LANGUAGE AND COGNITIVE PROCESSES
2008, 00 (00), 1–23

Psychology Press
Taylor & Francis Group

Effects of syllable preparation and syllable frequency in speech production: Further evidence for syllabic units at a post-lexical level

Joana Cholin
University of La Laguna, Spain

Willem J. M. Levelt
Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands

In the current paper, we asked at what level in the speech planning process speakers retrieve stored syllables. There is evidence that syllable structure plays an essential role in the phonological encoding of words (e.g., online syllabification and phonological word formation). There is also evidence that syllables are retrieved as whole units. However, findings that clearly pinpoint these effects to specific levels in speech planning are scarce. We used a naming variant of the implicit priming paradigm to contrast voice onset latencies for frequency-manipulated disyllabic Dutch pseudo-words. While prior implicit priming studies only manipulated the item's form and/or syllable structure overlap we introduced syllable frequency as an additional factor. If the preparation effect for syllables obtained in the implicit priming paradigm proceeds beyond phonological planning, i.e., includes the retrieval of stored syllables, then the preparation effect should differ for high- and low frequency syllables. The findings reported here confirm this prediction: Low-frequency syllables benefit significantly more from the preparation than high-frequency syllables. Our findings support the notion of a mental syllabary at a post-lexical level, between the levels of phonological and phonetic encoding.

Keywords: Phonological/phonetic encoding; Mental syllabary; Implicit priming; Syllable frequency.

Correspondence should be addressed to Joana Cholin, University of La Laguna, Department of Cognitive Psychology, Campus de Guajara s/n, 38205 La Laguna, Santa Cruz de Tenerife, Spain. E-mail: jcholin@ull.es

http://www.psypress.com/lcp DOI: 10.1080/01690960802348852

The mental syllabary is thought to be a store for whole gestural scores for at least the high-frequency syllables of a given language. The mental syllabary is an inherent part of the Levelt, Roelofs, and Meyer (1999) theory of spoken word production and is assumed to be located between the levels of phonological and phonetic encoding. At this interface, previously generated, abstract phonological syllables retrieve their phonetic matches from the syllabary. The retrieval of stored syllables facilitates the transformation from abstract phonological into context-dependent phonetic syllables as it reduces the workload relative to a segment-by-segment conversion. These phonetic syllables or motor programs will then guide the subsequent steps, including articulation, to produce spoken language.

In recent years, more and more evidence suggesting that speakers in fact retrieve stored syllabic units has been gathered (Carreiras & Perea, 2004; Cholin, Levelt, & Schiller, 2006; Laganaro & Alario, 2006; Levelt & Wheeldon, 1994; see also Aichert & Ziegler, 2004). However, findings that can clearly pinpoint the level at which those units are retrieved are very scarce. The Levelt et al. (1999) theory makes clear claims about the level where access to stored syllable programs ought to occur, namely at the interface of phonological and phonetic encoding.

The current paper aims to test whether the retrieval of syllabic units in fact takes place at this hypothesised location. So far, the available evidence is consistent with the notion of a separate retrieval of phonetic syllables at the phonetic/phonological interface but there is no direct evidence to support this claim. The present experiment aims to provide this direct evidence by examining the interaction between two independently established effects, one of which is relevant to identifying the relevant stage of processing and the other of which is relevant to the notion of stored representations. By examining the interaction between these two effects, we should find direct support for the claim that pre-compiled syllable representations are retrieved at this post-lexical level. We combined the *implicit priming paradigm* that has been successful in detecting the emergence of syllables during word-form encoding with a material set that manipulated the syllable-frequency of its items to directly test *when* syllabic units are retrieved during speech planning. The implicit priming technique makes use of the fact that implicit knowledge of certain aspects of an action accelerates the execution of this action (Rosenbaum, Inhoff, & Gordon, 1984). In implicit priming studies, participants learn sets of prompt-response pairs in which the responses are phonologically related to one another (e.g., *lo.tus*, *lo.ner*, *lo.cal*).[1] The relatedness may consist in segmental overlap, in overlap of syllabic structure or both (see above). The overlap between the responses within one set

[1] Dots indicate syllable boundaries.

functions as an implicit prime. Depending on the amount/quality of overlap, it can be tested what kind of information speakers need in order to best prepare for an utterance (in comparison to sets where there is no overlap between set members). Speakers can prepare an utterance more successfully when more information is given implicitly, i.e., when segmental and syllabic structures are shared between responses (as shown in the example above; Cholin, Schiller, & Levelt, 2004; Meyer, 1990, 1991; Roelofs & Meyer, 1998). As will be discussed in more detail below, implicit priming studies have established that knowledge about the syllable structure of a (to-be-prepared) utterance is relevant only *after* the word form are already retrieved from memory, which points towards the conclusion that syllables play a separate role (from words) during speech planning. However, even though these results testify to the relevance of syllabic information at this post-lexical level, these findings do not provide incontrovertible evidence for the assumption that syllables are in fact accessed as independent, pre-compiled units. On the other hand, the finding that high-frequency syllables yield faster production times than low-frequency ones strongly supports the notion that syllables are retrieved as whole units because only stored units are expected to exhibit frequency effects, however, a syllable-frequency effect per se is not informative with respect to the location of the assumed storage. Thus, a combination of the paradigm that has previously been found to be sensitive to effects occurring at this post-lexical level and an experimental factor that implicates stored syllabic units (i.e., syllable frequency) seems to be most promising in the endeavour to locate the mental syllabary. The virtue of the present experiment is that it will test for the interaction of the independently observed effects of syllable preparation and syllable frequency within one single experiment to provide evidence for the location of the mental syllabary.

Before we introduce the current study in more detail and further explain its logic, we present theories of word production and discuss the relevant evidence for and against their assumptions of an involvement of syllables at specific levels.

THEORIES OF WORD PRODUCTION AND THE INVOLVEMENT OF SYLLABLES

Theories of word production generally agree that syllables play a role in speech production planning (e.g., Dell, 1986, 1988; Levelt et al., 1999; Shattuck-Hufnagel, 1979, 1983), however, there are contrasting assumptions regarding the level at which syllables come into play. While some researchers (Dell, 1986, 1988; Shattuck-Hufnagel, 1979, 1983) assume that syllables are an inherent part of the lexicalized word forms, others (e.g. Cholin, Levelt, &

Schiller, 2006; Levelt et al., 1999; Schiller & Costa, 2006) argue that syllables (as abstract phonological units) emerge during context-dependent online syllabification processes and are separately stored and retrieved as phonetic syllable programs.

