Syllables in Psycholinguistic Theory:  
*Now You See Them, Now You Don’t*

Niels O. Schiller

*Department of Cognitive Neuroscience, Faculty of Psychology,*

*Maastricht University*

*Maastricht, The Netherlands*

Are syllables represented in the brain? In the following, I will summarize the views that psycholinguists and neurolinguists take on this question. Presumably, it might be more appropriate to ask “How are syllables represented in the brain?” because few people have disputed the psychological or psycholinguistic reality of the syllable (but see Panconcelli-Calzia, 1934, pp. 119–120). However, syllables have been proven to be notoriously difficult to capture in linguistic terms (for overviews see Bell, A. & Hooper, 1978; Blevins, 1995; Selkirk, 1982), and it has been suggested that “syllables are just epiphenomenal consequences of the necessity of making a succession of auditorily robust modulations in one or more acoustic parameters” (Ohalia, 1998, p. 526). Both phoneticians and phonologists (but not psycholinguists) have tried to define the syllable as a universal unit in the language system but “although nearly everybody can identify syllables, almost nobody can define them” (Ladefoged, 1982, p. 220). More than 30 years ago Liberman, I. Y., Shankweiler, Fischer, F. W., and Carter (1974) conducted a study to investigate children’s sensitivity to syllables. Results showed that about half of a group of four-year old children were able to
correctly determine the number of syllables in spoken words. The same proportion of six-year old children was able to determine the number of phonemes of spoken words correctly (Liberman, I. Y. et al., 1974). This has been interpreted as evidence that syllables are the first sublexical linguistic units that appear in the course of language acquisition and that syllables are accessible earlier than segments or phonemes. Furthermore, speakers can easily manipulate syllables in meta-linguistic laboratory tasks, e.g., syllable reversal (Schiller et al., 1997; Treiman & Danis, 1990), word games (Bagemihl, 1995; Fallows, 1981; Hombert, 1986; Lefkowitz, 1991; Treiman, 1983), or backward talking (Cowan, Braine, & Leavitt, 1985; Hombert; 1973; White, 1955). For instance, in the French secret language Verlan the word verité becomes terivé. Finally, the syllable forms the basis of many linguistic rules such as syllable-initial aspiration in English or syllable-final devoicing in Dutch or German (Kenstowicz, 1994). Thus, there is plenty of evidence supporting the view that syllables play a role in the speech production process. Now you see them, ... However, syllables rarely occur in speech errors (Fromkin, 1971). ... now you don’t. If syllables were independent units, one might expect that syllabic speech errors would regularly occur, just like segmental errors (left hemisphere $\rightarrow$ left hemisphere). Therefore, it is a legitimate question to ask which role the syllable plays in psycholinguistic models of word production.

**Phonological Encoding in Speech Production**

Psycholinguistic models of phonological encoding in speech production describe the processes of word form encoding that follow the selection of a word from the mental lexicon (Levett, 1989; Levett & Wheeldon, 1994; Lavelt, Roelofs, & Meyer, 1999). Once a word has been selected from the mental lexicon, it has to be encoded in a form that can finally be used to control neuromuscular commands necessary for the execution of articulatory movements (see Guenther, 2003, for a recent overview). When accessing a word’s form for phonological
encoding, speakers retrieve segmental and metrical information. During segmental encoding, the segments of a word and their order have to be retrieved. For example, for the word typist this would be the segments /t/, /AI/, /p/, /I/, /s/, /t/. During metrical retrieval, a metrical frame has to be retrieved, i.e., the number of syllables and the location of the lexical stress. For the example typist, the metrical frame would include two syllable slots, the first of which bears lexical stress (e.g., ‘– ’). Furthermore, the syllable or consonant-vowel (CV) structure of the individual syllables of the word may be retrieved (Dell, G. S., 1988; but see Roelofs & Meyer, 1998). Once the segmental and the metrical information has been retrieved, it is combined during a process called segment-to-frame association. During this process, the previously retrieved segments are combined from word beginning to end with their corresponding metrical frame. The resulting phonological string is syllabified according to universal and language-specific syllabification rules (Roelofs, 1997). A fully prosodified phonological word is generated, which forms the basis for the activation of syllables in a mental syllabary (Levelt, W. J. M., & Wheeldon, 1994). Presumably, the units in the syllabary can be conceived of as precompiled articulatory motor programs of syllabic size. These motor programs may be represented in terms of gestural scores, i.e., a phonetic plan that specifies the relevant articulatory gestures and their relative timing (see Goldstein & Fowler, 2003, for a review). The final step includes the execution of these gestures by the articulatory apparatus resulting in overtly produced speech (see Figure 7.1).

