dell (1988) model, but with a slight adaptation to the parameter that controls spreading to the current syllable (1.10 instead of 1.15) the distribution approximates the empirical data much closer (73% anticipations, 25% perseverations, and 2% exchanges).

In sum, the 1988 model and the 1986 model both capture important aspects of the distribution of substitution errors. We can therefore conclude that the 1988 model captures both the general patterns of substitution errors and the pattern of wordshape errors, using a single parameter set.

SIMULATION 3: THE ADDITION BIAS IN SPANISH

The results of the first simulations were quite straightforward. Not only did these simulations confirm that a model incorporating null-nodes produces far too many deletion errors, they also showed that a model containing syllabic structure nodes correctly predicts an addition bias. Furthermore,

the modified model still accounts adequately for the general pattern of substitution errors, which leave wordshape intact.

The present simulation tests a language that differs from Dutch with respect to the distribution of CV patterns: Spanish. Lexical statistics, reported in Table 3, show that the pattern CV is more frequent in Spanish than it is in Dutch and the pattern CVC is less frequent. Since these lexical statistics are mirrored in the composition of the model lexicon, one expects a different outcome in a Spanish simulation than in the simulation with Dutch. There are two reasons for this difference. First, since the pattern CVC is less frequent in Spanish than in Dutch, there will be fewer trials in which a CV syllable is presented in close vicinity to one or more CVC syllables, thus reducing the probability that a CVC syllable replaces a CV syllable. Second, reduced frequency of CVC syllables implies that such syllables receive less input than their Dutch counterparts. Thus, having a CVC syllable in close vicinity to a CV syllable, should result in an addition error less often. Therefore, the model should show a smaller addition bias in Spanish than in Dutch.

Method

Model construction and procedure. I selected 50 Spanish words from a Spanish lexical database (Sebastián-Gallés et al., 2000). From that database, which contains more than 4000 words, all words with clusters and all words consisting of more than two syllables were removed. There remained 3582 words. Similar to Simulation 1, the frequency of each wordshape in the corpus was determined, and 50 words were selected at random with the constraint that the overall distribution followed that of the overall corpus. Again, this led to the exclusion of very infrequent CV structures from the model lexicon. The distributions of wordshapes in the corpus and the model lexicon are listed in the final two columns of Table 3 and the selected words are listed in Appendix B.

A network was constructed which contained the relevant nodes and connections for the 50 words in the Spanish model lexicon and the associated syllables, syllabic structures, and segments. The model parameters were held at the same values as in the Dutch simulations. One thousand randomly drawn word pairs were selected as the target words, which the model had to produce.

Results and discussion

The model produced 241 additions (147 in the coda and 94 in the onset) and 163 deletions (all in the coda). Notice that there were many additions to the onset, even though only 8% of the words in the model lexicon allow an addition at that position (i.e., they contain a syllable with the structure

V or VC). This shows that infrequent syllable structures are relatively vulnerable. There are two reasons for this. First, an infrequent syllable structure such as VC will often be surrounded by other syllables that have consonants in the onset. The structure context hypothesis predicts that the structure of these syllables in the context will sometimes replace the correct syllable structure, leading to an addition in the onset. Second, since there are few syllables with structure VC in the syllable layer, feedback from the CV structure layer to the syllable layer will activate only few syllables with that particular structure. Therefore, the amount of subsequent feedforward to the corresponding CV structure node is small, resulting in a relatively low activation level for that node. These issues will be considered further in Simulations 4 and 5.

The addition bias was .60. This bias is consistent with the empirical distribution of wordshape errors in Spanish, although the model slightly underestimates the empirically observed addition bias (.72). Should we be concerned about this underestimation? I do not believe so. Construction of the model required some simplifications (e.g., the token frequency of syllables and syllabic structure is not taken into account, the model is limited to structures without clusters, and some structures with vowel onsets were excluded from the model lexicon)⁷. As a result of these simplifications, a more precise fit should not be expected.

What is important is that the model captures an interesting aspect of the data: the addition bias in Spanish was lower than in Dutch. The data from the two error corpora are certainly consistent with the hypothesis that there is a stronger addition bias in Dutch than in Spanish (the bias is 85% in Dutch and only 72% in Spanish). Recall that all aspects of the simulations in the two languages were similar, except for the structure of the model lexicon, which is dictated by the structure of the language. The simulation is thus able to capture a cross-linguistic difference in the distribution of wordshape errors, based solely on the languages' different distribution of syllable structure.

