

Research report

# Tracking the time course of phonological encoding in speech production: an event-related brain potential study

Niels O. Schiller<sup>a,b,\*</sup>, Mart Bles<sup>a</sup>, Bernadette M. Jansma<sup>a</sup>

<sup>a</sup>*Department of Neurocognition, Faculty of Psychology, University of Maastricht, Maastricht, The Netherlands*

<sup>b</sup>*Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands*

Accepted 15 August 2003

## Abstract

This study investigated the time course of phonological encoding during speech production planning. Previous research has shown that conceptual/semantic information precedes syntactic information in the planning of speech production and that syntactic information is available earlier than phonological information. Here, we studied the relative time courses of the two different processes within phonological encoding, i.e. metrical encoding and syllabification. According to one prominent theory of language production, metrical encoding involves the retrieval of the stress pattern of a word, while syllabification is carried out to construct the syllabic structure of a word. However, the relative timing of these two processes is underspecified in the theory. We employed an implicit picture naming task and recorded event-related brain potentials to obtain fine-grained temporal information about metrical encoding and syllabification. Results revealed that both tasks generated effects that fall within the time window of phonological encoding. However, there was no timing difference between the two effects, suggesting that they occur approximately at the same time.

© 2003 Elsevier B.V. All rights reserved.

*Theme:* Cognitive neuroscience

*Topic:* Language

*Keywords:* Psycholinguistics; Speech production; ERPs; Phonological encoding; Metrical structure; Syllabification; Internal monitoring

## 1. Introduction

The speech production process can be divided into several planning stages, such as conceptual, semantic, syntactic, and phonological encoding [24,25] (see Fig. 1). One central question in psycholinguistic research is the time course of these processes, i.e. which processes precede or follow other processes and how long do the processes approximately take to be completed. Levelt's theory of speech production makes explicit claims about the time course of these processes based on chronometric

evidence (see overview in Refs. [17,27]). For instance, by manipulating stimulus onset asynchronies (SOAs), Schriefers et al. [47] showed that semantically related prime words influenced the naming latencies of target pictures at an earlier point in time than phonologically related prime words (see also Ref. [6]). This led to the conclusion that semantic processing precedes phonological processing during speech production. Van Turenhout and colleagues were the first to test these claims with electrophysiological methods [50,51]. Using lateralized readiness potentials (LRPs) they were able to show that the processing of semantic information precedes the processing of phonological information by between 40 and 120 ms [50] when the initial and final phonemes of words with a mean length of 1.5 syllables were considered and that phonological processing follows syntactic processing by about 40 ms [51].

More recently, Schmitt and colleagues used another event-related potential (ERP) component to track the time

\*Corresponding author. Department of Cognitive Neuroscience, Faculty of Psychology, Universiteit Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Tel.: +31-43-388-4041; fax: +31-43-388-4125.

*E-mail addresses:* [n.schiller@psychology.unimaas.nl](mailto:n.schiller@psychology.unimaas.nl) (N.O. Schiller), <http://www.mpi.nl/world/persons/profession/schiller.html> (N.O. Schiller).