Generally, it is assumed that word production starts with the activation (Dell, 1986) or the selection (Levelt et al., 1999) of a word entry, the so-called lemma which, in turn, activates its corresponding word form. The different theories make different assumption with respect to the quality of the word form, or rather with respect to the kinds of information that are released upon retrieval of the word form. Dell (1986, 1988) assumes that the word's phonemic code is syllabified. In his theory, word-form retrieval makes two kinds of information accessible, on the hand, phonological syllabic units (bundle of segments); and on the other hand, syllabic frames or word-shape headers, that specify the C(onsonsant)-V(owel) (hereafter CV-)structure of the syllable and syllable-internal positions such as onset, nucleus, and coda (for similar assumptions see McNeilage, 1998; Shattuck-Hufnagel, 1979, 1983). The frames or word-shape headers serve as placeholders in which the segmental content will be filled in during the process of segment-to-frame-association.

Contrary to Dell (1986, 1988), Levelt and colleagues (1999) assume that the phonological code of a word form merely consists of an ordered set of phonemic segments. Crucially, at the stage of phonological encoding, phonological segments are not yet assigned to syllabic positions, nor is the CV-structure for the word specified. Similarly, while the metrical structure is an inherent feature of the retrieved word-shapes in Dell's model, the Levelt et al. theory assumes that the stress pattern for a given word is only stored in case of a non-default stress pattern. For monosyllabic words and for all other polysyllabic words with a default stress pattern (i.e., which in Dutch is the first syllable that carries stress), it is not stored but computed (see also Schiller, Fikkert, & Levelt, 2004).

The main argument for not assuming pre-determined syllable internal positions in the lexically stored phonological codes is based on 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 morphophonological processes like inflection or cliticisation (Booij, 1995). The ubiquity of 'resyllabifications' in the normal use of Dutch (see Schiller, Meyer, Baayen, & Levelt, 1996), renders pre-specification of segments to syllable positions highly inefficient.[2]

[2] 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).

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 online-syllabification, 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 maximisation of onsets and sonority gradations) and, on the other hand, language-specific rules, e.g., phonotactics. Together, these rules create easily pronounced syllables. The output of phonological encoding is a phonological word, specified for its metrical, syllabic, and segmental properties.

PHONETIC ENCODING AND ACCESS TO THE MENTAL SYLLABARY

The fairly abstract, syllabified phonological words are incrementally translated into articulatory-motor programs. The Levelt et al. theory assumes that as soon as a syllable emerges during incremental syllabification, the corresponding syllabic articulatory gesture will be selected from the repository possibly located in Broca's area or a pre-motor area (Dronkers, 1996; Indefrey & Levelt, 2000; Kerzel & Bekkering, 2000). The output of the mental syllabary in turn serves as input to phonetic encoding. During this latter step 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, for example, a glottal tier, a nasal tier, and an oral tier. The gestural scores are abstract in the sense that their execution is highly context-dependent (due to allophonic variation, coarticulation and, as a result of this, assimilation). The actual details of the movements in realising the scores, such as lip protrusion and jaw lowering, are 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. Finally, the articulatory network, a coordinative motor system that includes feedback mechanisms (Goldstein & Fowler, 2003; Saltzman, 1986; Saltzman & Kelso, 1987), transforms these articulatory plans into overt speech.

EVIDENCE FOR AND AGAINST THE INVOLVEMENT OF SYLLABLES AT SPECIFIC LEVELS

Syllable priming studies

Under the assumption that the syllable constitutes a relevant unit during speech planning, a series of studies in a number of different languages used a syllable priming task to identify syllabic units during word-form encoding (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 & Costa, 2006; 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 a prime for *ba*.sis or *bas* as a prime for *bas*.ket) or incongruent (e.g., *ba* as a prime for *bas*.ket or *bas* as a prime for *ba*.sis). 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 was, facilitated the response relative to unrelated control primes. The original segmental overlap effect (Schiller, 1998) has recently been specified in more detail (Schiller, 2004).[3] Schiller and Costa (2006) specifically asked whether the syllable priming method might be insensitive to syllabic effects and included more visible, that is, unmasked primes with a longer stimulus-onset-asynchrony (SOA) to allow for a longer and more explicit exposure of the prime. However, even this unmasked prime presentation did not lead to any syllable priming effects. Therefore, as Schiller and Costa concluded, we would like to deduce from these results that the reason why there was no syllable priming effect is that at the stage where the priming taps into speech planning there is only segmental but no syllabic structure available. The

[3] 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 a syllable priming effect (see also Evinck, 1997; and for a review Schiller et al., 2002). 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 low 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.

primes in the syllable-priming task never get overtly articulated; therefore, 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.

It should be noted that some studies have found effects of abstract syllable structure supporting Dell's idea (Dell, 1986, 1988) of word shape headers (Costa & Sebastián-Gallés, 1998; Sevald, Dell, & Cole, 1995; but see Roelofs & Meyer, 1998 for counter-evidence). However, as already discussed, in the case of stored *phonological* syllables, the various syllable priming studies should have shown a syllable priming effect instead of the repeatedly found segmental overlap effect. This latter finding seriously challenges the notion of syllabified word forms and leads us to the conclusion that there – in fact – might not be any syllabic information within the entries in the mental lexicon.