This suggestive result should however be accompanied with a word of warning. It may be unsound to directly compare the two error corpora. There may be possibly relevant differences in the way the errors were collected and scored and in the perceptual biases that the observers had in the different languages.

⁷ One of the reviewers suggested that incorporation of more structures with vowel onsets would have tended to increase the addition bias. However, note that given the constraint that the distribution of structures with vowel onsets in the model (presently 8%) follows that of the corpus (9.1%), any such increase is likely to be very small.

SIMULATION 4: CONTEXT EFFECTS

The previous simulations showed that the Dell (1988) model, which incorporates a representation for CV structure, predicted an addition bias, both in Dutch and in Spanish. My explanation for that bias is twofold: (1) the frequency with which CV structures occur determines both the activation level of CV structure nodes and the structural context in which syllables will occur; (2) feedback will tend to favour syllables with more segments. Evidence that feedback has this effect in the model will be presented in Simulation 5.

In the current simulation, I attempted to more straightforwardly confirm the role of structural context. If it is true that a main determinant of the addition bias is the frequency of syllabic structure contexts, one should observe that the distribution of additions and deletions varies with context. Furthermore, given the different structure of Spanish and Dutch, one expects that context has a different influence in those languages. In particular, in Dutch, the presence of a CVC syllable in the context should more often lead to an addition error than in Spanish.

Method

The following four disyllabic wordshapes occurred most frequently in both the Dutch and the Spanish lexicon: CV-CV, CV-CVC, CVC-CV, and CVC-CVC. Two simulations were conducted, one with the Dutch lexicon and one with the Spanish lexicon. In both simulations, I selected four words of each type, and systematically presented word pairs, consisting of each possible combination (including word pairs in which the same word is repeated). Each word pair was presented five times. The same parameter settings were used as in the previous simulations. However, to increase variability in the outcomes, this time random noise was added to the activation level of each node. The noise was drawn from a Gaussian distribution with mean 0 and a *SD* of .05 times the activation level of each node.

After the model produced each word pair, wordshape errors were scored. These were then divided into additions, deletions, and shifts (incidents in which a segment was deleted from one syllable and added to a different syllable). Finally, the proportion of these incidents was computed, relative to the total number of opportunities for each incident.

Results and discussion

The results for Dutch and Spanish are reported in Tables 6 and 7, below. Each cell in the table is based on 80 observations.

TABLE 6

Proportions of additions (A), deletions (D), and shifts (S), relative to the total number of opportunities in each condition in Dutch

<i>Word 1</i>	<i>Word 2</i>			
	<i>CV-CV</i>	<i>CV-CVC</i>	<i>CVC-CV</i>	<i>CVC-CVC</i>
CV-CV	—	—	A: 33%	A: 38%
CV-CVC	D: 99%	—	A: 50%	A: 100%
CVC-CV	S: 100%	A: 100%	A: 100%	A: 100%
CVC-CVC	—	—	A: 100%	—

As is shown in Tables 6 and 7, the model predicts that wordshape errors are triggered by the shape of syllables in the context. If all four syllables have the same CV structure, there are no wordshape errors. Wordshape errors are very likely when a particular CV structure is outnumbered three to one by the other CV structure. In six out of eight situations where that is the case, many wordshape errors occurred. If there are two CV syllables, and two CVC syllables, wordshape errors occurred less frequently. The wordshape errors that did occur with a two-two distribution are all additions, indicating that there is a general tendency in the model to add rather than delete.

As predicted, the impact of context differed between Spanish and Dutch. In both languages, addition errors occur in the same situations, but in most of these situations, there are fewer additions in Spanish than in Dutch. In one situation (CV-CV CVC-CVC) there were no additions in Spanish, but a fair number of additions in Dutch. This follows from the fact that there are more syllables with structure CVC in Dutch: therefore, the syllabic structure node for CVC will reach a relatively higher activation level in Dutch, triggering additions.