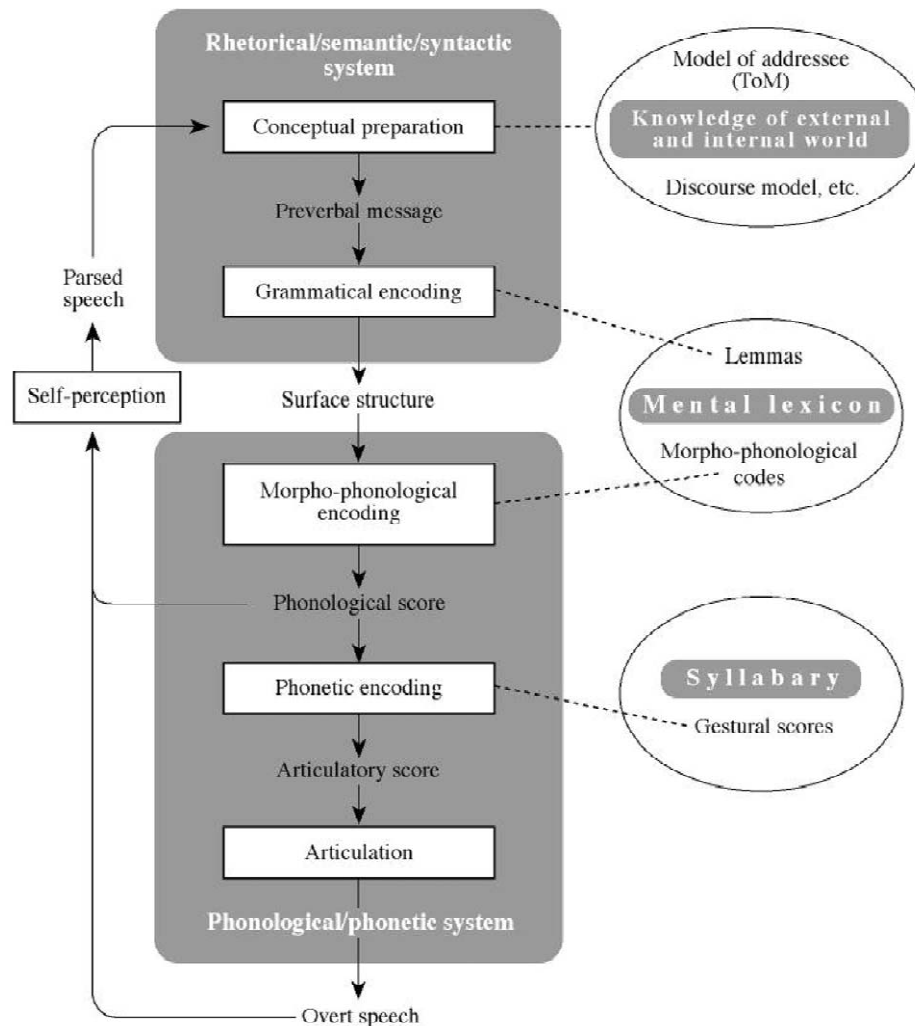


Fig. 1. Levelt's model of serial processing in speech production. The two gray boxes display the individual processing components (in white rectangles) and their output. The ellipses on the right display long-term memory components accessed during speech production, while the arrows on the left indicate the internal and the external loop for self-monitoring.

course of processing stages during speech production, namely the so-called N200 (see below). Schmitt et al. [44] showed that the peak latency of the N200 effect was 89 ms earlier when the decision process leading to the effect could be made on the basis of semantic information than when it was made on the basis of phonological information (see also Ref. [32]). This result replicated Van Turenout et al.'s [50] earlier LRP findings. Furthermore, Schmitt et al. [46] investigated the time course of conceptual and syntactic encoding during picture naming and found that conceptual information evoked an earlier N200 effect than syntactic information (by 73 ms). Finally, Schmitt et al. [45] estimated the time from semantic to syntactic encoding to be approximately 80 ms. Therefore, electrophysiological measurements have replicated earlier reaction time (RT) studies and extended those by providing fine-grained estimates of the temporal relationships between the processes involved in speech production. So far, we know that conceptual/semantic processing precedes syntactic pro-

cessing, which precedes phonological processing during tacit picture naming.

In this study, we try to track the time course of more specific processes within the phonological encoding module. Phonological encoding comprises a set of individual cognitive retrieval and encoding processes, which are involved in word form encoding during speech production. In Levelt's model, the most explicit model of phonological encoding to date, word form retrieval is divided into metrical spell-out and segmental spell-out (see Fig. 2). During segmental spell-out, the individual phonemes of a word and their ordering are retrieved. The number of syllables and the location of lexical stress form part of the information being retrieved during metrical spell-out, at least for words with irregular stress [27]. The stress pattern for regular words is presumably computed by means of a default rule (see also Ref. [34]; for a different perspective, see Ref. [41]). In a process called segment-to-frame association, segmental and metrical information is com-

