ELSEVIER

Effects of syllable frequency in speech production

Joana Cholin[a,*], Willem J. M. Levelt[a], Niels O. Schiller[a,b]

[a]*Max Planck Institute for Psycholinguistics, P. O. Box 310, 6500 AH, Nijmegen, The Netherlands*
[b]*Maastricht University, Faculty of Psychology, Department of Cognitive Neuroscience, The Netherlands*

Abstract

In the speech production model proposed by [Levelt, W. J. M., Roelofs, A., Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, **22**, pp. 1–75.], syllables play a crucial role at the interface of phonological and phonetic encoding. At this interface, abstract phonological syllables are translated into phonetic syllables. It is assumed that this translation process is mediated by a so-called *Mental Syllabary*. Rather than constructing the motor programs for each syllable on-line, the mental syllabary is hypothesized to provide pre-compiled gestural scores for the articulators. In order to find evidence for such a repository, we investigated syllable-frequency effects: If the mental syllabary consists of retrievable representations corresponding to syllables, then the retrieval process should be sensitive to frequency differences. In a series of experiments using a symbol-position association learning task, we tested whether high-frequency syllables are retrieved and produced faster compared to low-frequency syllables. We found significant syllable frequency effects with monosyllabic pseudo-words and disyllabic pseudo-words in which the first syllable bore the frequency manipulation; no effect was found when the frequency manipulation was on the second syllable. The implications of these results for the theory of word form encoding at the interface of phonological and phonetic encoding; especially with respect to the access mechanisms to the mental syllabary in the speech production model by (Levelt et al.) are discussed.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Language production; Word-form encoding; On-line syllabification; Mental syllabary; Syllable frequency

* Corresponding author. Tel.: +31 24 3521 303; fax: +31 24 3521 213.
  E-mail address: joana.cholin@mpi.nl (J. Cholin).

The aim of the present paper is to provide evidence for the assumption that speakers access a separate mental store containing pre-compiled motor-programs of syllabic size. The notion of such a *Mental Syllabary* that is accessed during speech production is an inherent part of the theory of spoken word production by Levelt et al. (1999) and is also implemented in its computer simulation WEAVER++ (Roelofs, 1997a,b, 1998, 1999).

In fluent speech, the individual sounds of a word are bundled together to form optimally pronounceable units, namely syllables, which serve as the basis for motor execution. The tens of thousands of words in any given language are composed of a relatively small inventory of syllables. In fact, 500 syllables from English, Dutch, and German, i.e. less than 5% of the entire syllable inventory in those languages, suffice to produce approximately 80% of all speech in those languages (Schiller, Meyer, Baayen, & Levelt, 1996). To account for the easy and rapid production of these frequently occurring syllables, Levelt and Wheeldon (1994) proposed a repository of ready-made syllabic motor programs, called the *Mental Syllabary*.

The idea of a phonetic store of syllables was originally proposed by Crompton (1981) and adopted in Levelt (1989). To account for certain types of speech errors (e.g. phoneme and syllable substitutions and blends), Crompton suggested that speakers retrieve syllable-sized programs from a library of articulatory routines. However, despite the general agreement that syllables constitute an important linguistic and phonological/phonetic unit (see Blevins, 1995; Fujimura & Lovins, 1978; Hooper, 1972; Kenstowicz, 1994; Selkirk, 1982) and the support from speech error analyses (Fromkin, 1971; Shattuck-Hufnagel, 1992) and especially meta-linguistic tasks (for Dutch: Schiller, Meyer, & Levelt, 1997; for English: Treiman & Danis, 1988), there is to date relatively little psycholinguistic on-line evidence that the syllable serves as a functional unit in speech production. This also holds for the data reported in Levelt and Wheeldon (1994)-see below. Furthermore, what evidence there is for the functional importance of the syllable is orthogonal to the issue of whether or not there is mental storage for precompiled syllabic programs.

First, we introduce the process of word-form encoding in the framework of the model proposed by Levelt et al. (1999) and its computer simulation WEAVER++ (Roelofs, 1997a,b, 1998, 1999) with particular attention to the encoding steps at the interface of phonological and phonetic encoding and crucially how access to the mental syllabary is involved. Second, an overview of the experimental evidence for syllabic representation in speech production and stored syllabic units at different processing levels is presented.

1. Phonological and phonetic encoding

As soon as a speaker has selected a target word from the mental lexicon, phonological encoding can be initiated. It largely consists of computing the word's syllabification and prosody in the larger context of the utterance. This, in turn, is incrementally followed by phonetic encoding, which includes the computation of the articulatory gestures for the target word's syllables in their phonetic context. Finally, the execution of these gestural scores by the laryngeal and supralaryngeal muscle systems produces the spoken word. The present paper exclusively concerns the stages of phonological and phonetic encoding.

2. Phonological encoding

The first operation in phonological encoding is the retrieval of the target word's *phonological code* from the mental lexicon. The code consists of an ordered set of phonemic segments. For stress-timed languages such as English and Dutch, the model also assumes the existence of sparse metrical markers in phonological codes. In these languages, the default position for lexical stress is defined as the first full-vowel syllable of a word. Thus, for polysyllabic words that do not have the stress on the first stressable syllable, the metrical structure is stored as part of the phonological code; but, for monosyllabic words and for all other polysyllabic words, it is not stored but computed. Metrical structures describe abstract groupings of syllables (σ) into feet (Σ) and feet into phonological words (ω). A phonological word was defined as the minimal unit above the metrical foot, to which clitics, such as unstressed words, can attach (as in 'she's' where the reduced form of 'is' attached to 'she', see Levelt, 1989; Wheeldon & Lahiri, 1997). Crucially, at the stage of phonological encoding, phonological segments are not yet assigned to syllabic positions, nor is the C(onsonsant)-V(owel) (hereafter CV-) structure for the word specified. This is in contrast to other models of spoken word production (in particular Dell, 1986, 1988), which assume that the retrieved phonological codes are pre-syllabified. Internal syllabic positions such as onset, rhyme and coda are pre-specified in Dell's model but not in the Levelt model. The main argument for not pre-determining syllable positions in the phonological codes stored in the lexicon results from the phenomenon of *resyllabification*. In connected speech, syllable boundaries often differ from a word's or morpheme's canonical syllabification. The domain of syllabification is the phonological word, which can be smaller or larger than the lexical word due to morpho-phonological processes like inflection or cliticization (Booij, 1995). If, for instance, the stored phonological code for the word *defend* would be syllabified (i.e. as *de-fend*), then the speaker must 'resyllabify' the word when used in a different context, such as the past tense (*de-fen-ded*) or cliticization (*defend it* - de - fen - *dit*). The ubiquity of such 'resyllabifications' in the normal use of English (or Dutch for that matter) renders pre-specification of segments to syllable positions highly inefficient.[1] The alternative assumption, therefore, is that a word's syllabification is not retrieved but computed on-line depending on the context in which the word appears. During this process, called 'prosodification', retrieved segments are incrementally combined to form successive syllables. Also, these successive syllables are incrementally assigned the appropriate metrical properties, either following default stress, or otherwise the retrieved non-default stress marking feature. The incremental composition of syllables follows, on the one hand, universal syllabification constraints (such as maximization of onsets and sonority gradations) and, on the other hand, language-specific rules, e.g. phonotactics. Together, these rules create well-pronounceable syllables. The output of phonological encoding is a phonological word, specified for its metrical, syllabic, and segmental properties.

---

[1] The claim that phonological codes are not pre-syllabified is, in part, a language-specific claim. For a language like Mandarin Chinese, which has a small set of syllables and limited resyllabification processes, the story might be different (see, Chen, Chen, & Dell, 2002).

Dell's (1986, 1988) model makes different assumptions with respect to the phonological encoding process. As already mentioned above, his model includes abstract phonological representations that are specified for internal syllabic positions, i.e. the word form retrieved from the mental lexicon activates not only segmental information but also syllabic frames. These syllabic frames serve as placeholders into which the retrieved segments are inserted during the process of segment-to-frame-association. The discussion of whether or not the syllabic structure is part of word form encoding will be revisited under the section 'Empirical evidence for syllabic units in speech production'.

3. Phonetic encoding

The output of the phonological encoding process consisting of fairly abstract, syllabified phonological words is incrementally translated into articulatory-motor programs. These programs consist in large part of specifications for subsequent syllabic gestures. A crucial assumption of Levelt's theory is that speakers have access to a repository of syllabic gestures. This repository, coined the 'mental syllabary' (Levelt, 1992; Levelt & Wheeldon, 1994), contains the articulatory scores for at least the high-frequency syllables of the language. The model assumes that as soon as a syllable emerges during incremental syllabification, the corresponding syllabic articulatory gesture will be selected from the repository in a pre-motor area (Dronkers, 1996; Indefrey & Levelt, 2000; Kerzel & Bekkering, 2000).

The output of the mental syllabary in turn serves as input to further phonetic encoding, at which time contextually driven phonetic fine-tuning of retrieved motor programs occurs: The motor programs are still rather abstract representations of the articulatory gestures which have to be performed at different articulatory tiers, a glottal tier, a nasal tier and an oral tier. The gestural scores are abstract in the way that their performance is highly context-dependent (due to allophonic variation, coarticulation and—as a result of this-assimilation). The actual details of the movements in realizing the scores, such as lip protrusion and jaw lowering, is within the domain of the articulatory system (Goldstein & Fowler, 2003). According to Levelt (1989), the stored syllable can be pronounced with more or less force, with shorter or longer duration, and different kinds of pitch movements. These are free parameters, which have to be set from case to case. For new or very low-frequency syllables it is proposed that articulatory plans are assembled using the segmental and metrical information specified in the phonological syllables. This route[2] can also be used for the assembly of high-frequency syllables, for instance in cases when more conscious on-line control of speech production is required, for example, when speakers give a lecture or self-correct *slips-of-the-tongue* or speech errors in general. However, the retrieval of gestural scores from the syllabary is usually faster and less error-prone.