One puzzling feature of this process is why segments and metrical frame are retrieved independently from memory when both types of information are reunified slightly later. However, this may only seem puzzling when considering single, isolated word production, but not when the production of words in context is taken into account. For instance, syllabification does not respect lexical boundaries since the domain of syllabification is the phonological word (not the lexical word; Booij, 1995). Let us take the example of the verb to type. In its citation form, type is a monosyllabic CVC word. Now consider the words typist (someone who types; dots indicate syllable boundaries), typing (the
gerund), or the phrase type it. In all of these examples, the coda /p/ of type /tap/ becomes the onset of a second syllable. In the example ty.pe it, it even straddles the lexical boundary between type and it. Therefore, it is important to bear in mind that segments are not inserted into a lexical word frame, but into a phonological word frame. The phonological word, however, is a context-dependent unit. It can solely consist of the lexical word type as in type faster or unstressed function words like it can cliticize to it as in type it faster, yielding ty.pe it /tap.itt/. A corollary of context-dependent syllabification in speech production is that it would not make sense to store syllable boundaries with the word forms in the mental lexicon because syllable boundaries will change as a function of the phonological context. The so-called syllable-position constraint observed in sound errors (i.e., onsets exchange with onsets, nuclei with nuclei, etc., Shattuck-Hufnagel, 1979) can probably not hold as an argument for stored syllable frames because it may just be a reflection of the general tendency of
segments to interact with phonemically similar segments. Therefore, it has been postulated that syllables are not stored with their lexical entries (Levelt, W. J. M. et al., 1999). Rather, syllable boundaries will be generated on-line during the construction of phonological words to yield maximally pronounceable syllables. This architecture lends maximal flexibility to the speech production system in all possible contexts.

However, there are models that assume syllables in the lexicon (Dell, G. S., 1986). In one version of Dell, G. S.’s (1988) theory, structural information about the distribution of consonants and vowels is explicitly represented in terms of so-called CV headers (see also Sevald, Dell, G. S., & Cole, 1995). However, whether or not CV structure information or abstract syllable frames are explicitly represented in the mental lexicon is still a matter of debate (Levelt, W. J. M. & Schiller, 1998; Roelofs & Meyer, 1998).

**A mental syllabary**

Although Levelt, W. J. M. et al. (1999) do not assume that syllables play a role at an abstract, early planning level in speech production syllables do play an important functional role in their theory. Levelt, W. J. M.’s theory assumes that a so-called *mental syllabary* is involved in speech production (see Figure 7.1). A mental syllabary is a “library of articulatory routines” (Crompton, 1981). The original hypothesis was that pre-compiled motor programs of syllable size could help reducing the computational load during speech production if they form the basic units of articulatory programming (for a critical note see McQueen, Dahan, & Cutler, 2003). From a lexico-statistical point of view this idea is attractive since about 85% of the speech in Dutch (and other Germanic languages such as German and English) can be produced with a relatively small set including less than 5% of all Dutch syllables (Schiller, Meyer, Baayen, & Levelt, 1996; see Figure 7.2). With less than 100 syllables, speakers of Dutch can produce 60% of their speech, and with the 500 most frequently occurring syllables this amounts even to 85% of their speech.
Cumulative syllable frequency

Dutch, rank 1-500

Figure 7.2.
Lexico-statistical distribution of Dutch syllables. Syllable rank is shown from position 1 to 500 on the x-axis; cumulative syllable proportion on the y-axis.