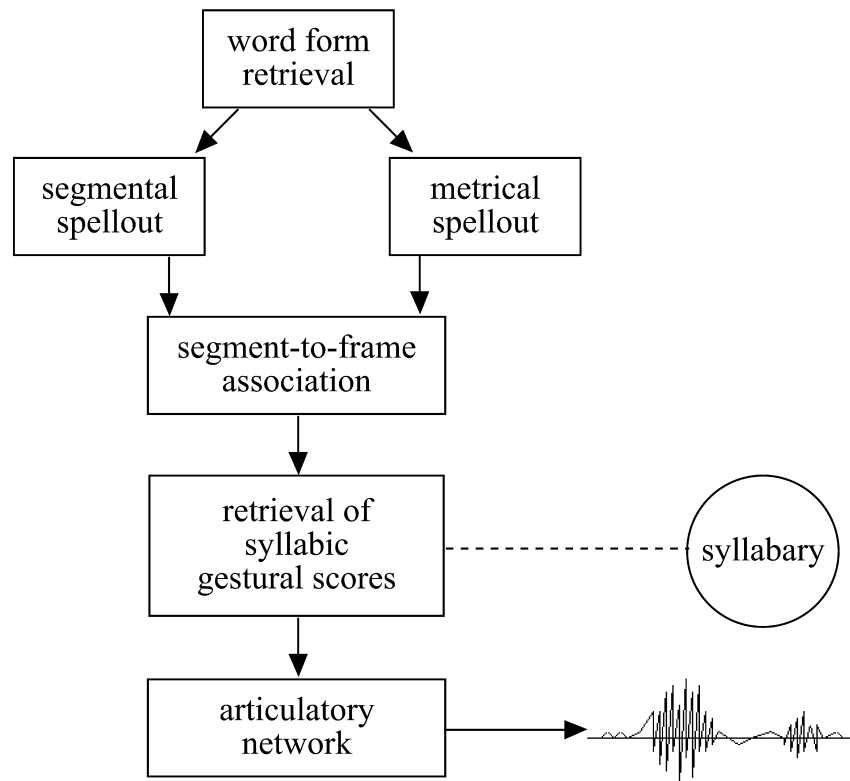


Fig. 2. A model of phonological encoding in speech production (after Levelt and Wheeldon [28]). The individual processing components are again displayed in rectangles, while the circle symbolizes a long-term memory component. The overt speech is indicated by the schematized acoustic waveform.

bined into a phonological word. During phonological word formation the previously retrieved segments are syllabified according to universal and language-specific syllabification rules (see Refs. [43,52] for overviews). The resulting phonological syllables are used to activate phonetic syllables in a so-called mental syllabary [28]. These phonetic syllables are sufficiently specified to control articulatory-motor movements necessary for articulation.

From Wheeldon and Levelt's [53] and Meyer's [29,30] work, we can assume that the segments and syllables of a word are encoded one-by-one in a rightward incremental fashion. Using a preparation paradigm, Meyer [29,30] had participants produce target words in response to a prompt word and found that RTs were faster when the beginning of the target words could be planned in advance. This preparation effect varied with the length of the string that could be planned: the longer the string that could be prepared the faster the RTs. However, there was no preparation effect when participants could only prepare the final part of the target words, suggesting that phonological words be planned from beginning to end.

Wheeldon and Levelt [53] provided additional evidence for the incremental nature of phonological encoding. In one experiment, they had bilingual participants generate internally Dutch translations to English prompt words. However, participants did not overtly produce the Dutch words but self-monitored them internally for previously specified segments. For example, participants would hear

the English prompt word *hitch hiker* and were required to press a button if the Dutch translation (*lifter*) contained the phoneme /t/. Thus, in the case of *hitch hiker (lifter)*, participants would press the button, but in the case of *cream cheese (roomkaas)* they would not. Results showed that button press latencies were dependent on the position of the pre-specified target segment in the translation word. Participants were faster to decide that the Dutch translation word contained a /t/ when the English word was *garden wall (tuinmuur)* than when it was *hitch hiker (lifter)* or *napkin (servet)*. The earlier the target segment occurred in the Dutch word, the shorter the decision latencies. Wheeldon and Levelt [53] interpreted their data to support the claim of rightward incremental phonological encoding. Phonological encoding is a strictly serial process that runs from the beginning to the end of words. The effect was located at the phonological word level, i.e. when segments and metrical frames are combined. Furthermore, these authors observed a significant increase in monitoring times when two segments were separated by a syllable boundary. Wheeldon and Levelt [53] suggested that the monitoring difference between the target segments at the syllable boundary (e.g., *fiet-ser* vs. *lif-ter*) might be due to the existence of a marked syllable boundary or a syllabification process that slows down the encoding of the second syllable.