---

[2] The terminology of 'indirect' or 'direct' route to refer to the computational process that bundles phonetic segments to form syllabic scores was avoided as there is no consistent use of these terms in the literature.

In the final step, the articulatory network, a coordinative motor system that includes feedback mechanisms (Saltzman, 1986; Saltzman and Kelso, 1981; Goldstein & Fowler, 2003) transforms these articulatory plans into overt speech.

4. Predictions of WEAVER++

WEAVER (*W*ord-form *E*ncoding by *A*ctivation and *VER*ification) is the spreading activation based computer network model developed by Roelofs (1992, 1996, 1997a, 1997b, 1998, 1999, 2002), which is based on Levelt's (1989, 1992) theory of speech production. WEAVER++ adopts Dell's (1986) assumption of word form retrieval by the spread of activation and Levelt's (1992) on-line syllabification and access to a syllabary (Levelt & Wheeldon, 1994).

WEAVER++ (Roelofs, 1997a,b, 2002) assumes that the syllabification of a word is computed on-line during the speech production process. In the WEAVER++ model, segments in the retrieved phonological code are not specified for their syllable position, but only for their serial order within a word. The actual syllabic position of a segment is determined by the syllabification process. Each retrieved segment in the phonological code spreads activation to all syllabic gestures in which it partakes. Hence, upon retrieval of a phonological code, there are always multiple phonetic syllable programs in a state of activation. How is the appropriate syllable program selected? There are, first, selection conditions. The crucial one is that the syllable matches the phonological syllable that is incrementally composed; this involves a procedure of verification. Second, each syllable in the syllabary has a frequency-dependent selection threshold. This causes the predicted syllable frequency effect on naming latencies. Notice, however, that the threshold assumption is a 'modular' one. Removing it does not affect the architecture of the system as a whole. Third, syllable selection is subject to Luce's (1959) choice rule. During any smallest interval, the probability of selecting the (verified) target syllable equals the ratio of its activation to the summed activation of all syllable nodes. Given the choice ratio, the expected selection latency can be computed.

5. Empirical evidence for syllabic units in speech production

The question of whether or not the syllable constitutes a relevant unit in speech production has been the subject of many psycholinguistic studies. Analyses of speech errors suggested that phonological speech errors could be explained to a large extent within a syllabic framework. It was often argued that segmental speech errors such as exchanges of segments only occur for identical syllable internal positions, i.e. onsets exchange with onsets, nuclei exchange with nuclei, etc. (Berg, 1988; MacKay, 1970; Nooteboom, 1969; Shattuck-Hufnagel, 1979; Stemberger, 1982); this is referred to as the *syllable position constraint*. A quantitative analysis, however, showed that the majority of such errors take place in the onset position. Thus, the syllable onset constraint may be a word-onset constraint (Shattuck-Hufnagel, 1987, 1992; see Meyer, 1992, for a critical review). Recent analyses of speech errors that have taken this factor into account

(Vousden, Brown, & Harley, 2000), however, do not support the view that the syllable-position constraint on errors is really just a word-onset constraint.

Another source of evidence are meta-linguistic tasks, which suggest that syllables play a role at some level of processing in speech production (Schiller et al., 1997; Treiman, 1983; Treiman & Danis, 1988; see Bagemihl, 1995 for a review) but no strong claims are made about which level is involved. Despite the support for the syllable and the relevance of syllables to linguistic phenomena, the support from reaction-time studies over the past decade is rather scarce.

A series of cross-linguistic studies used a syllable priming task to identify the syllable as a relevant unit in speech production (for Dutch: Baumann, 1995; Schiller, 1997, 1998; for Mandarin Chinese: Chen, Lin, & Ferrand, 2003; for French: Brand, Rey, & Peereman, 2003; Evinck, 1997; Ferrand, Segui, & Grainger, 1996; Schiller, Costa, & Colomé, 2002; for English: Ferrand, Segui, & Humphreys, 1997; Schiller, 1999, 2000; Schiller and Costa, submitted; for Spanish and an overview see Schiller et al., 2002).

In all these studies, a syllabic prime was given which was either congruent with the target's syllabic structure (e.g. *ba.* as prime for *ba.lade* or the prime *bal.* for *bal.con*; a dot indicates syllable boundaries) or incongruent (e.g. *ba.* as a prime for *bal.con* or the prime *bal.* for *ba.lade*). The majority of these studies found a segmental overlap effect rather than a syllable priming effect, i.e. phonologically related primes, whatever their syllabic relation to the target word, facilitated the response relative to unrelated control primes, with longer CVC-primes being more effective than CV-primes. Moreover, no specific syllable priming effects were obtained. The original segmental overlap effect (Schiller, 1998) has recently been specified in more detail (Schiller, 2004). Much discussion has been given to the results of the apparent syllable priming effect in French (Ferrand et al., 1997, Experiment 5). However, Brand et al.'s (2003) failure to replicate the Ferrand effects suggests that this should not be taken as strong evidence for the syllable effect (see also Evinck, 1997; and for a review Schiller et al., 2002). In fact, the (only) study that showed a clear syllable priming effect was a study by Chen et al., (2003) conducted in Mandarin Chinese. Mandarin Chinese compared to the other languages under investigation consists of a very reduced number of syllables. Mandarin Chinese has (not counting tone) a syllable inventory of 400 different syllables whereas languages as Dutch and English have more than 12,000 different syllables. Additionally, syllables in Mandarin Chinese are not resyllabified in connected speech. Thus, in languages with far less syllables that are not resyllabified the storage of syllables might be different. It might be the case that the syllable structure is in fact stored within the word-form that is retrieved from the mental lexicon. This would explain why a significant syllable priming effect could be found in Mandarin Chinese but not in Indo-European languages (see also O'Sheagdha, Chen, Shen, & Schuster, 2004). The issue of cross-linguistic differences has to be further investigated. So far, we would like to conclude that the reason why there was no syllable priming effect in Dutch, English, French and Spanish is that at the stage where the (explicit) priming taps into the speech production process, there is only segmental but no syllabic structure available. If one considers that the primes in the syllable-priming task never get overtly articulated, one can imagine that these syllabic primes may not reach the late stage where syllables are computed on-line. Instead, the primes are assumed to speed up the segmental retrieval from the mental lexicon. The finding that the magnitude of the priming effect

increases with an increase in the number of shared segments, independent of a syllable match or mismatch with the target's first syllable, confirms the assumption that only shared segments can be primed. The finding that the magnitude of the priming effect increases with an increase in the number of shared segments, independent of a syllable match or mismatch with the target's first syllable, confirms the assumption that only shared segments can be primed. However, there are also studies that suggest that the word-form in the mental lexicon is in fact richer than it is assumed within the Levelt et al. (1999) framework. These studies provide evidence that the word-form contains information about the internal syllabic structure such as the word's CV-structure as well as syllable internal positions (Costa & Sebastián-Gallés, 1998; Dell, 1986, 1988; Sevald, Dell, & Cole, 1995). But, as discussed, in case of stored *phonological* syllables, the various syllable priming studies should have shown different results for matching and mismatching prime-target pairs. The fact that there was a segmental overlap effect leads us to believe that the word-form does not contain any syllabic information. To trace the emergence of the syllable at the interface of phonological and phonetic encoding, we might have to opt for a paradigm which is known to be sensitive also to later stages of word form encoding: The 'implicit priming paradigm'. Whereas the explicit priming is sensitive only to early stages of phonological encoding, the implicit priming paradigm exhibits effects that emerge at these early stages but also comprise later stages at the interface of phonological and phonetic encoding, i.e. on-line syllabification, possibly including syllabary access (Meyer, 1990, 1991; Roelofs, 1996, 1998; Roelofs & Meyer, 1998).

In a recent study, Cholin, Schiller, and Levelt (2004) used an *odd man out* variant (Janssen, Roelofs, & Levelt, 2002) of the implicit priming technique to demonstrate that speakers can benefit from a shared syllable structure. In this study, subjects learned sets of prompt-target pairs. Two types of response sets were compared, namely constant and variable response sets: Constant sets had overlapping initial segments *and* a constant CV-structure (as in **spui**.en, *to drain*; **spui**.de, *drained;* **spui**.er, *person who drains*; **spui**.end, *draining)*. Variable sets had an overlap of the first segments but did not have a constant syllable structure (e.g. **spoe**.len, *to rinse*; **spoel**.de, *rinsed*; **spoe**.ler, *person who rinses*; **spoe**.lend, *rinsing)*. Note that the second item of this set shares the same initial segments but has a different initial syllable structure; it is the *odd-man out* for the set. The prediction was that-under the assumption that the syllable is a relevant processing unit in speech production and is stored and accessed independently during word form encoding-speakers need knowledge about the current syllabic structure in order to prepare for a target utterance. In two studies investigating two different CV-structures (CVV, Exp. 1; CCVV, Exp. 2), Cholin et al. (2004) found a significantly larger preparation effect for constant sets as compared to variable sets. The authors argue that this preparation effect, which cannot be attributed to a segmental overlap effect, offers strong evidence for the relevance of syllables in word form encoding.

The virtue of the implicit priming paradigm is that it taps into later stages of word-form encoding than explicit priming paradigms do, but its shortcoming is that it is sensitive to all processing stages prior to articulation, including segmental retrieval, on-line syllabification, and possibly access to the mental syllabary. Thus, the above effect could arise at any one of these stages or, possibly, at a combination of these stages. These results testify to the relevance of syllabic units at the interface of phonological and phonetic

encoding, however, they do not represent indisputable evidence for the notion of separately stored syllabic units. It cannot be concluded with certainty that access to the mental syllabary is involved. Effects that would certainly provide evidence for the existence of the mental syllabary are effects of syllable frequency because only the retrieval of stored units produces frequency effects.