Another aspect of the data in Tables 6 and 7 is also important. Both in Dutch and in Spanish, there were no wordshape errors in two situations

TABLE 7

Proportions of additions (A), deletions (D), and shifts (S), relative to the total number of opportunities in each condition in Spanish

<i>Word 1</i>	<i>Word 2</i>			
	<i>CV-CV</i>	<i>CV-CVC</i>	<i>CVC-CV</i>	<i>CVC-CVC</i>
CV-CV	—	—	A: 6%	A: 0%
CV-CVC	D: 100%	—	A: 4%	A: 100%
CVC-CV	D: 87.5% S: 12.5%	A: 0.5%	A: 7%	A: 100%
CVC-CVC	—	—	A: 11%	—

even though in both of these situations, a CV structure outnumbered the other one three to one: CV-CV CV-CVC and CVC-CVC CV-CVC. What prevented wordshape errors in those situations? Notice that the singleton CV structure occurred in Word 2 in both situations. The majority of wordshape errors occurred in Word 1. Apparently, the model has an anticipatory bias, both with respect to substitution errors and with respect to wordshape errors. Thus, even if a CVC syllable is outnumbered three to one, if it occurs in Word 2 it is relatively safe from wordshape errors.

In sum, this simulation shows that the model predicts effects of structural context and a different impact for context for different CV structure, depending on their frequency in the lexicon. As a result of these frequency differences the model predicts cross-linguistic differences in the strength of the context effects.

SIMULATION 5: THE ROLE OF FEEDBACK

The corpus analyses revealed an addition bias in Dutch and Spanish, and in Simulations 1 and 3 we managed to simulate that pattern. I proposed that an important determinant of this bias are the context and frequency of CV structures. Empirical evidence for context effects was obtained in Corpus analysis 2, and the model was able to simulate context effects.

There is an additional reason why the model produces more addition errors than deletion errors. This is a result of feedback in the model. Word nodes receive feedback from syllable nodes, and syllable nodes receive feedback, from both the syllabic structure layer and the segment layer. Feedback from segments to syllables will tend to favour larger syllables. The more segments a syllable contains, the more feedback it will receive and thus the higher its activation level and subsequent selection probability. If this explanation holds, the model should show a stronger addition bias if the amount of feedback is increased.

However, feedback from the syllable structure layer to the syllable layer has a different effect. The more connections there are to a certain syllable structure node, the more syllables with that structure will receive feedback. Since CVC is more frequent than CV in Dutch, such feedback will increase the addition bias. However, since CV is more frequent than CVC in Spanish, feedback from syllable structures to syllables will *decrease* the addition bias in that language. Thus, whereas one expects the same effect of segment-to-syllable feedback in Dutch and Spanish (increasing the addition bias), one expects a differential effect from syllable structure-to-syllable feedback in those languages.

Method

Two sets of simulations were run, one with the Dutch lexicon and one with the Spanish lexicon. The Dell (1988) model was used, and all parameters were set at the same values as in the previous simulations, with the exception of (1) the weights on feedback connections; and (2) whether there was feedback from the syllable structure to syllable layer or not. When this particular feedback was present, the amount of feedback was the same as for the segment-to-syllable feedback.

The model was tested with 1000 randomly drawn word pairs for each level of the design, obtained by crossing syllable structure-to-syllable feedback (present, absent), amount of feedback (0, 0.025, 0.05, 0.075, 0.1), and lexicon (Dutch, Spanish).

Results and discussion

The proportion of addition errors out of all wordshape errors (the addition bias) is graphed in Figure 3 for each cell of the design.

Figure 3 shows the predicted pattern. First, Figure 3 shows that the cross-linguistic difference in the strength of the addition bias holds for the whole range of the feedback parameter. For each parameter combination, the addition bias is stronger in Dutch than in Spanish. Second, Figure 3 shows that the addition bias depends on the amount of feedback. The larger the general feedback parameter is, the stronger the addition bias. Third, the presence of a specific type of feedback, from syllable structure units to syllable units, has different effects in Dutch and Spanish: In Dutch,