Syllables are chunks of segments forming minimal articulatory units [5,12]. Ferrand and colleagues claimed

that syllables are functional units of the output phonology in French and English [10,11]. One important feature of Levelt's model of phonological encoding is that syllables are not specified in the lexicon but instead generated 'on the fly' during segment-to-frame association [26,28,33] (but see Refs. [7,8] for a different view). Therefore, there should be no syllable priming in speech production since no stored phonological sequence can be pre-activated. Cross-linguistic evidence from several languages—including French and English—showed that this is in fact the case: syllables cannot be primed in speech production [37–40]. However, recent work by Cholin and colleagues using the preparation paradigm showed that syllables could be prepared, as predicted by Levelt's theory [4].

In a more recent study, Schiller et al. [42] investigated the time course of metrical encoding, i.e. stress. In a first experiment, participants were presented with pictures that had initial or final stress (*KAno* 'canoe' vs. *kaNON* 'cannon'; capital letters indicate stressed syllables). Picture names were bisyllabic and matched for frequency and object recognition latencies. A picture naming experiment revealed that the pictures with final stress were named marginally faster than picture names with initial stress. More interestingly, however, in a monitoring experiment the reversed pattern emerged: using an implicit picture-naming task, participants were asked to judge the stress position of the picture names. Participants saw the same pictures as in the naming experiment on a computer screen and decided for each picture whether its name had initial or final stress without overtly naming the pictures. Results showed significantly faster decision times for initially stressed targets than for targets with final stress. This effect was replicated with trisyllabic picture names (faster RTs for penultimate stress than for ultimate stress). These results reflect the incremental nature of the metrical encoding process, i.e. stress is also encoded from the beginning to the end of words. In a related study, Jansma and Schiller [18] investigated the monitoring of segments and found that the same phoneme /n/ in the same absolute position in a word is monitored significantly faster when it occurs before a syllable boundary (as in *kan-sel*) than when it occurs after a syllable boundary (as in *ka-no*), supporting Wheeldon and Levelt's [53] claim that inserting a syllable boundary takes time (see above).

Here, we investigate the time course of metrical encoding and syllabification with ERPs. According to Levelt's model [27], metrical information has to be retrieved (or computed) before segment-to-frame association, i.e. the level at which a word is syllabified (see Fig. 2). A series of time-course studies (see Refs. [17,25] for reviews) indicated that phonological encoding takes place in a time window between 275 and 450 ms after picture onset. We, therefore, expect any metrical and syllabic effect for tacit picture name encoding in this simple go/nogo task to occur in this time window. And, more specifically, if participants have access to metrical information at an early

level of metrical retrieval, peak latencies of the N200 related to metrical encoding should occur earlier than peak latencies of the N200 related to syllabification. However, there is also another possibility. If participants are only able to have access to phonological information like stress at the phonological word level (after segment-to-frame association), then the effects in the ERP should be visible later, i.e. after phonological encoding has taken place; that is, more than 450 ms after picture onset. In that case, we would be tapping internal self-monitoring of the phonological word and ERP effects related to metrical retrieval may either precede ERP effects related to syllabification or the two effects may occur at the same time. If, however, our experiment picks up both processes, i.e. the picture naming process and the self-monitoring process, we should be able to see early (between 275 and 450 ms) and late effects (after 450 ms).

### 1.1. The N200

The N200 is a negative-going deflection of the ERP waveform. When a participant in a go/nogo paradigm is asked to respond to one class of stimuli (*go* trials), e.g. by pressing a button, and not to respond to another class of stimuli (*nogo* trials), the ERP on *nogo* trials is characterized by a large negativity (1–4  $\mu$ V) compared to *go* trials between 100 and 300 ms after stimulus onset (N200). The N200 effect is especially marked over fronto-central electrode sites [13,21,31,36,48,49]. It has been suggested that the magnitude of the N200 effect is a function of the neural activity required for response inhibition [19,35].