6. Evidence for the mental syllabary and syllable frequency studies

As already stated, the mental syllabary is conceived as a store for abstract motor routines of syllabic size. The economic advantage of such a repository becomes apparent if one considers the computational load that would evolve when the single sub-syllabic or phonemic motor-patterns had to be computed anew each time in speech production. The above mentioned analyses by Schiller et al. (1996) demonstrated that speakers re-use a rather small set of syllables over and over again, even for languages such as English, German or Dutch which have more than 12,000 different syllables. It would be efficient to store these over learned syllabic patterns and retrieve them as wholes, instead of computing them anew time and again as they appear during syllabification.

Levelt and Wheeldon (1994) studied this storage hypothesis by comparing access latencies of high- versus low-frequency syllables. They had subjects produce disyllabic words consisting of high- versus low-frequency syllables. Participants learned to associate symbols with written target words (e.g. ###### = *koning* 'king'). During the test phase, the learned symbols were repeatedly presented in random order and each time participants responded with the associated target word.

Levelt and Wheeldon's core finding was that, when word frequency was controlled for, words consisting of high-frequency syllables were named faster than words consisting of low-frequency syllables. If all syllables are computed on-line rather than retrieved from a repository, their frequency of use should be irrelevant. The obtained syllable frequency effects therefore seemed to support the notion of the mental syllabary. However, a potential problem with this conclusion is that syllable frequency was correlated with segment frequency in some of Levelt and Wheeldon's experiments (Hendriks & McQueen, 1996).

Evidence for the assumption that syllables are retrieved as whole entities come from recent studies in Spanish and French. Carreiras and Perea (2004) reported syllable frequency effects in Spanish speech production tasks (see also Perea & Carreiras, 1996). Brand, Rey, Peereman, and Spieler (2002) found also significant syllable frequency effects in a larger-scale study. Laganaro (2003) investigated effects of syllable frequency in French from a psycho- and neurolinguistic perspective. In her psycholinguistic studies, she found that syllable frequency affected naming latencies in a picture-naming experiment (see also Alario, Ferrand, Laganaro, New, Frauenfelder, & Segui, 2004) and in non-word reading. In her neurolinguistic study, she obtained a significant frequency effect on segmental substitution errors and a syllable-frequency effect on non-word repetition performance in aphasics. Furthermore, Aichert and Ziegler (2004) found a syllable frequency effect in patients with apraxia of speech (but see Wilshire & Nespoulous, 2003; Howard & Smith, 2002).

To summarize, a speech production model was outlined that incorporates syllables as functional units in word form encoding. We presented empirical evidence supporting this assumption and moreover, we could make a good case for the stage at which syllables come into play. It was argued that the absence of syllable priming effects on the one hand and the positive syllable preparation effects on the other hand suggest that (phonologically abstract) syllables are not stored units (as argued by Dell, 1986, 1988) but rather generated by an on-line syllabification process.[3] Evidence that these abstract phonological syllables trigger the selection of their phonetic representations from a mental syllabary is still rather scarce. As already mentioned, the results by Levelt and Wheeldon (1994) have to be handled with care as some of the syllable frequency effects were confounded with segment frequency. The syllable frequency effects in Spanish (Carreiras & Perea, 2004; Perea & Carreiras, 1996) and also the positive results from French (Brand et al., 2002) provide evidence for the notion of the mental syllabary but have to be seen before the background that syllables in Romance languages might have a different processing status than syllables in Germanic languages such as Dutch, German, and English[4]. The question of whether or not speakers can access pre-compiled syllabic units from a repository is still under discussion. The aim of the present paper is to gather additional evidence for the existence of such a mental syllabary. An additional aim of the paper is to investigate whether or not the claim that was raised in the Levelt and Wheeldon paper (1994) that only the second syllable in disyllabic (pseudo-)words is sensitive to frequency manipulations can be maintained.

7. Symbol-Position Association Learning Task

A Symbol-Position Association Learning Task was used to investigate effects of syllable frequency, a technique which was comparable to the one used in the Levelt and Wheeldon (1994) study. However, in the Levelt and Wheeldon (1994) study, participants had to associate sets of four strings consisting of six non-alphabetic characters (as '%%%%%' or '&&&&&') to four written disyllabic Dutch words (containing high- and/or low-frequency syllables). In the present experiments, participants had to associate an auditorily presented word to one of two positions on the screen. In learning phases, participants were presented with a symbol, namely an icon of a little loudspeaker, in one of two potential positions, left or right, and were simultaneously presented with the to-be-associated word via headphones.

---

[3] Note that the idea of a mental syllabary is also compatible with Dell's syllabified phonological code, the existence of stored units does not deny a store for pre-compiled syllabic programs.

[4] Romance languages (or syllable-timed languages; Abercrombie, 1967; for psycholinguistic evidence see Cutler, 1997; Cutler, Mehler, Norris, & Segui, 1986) have syllables as their primary rhythmic units. Germanic languages (or stress-timed languages) have a stress-based rhythm and are less prone to use the syllable as primary segmentation unit in speech perception (however, see Zwitserlood, Schriefers, Lahiri, & van Donselaar, 1993, but also Vroomen & de Gelder, 1994; for a critical, cross-linguistic review see Cutler, McQueen, Norris, & Somejuan, 2001).

In production phases, i.e. practice and test phases, the symbol was presented either on the left, the right, the top or the bottom of the screen to prompt the previously associated target corresponding to the particular position. One of the main advantages of this technique compared to the one applied by Levelt and Wheeldon (1994) lies in the presentation of the spoken target item instead of a printed word during learning phases. The auditory presentation of the target ensures that potential confounds deriving from orthographic factors (e.g. grapheme frequency) can be excluded.

In order to test whether or not this method is at all sensitive to frequency effects, a pre-test was conducted comparing the production latencies of high-frequency words compared to low-frequency words. The fact that high-frequency words are produced significantly faster than low-frequency ones is a stable finding in the literature (e.g. Jescheniak & Levelt, 1994; Oldfield & Wingfield, 1965). The pre-test showed a clear word-frequency effect: High-frequency word-forms were produced on average 37 ms faster than the low-frequency word-forms: (694 ms for the high-frequency word-forms versus 731 ms for the low-frequency word-forms). This result can be taken as a replication of the standard finding that high-frequency words are produced faster than low-frequency words, which reflects faster retrieval processes for high-frequency words from the mental lexicon. As the pre-test demonstrated the Symbol-Position Association Learning Paradigm to be sensitive to *word* frequency, it is plausible that it would also be sensitive to *syllable* frequency, even though the two effects are assumed to arise at two different levels of processing.

8. Experiment 1: Monosyllabic pseudo words

In this experiment, the production of high-frequency versus low-frequency syllables is investigated. The high-frequency monosyllabic pseudo-words have all the same CV-structure, namely CVC. The outcome that high-frequency items are produced faster than low-frequency items would support the notion of a mental syllabary.

9. Method

9.1. Participants

Sixteen native speakers of Dutch participated in the Experiment. They were randomly taken from the pool of participants of the Max Planck Institute in Nijmegen, The Netherlands and were paid for their participation. They had no known hearing deficit, and they had normal or corrected-to-normal vision.

9.2. Materials

All frequency counts were obtained from the computer database CELEX (CEntre for LEXical Information), which has a Dutch lexicon based on 42 million word tokens. Syllable frequency was counted for phonetic syllables in Dutch. The phonetic script differentiates the reduced vowel schwa from full vowel forms, giving approximately

12,000 individual syllable types. Syllable frequencies were calculated for the database from the word form occurrences per one million. Two syllable frequency counts were calculated: The number of occurrences of each syllable within words (independent of the frequency of occurrence of the syllable in a particular word position), i.e. the type frequency and the number of the summed frequency of occurrence of each syllable (within words), i.e. the token frequency. The syllable frequency ranges from 0 to approximately 90,000 per one million words, with a mean frequency of 121.[5]

For the following experiments, the experimental high- and low-frequency items should only be different in their syllable frequency. Therefore, it was crucial to construct an experimental item set that was controlled for length in phonemes, phoneme frequency, CV structure and bigram frequency. We decided to administer only *CVC*-syllables to the list of experimental items for the following reasons: (i) the CV-structure *CVC* is one of the most frequent syllable structure in Dutch[6] and (ii) CVC syllables provided for the best match on various features between the high- and low-frequency sets. Furthermore, the CVC-syllables could fulfill all of the criteria listed below.

We applied the following search criteria to the list of syllables: For a given high-frequency syllable a low-frequency counterpart should be found that was different from the given syllable only in one segment. For example, if the high-frequency syllable is *kem* [kɛm], the corresponding low-frequency syllable should be different by only the last segment, thus, a possible counterpart is *kes* [kɛs]. Thus, we have one pair of syllables which has the same onset and the same nucleus, namely the same short vowel. The only difference lies in the deviating coda, that is, the final consonant differentiates the two high- and low-frequency syllables. The next search device was to take those two syllables and to look again for counterparts but this time for each syllable in the opposite direction: For the high-frequency syllable *kem* [kɛm] this involved looking for a low-frequency syllable that had an identical offset but a different onset. The last two positions should be matching. The low-frequency syllable *wem* [ʋɛm] has an offset-overlap with the high-frequency syllable *kem* [kɛm] and was taken as third member in the syllabic quadruple. The final step was to look for a high-frequency syllable that had a) the same onset as the low-frequency syllable *wem* [ʋɛm] and (b) the same offset as the other low-frequency syllable *kes* [kɛs]. A successful quadruple was constructed when a high-frequency syllable was found which fulfilled this criterion, which in this case was the high-frequency syllable *wes* [ʋɛs]. See Table 1 for a schematic depiction of the described quadruple.

––––––––––

[5] The CELEX lexical database includes a list of Dutch syllables and their frequencies, based on syllabification of isolated word forms. In connected speech, as discussed above however, context-dependent phonological rules can modify the syllables and accordingly their token frequency. Schiller et al. (1996) carried out an empirical investigation in order to estimate the changes syllables may undergo in connected speech. A large Dutch newspaper corpus (TROUW) was transcribed, word-level rules were applied and then syllabified. In a first step the resulting lexeme syllables were compared to the corresponding entries in the CELEX database. In a second step, additional phonological sentence-level rules were applied to the TROUW corpus and then the frequencies of the resulting connected-speech syllables were compared to the lexeme syllables again. The overall correlation between lexeme and connected speech syllables was very high.