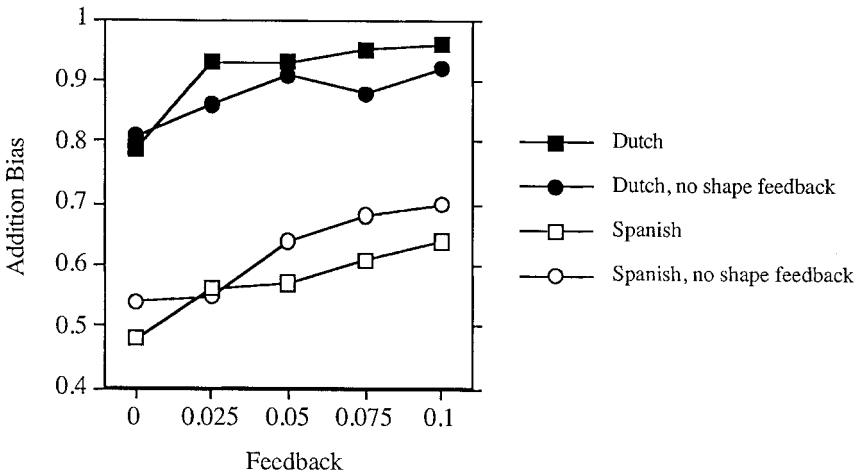


Figure 3. The addition bias in Dutch and Spanish as a function of the strength of feedback and the presence or absence of feedback from syllabic structure units to syllable units.

presence of this particular kind of feedback tends to increase the addition bias, but in Spanish it tends to decrease the addition bias.

The simulations thus show that feedback is an important determinant of the addition bias. In particular, feedback from segments to syllables tends to favour larger syllables, resulting in an increased addition bias. Feedback from syllable structures to syllables tends to favour *canonical* syllables, thus resulting in either an increased or a decreased addition bias, depending on which syllable structure is canonical in a particular language.

GENERAL DISCUSSION

Let me summarise the findings. A number of predictions were derived from the hypothesis that in speech production, a unit of processing is selected that represents a syllables' CV structure. First and foremost, a model of phonological encoding that incorporates such representations should be able to simulate the empirical distribution of speech errors in which CV structure is changed (wordshape errors). Second, one should expect effects of CV structures in the context. In particular, contextually present CV structures should often replace the target CV structure.

In two corpus analyses the distribution of wordshape errors was determined in Dutch and in Spanish. As opposed to earlier reports, I included both contextual and non-contextual additions and deletions. If the hypothesis is correct that CV structures are retrieved in phonological encoding, many apparently non-contextual errors may be considered contextual: The contextual source is the CV structure of neighbouring syllables. The first corpus analysis showed that there is a bias towards adding, rather than deleting segments, in both Dutch and Spanish. This bias was stronger in the Dutch corpus. The second corpus analyses confirmed the context hypothesis.

Subsequently, a number of simulation studies were conducted with localist interactive activation network models, based on the work by Dell (1986, 1988). The simulations showed that a model that explicitly incorporates a representation for CV structure (Dell, 1988) successfully generated the addition bias, both in Dutch and in Spanish. A model that does not incorporate such representations, but rather relies on "cluster nodes" and "null nodes" generated the opposite pattern, a strong deletion bias. Both models managed to simulate the pattern of substitution errors.

Further simulations confirmed the existence of syllable structure context effects and showed cross-linguistic differences in these context effects. The likelihood that a source syllable structure would intrude depended on the context (number of CV structures of each kind) and on the frequency of each structure in the model lexicon. The model made differential

predictions with respect to the size of context effects in Dutch and Spanish (see below).

In a final simulation, an additional explanatory mechanism for the addition bias was tested, feedback between segments and syllables. Such feedback tends to favour larger syllables, increasing the probability of additions. The simulations confirmed that feedback has this effect in the model, and furthermore that a specific kind of feedback, from CV structures to syllables, has different effects in the two languages.

An interesting and suggestive finding was a smaller addition bias in Spanish than in Dutch, both in the empirical data and in the simulations. This was found to be the case, even though the Dutch and Spanish models differed only with respect to lexical structure. Why did the model predict a larger addition bias in Dutch? There are two related reasons. Let us consider the two most frequent syllable types, CVC and CV. In Dutch, CVC is much more frequent than CV. Therefore, the situation that a syllable of type CV is surrounded by CVC syllables occurs more frequently, leading to addition errors. Second, because CVC is more frequent than CV, feedback from the CVC syllable structure node to the syllable layer will activate more syllables with that structure. Subsequent feedforward from these syllables will then result in the node for the structure CVC to obtain a relatively higher activation than its counterpart in Spanish. Thus, crosslinguistic differences in the CV structure of syllables lead to differences in the type of context a given syllable will tend to encounter and in the amount of input from the syllable layer.