Implicit priming studies

Another paradigm that uses a priming procedure is the already mentioned implicit priming paradigm. Here, the phonological overlap between responses within one set, the homogeneous set, is used to test what information speakers need in order to prepare for an utterance. In the Meyer (1990, 1991) studies, participants learned sets of semantically related prompt-response pairs in Dutch. In homogeneous sets, the overlap ranged from an overlap of the first segment (e.g., dijk [dike] – ***p****ol.der* [polder]; nootje [nut] – ***p****in.da* [peanut]; tijger [tiger] – ***p****an.ter* [panther]) to an overlap of the first two (e.g., podium [platform] – ***to***.neel [stage]; geheel [whole] – ***to***.*taal* [total]; komkommer [cucumber] – ***to***.*maat* [tomato]) and three segments (e.g., kruid [herb] – ***ker***.*vel* [chervil]; specerij [spice] – ***ker***.*mis* [fair]; gevangenis [prison] – ***ker***.*ker* [dungeon]). Heterogeneous sets are created by regrouping the prompt-response pairs from the homogeneous sets; as a consequence, there is no shared phonological property among responses in those sets. Production latency (the time between onset of prompt and speech onset) is the dependent variable. The standard effect in the implicit priming paradigm is faster production latencies for homogeneous than for heterogeneous blocks. Meyer (1990, 1991) reported such an effect only when the response words in the homogeneous sets shared one or more word-initial segment(s). No effect was shown for shared word-final segments indicating that phonological encoding is a serial process that proceeds from the beginning to the end of a lexical item. Furthermore and more importantly, the preparation effect was found to increase with the number of shared segments.[4] That is, even though syllable structure was not directly manipulated in Meyer's experiments, the fact that the magnitude of the effect increases with the number of shared segments strongly points towards the same conclusion that we drew from the previously reported syllable priming studies: There is no syllable information stored within the word form retrieved from the mental lexicon. If word forms would already be chunked into syllables one would expect that syllables rather than shared segments would be sensitive to the preparation and accordingly there should be no difference whether there is an overlap of the first two or three segments.

But can we find evidence for the alternative assumption by Levelt and colleagues (1999) that syllables first surface during online-syllabification? In order to test whether the emergence of syllables during this stage of word-form encoding can be detected, the study by Cholin, Schiller, and Levelt (2004) directly manipulated syllable structure. An odd-man-out variant of the implicit priming technique (Janssen, Roelofs, & Levelt, 2002) was used to test whether speakers can benefit from a shared syllable structure. 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 underlying hypothesis of this study was that – under the assumption that the syllable is a relevant processing unit in speech production – speakers need knowledge about the current syllabic structure in order to prepare for a target utterance. Thus, it was predicted that speakers can prepare their utterance more efficiently when they know what the structure of the initial syllable will be (as they do in constant sets) then when the initial syllable varies within the set of words (as they do in the variable sets). In other words, if syllables are encoded during online-syllabification and syllable structure is indeed a crucial piece of information then we should find a larger preparation effect for constant sets since speakers can proceed with their preparation beyond the level where syllables are encoded. Crucially, the deviant syllable structure was expected to

---

[4] For one experiment, Meyer (1991) reports that sets with open initial syllables (CV) that share only those two initial segments produced preparation effects that were equivalent to effects produced for sets with closed syllables (CVC) that shared three initial segments. This result was surprising because a pure segmental length effect would predict larger preparation effects in the CVC sets since they comprise one more shared segment. This finding supports the possibility of syllabic effects that are independent of segmental length. However, contrary to this result, Roelofs (1996), Experiment 6) showed that the size of the preparation effect depends on the length of the shared syllable in terms of number of segments.

spoil the preparation effect for the entire set, that is, even after removing the odd-man-out from the analysis of reaction times, the preparation effect for the remaining responses of an entire variable set should be reduced (see Cholin et al., 2004 for the details of the statistical analyses). Two different CV-structures (CVV, Exp. 1 and CCVV, Exp. 2) were investigated. The results in fact showed a significantly larger preparation effect for constant sets. The constant sets were on average 64 ms faster produced than the variable sets. Control studies showed that variable sets also yielded a preparation effect in comparison to a baseline condition where no phonological property was shared between members in a response set. Thus, responses in variable sets can also be prepared but to a lesser extent, indicating that syllables do play a functionally important role during speech planning. Of course, the crucial question is at what level. Recall that we attributed the segmental overlap effect that was replicated time and again in different languages and paradigms to the fact that those tasks do not tap into the level where syllables are retrieved or encoded. Since we find a graded preparation effect also for variable sets, we concluded that the segmental overlap in variable sets allows for a preparation of the shared segments but not for the syllable structure which is needed during online-syllabification. Thus, in variable sets, all stages preceding on-line syllabification contribute to the preparation effect. In contrast, in constant sets, the retrieval of stored syllabic units, phonetic encoding through articulation can contribute to the preparation effect. Thus, these results strongly support the assumption that syllables are computed online. However, they do not represent indisputable evidence for the notion of a separate retrieval of stored syllabic units. Effects that do represent strong evidence for the existence of the mental syllabary are effects of syllable frequency because only stored units are thought to exhibit frequency effects.

Syllable-frequency effects

In analogy to the findings of word frequency, that high-frequency words are retrieved and produced faster than low-frequency words (Jescheniak & Levelt, 1994; Oldfield & Wingfield, 1965), stored syllables should exhibit (syllable) frequency effects. If the mental syllabary consists of retrievable representations corresponding to syllables, than the retrieval process should be faster for high-frequency syllables than for low-frequency syllables. Syllable-frequency effects were found in a number of studies investigating syllable-frequency effects in different languages with words and pseudo-words (German: Aichert & Ziegler, 2004; Dutch: Cholin et al., 2006; Levelt & Wheeldon, 1994; Spanish: Carreiras & Perea, 2004; Perea & Carreiras, 1996; French: Brand, Rey, Peereman, & Spieler, 2002; Laganaro & Alario, 2006; but see also Croot & Rastle, 2004; Monsell, van der Lugt, & Jessiman, 2002). These studies strongly support the claim that syllables are retrieved as whole units that are separately stored from the word forms, most likely in a mental syllabary.[5] Two studies shall be mentioned in detail as they have further implications for the current study. Cholin et al. (2006) used a symbol-position association learning task to contrast the production of high- and low-frequency syllables in mono- and disyllabic pseudo-words. The material set was very carefully chosen to control for any potential confound. Since the materials from this study will provide a crucial piece for the material set of the current study it will be described in further detail. Quartets of four CVC-syllables, two of high- and two of low-frequency, were selected sharing the same nucleus (e.g., /ɛ/), two different onsets (e.g., /ʋ/ and /k/), and two different offsets (e.g., /m/ and /s/). The two syllables sharing the same onset were of different frequency (high vs. low) such as 'wem' (low-frequency) and 'wes' (high-frequency). The same holds for the two syllables sharing the same offset, such as 'wem' (low-frequency) and 'kem' (high-frequency). See Table 1 for a depiction.