[6] In the inventories (CELEX and TROUW) the three most frequently occurring syllable types are, in order of frequency, CVV, CVC and CVVC, together accounting for more than 70% of all syllables. For more details and exact values of the two inventories, please see Schiller et al. (1996).

Table 1
Example of an experimental quadruple consisting of two high- and two low-frequency syllables, onsets and offsets are frequency controlled

Syllables within one quadruple	
High-frequency	Low-frequency
kem [kɛm]	kes [kɛs]
wes [ʋɛs]	wem [ʋɛm]

One remaining criterion had to be taken into account before such a quadruple was actually chosen to serve as an experimental quadruple. It was carefully controlled that none of the syllables within one quadruple were in fact existing words as it was crucial to investigate syllable frequency that was independent of word frequency. As already mentioned, two different syllable frequency counts were calculated, the number of occurrences and the summed frequency of occurrences; only instances that had in both scores comparable values were taken. Eventually, eight quadruples were taken as experimental item set. Thus, as every quadruple consisted of two high- and two low-frequency syllables each, the item set contained sixteen high-frequency and sixteen low-frequency items. The low-frequency items ranged in the count for the number of occurrence (per one million words) from a value of 0.02 to 1.19 with an average of 0.31 (SD = 0.36) and for the count of the summed frequency of occurrence (per one million words) from a value between zero and 3.67 with an average of 1.28 (SD = 1.28). For the high-frequency-items the values in both counts were as follows: For the count number of occurrence (per one million words), the high-frequency items ranged from 1.48 to 23.50 with an average of 5.32 (SD = 5.23). For the count of summed frequency (per one million words) of occurrence, high-frequency items ranged from 34.12 to 1,192.57 with an average of 242.32 (SD = 287.22). For an overview of all experimental quadruples and their values in both frequency counts see the Appendix A. We decided not to present an experimental quadruple as one experimental item set. Rather, all experimental sets were composed of syllables from different quadruples, this to avoid phonological overlap between the target items in one experimental set.

Since the immediate repetition of items could not be prevented in the present two-item design, we introduced fillers in the form of different numbers presented in the center of the computer screen that have to be named in between trials. This number naming should distract participants' attentiveness to any expectation of trial succession and beyond that it should help to 'neutralize' the articulators as the immediate repetition of two identical items could have huge facilitation effects. Four numbers were taken to serve as 'fillers': At the beginning of each set and between two subsequent targets a number which appeared in the middle of the screen had to be named. The following numbers were used: 1 (*een*), 3 (*drie*), 5 (*vijf*), and 8 (*acht*). These numbers were all monosyllabic and there was no phonological overlap between the numbers and the experimental test items.

Acoustic versions of the syllables were spoken by a female native speaker of Dutch. The spoken syllables were digitized at a sampling rate of 22 kHz, to be used during the learning phase of the experiment. They varied in duration from 382 ms to 586 ms with an

average of 484.69 ms (SD=48.23 ms). There was no difference in duration between high-frequency and low-frequency syllables (both ts<1).

9.3. Procedure and Apparatus

The experiment consisted of alternating learning, practice and test phases. Participants were tested individually in a quiet room. They were given a detailed written instruction specifying that they had to respond as accurately and as quickly as possible.

In the learning phases, the participant's task was to associate an auditorily presented target-word to one of two positions on the computer screen. An icon of a little white loudspeaker (4 by 4 cm) was presented on one of two possible positions on a black computer screen (iiyamaLM704UT) while the to-be-learned target word was simultaneously presented auditorily via headphones (SennheiserHD250). The two positions of the loudspeaker icons were on the left or the right side on the screen. Participants were instructed to listen carefully to the spoken pseudo-words and to memorize the position on which the current two target items were presented. Each target word was presented two times on its specific position.

In practice phases, the loudspeaker symbols were simultaneously presented on both sides, namely the left and right position on the screen. Both icons were presented as interactive fields. While displaying the two loudspeakers, one of the target items was presented via headphones and participants were instructed to associate this target item to its position by clicking with the left mouse button (cordless Wheel Mouse, Logitech) on the correct loudspeaker on one of the two positions. The practice phase contained eight trials in which the target items had to be correctly associated to their corresponding position; each target was auditorily presented four times. During practice phases, erroneous trials were counted and were displayed by a little white number on a black computer screen succeeding the practice phase. Only if participants could pass the practice phase with zero errors, the experimenter started the test phase. When they failed, the learning phase was started again and the session was rehearsed to ensure that they learn the sets accurately. Participants were explicitly asked to refrain from rehearsing the target items by articulating them.

In test phases, one position (left or right) was presented on the screen and had to be named as accurately and fast as possible. At the beginning of each test phase and between two subsequent target trials, i.e. responding to the left or the right position on the screen, one out of four numbers (1, 3, 5 or 8) was presented in the middle of the computer screen, white on black, which had to be named. Simultaneously with the presentation of a loudspeaker symbol on the left or the right side of the screen the voice key was activated for 1,500 ms. The symbol disappeared after the response with a delay of 500 ms. One of the numbers appeared after 100 ms. Loudspeaker symbols triggering target items within one set were repeated eight times in a random order, resulting in response blocks of 16 items plus the additional four number items, which were repeated four times each. Thus, one response block involved 32 items in total. The experiment started with a short practice session of two sets, practice items mimicked the experimental materials, in which participants could get acquainted with different phases and the experimental setting.

The presentation of the stimuli and the measuring of the reaction times were controlled by the NESU software package. The spoken reactions were registered by a Sennheiser MD211N microphone, which fed into a NESU-box voice key device and a DAT recorder (Sony DTC-55ES). The experimenter sat in the same room and took note of hesitations, voice key errors, wrong naming responses, and time outs. On average, an experimental session lasted 30 min.

9.4. Design

The two-level variable *Frequency* (high-frequency vs. low-frequency) was tested within participants and within quadruples. Each participant produced each of the 32 syllables eight times, resulting in a total of 256 experimental trials. One experimental set consisted of two syllables that were presented as a pair. These pairs were constructed as follows: Only items of the same frequency, i.e. either high- or low-frequency items, were combined to build one item pair. These frequency-homogeneous pairs were constructed on the basis that they were as distinct as possible from each other, that is, it was taken care of that the two syllables within each set had no segmental overlap. That implied that no two items were taken from the same quadruple. The pairing of two syllables resulted in 16 sets, that is, eight high- and eight low-frequency sets. The pairing of two items into one frequency-homogeneous set was the same for high- and low-frequency pairs, that is, once a pair was built, take for example the high-frequency pair *lug* [lʏx]—*bin* [bɪn], this pairing determined also the pairing of their counterparts, which is in this case the low-frequency pair *lur* [lʏr]—*bing* [bɪŋ]. A full list of experimental item pairs is given in the Appendix B.

For the order of the 16 experimental sets two constraints had to be respected: 1) High- and low-frequency counterparts (e.g. *lug* [lʏx]—*bin* [bɪn]; *lur* [lʏr]—*bing* [bɪŋ]) should be presented with a maximum distance of trials between them. 2) Items or sets which had an overlapping onset with other items or sets should not be presented as trials in direct succession (e.g. the item set *lug* [lʏx]—*bin* [bɪn]) should not be presented in direct succession to the other item set *luk* [lʏk]—*teg* [tɛx]). High- and low-frequency item sets alternated across the 16 item sets. Taking these constraints into account, two separate Latin Squares were designed that contained only those four item sets, which did not overlap with each other. These two Latin Squares were then combined in such a way that the transitions between them were also controlled for initial overlap. The order of the first eight trials was reflected by the remaining eight item sets. Using this Latin Square procedure, sixteen experimental versions resulted: Every item set occurred at each position across all experimental versions. The position of the production cue (left vs. right) was also counterbalanced across participants.

10. Results

Test items leading to wrong or invalid responses (mispronunciations, voice key errors, or hesitations) were not included in the reaction time analysis. They were coded as errors. Reaction times above 800 ms and below 200 ms were considered as invalid and did not enter the reaction time analysis. Observations deviating from a participant's and an item's

Table 2
Mean voice onset latencies (in ms), percentage errors, and standard deviations (in parentheses) in Experiment 1

Frequency	M	(SD)	% Err	(SD)
High	436	(41)	3.2	(2.2)
Low	445	(48)	4.6	(3.5)
Difference scores	−9		−1.4	

mean by more than two standard deviations were considered as outliers and also discarded from the reaction time analysis. One hundred and sixty-one (3.9%) trials were treated as errors and 67 (1.6%) as outliers.

The mean of the high-frequency items and the mean of the low-frequency items were submitted to *t*-tests. Two complementary analyses were computed, one treating participants (t_1) and one treating quadruples (t_2) as random factor (Clark, 1973). The mean voice onset latencies, standard deviations and error rates for Experiment 1 are summarized in Table 2.

The analysis of reaction times showed that high-frequency monosyllabic pseudo-words were produced significantly faster than low-frequency monosyllabic pseudo-words (t_1 (15) = 2.651, $P < 0.05$; t_2 (7) = 2.904, $P < 0.05$). The analysis of errors revealed a tendency towards more errors in the low-frequency condition than in the high-frequency condition ($t_1(15) = 1.689$, $P = 0.11$; t_2 (7) = 2.418, $P < 0.05$).

11. Discussion

The results of Experiment 1 show a significant frequency effect for monosyllabic Dutch pseudo-words. High-frequency syllables are faster produced than low-frequency syllables. As potential confounds were carefully controlled for, the present effect might be interpreted as a small but significant syllable frequency effect. As already pointed out, syllable-frequency effects are expected to be found only for stored units. Thus, these results clearly support the notion of the mental syllabary. In consideration of the negative results of the syllable priming studies (see above), we conclude that this frequency has to be attributed to the frequency of stored *phonetic* syllables rather than to stored *phonological* syllables.