In sum, the results suggest that there are multiple determinants for the addition bias: (1) the presence of other CV structures in the context; (2) the frequency of CV structures in the lexicon; (3) the presence of feedback, in particular from segments to syllables. The implication of these results is that the CV structure of syllables is a relevant unit of processing in speech production. I have reported empirical data that constitute a benchmark for this claim and I have reported simulations that show that models incorporating CV structure representations adequately simulate these data.

The present results corroborate earlier findings (e.g. Costa & Sebastián, 1998; Meijer, 1996; Sevald et al., 1995) but contrast with the findings of Meijer (1994) and Roelofs and Meyer (1998). How should we reconcile the discrepancy in the empirical results? Meijer (1994) found priming effects of CV structure in two experiments that used a word-translation paradigm, but he failed to find a priming effect in another experiment (his Experiment 6). In this final experiment he contrasted words with short and long vowels, which according to some theorists (e.g., Clements & Keyser, 1983) have different CV structures (CVC and CVVC). A null effect in this experiment can be taken to mean that there is no priming of

CV structure, but as Meijer argues, it can as easily be taken to mean that long vowels take only a single position in the CV structure. It should be noted that there is no consensus in the phonological literature about the representation of long vowels in Dutch (see Gilbers, Van der Linde, & Bastiaanse, 1997 for a discussion of this issue).

What about the results of Roelofs and Meyer (1998)? These authors used the implicit priming paradigm (Meyer, 1990) and contrasted homogeneous sets, in which all the words had the structure CCVC, with heterogeneous sets in which CV structure was CCVC, CCVVC, CCV, or CCVCC. Within these conditions, they varied whether onset clusters were shared or not. They found identical priming effects of shared onsets in a set, regardless of homogeneity of CV structure in the set and claimed that therefore CV structure is not part of the metrical representation. But notice that this is an indirect test: The claim depends on the assumption that metrical structures need to be exactly identical, in order for a priming effect to surface. Furthermore, the structures in the heterogeneous set were all very similar. In fact, half of the structures differed only with respect to the length of the vowel (CCVVC and CCVC). Given the lack of consensus in phonology with respect to the representation of long vowels and given Meijer's (1994) results, it is possible that half of the structures in the heterogeneous set were identical (in any case with respect to processing). In sum, implicit priming might not be the most ideal instrument to uncover effects of CV structure: the test is indirect and only a limited range of CV structures can be studied. In Roelofs and Meyer's study, that range was possibly further limited by including structures that differed only with respect to vowel length.

In the next section, I first discuss two alternative explanations for the addition bias. I will conclude by discussing a number of limitations of the present model.

Alternative explanations for the addition bias

In this section, I will discuss two alternative explanations for the addition bias: An explanation proposed by Stemberger (1991), and an explanation which can be derived from the WEAVER model (Roelofs, 1997). I will argue that both of these explanations correctly predict that additions are more frequent than deletions. However, neither explanation can account for the contextual effects we observed in our corpus analysis.

Stemberger (1991) observed an addition bias in English and provided an explanation for speech errors in terms of competition and inhibition. This explanation holds that speech errors arise because of competition between an intended segment and a competing segment. The intended segment will usually inhibit the competitor, but if the inhibition is too slight, the

competitor wins, and is erroneously selected. Thus, if both a target segment /r/ and a competitor segment /s/ are activated, /s/ may sometimes win the competition, resulting in a substitution error (e.g., bar → bas). The mechanism that yields wordshape errors is similar. If the target is an empty coda (i.e., no segment should be selected), but some segment happens to be active, say /s/, then /s/ may be inappropriately selected, yielding an addition error. If the target is /s/, but no segment is selected, a deletion error would occur.

If this explanation is correct, additions would be more likely than deletions, because in the case of an addition, there is a competitor element that receives no inhibition from a target representation (there is no *intended* segment for that position). Therefore, the competitor is likely to be selected, resulting in an addition. On the other hand, in a deletion error there is an intended element that is not selected. However, there is no competitor for the target. Therefore, the target is likely to be selected, resulting in a correct production, rather than in a deletion. Stemberger's explanation is similar to mine in its appeal to competition and activation (but notice that there is no explicit inhibition in the Dell, 1986, 1988 models). However, it is hard to see how Stemberger's explanation could account for the effects of syllabic structure context I presented in corpus analysis 2, since this explanation focuses on competition at the segment level rather than at a structural level.