Levett and Wheeldon (1994) tested the hypothesis of a mental syllabary in an experiment comparing the production latencies of words differing in syllable frequencies. For instance, there were words in the experiment that consisted of high-frequency syllables (e.g., bo.te.r “butter”) and words that were made-up from low-frequency syllables (e.g., gi.ra.ff “giraffe”) while word frequency was controlled. Results showed that words with high-frequency syllables were named significantly faster than words with low-frequency syllables, independent of word frequency. Levett and Wheeldon (1994) took this finding as evidence for a separate store from which syllabic units can be recruited during speech production. Now you see them, ... Previous evidence, however, indicated that articulation only starts after the whole phonological word has been encoded (Wheeldon & Lahiri, 1997). Thus, the frequency of the first syllable may only have a limited impact because the response to a bisyllabic word can only start after the second syllable has been retrieved. Therefore, Levett and Wheeldon manipulated the frequency of the second syllable while controlling the frequency of the first and found that the frequency of the second
syllable was crucial. However, syllable frequency correlated highly with segment or phoneme frequency. Therefore, Leevet, and Wheeldon’s (1994) syllable frequency effect could as well be attributed to segment frequency. When segment frequency was controlled in subsequent experiments, a small set of awkward word stimuli remained and the syllable frequency effect disappeared. ... now you don’t.

Recently several investigators looked at syllable frequency from different perspectives. Wilshire and Nespoulos (2003) looked at the influence of syllable frequency on phonological errors produced by two French-speaking aphasic patients. The patients were asked to read aloud and repeat 110 polysyllabic word pairs matched for word frequency and syllabic complexity, but differing in word-final syllable frequency. The authors found that performance was not more accurate with words consisting of high-frequency than words consisting of low-frequency syllables. Wilshire and Nespoulos concluded that a “plausible interpretation of the absence of syllable frequency effects [...] is that this variable has no influence on the process of phonological encoding” (Wilshire & Nespoulos, 2003, p. 441). Moreover, they state that their failure to obtain a syllable frequency effect indicates that syllables are not retrieved as representational units nor accessed at the level of phonological encoding.

Similarly, Aichert and Ziegler (2004) investigated the influence of syllable frequency on speech production of ten German patients with apraxia of speech (AOS) using a word repetition task. The materials consisted of 40 bisyllabic words, matched for word frequency and syllabic complexity, but differing in initial syllable frequency. Unlike the previous study, these authors found a tendency towards more segmental errors on words made up of low-frequency syllables. This effect became significant when the most extreme frequency regions were analyzed separately. Aichert and Ziegler (2004) concluded that their findings “indicate that patients with AOS must have access to the syllabary” (Aichert & Ziegler, 2004, p. 153). It seems that patients with an apraxic impairment of speech have access to a mental syllabary, but either the
retrieval of the syllabic motor programs or the motor programs themselves are defective.

Recently, Cholin, Levelt & Schiller (2006) conducted a naming study with Dutch speakers using pseudo-words. The reason to employ pseudo-words was that segment frequency could not be controlled when using existing words (see above). All pseudo-words consisted of existing Dutch syllables. In the first experiment, monosyllabic pseudo-words were named in a newly designed symbol-association-naming task. In this task, participants were required to learn to associate a particular pseudo-word either with a left or a right position on the screen. Then a symbol would appear on the screen, and depending on its position (left or right) the corresponding pseudo-word was to be produced. The authors observed a small but significant syllable-frequency effect: pseudo-words consisting of a high-frequency syllables were produced faster than pseudo-words consisting of a low-frequency syllables (Cholin et al., 2006). In a second experiment, the second syllable of bisyllabic pseudo-words was frequency-manipulated, and the naming latencies between pseudo-words including high- or low-frequency second syllables did not differ. Finally, in the third experiment, the first syllable of bisyllabic pseudo-words varied in frequency. A syllable-frequency effect emerged similar to the one in the first experiment. These results strongly support the notion of a mental syllabary to mediate between abstract phonological syllables and phonetic syllables, conceived of as precompiled gestural scores controlling execution of an articulatory motor program. Cholin et al. (2006) concluded that speakers can start articulating a target word when the first syllable is phonetically encoded.