The construction of those syllabic quartets guaranteed a control not only for onsets and offset within one quartet but also for CV structure (which had to be CVC for that matter), phoneme and biphone-frequency and also the transitional probabilities between the single phonemes of the syllables.

The participants' task was to respond as fast as possible with a previously learned associated target-word when a production cue was presented on the screen. In learning phases, participants were presented a symbol on one of two potential positions (the left or right position on a computer screen) and were simultaneously presented with the to-be-associated word via headphones. In the test phase, the same symbol was shown on either the right or the left side of the computer screen to prompt the previously associated target utterance. The auditory presentation of the target ensured that potential confounds deriving from orthographic factors could be excluded. The results can be summarised as follows: A small but highly significant syllable-frequency effect of 9 ms was obtained in testing monosyllabic pseudo-words. This effect was replicated investigating disyllabic pseudo-words bearing the frequency-manipulation on the first syllable. Both effects strongly support the notion of stored syllables, especially because all potentially confounding factors have been carefully controlled for. The material that was used in this study will serve as the basis for constructing the material set in the current experiment.

Laganaro and Alario (2006) employed a different paradigm to investigate the assumption that stored syllables are retrieved during phonetic encoding. In six experiments involving immediate and delayed production, with or without an interfering task (articulatory suppression) they tested the

[5] Note that the idea of a mental syllabary is also in principle compatible with Dell's syllabified phonological code, the idea of syllabified word-forms itself does not deny the existence of a syllabary storing phonetic syllables.

TABLE 1
High- and low-frequency syllables within one quartet

High-frequency	*Low-frequency*
kem [kɛm]	kes [kɛs]
wes [ʋɛs]	wem [ʋɛm]

production of picture names and pseudo words that consisted of high- and/or low-frequency syllables. They found syllable-frequency effects in immediate pseudo-word production and picture naming and in a delayed naming task with articulatory suppression but not in a delayed naming task without articulatory suppression. The authors' interpretation that the syllable-frequency effect has its origin at the level of phonetic encoding mainly stems from the comparison of the delayed naming experiments. In both experiments, participants had to name targets with high- and low-frequency syllables not upon target presentation but in response to a (delayed) cue. Syllable-frequency effects were only obtained when the delay was filled with an articulatory suppression task (repetition of the syllable 'ba'). The authors assume that the suppression task only disrupts phonetic processing but not phonological encoding and that the results therefore indicate that syllable-frequency effects are likely to arise at the level of phonetic encoding.

To summarise, the available evidence strongly points towards the conclusion that syllables play a role during later stages of word-form encoding and, most likely, in form of separately stored syllabic units.

The present study combines two factors, preparation (overlapping segments and syllabic structure) and syllable frequency (high- versus low-frequency), that have previously been used only in separate experiments to specifically investigate the involvement of syllables in speech production. The combination of these two factors might offer a verification of inferences that, so far, could only be drawn by pulling threads from different paradigms across experiments. An interaction of syllable-preparation and syllable-frequency effects would provide convincing support for the assumption that the retrieval of stored syllables directly follows phonological encoding. The virtue of the current experiment is that it tests the critical conditions as within-subject factors in a single experiment.

EXPERIMENT

We used a naming variant of the implicit priming paradigm (Roelofs, 2004) and contrasted production times for sets consisting of four Dutch disyllabic (pseudo-)words that had identical first syllables (i.e., the homogeneous sets), with sets consisting of four different Dutch disyllabic (pseudo-)words (i.e., the heterogeneous sets). Thus, the first factor is the standard preparation effect: Sets with syllabic overlap are predicted to be produced faster as speakers can prepare part of the utterance due to the implicit primes in homogeneous sets. The second factor is syllable frequency. The first syllables within both sets were either (all) high- or low-frequent. The critical prediction in this experiment is an interaction of these two factors: If speakers access stored syllable units at the level where syllables are prepared, and thus, the retrieval of those units contributes to the preparation effect, then we should find a difference in the preparation effects for high- and low-frequency syllables. In other words, low-frequency syllables should benefit more from the preparation that is possible in homogeneous sets than high-frequency syllables. As a result, the advantage that high-frequency syllables should have over low-frequency syllables should only be visible in heterogeneous sets (where no preparation is possible). In homogeneous sets, however, the effect of syllable-frequency should be levelled out by the preparation effect. The frequency difference should no longer have an effect since high- and low-frequency syllables can be equally prepared for. The predicted interaction of the two effects would (a) indicate that the implicit priming paradigm entails syllabic processing and (b) support the assumption of a mental syllabary between the levels of phonological and phonetic planning levels.

Method

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.

Materials. Four of the previously described and tested syllabic quartets from the Cholin et al. (2006) study (see section ‘Syllable-frequency effects’) were used as the basis for constructing the materials of the present experiment. By a very specific pairing, high- and low-frequency syllables within one quartet were controlled for the following factors: CV-structure, number of phonemes, phoneme frequency, bigram frequency and the transitional probabilities of phonemes within syllables. A full list of experimental quartets and their frequency counts is given in Appendix A. In order to create disyllabic pseudo-words out of these quartets, we chose four high-frequency CV syllables to serve as second syllables (li [li:], ta [ta:], wa [ʋa:], and jo [jo:]).[6] Each of these four CV syllables was combined with each member of the quartets thereby ensuring that the transitions (and the transitional probabilities) between first and second syllables were also controlled for, as in we**m.t**a [ʋɛm.ta], low-frequency, ke**m.t**a [kɛm.ta], high-frequency, and we**s.t**a [ʋɛs.ta], high-frequency, ke**s.t**a [kɛs.ta], low-frequency; see also Table 2 for a schematic depiction of the combination of all four CV syllables with one quartet.

All of the pseudo-words were phonotactically legal strings of Dutch but none of those words was an existing Dutch word. By pairing each quartet, i.e., all four members of one quartet with four different high-frequency syllables, 16 different disyllabic pseudo-words were created, that shared the first syllable.