Roelofs (1997; see also Levelt et al., 1999) presented a model of word form encoding, from which one can derive another explanation for the addition bias. This model (WEAVER) is a mixture of activation models (such as the ones reported in the present paper) and classical AI-models with production rules. Its main goal is to simulate chronometrical data. The model has not been worked out with respect to the simulation of speech errors, but some initial attempts have been made (Levelt et al., 1999). In WEAVER, phonological encoding comprises three levels: the morpheme and its metrical structure, the segment, and the syllable program. Segments activate syllable programs, and these in turn provide access to a "syllabary", a store of motor programs for each syllable. Units are selected through a dual mechanism: first, activation spreads feedforward through the network. If a node is activated above threshold, a production rule fires which checks whether the node is linked to the appropriate nodes one level up. For an error to occur in WEAVER, two conditions have to be met: (1) the wrong node is activated above threshold; and (2) the verification mechanism that is subsequently triggered fails. In a simulation with a modified version of WEAVER (Levelt et al., 1999), in which the verification procedure was programmed to fail a random number of times, the pattern of substitution errors was correctly captured.

Wordshape errors would occur if the wrong syllable program node were activated, and that syllable program node would consist of a different number of phonemes than the correct one. Thus, if the target syllable program were [t@r], the erroneous syllable programs [t@] and [@r] might also become somewhat active, since they receive input from some of the same segments. If one of these alternative syllable program nodes would be selected, a deletion error would occur. It is conceivable that in this system additions occur more frequently than deletions. This is because larger syllable program nodes receive input from more segment nodes. Thus their activation tends to be higher than that of smaller syllable program nodes. Notice that this mechanism is similar to the feedback mechanism proposed in the present study. The difference is that in *WEAVER*, it is feedforward from segments to syllabic representations, rather than feedback to syllabic representations that contributes to an addition bias.

Thus, it is conceivable that *WEAVER* would predict an addition bias, although it is unclear whether the strength of this bias would correspond to the empirical findings. However, a more difficult challenge for *WEAVER* would be to predict effects of structural context. Given the mechanism for wordshape errors in *WEAVER* (selection of a syllable program node that is highly active, as a result of segmental input) a requirement for a context effect would be the presence in the context of a syllable that shares segmental structure with the target syllable. The implication is that wordshape errors in *WEAVER* would always be (segmental) contextual errors. An important point that follows from the analysis presented here, is that this view is incorrect: Wordshape errors are often contextual in the sense of structural context.

Some limitations of our approach

There are some obvious limitations to the present approach. One limitation of the simulations is that the model was restricted to produce only syllables of four types: V, VC, CV, and CVC. More complex syllable types were ignored, although many wordshape errors involve clusters, both additions and deletions (Stemberger, 1990). The reason clusters were ignored is that it is unclear how the model should be modified in order to handle more complex syllable types. Because there are separate pools of segment nodes for each position, extending the model leads to (1) an increase in the number of pools of segments; (2) a representational problem: if the model were able to handle syllables with, for example, structure CCVC, we need at least two pools of onset consonants. The question is then how syllables with structure CVC should be handled: From which of the two onset consonant pools should the onset be

retrieved? There are two possible answers. First, the structure CVC could be connected to the first onset consonant pool and this pool would contain all possible onset consonants. Second, there could be a C1 pool, containing e.g., obstruents and a C2 pool containing, e.g., liquids, nasals, and glides. This avoids the problem of reduplicating representations for segments, but instead of a single structure CVC, we would now need separate structures (e.g., C1VC and C2VC). For now, I have avoided these issues by restricting the model to structures without clusters (see Hartley & Houghton, 1996 for a computational model that does allow production of more complex syllable types). The goal here was to reveal effects of CV structure, focusing on just the most frequent syllable types.