Syllable Priming

The idea of a mental syllabary stimulated even more research. Ferrand, Segui and Grainger (1996) carried out a series of syllable priming experiments and reported a syllable priming effect in French speech production. Now you see them, ... These authors used the masked priming procedure and found that a
Figure 7.3.
Results of Experiment 4 from Ferrand et al. (1996). Structure of the visual prime (CV or CVC) is shaded. Target picture categories are depicted. On the left are CV words such as ca.rotte. On the right are CVC words such as car.table. On the y-axis are mean reaction times. Naming latencies are in milliseconds on the x-axis.

of car.table. Similarly, the prime car primed the naming of car.table better than the naming of ca.rotte (see Figure 7.3).

Ferrand et al. (1996) concluded that output phonology must be syllabically structured since the syllable priming effect disappeared in a task that does not make a phonological representation necessary, such as a lexical decision task. This fit nicely with the evidence about syllables as pre-lexical processing units in French speech comprehension (see below). Ferrand et al. (1996) claimed that the mental syllabary model predicts a syllable-priming effect. This account presumed that the prime could pre-activate syllables in the syllabary. Interestingly, Ferrand, Segui, and Humphreys (1997) also report a syllable priming effect for English, although the evidence for syllables as processing units in speech comprehension in English is rather weak (Cutler, Mehler, Norris, & Segui, 1983; Cutler, 1997). Recently, this syllabic effect in English has been replicated in a silent word reading study by Ashby and Rayner (2004) using a parafoveal preview technique although Ferrand et al. (1996, 1997) claimed that the speech-output phonology syllabically structured and did not
find a syllabic effect in lexical decision tasks. It should be noted, however, that when using a different technique, i.e., fast priming, no syllabic effect could be demonstrated in reading (Ashby & Rayner, 2004).

However, when Schiller (1998) tried to replicate these syllable priming effects in Dutch speech production, he failed to find syllabic effects. ... now you don't. Instead, what he obtained was a clear segmental overlap effect, i.e., the more segmental overlap between prime and target picture name, the faster the naming latencies (see also Schiller, 2004). That is, the prime *kan* yielded not only faster responses than *ka* for the picture of a “pulpit” (*kan.sel*) but also for the picture of a “canoe” (*ka.no*) (see Figure 7.4).

Similar results were obtained in the auditory modality, i.e., presenting either /ro/ or /rok/ when Dutch participants were requested to produce either
segmental overlap effect was obtained, i.e., /rok/ was a better prime than /ro/ independent of the target (Baumann, 1995). The failure to find a syllable priming effect in Dutch is in agreement with the statement that syllables are never retrieved during phonological encoding (Levelt, et al., 1999). The syllable priming effect found by Ferrand and colleagues (1996) in French can be accounted for by assuming that segments in the prime are coded with their corresponding syllable structure information. For instance, the prime *pal* pre-activates segments specified for syllable position in the perceptual network, e.g., \( p_{onsets} \), \( a_{nucleus} \), and \( l_{coda} \). Active phonological segments in the perceptual network can directly affect the corresponding segment nodes in the production lexicon. Therefore, the prime *pal* matches with the target *pal.mier*, but not with *pa.lace* because the /l/ in *pal* is specified for coda and not for onset.