Design. The 64 disyllabic pseudo-words were grouped into homogeneous (high- and low-frequency) and heterogeneous (high- and low-frequency) sets. Homogeneous sets consisted of four pseudo words sharing the first syllable (e.g., wemta, wemli, wemwa, and wemjo, low-frequency). There were homogeneous high-frequency sets as well as homogeneous low-frequency sets, the four pseudo-words again 'derived' from one base syllable (as in westa, wesli, weswa, and wesjo, high-frequency). Following this procedure, eight homogeneous high-frequency sets were created as well as eight homogeneous low-frequency sets, each of them containing four items. Sixteen heterogeneous sets were created by regrouping the items from the homogeneous sets, again – depending on the frequency of the first syllables – divided into eight low-frequency and eight high-frequency heterogeneous sets. Thus, the two high- and the two low-frequency syllables of each quartet appeared in a homogeneous as well as in a heterogeneous context; every item served as its own control. The experiment consisted of two crossed within-subject factors, the factor *Preparation* (two levels, homogeneous versus heterogeneous) and the factor *Frequency* (two levels, high versus low). In total, the experiment consisted of 32 different item sets, containing a total of 64 different items. For a full list of items and their distribution over the four conditions see Appendix B.

The presentation of blocks of homogeneous and heterogeneous sets alternated in blocks of four sets within the same frequency condition. The

[6] The four high-frequency syllables that were used to construct the second syllables in the disyllabic pseudo-words were among the first percentile of the most high-frequency Dutch syllables; for their frequency values see notes of Appendix A. We opted for these very high-frequency syllables because these syllables are most likely to be stored within the syllabary. The retrieval of those high-frequency syllables should be fast and least error-prone. Furthermore, we decided to have frequency-constant second syllable in order to keep this condition equal across items.

TABLE 2
Example for the learning sets in the four different conditions

Homogeneous sets		*Heterogeneous sets*	
High-frequency	*Low-frequency*	*High-frequency*	*Low-frequency*
kem.ta [kɛm.ta]	kes.ta [kɛs.ta]	kem.ta [kɛm.ta]	kes.ta [kɛs.ta]
kem.li [kɛm.li]	kes.li [kɛs.li]	bin.li [bɪn.li]	bing.li [bɪŋ.li]
kem.wa [kɛm.ʋa]	kes.wa [kɛs.ʋa]	mer.wa [mɛr.ʋa]	meg.wa [mɛx.ʋa]
kem.jo [kɛm.jo]	kes.jo [kɛs.jo]	sup.jo [sʏp.jo]	suk.jo [sʏk.jo]
wes.ta [ʋɛs.ta]	wem.ta [ʋɛm.ta]	wes.ta [ʋɛs.ta]	wem.ta [ʋɛm.ta]
wes.li [ʋɛs.li]	wem.li [ʋɛm.li]	ning.li [nɪŋ.li]	nin.li [nɪn.li]
wes.wa [ʋɛs.ʋa]	wem.wa [ʋɛm.ʋa]	reg.wa [rɛx.ʋa]	rer.wa [rɛr.ʋa]
wes.jo [ʋɛs.jo]	wem.jo [ʋɛm.jo]	luk.jo [lʏk.jo]	lup.jo [lʏp.jo]

succession of the different sets was fully counterbalanced over 32 participants by applying a Latin-Square design.

Procedure and apparatus. The participants were seated in a quiet room in front of a 17-inch computer screen (iiyamaLM704UT). The distance between participants and screen was approximately 50 cm. Before the experiment started, participants received written instructions about the task. The experiment consisted of alternating learning and test phases. In the learning phase, participants were shown sets of four words on the computer screen. The participants were instructed to look silently over the four target words. Words were presented for 9 seconds with a vertical spacing in 24-points lowercase Arial, white on a black screen; see Figure 1.

In the test phase, the words were presented one by one on the screen and participants were instructed to read aloud the target words as fast and as accurately as possible. A trial was structured as follows: participants saw a warning signal (an asterisk) for 500 ms. Next, the screen was cleared for 500 ms, followed by the display of the target word for 1500 ms. Finally, before the start of the next trial there was a blank interval of 500 ms. Simultaneously with target onset the voice key was activated for 1500 ms. The target word disappeared after the response with a delay of 500 ms. The asterisk of the next trial appeared after 100 ms. All the items within each set were repeated five times in a random order, resulting in 20 naming responses per set. On average, an experimental session lasted about 30 minutes.

The presentation of the stimuli and the measuring of the reaction times were controlled by the NESU2000 software package. The spoken reactions were registered by a Sennheiser MD211N microphone, which fed into a NESU-Box2 voice key device and a DAT recorder (Sony DTC-55ES). The experimenter sat in the same room as the participant, separated by a

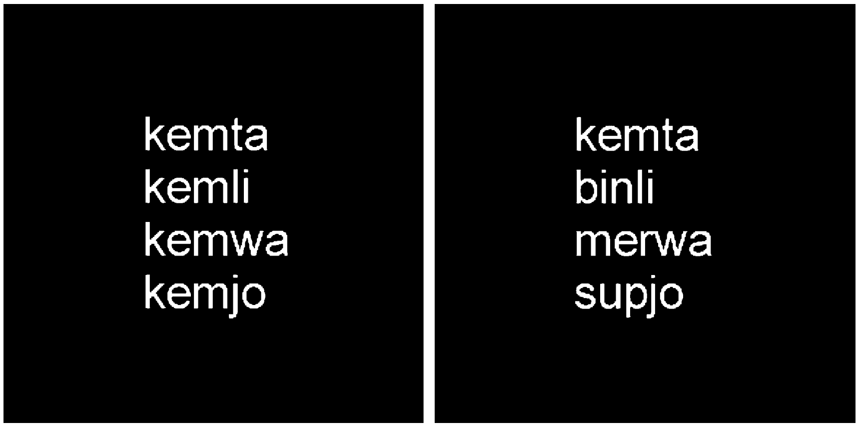

Figure 1. Example for a homogeneous (left) and heterogeneous (right) learning set.

partition-wall, and took note of hesitations, voice key errors, wrong or no naming responses.

Results and discussion

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 mean by more than 2 standard deviations were considered as outliers and also discarded from the reaction time analysis. 347 (1.7%) trials were treated as errors and 390 (1.9%) as outliers. Analyses of variance were run with *Preparation* (homogeneous versus heterogeneous) and *Frequency* (high versus low) as independent variables. Two complementary analyses were computed, one treating participants (F_1) and one treating item-quartets (F_2) as random factor (Clark, 1973).[7] The mean voice onset latencies, standard deviations and error rates are summarised in Table 3.