After having demonstrated a syllable frequency effect on mono-syllables, we now turn to disyllabic targets. This can provide a follow-up to another finding of Levelt and Wheeldon's (1994) results. By using disyllabic Dutch words to investigate effects of syllable frequency they factorially manipulated first versus second syllable frequency. The overall word-form frequency was controlled for. An interesting outcome of their experiments was that there was only a frequency effect for the second syllable but not for the first one. They argued that this result reflects an underlying principle namely that '[…] the speaker cannot or will not begin to articulate the word before its phonetic encoding is complete.' (Levelt & Wheeldon, 1994, p. 254).

This conclusion depends crucially on the assumption that the retrieval of a phonetic

syllable from the mental syllabary is independent of the processing status of the preceding syllable. For example, as soon as the second syllable of a disyllabic word has been constructed at the level of phonological encoding, the retrieval of its gestural score will be initiated, even if the retrieval of the gestural score for the first syllable is not yet complete. As a consequence there can be some temporal overlap between the retrieval of the second and the first syllable. But even if the first syllable's gestural score has been retrieved before phonetic encoding of the second syllable has been completed, articulation will wait until both syllabic programs of the target word have entered the output buffer. Any speed benefit gained from a high-frequency first syllable will thus not become apparent.[7]

In order to replicate the frequency effect obtained with mono-syllabic pseudo-words and to test the prediction that (only) the second syllable in a disyllabic word will exhibit frequency effects we investigated disyllabic (Dutch) pseudo-words having the frequency-manipulation on their second syllables.

12. Experiment 2: Disyllabic pseudo-words, frequency-manipulation on the second syllable

In Experiment 2, disyllabic Dutch pseudo-words were used to investigate effects of syllable frequency. High- and low-frequency syllables were embedded in pseudo-words, the second syllable in these pseudo-words was frequency-manipulated. All of the used pseudo-words were phonotactically possible strings of Dutch. Under the assumption that articulation waits for the retrieval of all gestural scores of the phonological word, the syllable frequency effect should be replicated.

13. Method

13.1. Participants

Thirty-two native speakers of Dutch participated in the Experiment. They were randomly taken from the pool of participants of the Max Planck Institute in Nijmegen, The Netherlands and were paid for their participation. They had no known hearing deficit, and they had normal or corrected-to-normal vision.

---

[7] Notice that this assumption presupposes that the processing duration of the second syllable is independent of the processing of the preceding syllable(s). Moreover-according to Levelt and Wheeldon-the first syllable could only have an effect when the retrieval of the first syllable is completed only after the second syllable is retrieved, which is very unlikely. Levelt and Wheeldon state that it takes approximately 200 ms to articulate a syllable; whereas Wheeldon and Levelt (1995) showed that the generation of phonological syllables only needs half of that time, some 100 ms. Thus, they argue-as their reported syllable frequency effect amounts to only 15 ms-'[…] it seems implausible that phonetic encoding of the second syllable can 'overtake' encoding of the first one due the advantageous frequency conditions.' (Levelt & Wheeldon, 1994, p. 254).

13.2. Procedure and apparatus

The procedure and apparatus were the same as in Experiment 1.

13.3. Materials

All 32 items were disyllabic pseudo-words obeying Dutch phonotactics. In this experiment, the same syllables as in Experiment 1 were used which served as second syllables. In total, there were again 16 high-frequency and 16 low-frequency disyllabic items. As first syllables, eight high-frequency CV-syllables were selected. For the count *number of occurrences*, the CV-syllables ranged in frequency from 54.43 to 137.74 with an average of 97.34 (SD=33.01). For the count *summed frequency of occurrence*, the CV-syllables ranged in frequency from 1,101.76 to 3,475.79 with an average of 2,900.53 (SD= 924.34). One high-frequency syllable was assigned to one quadruple to serve as first syllable for all four members of this quadruple in order to keep the comparability within the quadruples (e.g. li.kem [li.kɛm], high-frequency; li.kes [li.kɛs], low-frequency; li.wes [li.ʋɛs], high-frequency; li.wem [li.ʋɛm], low-frequency). It was taken care of that none of the resulting disyllabic pseudo-words would form any existing word form in Dutch. A full list of items used in this experiment can be found in the Appendix C. The assignment of items to experimental pairs or sets remained the same, i.e. the pairing of two items taken from different quadruples, which were presented as one experimental set was the same as in Experiment 1. A list showing these pairs is also given in the Appendix D. The same numbers as in Experiment 1 were used which were presented in between two subsequent test items.

Acoustic versions of the disyllabic pseudo-words were spoken by a female native speaker of Dutch. The spoken pseudo-words were digitized at a sampling rate of 22 kHz, to be used during the learning phase of the experiment. They varied in duration from 567 ms to 799 ms with an average of 644.78 ms (SD=57.42). There was no difference in duration between high-frequency and low-frequency items (both $ts<1$). The disyllabic pseudo-words were spoken with stress on the second syllable. It has been suggested that articulatory routines for stressed and unstressed syllables are independently represented in the repository (Crompton, 1981; Levelt, 1989). Thus, in order to keep the basic syllable material between Experiments 1 and 2 as consistent as possible, we opted for the non-default stress-pattern in Dutch, that is, the second syllable carries the stress.

13.4. Design

Thirty-two different experimental versions were constructed applying the same principles as in Experiment 1.

14. Results

The raw data were treated in the same way as in the previous experiments. Four hundred and ten (5%) trials were treated as errors and 156 (1.9%) as outliers. The mean

Table 3
Mean voice onset latencies (in ms), percentage errors, and standard deviations (in parentheses) in Experiment 2

Frequency	M	(SD)	% Err	(SD)
High	435	(51)	5.2	(3.5)
Low	435	(51)	4.8	(4.0)
Difference scores	0		−0.4	

voice onset latencies, standard deviations, and error rates of Experiment 2 are summarized in Table 3.

The frequency of the second syllable in a disyllabic pseudo-word did not affect naming latencies in this experiment as demonstrated by t-tests (both t's <1). The comparison of error rates did not show a significant difference either ($t_1 < 1$; $t_2(7) = 1.049$, $P > 0.3$).

15. Discussion

The results of Experiment 2 did not show any second syllable frequency-effect for disyllabic Dutch pseudo-words. However, before concluding that disyllabic pseudo-words containing high- and low-frequency syllables are not sensitive to frequency-effects, let us carefully consider what could have possibly caused the present null-effect.

By comparing Levelt and Wheeldon's (1994) results to the present experiments, we can distinguish a few differences that might be crucial in explaining the (divergent) data pattern. The fact that their experiment used existing disyllabic Dutch words that had to be learned in the context of four other targets, whereas participants in the present experiments had to memorize only two disyllabic pseudo-words could be responsible for the different results. The theoretical assumption that led to the hypothesis that only the second syllable should exhibit frequency effects was that speakers do not initiate articulation before all phonetic syllables have been retrieved. Could it be the case that this assumption is more plausible for existing words where linguistic integrity plays a role than for pseudo-words where integrity does not play such a prominent role? Do speakers set another response criterion? Under the present high-speed conditions participants may have started articulation as soon as the gestural score for the initial syllable was available for execution.

Results that point to the assumption that not all of the phonological word's components have to be phonetically encoded before articulation is initiated stem from a study by Schriefers and Teruel (1999). They investigated the production of German adjective-noun phrases while hearing prime syllables that were identical to one of the syllables in the target phrase (as in 'lila Säge' [purple saw]). Overall, they obtained priming effects of the first syllable of the first word, weak priming effects for the second syllable of that word but no priming effects for the second word. However, based on the number of restarts and hesitations, participants were divided into two groups for post-hoc analyses. Participants in the 'careful' group (mean of three restarts and hesitation errors) showed priming for both, the first and the second syllable of the first word, whereas participants in the 'hasty' group (mean of eight restarts and hesitation errors) only showed priming effects for

the first syllable. Schriefers and Teruel argued that this reflects speakers' ability to adjust the size of the planning unit and that planning of the first syllable might suffice to initiate articulation.[8]

The assumption that speakers in general 'wait' until all components of a polysyllabic phonological word is phonetically encoded as proposed by Levelt and Wheeldon (1994) is apparently incorrect. It rather seems likely as argued in Meyer, Roelofs, and Levelt (2003) that articulation cannot start before completion of *phonological word encoding*, but it can start after *phonetic encoding* of the first syllable of a disyllabic word. Then, if this is indeed what happens here, one should predict a first-syllable frequency effect. The next experiment was designed to determine just that.

16. Experiment 3: Disyllabic pseudo-words, frequency-manipulation on the first syllable

In Experiment 3, the *first* syllable of the disyllabic Dutch pseudo-words was frequency-manipulated. Since there was no frequency effect for the second syllable in a disyllabic pseudo-word, the aim of this experiment is to test whether or not the first syllable in this pseudo-word context is sensitive to frequency effects. According to the explanation that the absence of the syllable frequency effect for the second syllable is due to the initiation of articulation upon phonetic completion of the first syllable of the disyllabic targets, we expect-contrary to the original hypothesis-a syllable frequency effect for the first syllable.

17. Method

17.1. Participants

Sixteen participants were paid for their participation. None of them took part in any of the previous experiments.

17.2. Procedure and apparatus

The procedure and apparatus were the same as in Experiment 1.