A general limitation of models in the tradition of Dell (1986, 1988) is that these models are designed for the simulation of speech error data and not for the simulation of production latencies. There is now a large body of research on phonological encoding in the so-called chronometrical tradition (see Levelt et al., 1999 for an overview). A challenge for further research is the construction of a model that can handle both types of data. Needless to say, the present goals were much more modest. Another limitation of the Dell (1986, 1988) model is that it produces words only in citation form. Yet as Levelt et al. (1999) correctly remark, in fast running speech there is resyllabification across word boundaries, e.g., /give it/ would become syllabified /gi vit/. The model is not able to account for such phenomena (as opposed to WEAVER, see Roelofs, 1997).

A final limitation of the model concerns two types of speech errors that the model cannot generate. The first error type cannot be produced, because the model produces a single syllable every fixed number of time steps. The implication is that the model does not allow for the addition or deletion of syllables. An example of a deletion error that alters the number of syllables is (15), from the Del Viso et al. (1987) corpus.

- (15) el periódico *Britano* (Intended: Británico)
The British newspaper

The model also does not allow for shift errors in which onset segments move to the coda or *vice versa*, e.g. (16), also taken from Del Viso et al. (1987).

- (16) *persumo* que sí (Intended: presumo)
I *assume* yes

In (16) the onset segment /r/ moves to the coda of the first syllable. The reason the model does not handle such situations is a result of the position-specific coding of segments in the model (i.e., segments can only substitute for other segments in the same pool).

In sum, there are obvious limitations to the present work, both in the scope of the model, in the restricted number of syllable types the model handles, and in the types of speech error the model can produce. We are still far removed from the ultimate model, which according to Levelt et al. (1999) should handle production latencies as well as the intricate patterns of speech errors.

Conclusions

The main contribution of this study is that it shows incorporation of CV structure representations in an activation-based model allows adequate simulation of empirically obtained speech error data. As a result of the incorporation of such representations in the model, effects of structural context as well as of segmental context are predicted. The data confirmed these effects. Both the model and the corpus data suggest cross-linguistic differences. This finding is consistent with our explanation for the addition bias as multiply determined by the CV structure context, the frequency of each CV structure, and the impact of feedback.

Manuscript received December 1999

Revised manuscript accepted November 2000

REFERENCES

- Baayen, R.H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database*. Linguistic Data Consortium, University of Pennsylvania, PA.
- Bock, J.K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355–387.
- Bock, J.K. (1989). Closed-class immanence in sentence production. *Cognition*, *31*, 163–186.
- Bock, J.K., & Loebell, H. (1990). Framing sentences. *Cognition*, *35*, 1–39.
- Booij, G. (1995). *The phonology of Dutch*. Oxford: Clarendon Press.
- Branigan, H.P., Pickering, M.J., & Cleland, A.A. (2000). Syntactic coordination in dialogue. *Cognition*, *75*, 1–13.
- Clements, G.N., & Keyser, S.J. (1983). *CV phonology: a generative theory of the syllable*. Cambridge, MA: MIT Press.
- Costa, A., & Sebastián-Gallés, N. (1998). Abstract phonological structure in language production: Evidence from Spanish. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 886–903.
- Del Viso, S., Igoa, J.M., & García-Albea, J.E. (1987). Corpus of spontaneous slips of the tongue in Spanish. Unpublished manuscript.
- Del Viso, S., Igoa, J.M., & García-Albea, J.E. (1991). 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., Schwartz, M., Martin, M., Saffran, E.M., & Gagnon, D.A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*, 801–838.
- Gilbers, D., Van der Linde, K., & Bastiaanse, R. (1997). Phonological length, phonetic duration and aphasia. *Clinical Linguistics and Phonetics*, *11*, 411–422.
- Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language*, *35*, 1–31.
- Hartsuiker, R.J., & Kolk, H.H.J. (1998). Syntactic persistence in Dutch. *Language and Speech*, *41*, 143–184.
- Hartsuiker, R.J., Kolk, H.H.J., & Huiskamp, P. (1999). Priming word order in sentence production. *Quarterly Journal of Experimental Psychology*, *52A*, 129–147.
- Hartsuiker, R.J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition*, *75*, 27–39.
- Jacobs, A.M., & Grainger, J. (1994). Models of visual word recognition—Sampling the state of the art. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1311–1334.
- La Heij, W., De Bruyn, E., Elens, E., Hartsuiker, R., Helaha, D., & Van Schelven, L. (1990). Orthographic facilitation and categorical interference in a word-translation variant of the Stroop task. *Canadian Journal of Psychology*, *44*, 76–83.
- Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioural and Brain Sciences*, *22*, 1–75.
- Levelt, W.J.M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, *50*, 239–269.
- Martin, N., Dell, G.S., Saffran, E.M., & Schwartz, M.F. (1994). Origins of paraphasias in deep dysphasia: Testing the consequences of a decay impairment to an interactive spreading activation model of lexical retrieval. *Brain and Language*, *47*, 609–660.
- McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, *88*, 375–407.
- Meijer, P.J.A. (1994). *Phonological encoding: The role of suprasegmental structures*. Unpublished dissertation, University of Nijmegen.
- Meijer, P.J.A. (1996). Suprasegmental structures in phonological encoding: The CV structure. *Journal of Memory and Language*, *35*, 840–853.
- Meyer, A.S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language*, *29*, 524–545.
- Navarro, T. (1966). *Estudios de fonología Española*. New York: Las Americas.
- Nooteboom, S.G. (1969). The tongue slips into patterns. In A.G. Sciarone, A.J. Van Essen, & A.A. Van Raad (Eds.), *Nomen. Leyden studies in linguistics and phonetics*. The Hague: Mouton.
- Pickering, M.J., & Branigan, H.P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*, 633–651.
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*, 249–284.
- Roelofs, A., & Meyers, A.S. (1998). Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *24*, 922–939.
- Schelvis, M. (1985). *The collection, categorizations, storage, and retrieval of spontaneous speech error material at the Institute of Phonetics, Utrecht*. Progress report, Institute of Phonetics, University of Utrecht.