There is a significant effect of Preparation in the analysis of reaction times, $F_1(1, 31) = 122.99$, $MSE = 857.98$, $p < .001$; $F_2(1, 3) = 480.67$, $MSE = 27.48$, $p < .001$, reflecting faster naming latencies for homogeneous than for heterogeneous sets. Frequency did not show a significant main effect, $F_1(1, 31) = 1.69$, $MSE = 317.90$, $p = .204$; $F_2(1, 3) = 5.22$, $MSE = 11.16$, $p = .106$. Most interestingly, a significant interaction between Preparation and

[7] Since the selection of one item determined the selection of three remaining items within one quartet, items cannot be considered a random factor in this design.

TABLE 3
Mean voice onset latencies (in ms), percentage errors, and standard deviations (in parentheses).

	Frequency							
	High-frequency				Low-frequency			
	M	(*SD*)	% *Err*	(*SD*)	*M*	(*SD*)	% *Err*	(*SD*)
Homogeneous sets	386	(40)	2.0	(1.5)	386	(44)	1.7	(1.5)
Heterogeneous sets	439	(37)	1.5	(1.6)	447	(38)	1.5	(1.5)
Difference scores	−53		0.5		−61		0.2	

Frequency was obtained, $F_1(1, 31) = 5.21$, $MSE = 101.47$, $p < .05$; $F_2(1, 3) = 14.34$, $MSE = 4.66$, $p < .05$, reflecting a significantly larger preparation effect for the low-frequency condition (61 ms) than for the high-frequency condition (53 ms). *t*-tests revealed a significant difference between the high- and the low-frequency sets in the heterogeneous context, $t_1(31) = 2.701$, $p < .05$; $t_2(3) = 4.210$, $p < .05$. Here, no advanced preparation was possible as the four items within these sets had no initial syllable overlap. Thus, the effect of 8 ms reflects a clear syllable-frequency effect. Reaction times for the high- and low-frequency sets in the homogeneous context had identical values; the preparation effect wiped out the frequency effect (both ts < 1). The analysis of error rates yielded no significant effects.

An alternative explanation for the observed interaction could be that, in homogeneous sets, the first syllable of the disyllabic target word was repeated 20 times (five repetitions of each of the four target words) while in heterogeneous sets the same syllable was only repeated five times. The absence of a syllable-frequency effect in homogeneous sets might thus be attributed to the higher number of repetitions in those sets. In order to test for this possibility, an analysis including the factor Repetition was carried out. No interaction of Repetition, Frequency, and Preparation, $F_1(4, 124) = 1.55$, $MSE = 313.28$, $p = .193$; $F_2 < 1$, was found, indicating that the factor Repetition does not alternate the observed interaction of Frequency and Preparation. Additionally, inspection of the first item production only shows a data pattern similar to that of the overall data set: a significant syllable-frequency effect for heterogeneous blocks (10 ms, $p < .05$), while there was no frequency effect for homogeneous blocks (4 ms, $t < 1$).

GENERAL DISCUSSION

The aim of the present study was to gain further insights into the question where access to stored syllables occurs during speech planning. Two experimental paradigms that have previously been employed to investigate the role of the syllable in speech production were combined in a single experiment. The implicit priming paradigm provided an insight with respect to the level where syllables are encoded. The syllable-frequency effect provided evidence as to whether or not syllables constitute separate units. A naming variant of the implicit priming paradigm with a syllable-frequency manipulation was used to investigate whether or not access to the mental syllabary is involved and is thereby contributing to the syllable-preparation effect. We examined the voice-onset latencies for frequency-manipulated disyllabic pseudo-words in Dutch.

Disyllabic pseudo-words were named faster if presented in homogeneous sets consisting of items sharing the first syllable compared with heterogeneous sets that consisted of non-overlapping items. The classic preparation effect for shared versus non-shared phonological properties (Meyer, 1990, 1991) was replicated: the homogeneous sets were on average 57 ms faster produced than the heterogeneous sets. Moreover, the difference between homogeneous and heterogeneous item sets was larger for item sets consisting of low-frequency first syllables than for item sets consisting of high-frequency first syllables. The replication of the syllable-frequency effect within the implicit priming paradigm indicates not only that the observed preparation effect includes access to the mental syllabary but also supports the assumption that the mental syllabary is to be located at a post-lexical level. In homogeneous sets high- and low-frequency syllables could be produced equally fast: The head start that high-frequency syllables have by a faster syllabary access compared with low-frequency syllables was annihilated by the advanced preparation in the present experiment.

Only heterogeneous sets yielded a syllable-frequency. Here, word-form encoding processes could first start with target presentation and thus, high-frequency syllables had an advantage over low-frequency syllables. From the former syllable-frequency studies (Cholin et al., 2006; see also Carreiras & Perea, 2004) we knew that a rather small effect had to be expected. In fact, the effect obtained in this study amounted to 8 ms, mirroring the syllable-frequency effect from the former syllable-frequency study (9 ms and 10 ms testing monosyllabic and disyllabic pseudo-words with the frequency-manipulation on the first syllable).

The outcome of the current experiment reveals two important findings: Firstly, the preparation effect in the implicit priming paradigm includes all stages of word-form encoding prior to articulation. Implicit primes that entail information that allow for the preparation of the full (first) syllable will prompt speakers to proceed with the item's preparation beyond on-line syllabification (see Cholin et al., 2004). This finding significantly contributes to our understanding of the implicit priming paradigm: It can serve as a task that can be used to investigate all stages of the word-form encoding process starting from the first stages of morpho-phonological encoding until and including syllabary access. Secondly, the results of this experiment provide further evidence for syllable-frequency effects and confirm thereby the notion of a mental syllabary using a different paradigm, a naming version of the implicit priming paradigm. Since this version of the paradigm does not require an intensive learning phase as well as a rehearsal of previously learned items, a memory effect as the basis for the syllable-frequency effect can be excluded. This result is particularly relevant as the other paradigms that have been used to investigate syllabic effects did involve memorisation of auditory or visual syllables (Cholin et al., 2004, 2006; Laganaro & Alario, 2006).