This fits well with the assumption that the syllable is an important pre-lexical processing unit in French. A classical study by Mehler, Dommergues, Frauenfelder, and Segui (1981) showed that when French native speakers are requested to detect a particular target syllable (e.g., *pal*) in an auditorily presented word, they are faster in responding when the target syllable corresponds to the first syllable of the word (e.g., *pal.mier*) than when it does not (e.g., *pa.lace*). Similarly, when the target is a CV syllable (e.g., *pa*), participants are faster in responding to *pa.lace* than *pal.mier* (see Figure 7.5). Mehler and colleagues concluded that the syllable is a pre-lexical processing unit in speech perception. More specifically, syllables may be useful in the segmentation of the incoming speech signal. Now you see them, ...
Figure 7.5.
Results of Experiment 1 from Mehler et al. (1981). Structure of the carrier words (CV such as ba.lade or CVC such as bal.con) is color-coded. Target syllables are depicted (left: CV such as ba; right: CVC such as bal) on the x-axis; Mean reaction times (monitoring latencies) in milliseconds on the y-axis.

This syllable match effect has been replicated in many other Romance languages, but not in Germanic languages (see Cutler, 1997 for a summary). Subsequent research showed that at least the following two factors play a crucial role for obtaining a syllable match effect, i.e., the rhythmic structure of the language and the quality of the pivotal consonant, which follows or precedes the syllable boundary. Germanic languages are stress-timed. That may be the reason that the syllable match effect could not be replicated for these languages. Syllables simply do not play a prominent role for the rhythmic pattern in those languages and therefore they are not used for pre-lexical segmentation; instead, stress is used to segment the incoming speech stream. In contrast, Romance languages have more or less clear syllable boundaries (syllable-timed languages), and therefore syllable match effects can be obtained in those languages. However, it has been shown that the pattern of effects also depends on the
nature of the pivotal consonant, which can interact with syllable structure. Content, Meunier, Kearns, and Frauenfelder (2001) assessed the role of the characteristics of the pivotal consonant for the syllable priming effect with pseudo-words. They only found the syllabic effect for liquids (e.g., /l/, /r/, etc.) in a blocked condition, but not for stops (e.g., /l/, /k/, /p/, etc.). Liquids are known to show a higher degree of coarticulation with preceding vowels than stops, and listeners might be able to exploit this when detecting syllables because when they perceive the vowel /a/ in *pal.mier*, they know relatively earlier whether or not a syllable boundary intervenes between the /a/ and the /l/. Therefore, they can detect the syllable *pal* or reject the syllable *pa* relatively quickly. Rietveld and Schiller (submitted) recently obtained similar results for Dutch. The Content et al. (2001) as well as the Rietveld and Schiller (submitted) results argue against a role for the syllable as a pre-lexical segmentation unit and more for a direct phonemic input processing. ... *now you don’t.*

The segmental overlap effect is not restricted to Dutch. When Schiller (2000) tried to replicate the Ferrand et al. (1997) results for English with better-controlled material, no syllabic effect was obtained. A segmental overlap effect was found. These English data are interesting. In English there is phonological equivalence between corresponding syllable structures. For example, *pi* /pat/ matches phonologically the first syllable in *pi.lot* but not in *pi[l.l]ow* (square brackets indicate ambisyllabicity of the intervocalic consonants), and *pi.l* /pi/ matches phonologically the first syllable in *pi[l.l]ow* but not in *pi.lot*. Nevertheless, the prime *pi.l* yielded faster responses than *pi* for both *pi.lot* and *pi.low* (see Figure 7.6).

Either the contribution of vowels is less important in segmental priming (Schiller & Costa, submitted) or consonants and vowels have different time courses of activation (Berent & Perfetti, 1995), consonants being faster than vowels and therefore more effective. Further testing revealed no syllable effect in Spanish, but a small segmental overlap effect (Schiller, Costa, & Colomé, 2002), and no syllabic effect in French when a larger set of materials
Figure 7.6. Overview of the results of Experiment 2 from Schiller (2000). The structure of the visual prime (CV, CVC, or control, i.e. non-linguistic characters) is color-coded. Target picture categories are depicted (left: CV words such as *pi*lot; middle: CVC words such as *pic*nic; right: CV[C] words such as *pi[l]*low; square brackets indicate ambisyllabic consonants on the x-axis (the); mean reaction times (naming latencies) in milliseconds on the y-axis.

was tested (Schiller et al., 2002). Taken together, these results support the idea that syllables are not retrieved, but created on-line during phonological encoding.