17.3. Materials

Materials in this experiment consisted of disyllabic pseudo-words, all of which were possible strings in Dutch. Most of the syllables used for building the second syllables were the same high-frequency CV-syllables used in Experiment 2. In three cases, these second

---

[8] The fact that the first syllable for all items in the present experiment consisted of only eight syllables which were thus highly repeated might also have contributed to the initiation of articulation after the retrieval of the first syllable.

syllables were exchanged by other high-frequency syllables as they fitted the base material better than the original ones. In these cases the regrouping of syllable position resulted in either existing words or in words very similar to existing words. Thus, we exchanged two of the previously taken CV-syllables ('li' and 'ra') to appear as part of another quadruple and introduced one new CV-syllable ('mo' to replace 'si') (see item list in the Appendix E; for the pairing of items into frequency-homogeneous sets see Appendix F). For the count *number of occurrences*, the CV-syllables ranged from a frequency of 54.43 to 135.38 per one million words with an average of 88.89 (SD=29.68). For the count *summed frequency of occurrence*, the CV-syllables ranged from a frequency of 1,101.76 to 3,475.79 per one million words with an average of 2,845.28 (SD=910.09).

Acoustic versions of the pseudo-words were spoken by a female native speaker of Dutch. The spoken targets were digitized at a sampling rate of 22 kHz, to be used during the learning phase of the experiment. They varied in duration from 491 ms to 722 ms with an average of 622.69 ms (SD=55.86). There was no difference in duration between high-frequency and low-frequency syllables (both ts < 1).

17.4. Design

The design of Experiment 3 was identical to the one used in Experiment 1.

18. Results

The raw data were treated in the same way as in the previous experiments. There were 148 (3.6%) errors and 82 (2.0%) outliers. The mean voice onset latencies, standard deviations and error rates for Experiment 3 are summarized in Table 4.

Disyllabic pseudo-words were produced significantly faster when the first syllable was high-frequency than when the first syllable was low-frequency as shown by t-tests (t_1 (15) = 3.508, $P < 0.01$; t_2 (7) = 7.320, $P < 0.001$). The analysis of error rates did not show a significant difference (t_1 (15) < 1; t_2(7) = 1.008, $P > 0.3$).

19. Discussion

The results of Experiment 3 revealed a significant syllable frequency effect for disyllabic Dutch pseudo-words in which the first syllable is frequency-manipulated whereas the second syllable is consistently high-frequency. The magnitude of the effect

Table 4
Mean voice onset latencies (in ms), percentage errors, and standard deviations (in parentheses) in Experiment 3

Frequency	M	(SD)	% Err	(SD)
High	417	(50)	3.9	(4.2)
Low	427	(55)	3.3	(2.7)
Difference scores	−10		0.6	

(10 ms) is comparable to the size of the effect in Experiment 1 (9 ms). Thus, this experiment replicates the results of the experiment that investigated the monosyllabic pseudo-words. Again, the results of this experiment support the notion of a mental syllabary.

Initially, it was predicted that (only) the second syllable in a disyllabic (pseudo-)word would be frequency-sensitive, which we did not find in Experiment 2. We cannot attribute this difference to differences in materials between the experiments, because the relevant syllables were essentially the same. The obvious conclusion is that under the present experimental conditions subjects were able to initiate articulation of the first syllable right upon completion of its phonetic encoding. In that case frequency properties of the first, but not the second syllable will affect onset latencies.

20. General discussion

The aim of the present study was to test the notion of a mental syllabary. The assumption that speakers access a mental storage of precompiled articulatory gestures at the interface of phonological and phonetic encoding was described as an inherent part of the speech production model by Levelt et al. (1999). Such a mental storage implies different retrieval times for high- and low-frequency syllables. In three experiments using a symbol-position association learning task we investigated effects of syllable frequency by contrasting the production of high- and low-frequency syllables in mono- and disyllabic pseudo-words. The syllabic material used in these experiments was carefully controlled for any potential confounds by applying specific search routines. Two high- and low-frequency syllables were each assigned to one quadruple thereby guaranteeing that onsets, offsets, phoneme and bigram frequency and also the transitional probabilities between the single phonemes were controlled for within each of these quadruples. In Experiment 1, a significant syllable frequency effect was obtained. We therefore attributed the observed syllable-frequency effect to the retrieval of stored, pre-compiled syllabic gestural scores.

In all experiments reported here we exclusively used existing Dutch syllables. Speakers are, however, able to produce new syllables as well. Participants will, for instance, rather fluently read materials containing pseudo-words consisting of phonotactically legal but non-existing syllables. Apparently, we can assemble syllables on-line. This means that there are two possible operation routes in the phonetic encoding of syllables: Retrieving pre-compiled syllabic programs and on-line assembly of phonetic programs. Although the present experiments provide support for the first route, retrieval of pre-compiled gestural programs, they are neutral with respect to the existence and ubiquity of the second, on-line assembly route. Assembly may be limited to just new syllables. It may also be the dominant operation for low-frequency syllables. But it cannot be excluded that the assembly route is *always* active, running in parallel to the retrieval route. Speech latencies are then determined by whichever operation is fastest.

20.1. Disyllabic pseudo-words with frequency manipulated syllables

In Experiments 2 and 3, disyllabic pseudo-words were tested that had the frequency-manipulation on the second syllable (Exp. 2) and first syllable (Exp. 3), respectively. Both experiments were carried out in order to replicate the syllable-frequency effect previously obtained for mono-syllabic pseudo-words and to test the theory's prediction for syllable-frequency effects in disyllabic (pseudo-)words.

Experiment 2 using disyllabic pseudo-words containing frequency-manipulated syllables on the second position attempted to find a frequency effect as predicted by the outcome of prior experiments (Levelt & Wheeldon, 1994). As Levelt and Wheeldon (1994) claimed, the initiation of articulation itself is dependent on the completeness of the phonetic encoding of the corresponding planning unit, namely the phonological word. In case of a disyllabic word, phonetic encoding is completed as soon as both syllables are ready for execution. The outcome of Experiment 2 contradicts this assumption. No syllable frequency effect was obtained for the second syllable of a disyllabic (pseudo-)word. Apparently, articulation was initiated before the frequency-manipulated second syllable was retrieved from the mental syllabary.

There are other findings in conflict with the assumption that speakers wait with speech onset until both syllables of a disyllabic word are phonetically encoded. For example, Bachoud-Lévi, Dupoux, Cohen, and Mehler (1998) failed to find different production latencies for mono- and disyllabic target words in a series of picture-naming experiments in French and English. The authors concluded that either a) word forms are not generated sequentially or b) that speakers in fact start articulation before completion of the phonological word's phonetic encoding. The assumption that phonological encoding is a sequential process is supported by several studies (Meyer, 1990, 1991; Meyer & Schriefers, 1991; Roelofs, 1998), thus, it rather has to be concluded that speakers do start articulation before the whole phonological word is phonetically generated. However, by presenting results from a series of experiments investigating effects of word length, Meyer et al. (2003) offer a convincing explanation for the results of Bachoud-Lévi et al. (1998). In Meyer et al.'s experiments, Dutch participants had to name objects with short, monosyllabic and long, disyllabic names. Interestingly, they found effects of word length only when long and short target words were presented in separate blocks, not when they were presented in mixed blocks. They argued that speakers used different response criteria in pure and mixed blocks: Participants tried to meet different response deadlines by either generating the motor program for one syllable of monosyllabic target words or the motor programs for both syllables of disyllabic target words before speech onset.

Several studies suggest that participants can adopt different time criteria when to respond contingent on task difficulty (Lupker, Brown, & Colombo, 1997; Monsell, Patterson, Graham, Hughes, & Milroy, 1992; Taylor & Lupker, 2001). For example, Lupker et al. (1997) found that participants can adjust their performance in relation to the preceding trial. Therefore, Meyer et al. (2003) suggest

that the absence of a word length effect reported in the study by Bachoud-Lévi et al. (1998) could be due to their stimuli presentation only in mixed blocks. In mixed blocks, participants might set an intermediate response criterion and start with the first completed syllable also in disyllabic words. Schriefers and Teruel (1999) also argued for strategic control of response initiation by speakers. They claim that the actual response criterion is apparently influenced by various factors and that the amount of what has to be (phonetically) completed before speech onset is variable.

Taken the results from the latter studies together, speakers seem to be able to adjust their speech coordination to the difficulty and speed of the current experimental task, and furthermore different speakers generally seem to have a preference for faster or slower responses or for smaller or larger planning units. The actual response criterion is apparently influenced by various factors and the amount of what has to be phonetically completed before speech onset seems variable. We interpreted the results of Experiment 2 in that light and concluded that the absence of the second syllable's frequency effect was due to the early initiation of articulation, namely after completion of phonetically preparing the first syllable. This account was confirmed by the results observed in Experiment 3. Here, the frequency of the first syllable of a disyllabic pseudo-word was manipulated. We found a significant syllable frequency effect (which was of the same size as the effect observed in Experiment 1 and thereby replicated the results with the monosyllabic pseudo-words). Thus, taken the results of Experiments 2 and 3 together, the assumption that speakers can start articulation as soon as the first syllable is fully encoded is supported.

In line with the assumption that speakers can start articulation upon phonetic completion of the first syllable in a disyllabic pseudo-word are recent results from Spanish and French. Carreiras and Perea (2004) investigated disyllabic pseudo-word naming in Spanish, manipulating first and second syllable frequency. They obtained significant syllable-frequency effects for the first but not for the second syllable. This result also suggests that speakers start their articulation before having retrieved or computed all of the word's syllables. This result is not restricted to the experimental use of pseudo-words[9] as shown by not only the study by Schriefers and Teruel (1999) but also by Brand et al. (2002). These latter authors collected the naming latencies of 600 French frequency-manipulated disyllabic words from 100 participants in a larger-scale study. A facilitative effect for the frequency of the first syllable but not for the second syllable was obtained.