Furthermore, it can be excluded that the observed effects are due to grapheme frequencies because orthographic factors have been carefully controlled for by a specific pairing of high- and low-frequency onsets and offsets within one syllabic quartet. Moreover, syllable-frequency effects have been found with the very same material when it was presented exclusively auditorily (see Cholin et al., 2006).

Could the current findings possibly be explained within another account? In principle, the current data by itself do not speak against other theories. For example, Dell's model also assumes that syllabic nodes are sensitive to frequency (Dell, 1986, 1988). A larger gain for low-frequency syllables than for high-frequency syllables due to preparation would therefore be expected. As already stated, syllable-frequency effects as well as the notion of a mental syllabary are compatible with Dell's model. However, the repeated finding of a segmental overlap effect in more than a dozen of priming studies, while no syllabic-priming effect could be obtained, weakens the assumption of syllabified word form at the phonological level. In this sense, the current results support the assumption that the retrieval of stored syllabic units occurs *after* phonological encoding is completed.

Our results corroborate the recent findings by Laganaro and Alario (2006) who found syllable-frequency effects in delayed naming conditions only when the delay was filled with an articulatory suppression task that prevented speakers from phonetically preparing the target utterances. These authors therefore concluded that syllabary access takes place during phonetic encoding.

It should be noted though, that whereas syllable-frequency effects provide evidence for independently stored syllable units that facilitate the late planning processes, these effects remain neutral as to the existence of other, alternative processes that assemble phonetic syllables segment by segment. Apparently, speakers are able to produce syllables they have neither heard nor produced before and that therefore cannot be part of the stored syllable inventory. The alternative assembly route might be limited to new syllables but it may also be the dominant operation for low-frequency syllables. 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. The question of how we can envision different routes of phonetic encoding and how the alternative routes interact with one another remains subject to further research. The related question of whether, in fact, only the high-frequency syllables (of a given language) are stored within the syllabary or whether there are additional entries, for units that are larger and smaller than the syllable, must also remain open for the moment. Unequivocally, the core of the mental syllabary consists of early acquired, high-frequency syllables that are embedded in a network with syllabic neighbours; the number of these high-frequency syllables might be correctly estimated at 500 syllables (Schiller et al., 1996). Though it cannot be excluded that all syllables, once produced, will find their entry in the syllabary, it seems unlikely that all of the remaining 11,000–12,000 syllables (the total number of syllables in languages such as English, Dutch, and German) are stored. Ultimately, the frequency of any given unit might be the decisive criterion for the potential storage within the mental syllabary. The investigation of factors such as (syllable) frequency, age of acquisition, and neighbourhood density (see Vitevitch, Armbruster, & Chu, 2004; Vitevitch & Sommers, 2003) can possibly tell us more about the internal organisation of the syllabary.

To conclude, the location of the mental syllabary at a post-lexical encoding level is strongly supported by these results. Further research will have to clarify the exact nature of retrieval mechanisms of stored syllables and how different planning procedures might interact during phonetic encoding.

Manuscript received January 2008
Revised manuscript received July 2008
First published online month/year

REFERENCES

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

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

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.

Cholin, J., Levelt, W. J. M., & Schiller, N. O. (2006). Effects of syllable frequency in speech production. *Cognition*, *99*, 205–235.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335–359.

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.

Croot, K., & Rastle, K. (2004). Is there a syllabary containing stored articulatory plans for speech production in English? *Proceedings of the 10th Australian International Conference on Speech Science and Technology*, Macquarie University, Sydney, 8–10 December 2004.

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.

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 Ph.D. dissertation, Brussels 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.

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.

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.

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.

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.

Laganaro, M., & Alario, F.-X. (2006). On the locus of the syllable frequency effect in speech production. *Journal of Memory and Language*, *55*, 178–196.

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

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *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.

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, *21*, 499–546.

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.

Monsell, S., van der Lugt, A., & Jessiman, T. (2002, April). *In pursuit of the Syllabary: This Snark is a Boojum!* Paper presented at the Experimental Psychology Society Meeting, Leuven, Belgium.

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.

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. (2004). Seriality of phonological encoding in naming objects and reading their names. *Memory and Cognition*, *32*, 212–222.

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.

Rosenbaum, D. A., Inhoff, A. W., & Gordon, A. M. (1984). Choosing between movement sequences: A hierarchical editor model. *Journal of Experimental Psychology: General*, *113*, 372–393.

Saltzman, E. (1986). Task dynamic coordination of the speech articulators: A preliminary model. In H. Heuer & C. Fromm (Eds.), *Generation and modulation of action patterns* (pp. 129–144). New York: Springer-Verlag.

Saltzman, E., & Kelso, J. A. S. (1987). Skilled actions: A task-dynamic approach. *Psychological Review*, *94*, 84–106.

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. (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. (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. (2006). Activation of segments, not syllables, during phonological encoding in speech production. *The Mental Lexicon*, *1*, 231–250.

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

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.

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.

Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In P. F. MacNeilage (Ed.), *The production of speech*. New York: Springer.

Vitevitch, M. S., Armbruster, J., & Chu, S. (2004). Sub-lexical and lexical representations in speech production: Effects of phonotactic probability and onset density. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 514–529.

Vitevitch, M. S., & Sommers, M. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production. *Memory and Cognition*, *31*, 491–504.

Appendix A

Experimental item-quartets

		Frequency counts			*Frequency counts*	
Quartet No.	*High-freq. sets*	*No. of occurrence*	*No. of summed frequency*	*Low-freq. sets*	*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	1,192.57	nin [nɪn]	0.02	0.00
2	kem [kɛm]	1.48	62.24	kes [kɛs]	1.19	3.1
2	wes [ʋɛs]	3.24	162.60	wem [ʋɛm]	0.02	0.1
3	luk [lʏk]	1.74	209.14	lup [lʏp]	0.19	0.67
3	sup [sʏp]	4.17	82.55	suk [sʏk]	0.26	3.02
4	mer [mɛr]	5	313.12	meg [mɛx]	0.07	1.4
4	reg [rɛx]	8.19	339.86	rer [rɛr]	0.05	0.00