The presence of an N200 can be used as an indicator that the information necessary to determine whether or not to respond must have been available. One can manipulate the information on which a go/nogo decision is based and use the peak latency of the N200 effect (difference between go and nogo ERPs) as an upper estimate of when in time the specific information must have been encoded. In the present study, participants were to make a binary decision, i.e. classify picture names according to their lexical stress (does the target have initial or final stress?) or the syllable affiliation of their first post-vocalic consonant (does this consonant belong to the first or second syllable?).

### 1.2. The experimental paradigm

We measured the N200 in a go/nogo paradigm to determine the time course of phonological encoding. More specifically, we looked at whether metrical encoding precedes syllabification as predicted by the general architecture of the model of Levelt et al. [27] or whether the encoding of stress coincides with the encoding of syllable boundaries. Note that one alternative model of phonological encoding [7,8] assumes that the syllable structure of a

word is stored in the lexicon, but Dell’s model is silent about metrical encoding. Therefore, it is difficult to state a precise hypothesis about metrical encoding based on Dell’s model.

The experiment was carried out in Dutch. Participants were required to name internally a set of pictures and then carry out a binary decision task, i.e. classifying the picture names with respect to their metrical or their syllabification properties. The metrical task involved a decision about the location of the lexical stress of the target word. In the Dutch lexicon, lexical stress is not fixed—in principle, it can fall on every syllable with a full vowel (i.e. not a schwa). The Dutch stress system is a mixture of a Germanic initial stress pattern, a French final stress pattern, and a Latin penultimate stress pattern [3]. However, there is a strong bias towards initial stress in Dutch. More than 90% of the word form tokens have stress on the first syllable containing a full vowel [26]. In our experiment, participants were required to decide whether a bisyllabic picture name had initial (e.g., *LEpel* ‘spoon’) or final stress (e.g., *liBEL* ‘dragonfly’) (see Fig. 3). Note that information about the location of lexical stress is not reliably represented in the orthographic form of words in Dutch. Participants need to generate the phonological output form of the target in order to be able to make the correct decision. The syllabification decision also involved generation of the phonological code of the target. The participants’ task was to decide whether the first post-vocalic consonant belonged to the first (e.g., *kaN.sel* ‘pulpit’) or second syllable (e.g., *ka.No* ‘canoe’; syllable boundaries are indicated by dots and pivotal consonants are in upper case) of the target (see Fig. 3). To be able to make this

decision, participants have to access the syllabification of the target word. According to Levelt’s model, this information becomes available at the phonological word level, i.e. when segmental and metrical information is combined yielding a phonological word (see above).

The logic of the paradigm with regard to the N200 is as follows. In the metrical condition, the key press is contingent on metrical information while in the syllabification condition, the response is contingent on information about the syllable affiliation of a particular consonant. The timing of the N200 effect (i.e. the difference between go and nogo responses) provides an upper limit of the moment in time when the respective information must be available for determining whether or not to respond. According to Levelt’s model of phonological encoding [25,27], metrical information is available before information about the syllabification or both types of information are available at the same time. Therefore, the information to inhibit a metrical response should never be available later than the syllabification response. We would expect to see a potential difference in availability in an earlier N200 when the go/nogo decision is based on metrical information than when it is based on syllabic information. If, however, both types of information become available at the same time, then we would expect to find no difference in the timing of their associated N200.

A potential concern about the go/nogo task used here to probe speech production processes is that during the recording session participants merely responded to the pictures with a key press instead of naming the pictures aloud. But, as the key press responses are contingent on phonological information of the target picture name, they