Although syllables cannot be primed in Dutch, Cholin, Schiller, and Levelt, W. J. M. (2004) found that syllable structure can be prepared in the planning of speech production. These authors hypothesized that masked priming taps a planning level in speech production that is not sensitive to syllabic units. Priming speeds up the spell-out of information from memory, such as segments. However, since syllables are presumably not stored with the word form because syllable boundaries are computed on the fly, syllables cannot be primed. Therefore, Cholin et al. (2004) used a preparation paradigm to investigate whether or not syllables are constructed as part of the phonological word and used to access the mental syllabary. The preparation paradigm was introduced
W. J. M. (2002). Participants learned two different categories of item sets, i.e., so-called constant and variable sets. Both sets contained words with initial segmental overlap (e.g., spui.en “to drain”; spui.de “drained”; spui.er “person who drains”; spui.end “draining”), but in the variable set, syllable structure varied (e.g., ro.ken “to smoke”; rook.te “smoked”; ro.ker “smoker”; ro.kend “smoking”). The past tense form rookte is called the odd-man-out because its first syllable has the structure CVVC, while all other syllables are CVV. Therefore, in the constant sets, participants can prepare not only the initial segments but also the syllable structure of the first syllable. In the variable sets, in contrary, it was assumed that the odd-man-out would spoil this preparation effect.

Apart from naming the words in a four-item-set, i.e., including four different words (see above), participants named them in a three-item-set which consisted of the same items except for the past tense form. Thus, the three-item-sets were constant for both types of stimuli (e.g., spui.en, spui.er, spui.end vs. ro.ken, ro.ker, ro.kend). Since different item sets cannot be compared directly, the mean naming latencies of all three-item-sets were subtracted from the corresponding four-item-sets and the difference scores were compared. We expected that the four-item-set was relatively more difficult to name in the variable than in the constant condition, even when removing the odd-man-out from the data before conducting the analyses. The reason for this hypothesized outcome is that the syllable structure could be prepared, i.e., the syllabification process can begin to incrementally put the initial segments together and create the first syllable, in the constant but not in the variable sets. In variable sets, the preparation cannot go beyond the retrieval of the initial segments. This is exactly what was found in two experiments (Cholin et al., 2004).

Possibly, the first syllable can be fully prepared for articulation in the constant sets, including the retrieval of the corresponding gestural score from the mental syllabary. In the constant sets all stages prior to articulation, including segmental spell-out, on-line syllabification, and possibly access to the mental syllabary, might contribute to the preparation effect, whereas in variable sets
responses can only be prepared up to on-line syllabification. Thus, Cholin et al. (2004) concluded that syllables are probably encoded at the interface between phonological and phonetic encoding.

**Metrical Encoding**

Syllables may not only be important for segmental encoding, i.e., grouping segments into larger units (syllable priming and mental syllabary), but also for metrical encoding, i.e., encoding of prosodic features of words such as metrical stress. Probably the most important function of syllables is to bear stress. Metrical stress is a suprasegmental feature that is usually not realized on a single segment but on a number of segments that belong together, i.e., a syllable.

Roelofs and Meyer (1998) investigated how much information about the metrical structure of words is stored in memory. Possible candidates are lexical stress, number of syllables, and syllable structure. In one experiment, for instance, they compared the production latencies for sets of homogeneous bisyllabic words such as *ma.NIER* ("manner"; capital letters indicate stressed syllables), *ma.TRAS* ("mattress"), and *ma.KREEL* ("mackerel") with sets including words with a variable number of syllables such as *ma.JOOR* ("major"), *ma.TE.rie* ("matter"), and *ma.LA.ri.a* ("malaria"). Lexical stress was kept constant (always on the second syllable). Relative to a heterogeneous control condition, there was strong and reliable facilitation for the bisyllabic sets, but not for the sets with a variable number of syllables. This showed that the number of syllables of a word must be known to the phonological encoding system. Hence, this information must be part of the metrical representation of words.