All in all, the moment speakers actually initiate articulation apparently depends on various factors. It does not seem to be the case that there is a fixed planning unit here. We conclude that the present results as well as the findings reported in the literature can be explained by an advanced and completed *phonological* planning on the one hand and (more importantly) the incremental nature of the subsequent

---

[9] The experimental use of pseudo-words is unavoidable when we have to control for a great number of potentially confounding variables as in the present experiment.

processes on the other hand. The phonological encoding procedures must operate on the basis of the whole phonological, i.e. the on-line syllabification process must be applied to the planning unit as a whole. The further processes that are the transformation of the single phonological syllables into phonetic representations might take place in a piecemeal fashion, syllable by syllable. This incremental, left-to-right procedure in which successive syllables are generated leads to some temporal overlap of processes: The first syllable's gestural score has been successfully retrieved from storage, while the second and potential later syllables are still under construction. Furthermore, the fact that words with more syllables have longer speech onset latencies than words with fewer syllables when other factors, such as number of segments or word frequency are controlled for does not pose a problem for this proposal. These effects of word length can be explained by assuming that the *phonological* rather than the *phonetic* encoding of a phonological word must be completed before speech onset (see Eriksen, Pollack, & Montague, 1970; Klapp, Anderson, & Berrian, 1973; Meyer et al., 2003; Santiago, MacKay, Palma, & Rho, 2000, but see also Roelofs, 2002; Santiago, MacKay, & Palma, 2002; Wheeldon & Lahiri, 1997). Further research is needed to investigate which planning unit is used in a given speech context and what factors release the articulation initiation.

21. Conclusion

We found a clear and significant syllable-frequency effect in monosyllabic as well as in disyllabic pseudo-words. This was interpreted as evidence for the existence of the mental syllabary; access to pre-compiled gestural scores for high-frequency syllables from storage being faster than access to low-frequency syllables. Such pre-compiling will contribute to the speed and fluency of spoken language production.

Furthermore, only in the disyllabic pseudo-words that had the frequency-manipulation on the *first* syllable we obtained a significant syllable frequency-effect but not when the *second* syllable was manipulated. We interpreted this finding as a result of speakers' flexibility to initiate articulation when the phonetic planning of the first syllable is completed for execution.

Appendix A

Experimental syllable quadruples

Quad. nr.	High-freq. sets	Frequency counts		Low-freq. sets	Frequency counts	
		No. of occurrence	No. of summed frequency		No. of occurrence	No. of summed frequency
1	bin [bɪn]	6.4	127.26	bing [bɪŋ]	0.17	0.48
1	ning [nɪŋ]	23.5	1192.57	nin [nɪn]	0.02	0.00
2	bur [bʏr]	4.76	214.79	bug [bʏx]	0.02	0.38
2	lug [lʏx]	2.36	49.71	lur [lʏr]	0.21	0.67
3	ket [kɛt]	4.86	60.81	keg [kɛx]	0.07	0.24
3	teg [tɛx]	6.71	190.57	tet [tɛt]	0.67	2.43
4	kem [kɛm]	1.48	62.24	kes [kɛs]	1.19	3.1
4	wes [ʋɛs]	3.24	162.60	wem [ʋɛm]	0.02	0.1
5	luk [lʏk]	1.74	209.14	lup [lʏp]	0.19	0.67
5	sup [sʏp]	4.17	82.55	suk [sʏk]	0.26	3.02
6	mer [mɛr]	5	313.12	meg [mɛx]	0.07	1.4
6	reg [rɛx]	8.19	339.86	rer [rɛr]	0.05	0.00
7	sjam [ʃɑm]	1.67	34.12	sjag [ʃɑx]	0.05	0.07
7	wag [ʋɑx]	4.31	546.24	wam [ʋɑm]	0.86	3.67
8	tur [tʏr]	4.95	227.81	tug [tʏx]	0.71	1.98
8	zug [zʏx]	1.86	63.81	zur [zʏr]	0.33	2.33

Appendix B

Pairing of syllables into frequency-homogeneous sets in Experiment 1

Set number	High-frequency sets	Low-frequency sets
1	lug-bin	lur-bing
2	mer-sjam	meg-sjag
3	ning-reg	nin-rer
4	wag-bur	wam-bug
5	luk-teg	lup-tet
6	wes-sup	wem-suk
7	ket-zug	keg-zur
8	tur-kem	tug-kes

Appendix C

Materials for Experiment 2

Quadruple nr.	High-frequency sets	Low-frequency sets
1	ta.bin [ta.bɪŋ]	ta.bing [ta.bɪŋ]
1	ta.ning [ta.nɪŋ]	ta.nin [ta.nɪŋ]

Quadruple nr.	High-frequency sets	Low-frequency sets
2	ko.bur [ko.bʏr]	ko.bug [ko.bʏx]
2	ko.lug [ko.lʏx]	ko.lur [ko.lʏr]
3	si.ket [si.kɛt]	si.keg [si.kɛx]
3	si.teg [si.tɛx]	si.tet [si.tɛt]
4	li.kem [li.kɛm]	li.kes [li.kɛs]
4	li.wes [li.ʋɛs]	li.wem [li.ʋɛm]
5	ra.luk [ra.lʏk]	ra.lup [ra.lʏp]
5	ra.sup [ra.sʏp]	ra.suk [ra.sʏk]
6	wa.mer [ʋa.mɛr]	wa.meg [ʋa.mɛx]
6	wa.reg [ʋa.rɛx]	wa.rer [ʋa.rɛr]
7	ti.sjam [ti.ʃɑm]	ti.sjag [ti.ʃɑx]
7	ti.wag [ti.ʋɑx]	ti.wam [ti.ʋɑm]
8	jo.tur [jo.tʏr]	jo.tug [jo.tʏx]
8	jo.zug [jo.zʏx]	jo.zur [jo.zʏr]

Appendix D

Pairing of items into frequency-homogeneous sets in Experiment 2

Set number	High-frequency sets	Low-frequency sets
1	kolug–tabin	kolur–tabing
2	wamer–tisjam	wameg–tisjag
3	taning–wareg	tanin–warer
4	tiwag–kobur	tiwam–kobug
5	raluk–siteg	ralup–sitet
6	liwes–rasup	liwem–rasuk
7	siket–jozug	sikeg–jozur
8	jotur–likem	jotug–likes

Appendix E

Materials for Experiment 3

Quadruple nr.	High-frequency sets	Low-frequency sets
1	bin.ta [bɪn.ta]	bing.ta [bɪŋ.ta]
1	ning.ta [nɪŋ.ta]	nin.ta [nɪn.ta]
2	bur.ko [bʏr.ko]	bug.ko [bʏx.ko]
2	lug.ko [lʏx.ko]	lur.ko [lʏr.ko]
3	ket.li [kɛt.li]	keg.li [kɛx.li]
3	teg.li [tɛx.li]	tet.li [tɛt.li]
4	kem.ra [kɛm.ra]	kes.ra [kɛs.ra]
4	wes.ra [ʋɛs.ra]	wem.ra [ʋɛm.ra]
5	luk.mo [lʏk.mo]	lup.mo [lʏp.mo]
5	sup.mo [sʏp.mo]	suk.mo [sʏk.mo]

(continued on next page)

Quadruple nr.	High-frequency sets	Low-frequency sets
6	mer.wa [mɛr.ʊa]	meg.wa [mɛx.ʊa]
6	reg.wa [rɛx.ʊa]	rer.wa [rɛr.ʊa]
7	sjam.ti [ʃɑm.ti]	sjag.ti [ʃɑx.ti]
7	wag.ti [ʊɑx.ti]	wam.ti [ʊɑm.ti]
8	tur.jo [tʏr.jo]	tug.jo [tʏx.jo]
8	zug.jo [zʏx.jo]	zur.jo [zʏr.jo]

Appendix F

Pairing of items into frequency-homogeneous sets in Experiment 3

Set number	High-frequency sets	Low-frequency sets
1	lugko-binta	lurko-bingta
2	merwa-sjamti	megwa-sjagti
3	ningta-regwa	ninta-rerwa
4	wagti-burko	wamti-bugko
5	lukmo-tegli	lupmo-tetli
6	wesra-supmo	wemra-sukmo
7	ketli-zugjo	kegli-zurjo
8	turjo-kemra	tugjo-kesra

References

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language*, *88*, 148–159.

Alario, F.-X., Costa, A., & Caramazza, A. (2002). Frequency effects in noun phrase production: Implications for models of lexical access. *Language and Cognitive Processes*, *17*, 299–319.

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U.H., Segui, J. (in press). Predictors of picture naming speed. Behavior Research Methods, Instruments, and Computers.

Bachoud-Lévi, A.-C., Dupoux, E., Cohen, L., & Mehler, J. (1998). Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory and Language*, *39*, 331–346.

Bagemihl, B. (1995). Language games and related areas. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 697–712). Cambridge, MA: Blackwell, 697–712.

Baumann, M. (1995). The production of syllables in connected speech. PhD dissertation, Nijmegen University.

Berg, T. (1988). Die Abbildung des Sprachproduktionsprozesses in einem Aktivationsflußmodell. Untersuchungen an englischen und deutschen Versprechern [The representation of the speech production process in a spreading activation model: Studies of German and English speech errors]. Tübingen: Niemeyer.

Blevins, J. (1995). The syllable in phonological theory. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 206–244). Cambridge, MA: Blackwell, 206–244.

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

Brand, M., Rey, A., & Peereman, R. (2003). Where is the syllable priming effect in visual word recognition? *Journal of Memory and Language*, *48*, 435–443.

Brand, M., Rey, A., Peereman, R., & Spieler, D. (2002). Naming bisyllabic words: A large scale study. *Abstracts of the Psychonomic Society*, *7*, 94.

Carreiras, M., & Perea, M. (2004). Naming pseudowords in Spanish: Effects of syllable frequency in production. *Brain and Language*, *90*, 393–400.

Chen, J.-Y., Chen, T.-M., & Dell, G. S. (2002). Word form encoding in Mandarin Chinese as assessed by the implicit priming paradigm. *Journal of Memory and Language*, *46*, 751–781.

Chen, J.-Y., Lin, W.-C., & Ferrand, L. (2003). Masked priming of the syllable in Mandarin Chinese speech production. *Chinese Journal of Psychology*, *45*, 107–120.

Cholin, J., Schiller, N. O., & Levelt, W. J. M. (2004). The preparation of syllables in speech production. *Journal of Memory and Language*, *50*, 47–61.