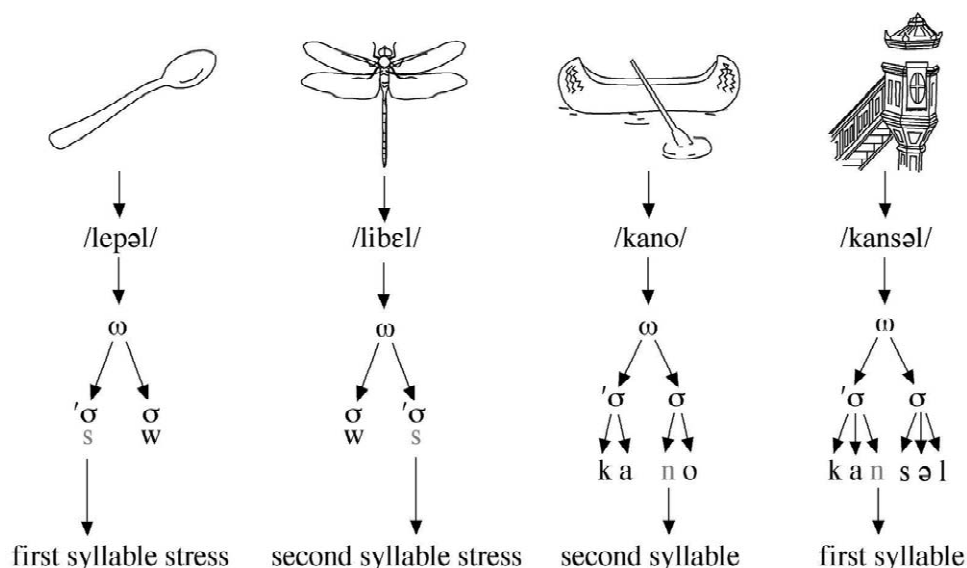


Fig. 3. Examples of the four conditions in the experiment. The spoon on the left has initial stress in Dutch (*lepel*), while dragonfly has final stress in Dutch (*libel*). Similarly, the first post-vocalic consonant in pulpit belongs to the first syllable in Dutch (*kansel*), while it is in the first syllable in canoe in Dutch (*kano*).

must necessarily probe the availability of these two kinds of information. In order to have this information available, the subjects have to silently generate the name. We now assume that the silent or tacit generation of a name is similar to that of overt production.

## 2. Materials and method

### 2.1. Participants

Twenty-seven native speakers of Dutch took part in the experiment. All participants but one were right handed. All had normal or corrected-to-normal vision. Participants were paid for their participation in the experiment. They were informed that they would take part in an ERP study on picture naming and gave written consent.

### 2.2. Materials

A set of 96 simple white-on-black line drawings was used as target pictures. All items corresponded to monomorphemic, bisyllabic Dutch nouns. They were taken from the picture database at the Max Planck Institute for Psycholinguistics in Nijmegen. The two factors being manipulated in the experiment, i.e. stress and syllable affiliation, were completely crossed, resulting in four categories of picture names: (1) picture names with initial stress and initial syllable affiliation (e.g., *kan.sel* ‘pulpit’), (2) picture names with initial stress and second syllable affiliation (e.g., *ka.no* ‘canoe’), (3) picture names with final stress and initial syllable affiliation (e.g., *kal.koen* ‘turkey’), and (4) picture names with final stress and second syllable affiliation (e.g., *ka.non* ‘cannon’) (see Appendix A for the whole list of items). All items were between four and seven segments (phonemes) long and the item categories had a mean frequency of occurrence between 15 and 32 per million as determined by CELEX (see Ref. [1]), i.e. all item categories were of moderate frequency (for details, see Table 1)<sup>1</sup>.

### 2.3. Design

Each participant received four different instruction sets, i.e. conditions, altogether. In two instruction sets, the key press response (go/nogo) was contingent on metrical information; in the other two sets, the response was

Table 1

Lexico-statistical characteristics of the target words

Stress location	CV structure of the first syllable	Example	Mean CELEX frequency (per one million words)	Mean length in segments
Initial	CV	<i>kano</i>	31.1	5.1
Initial	CVC	<i>kansel</i>	19.5	6.2
Final	CV	<i>kanon</i>	19.0	5.0
Final	CVC	<i>kalkoen</i>	15.6	6.2

Note. The mean CELEX frequency for the CV items with initial stress (31.1 per one million words) is slightly higher than for the other three categories because one item, i.e. *tafel* ‘table’, has a frequency of 247.4 per one million words, by far the highest frequency of all items. Discarding the item *tafel*, this category has a mean frequency of 21.7.

contingent on the syllabic information. Each picture was presented four times to each participant, i.e. once per condition. The order of conditions was counter-balanced across participants.