Note: All frequencies 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 forms. Syllable frequencies were calculated for the database from the word-form occurrences per million count. The syllable frequency ranges from 0 to approximately 90,000 per million words, with a mean frequency of 121. Two syllable frequency counts were calculated: The number of occurrences of each syllable (independent of the frequency of occurrence of the syllable in a particular word position, i.e., first or second syllable position within a word) and the number of the summed frequency of occurrence of each syllable (within words). Only instances that had in both scores comparable values were taken. The 8 high-frequency items ranged in the count for the number of occurrence (per one million words) from a value of 1.48 to 23.50 with an average of 6.72 ($SD = 7.15$) and for the count of the summed frequency of occurrence (per one million words) from a value from 62 to 1192.57 with an average of 311.18 ($SD = 369.99$). For the low-frequency-items the values in both counts were a follows: For the count number of occurrence (per one million words), low-frequency-items ranged from 0.02 to 1.19 with an average of 0.25 ($SD = 0.39$). For the count of summed frequency of occurrence (per million words), low-frequency items ranged from zero to 3.1 with an average of 1.1 ($SD = 1.3$). The four high-frequency syllables that served as second syllables had the following values in the two counts: li: 131 3476, ta: 89 2069, wa: 64 3254, jo: 54 1102. The first value gives the number of occurrence (per one million words), the second the summed frequency (per one million words).

Appendix B

Materials for experiment

Homogenous sets		*Heterogeneous sets*	
High-frequency	*Low-frequency*	*High-frequency*	*Low-frequency*
bin.ta [bɪn.ta]	bing.ta [bɪŋ.ta]	bin.ta [bɪn.ta]	bing.ta [bɪŋ.ta]
bin.li [bɪn.li]	bing.li [bɪŋ.li]	kem.li [kɛm.li]	kes.li [kɛs.li]
bin.wa [bɪn.ʋa]	bing.wa [bɪŋ.ʋa]	sup.wa [sʏp.ʋa]	suk.wa [sʏk.ʋa]
bin.jo [bɪn.jo]	bing.jo [bɪŋ.jo]	mer.jo [mɛr.jo]	meg.jo [mɛx.jo]
kem.ta [kɛm.ta]	kes.ta [kɛs.ta]	kem.ta [kɛm.ta]	kes.ta [kɛs.ta]
kem.li [kɛm.li]	kes.li [kɛs.li]	bin.li [bɪn.li]	bing.li [bɪŋ.li]
kem.wa [kɛm.ʋa]	kes.wa [kɛs.ʋa]	mer.wa [mɛr.ʋa]	meg.wa [mɛx.ʋa]
kem.jo [kɛm.jo]	kes.jo [kɛs.jo]	sup.jo [sʏp.jo]	suk.jo [sʏk.jo]
sup.ta [sʏp.ta]	suk.ta [sʏk.ta]	sup.ta [sʏp.ta]	suk.ta [sʏk.ta]
sup.li [sʏp.li]	suk.li [sʏk.li]	mer.li [mɛr.li]	meg.li [mɛx.li]
sup.wa [sʏp.ʋa]	suk.wa [sʏk.ʋa]	kem.wa [kɛm.ʋa]	kes.wa [kɛs.ʋa]
sup.jo [sʏp.jo]	suk.jo [sʏk.jo]	bin.jo [bɪn.jo]	bing.jo [bɪŋ.jo]
mer.ta [mɛr.ta]	meg.ta [mɛx.ta]	mer.ta [mɛr.ta]	meg.ta [mɛx.ta]
mer.li [mɛr.li]	meg.li [mɛx.li]	sup.li [sʏp.li]	suk.li [sʏk.li]
mer.wa [mɛr.ʋa]	meg.wa [mɛx.ʋa]	bin.wa [bɪn.ʋa]	bing.wa [bɪŋ.ʋa]
mer.jo [mɛr.jo]	meg.jo [mɛx.jo]	kem.jo [kɛm.jo]	kes.jo [kɛs.jo]
ning.ta [nɪŋ.ta]	nin.ta [nɪn.ta]	ning.ta [nɪŋ.ta]	nin.ta [nɪn.ta]
ning.li [nɪŋ.li]	nin.li [nɪn.li]	wes.li [ʋɛs.li]	wem.li [ʋɛm.li]
ning.wa [nɪŋ.ʋa]	nin.wa [nɪn.ʋa]	luk.wa [lʏk.ʋa]	lup.wa [lʏp.ʋa]
ning.jo [nɪŋ.jo]	nin.jo [nɪn.jo]	reg.jo [rɛx.jo]	rer.jo [rɛr.jo]
wes.ta [ʋɛs.ta]	wem.ta [ʋɛm.ta]	wes.ta [ʋɛs.ta]	wem.ta [ʋɛm.ta]
wes.li [ʋɛs.li]	wem.li [ʋɛm.li]	ning.li [nɪŋ.li]	nin.li [nɪn.li]
wes.wa [ʋɛs.ʋa]	wem.wa [ʋɛm.ʋa]	reg.wa [rɛx.ʋa]	rer.wa [rɛr.ʋa]
wes.jo [ʋɛs.jo]	wem.jo [ʋɛm.jo]	luk.jo [lʏk.jo]	lup.jo [lʏp.jo]
luk.ta [lʏk.ta]	lup.ta [lʏp.ta]	luk.ta [lʏk.ta]	lup.ta [lʏp.ta]
luk.li [lʏk.li]	lup.li [lʏp.li]	reg.li [rɛx.li]	rer.li [rɛr.li]
luk.wa [lʏk.ʋa]	lup.wa [lʏp.ʋa]	wes.wa [ʋɛs.ʋa]	wem.wa [ʋɛm.ʋa]
luk.jo [lʏk.jo]	lup.jo [lʏp.jo]	ning.jo [nɪŋ.jo]	nin.jo [nɪn.jo]
reg.ta [rɛx.ta]	rer.ta [rɛr.ta]	reg.ta [rɛx.ta]	rer.ta [rɛr.ta]
reg.li [rɛx.li]	rer.li [rɛr.li]	luk.li [lʏk.li]	lup.li [lʏp.li]
reg.wa [rɛx.ʋa]	rer.wa [rɛr.ʋa]	ning.wa [nɪŋ.ʋa]	nin.wa [nɪn.ʋa]
reg.jo [rɛx.jo]	rer.jo [rɛr.jo]	wes.jo [ʋɛs.jo]	wem.jo [ʋɛm.jo]