Conrad, M., & Arthur, J. (2004). Replicating syllable frequency effects in Spanish in German: One more challenge to computational models of visual word recognition. *Language and Cognitive Processes*, *19*, 369–390.

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.

Crompton, A. (1981). Syllables and segments in speech production. *Linguistics*, *19*, 663–716.

Cutler, A. (1997). The syllable's role in the segmentation of stress languages. *Language and Cognitive Processes*, *12*, 839–845.

Cutler, A., McQueen, J., Norris, D., & Somejuan, A. (2001). The roll of the silly ball. In E. Dupoux (Ed.), *Language, brain and cognitive development: Essays in honor of Jacques Mehler* (pp. 181–194). Cambridge, MA: MIT Press, 181–194.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, *25*, 385–400.

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.

Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature*, *384*, 159–161.

Eriksen, C. W., Pollack, M. D., & Montague, W. E. (1970). Implicit speech: Mechanisms in perceptual encoding? *Journal of Experimental Psychology*, *84*, 502–507.

Evinck, S. (1997). Production de la parole en français: Investigation des unités impliquées dans l'encodage phonologique des mots [Speech production in French: Investigation of the units implied during the phonological encoding of words]. Unpublished PhD dissertation, Bruxelles University.

Ferrand, L., Segui, J., & Grainger, J. (1996). Masked priming of word and picture naming: The role of syllable units. *Journal of Memory and Language*, *35*, 708–723.

Ferrand, L., Segui, J., & Humphreys, G. W. (1997). The syllable's role in word naming. *Memory and Cognition*, *25*, 458–470.

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, *47*, 27–52.

Fujimura, O., & Lovins, J. B. (1978). Syllables as concatenative phonetic units. In A. Bell, & J. B. Hooper (Eds.), *Syllables and segments* (pp. 107–130). Amsterdam: North-Holland, 107–130.

Goldstein, L., & Fowler, C. A. (2003). Articulatory phonology: A phonology for public language use. In N. O. Schiller, & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 159–207). Berlin: Mouton de Gruyter, 159–207.

Hendriks, H., & McQueen, J. (Eds.). (1996). *Annual Report 1995*. Nijmegen, The Netherlands: Max Planck Institute for Psycholinguistics.

Hooper, J. B. (1972). The syllable in phonological theory. *Language*, *48*, 525–540.

Howard, D., & Smith, K. (2002). The effects of lexical stress in aphasic word production. *Aphasiology*, *16*, 198–237.

Indefrey, P., & Levelt, W. J. M. (2000). The neural correlates of languages production. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 845–865). Cambridge, MA: MIT Press, 845–865.

Janssen, P. D., Roelofs, A., & Levelt, W. J. M. (2002). Inflectional frames in language production. *Language and Cognitive Processes*, *17*, 209–236.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 824–843.

Kenstowicz, M. (1994). *Phonology in generative grammar*. Cambridge, MA: Blackwell.

Kerzel, D., & Bekkering, H. (2000). Motor activation from visible speech: Evidence from stimulus-response compatibility. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 634–647.

Klapp, S. T., Anderson, W. G., & Berrian, R. W. (1973). Implicit speech in reading, reconsidered. *Journal of Experimental Psychology*, *100*, 368–374.

Laganaro, M. (2003). *Are syllables represented and retrieved during phonological encoding? An integration of psycholinguistic and neurolinguistic studies Poster presented at Euresco conference on The Science of Aphasia: Functional Neuroimaging Studies of Language and its Impairments*. Italy: Acquafredda di Maratea.

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

Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, *42*, 1–22.

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.

Levelt, W. J. M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, *50*, 239–269.

Luce, R. A. (1959). *Individual choice behavior*. New York: Wiley.

Lupker, S. J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*, 570–590.

MacKay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia*, *8*, 323–350.

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.

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.

Meyer, A. S. (1992). Investigation of phonological encoding through speech error analyses: Achievements, limitations, and alternatives. *Cognition*, *42*, 181–211.

Meyer, A., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, *48*, 131–147.

Meyer, A., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *17*, 1146–1160.

Monsell, S., Patterson, K. E., Graham, A., Hughes, C. H., & Milroy, R. (1992). Lexical and sublexical translation of spelling to sound: Strategic anticipation of lexical status. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *18*, 452–467.

Nooteboom, S. (1969). The tongue slips into pattern. In A. G. Sciarone, A. J. von Essen, & A. A. van Raad (Eds.), *Nomen: Leyden studies in linguistics and phonetics* (pp. 114–132). The Hague: Mouton, 114–132.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*, 273–281.

O'Sheagdha, P. G., Chen, J.-Y., Shen, Z. Y., & Schuster, K. (2004). *The role of the syllable is distinctly different in Mandarin and English Poster presented at Architectures and Mechanisms for Language Processing (AMLaP), Aix-en-Provence, France*.

Perea, M., Carreiras, M. (1996). Efectos de frecuencia silábica en la tarea de pronunciación mixta ['Effects of syllable frequency in the mixed pronunciation task']. Psicológica, 17, 425–440.

Perea, M., & Carreiras, M. (1998). Effects of syllable frequency and syllable neighbourhood frequency in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 134–144.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*, 107–142.

Roelofs, A. (1996). Serial order in planning the production of successive morphemes of a word. *Journal of Memory and Language*, *35*, 854–876.

Roelofs, A. (1997a). Syllabification in speech production: Evaluation of WEAVER. *Language and Cognitive Processes*, *12*, 657–693.

Roelofs, A. (1997b). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*, 249–284.

Roelofs, A. (1998). Rightward incrementality in encoding simple phrasal forms in speech production: Verb-particle combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 904–921.

Roelofs, A. (1999). Phonological segments and features as planning units in speech production. *Language and Cognitive Processes*, *14*, 173–200.

Roelofs, A., et al. (2002). Syllable structure effects turn out to be word length effects; Comment on Santiago (2000). *Language and Cognitive Processes*, *17*, 1–13.

Roelofs, A., & Meyer, A. S. (1998). Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *24*, 922–939.

Santiago, J., MacKay, D. G., & Palma, A. (2002). Length effects turn out to be syllable structure effects: Response to Roelofs (2002). *Language and Cognitive Processes*, *17*, 15–29.

Santiago, J., MacKay, D. G., Palma, A., & Rho, C. (2000). Sequential activation processes in producing words and syllables: Evidence from picture naming. *Language and Cognitive Processes*, *15*, 1–44.

Schiller, N.O. (1997). The role of the syllable in speech production. Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. PhD dissertation, Nijmegen University (MPI series; 2).

Schiller, N. O. (1998). The effect of visually masked primes on the naming latencies of words and pictures. *Journal of Memory and Language*, *39*, 484–507.

Schiller, N. O. (1999). Masked syllable priming of English nouns. *Brain and Language*, *68*, 300–305.

Schiller, N. O. (2000). Single word production in English: The role of subsyllabic units during speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *26*, 512–528.

Schiller, N. O. (2004). The onset effect in word naming. *Journal of Memory and Language*, *50*, 477–490.

Schiller, N.O., Costa, A. The role of the syllable in phonological encoding: Evidence from masked priming? (submitted for publication).

Schiller, N. O., Costa, A., & Colomé, A. (2002). Phonological encoding of single words: In search of the lost syllable. In C. Gussenhoven, & N. Warner, *Papers in Laboratory Phonology 7* (pp. 35–59). Berlin: Mouton de Gruyter, 35–59.

Schiller, N. O., Fikkert, P., & Levelt, C. C. (2004). Stress priming in picture naming: An SOA study. *Brain and Language*, *90*, 231–240.

Schiller, N. O., Meyer, A. S., Baayen, R. H., & Levelt, W. J. M. (1996). A comparison of lexeme and speech syllables in Dutch. *Journal of Quantitative Linguistics*, *3*, 8–28.

Schiller, N. O., Meyer, A. S., & Levelt, W. J. M. (1997). The syllabic structure of spoken words: Evidence form the syllabification of intervocalic consonants. *Language and Speech*, *40*, 103–140.

Schriefers, H., & Teruel, E. (1999). Phonological facilitation in the production of two-word utterances. *European Journal of Cognitive Psychology*, *11*, 17–50.

Sebastián-Gallés, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language*, *31*, 18–32.

Selkirk, E. O. (1982). The syllable. In H. van der Hulst, & N. Smith (Eds.), *The structure of phonological representations. Part II* (pp. 337–383). Dordrecht: Foris, 337–383.

Sevald, C. A., Dell, G., & 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. (1979). Speech errors as evidence for a serial ordering mechanism in sentence production. In W. E. Cooper, & E. C. T. Walker (Eds.), *Sentence processing* (pp. 295–342). New York: Halsted Press, 295–342.

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 processing in language* (pp. 17–51). Hilllsdale, NJ: Erlbaum, 17–51.

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

Stemberger, J. P. (1982). The nature of segments in the lexicon: Evidence from speech errors. *Lingua*, *56*, 235–259.

Taylor, T. E., & Lupker, S. J. (2001). Sequential effects in naming: A time-criterion account. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*, 117–138.

Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition*, *15*, 49–74.

Treiman, R., & Danis, C. (1988). Syllabification of intervocalic consonants. *Journal of Memory and Language*, *27*, 87–104.

Vousden, J., Brown, G. D. A., & Harley, T. A. (2000). Oscillator-based control of the serial ordering of phonology in speech production. *Cognitive Psychology*, *41*, 101–175.

Vroomen, J., de Gelder, B. (1994). Speech segmentation in Dutch: No role for the syllable. In: Proceedings of the 1994 international conference on spoken language processing, Yokohama, September 18–22, pp. 1135–1138

Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language*, *37*, 356–381.

Wilshire, C. E., & Nespoulous, J.-L. (2003). Syllables as units in speech production: Data from aphasia. *Brain and Language*, *84*, 424–447.

Zwitserlood, P., Schriefers, H., Lahiri, A., & van Donselaar, W. (1993). The role of syllables in the perception of spoken Dutch. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 260–271.