### 2.4. Procedure

Participants were tested individually while seated in a soundproof chamber in front of a computer screen. They were first familiarized with the pictures during a learning block. In a learning block, each picture appeared on the screen as a white-on-black line drawing with the designated name added below the picture. Participants were asked to use the designated name for each picture in the experiment. The learning block was followed by a practice block, during which each picture was presented once in the center of the screen preceded by a fixation point. Participants’ task was to name the picture as quickly and as accurately as possible using the designated picture name. This procedure assured that each participant knew and used the designated names of the pictures during the experiment.

During the experiment proper, participants did not name the pictures aloud. Rather, they were asked to carry out a go/nogo task. Metrical and syllabic decision tasks were blocked. In the metrical decision, participants were asked in each experimental trial to press a key on a keyboard when the picture name had initial stress (e.g., *LEpel* ‘spoon’). In case the picture name had final stress (e.g., *liBEL* ‘dragonfly’), they were required to withhold the key press. In a second block, instructions were switched and the same pictures were shown again in order to get a response for every item (once as a go and once as a nogo response item). The metrical decision was run to obtain temporal information about metrical encoding. Alternatively, participants were asked to press the key when the first post-vocalic consonant belonged to the first syllable (e.g., *kaN.sel* ‘pulpit’) and withhold the key press if the pivotal consonant belonged to the second syllable (e.g., *ka.No*

<sup>1</sup>As mentioned above, reduced vowels (i.e., schwas) cannot bear stress in Dutch. We would like to note that for initial stress targets there is a high but far from perfect correlation with reduced vowels in the second syllable. Therefore, a strategy of searching for reduced vowels in the target picture names to determine the stress location would not be very efficient.

‘canoe’). Again, instructions were swapped in another block afterwards. The syllabic decision had the purpose of investigating the role of syllable boundaries in phonological encoding<sup>2</sup>.

There were four different instructions, i.e. each participant performed four different tasks, one per experimental condition. Each condition began with eight practice trials (two pictures from each metrical/syllabic category), followed by 96 experimental trials. Each condition was blocked. Before the beginning of a new block, there was a short break. The sequence of pictures was randomized in every block and for every participant. Each experimental block lasted about 10 min. The entire experiment (including the placement of the electro cap and the learning/practicing of the picture names) lasted about 2 h.

A trial began with the presentation of a fixation cross (size 14 pt.) in the middle of a computer screen for 500 ms, followed after 300 ms by the picture. Pictures were of approximately equal size. They all fitted into a 7×7 cm square. As soon as possible after the picture appeared on the screen participants were required to give their response. RTs were registered automatically. The picture disappeared from the screen when participants responded or after 2000 ms. The following trial began after an inter-trial interval of 1000 ms.

Participants were instructed to rest their arms and hands on the elbow rest of the armchair and put the index finger of their right hand on the right shift key of a keyboard in front of them. In go trials, participants were expected to respond by pressing the key as fast as possible. Participants were instructed not to speak, blink, or move their eyes while a picture was on the screen.

## 2.5. Apparatus and recordings

Key-press responses were measured from picture onset with a time-out limit of 2000 ms. Time-outs and wrong responses were considered as errors and excluded from the analyses. The electroencephalogram (EEG) was recorded from 29 scalp sites using tin electrodes mounted on an electro cap with reference electrodes placed at both mastoids. The EEG signal was collected using the left mastoid as an on-line reference and it was re-referenced

off-line to the mean of the activity at the two mastoids. Bipolar electrodes placed on the right and left lower orbital ridge monitored eye blinks and vertical eye movements. A bipolar montage using two electrodes placed on the right and left external canthus monitored lateral eye movements. Eye movements were recorded for later off-line rejection of trials including eye movements. Electrode impedance was kept below 5 k $\Omega$  for the EEG and eye movement recordings.