Similarly, the production of sets of homogeneous trisyllabic words with constant stress (e.g., *ma.RI.ne* "navy," *ma.TE.rie* "matter," *ma.LAI.se" depression," *ma.DON.na" madonna") and variable stress (e.g., *ma.RI.ne* "navy," *ma.nus.CRIPT* "manuscript," *ma.TE.rie* "matter," *ma.de.LIEF" daisy") was measured and compared to the corresponding heterogeneous sets. Again, facilitation was obtained for the constant but not for the variable sets. Therefore,
planning of polysyllabic words – at least when stress is in non-default position (see Schiller, Fikkert, & Levelt, W. J. M., 2004 for an alternative). However, CV structure did not yield an effect. When the production latencies for words with a constant CV structure (e.g., *bres* “breach,” *bril* “glasses,” *brok* “piece,” *brug* “bridge”; all CCVC) were compared to words with a variable CV structure (e.g., *brij* “porridge”; CCVV, *brief* “letter”; CCVVC, *bron* “source”; CCVC, *brand* “fire”; CCVCC), relative to the corresponding heterogeneous condition, no difference was found suggesting that the metrical structure speakers retrieve does not contain information about the CV or syllable structure of a word (but see Costa & Sebastián-Gallés, 1998).

To investigate the time course of metrical processing, Schiller, Jansma, Peters, and Levelt, (2005) employed a tacit naming task and asked participants to decide whether the bisyllabic name of a visually presented picture had initial or final stress. Their hypothesis was that if metrical encoding is a parallel process, then there should not be any differences between the decision latencies for initial and final stress. If, however, metrical encoding is a rightward incremental process—just like segmental encoding (Meyer, 1991; Schiller, 2005; van Turennout, Hagoort, & Brown, 1997; Wheeldon & Levelt, 1995)—then decisions to picture names with initial stress should be faster than decision latencies to picture names with final stress. The latter turned out to be the case (Schiller et al., 2005). However, Dutch—like other Germanic languages—has a strong preference for initial stress. More than 90% of the words occurring in Dutch have stress on the first syllable. Therefore, this stress effect might have been due to a default strategy. However, when pictures with trisyllabic names were tested, participants were still faster to decide that a picture name had penultimate stress (e.g., *asPERge* “asparagus”) than that it had ultimate stress (e.g., *artiSJOK* “artichoke”; Schiller et al., 2005). This result suggests that metrical encoding proceeds from the beginning to the end of words, just like segmental encoding (see Figure 7.7).
Figure 7.7.
Main results of Experiments 1 and 2 from Schiller et al. (2005). Symbols indicate different experiments. Target picture categories are depicted (left: first syllable stress such as *to.ren*; middle: second syllable stress such as *to.maat* or *as.per.ge*; right: third syllable stress such as *ar.ti.sjok*) on the x-axis; mean reaction times (monitoring latencies) in milliseconds on the y-axis.

Summary and Conclusion

This chapter summarized the role of the syllable in psycholinguistic models of speech production. A detailed model of phonological encoding was described. Evidence in favor of a model where syllables come into play at a relatively late stage in phonological encoding when syllable-sized units are accessed in a mental syllabary was presented. It was argued on the basis of reaction time data that syllables play a role in production planning. Furthermore, the nature of metrical frames was described and it was argued that lexical stress (realized on syllables) is encoded rightward incrementally. It is concluded that syllables fulfill their functional role as stress-bearing units at a late stage in speech production planning, i.e., when the mental syllabary is accessed. However, more research on phonological encoding and the syllabary is necessary to further specify aspects of syllabic representation in the syllabary.
References


Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language,


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