- Schriefers, H., Meyers, A.S., & Levelt, W.J.M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, 29, 86–102.
- Sebastián-Gallés, N., Martí, A., Carreiras, M., & Cuetos, F. (2000). *LEXESP: Léxico informatizado del español*. Barcelona: Edicions de la Universitat de Barcelona.
- Sevald, C.A., Dell, G.S., & Cole, J.S. (1995). Syllable structure in speech production. Are syllables chunks or schemas? *Journal of Memory and Language*, 34, 807–820.
- Shattuck-Hufnagel, S. (1983). 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. Neuropsychology and Neurolinguistics*, pp. 17–51. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Shattuck-Hufnagel, S., & Klatt, D.H. (1980). How single segment error data rule out two models of error generation. In V.A. Fromkin (Ed.), *Errors in linguistic performance. Slips of the tongue, ear, pen, and hand*. pp. 35–46. New York: Academic Press.
- Stemberger, J.P. (1990). Wordshape errors in language production. *Cognition*, 35, 123–157.
- Stemberger, J.P. (1991). Apparent anti-frequency effects in language production: The addition bias and phonological underspecification. *Journal of Memory and Language*, 30, 161–185.
- Stemberger, J.P., & Treiman, R. (1986). The internal structure of word-initial consonant clusters. *Journal of Memory and Language*, 25, 163–180.
- Van der Hulst, H.G. (1984) *Syllable structure and stress in Dutch*. Dordrecht: Foris.

APPENDIX A: DUTCH LEXICON

jee	lok	sjoel	heel	wijs	puik
riep	puren*	gejij	wieren*	builen*	boegen*
boernoës	Malta's	duchtig	daagjes	fantom	koopman
heitjes	zuignap	tompouce	signet	tamboer	gander
wektoon	ronker	voordoën	paapje	leemde	nikte
garde	woorden*	kookten*	faktie	dorpen*	deisde
deerden*	fysisch	dading	bereid	mijner	mousseer
meereed	pijnig	gemuit	uitkeek	aanloodt	oester
oorzaak	aju				

Note: The items marked with a “*” have an orthographical final CVC syllable, but a phonological CV syllable, since word-final /n/ following a schwa is deleted in standard Dutch. The phonological, rather than orthographical, structure of the words was used in the simulations.

APPENDIX B: SPANISH LEXICON

daca	pupa	papa	coso	tuna	arder
pomo	nono	gajo	befa	dote	bebe
peca	zorro	sota	terrón	layar	lápiz
torrar	tifus	quídám	chocar	gorrón	rajár
saber	rabión	busto	jerga	mango	mismo
jaspe	silva	parca	yente	nervio	raspar
bordar	zumbar	rondon	tender	pensar	después
la	San	par	huso	erre	usar
erial	hernia				