Signals were amplified with a band pass filter from 0 to 50 Hz and digitized at 250 Hz. Averages were obtained for 1000 ms (–100 to +900 ms) epochs including a 100 ms pre-stimulus baseline. Correct response trials were visually inspected, and trials contaminated by eye movements within the critical time window were rejected and excluded from averaging. On average, 16.4% of the trials in the metrical condition and 11.5% of the trials in the syllabic condition were excluded from further analysis (including ERP artifacts and incorrect responses). The N200 was calculated for all electrode sites. For the N200 ERP peak analysis only frontal midline electrode sites were investigated, as for these sites the N200 effect is generally largest.

## 2.6. Pretest

### 2.6.1. Task difficulty

Furthermore, it is important to show that the two tasks employed in this study are approximately equally difficult to perform. If the metrical decision and the syllable affiliation decision were different with respect to task difficulty, any time-course differences between the two tasks would be difficult to interpret. Therefore, we ran a behavioral pretest in which 20 different participants decided for 80 pictures<sup>3</sup> (20 from each of the experimental conditions) with a right/left decision where the main stress of a picture name was. In another block, the same participants were required to make a syllable affiliation decision. Correct responses were averaged for both conditions. The mean RT for the metrical decision was 1184 ms and for the syllabic decision it was 1198 ms. The difference did not yield significance ( $t_1(19) < 1$ ,  $t_2(79) = 2.06$ ,  $P > 0.05$ ). The error analysis yielded similar results. This showed that the metrical and the syllable affiliation task were approximately equally difficult to perform.

<sup>2</sup>Wheeldon and Levelt [53] found approximately the same difference in monitoring latencies between the first and the second consonant of a bisyllabic word (55 ms) as between the second and the third consonant (56 ms). However, in the former case there was an intervening vowel between C1 and C2, whereas in the latter case there was no vowel, but a syllable boundary in between. Wheeldon and Levelt [53] explained this constant effect by proposing two different processes: a) intervening segments, i.e. the vowel between C1 and C2, and b) syllable boundaries, i.e. the boundary between C2 and C3. However, the fact that C3 also occurs at a later position in the word than C2 is confounded with the syllable boundary position.

<sup>3</sup>Initially, we used 80 pictures, to which 16 (equally distributed over the experimental conditions) were added in the main experiment to increase the signal-to-noise ratio in the ERP measurements. Since we replicate the result from the pre-test in the key-press RTs recorded in the main experiment, we can be confident that adding the 16 pictures to the original set of 80 pictures did not alter the difficulty balance.

### 3. Results

#### 3.1. Key-press RTs

Nine participants were excluded due to extremely high error rates (>35%) and two additional participants had to be excluded due to excessive eye blinking. Wrong key presses and time-outs were counted as errors (16.8%) and discarded from the RT analysis. Furthermore, for the RTs only latencies above 350 ms and below 1500 ms were taken into account. The mean RTs were 1122 ms (S.D. 105) for the metrical decision and 1080 ms (S.D. 109) for the syllable affiliation decision. This difference was only significant by items, but not by participants ( $t_1(15) = 1.86$ , n.s.;  $t_2(95) = 3.56$ ,  $P < 0.01$ ). These RTs might seem relatively long, but other language-related N200 studies have obtained similar RTs for go-responses [44–46]. Therefore, we can conclude that there was no difference between the two conditions in the current experiment, and this result replicates the outcome of the pretest on task difficulty (see above).

#### 3.2. N200 analysis

The N200 analysis is based on the assumption that increased negativity for nogo trials relative to go trials reflects the moment in time by which the relevant information necessary to withhold a key-press response must have been encoded. The time it takes to encode the relevant information might, therefore, be seen in the peak latencies and the peak amplitudes of the N200 effects.

First of all, we looked at whether or not the two main conditions (i.e. metrical vs. syllabic) showed a difference. As can be seen in the left column of Fig. 4, the two grand average ERP waveforms for 16 participants at midline sites (Fz, FCz, and Cz) lie almost exactly on top of each other and serial  $t$ -tests revealed no significant differences between the two curves at any time. This means that there was no task effect in the data, which is what we expected since the same items were used and task difficulty turned out to be the same as shown in the pretest.

Our main interest in this study was whether the latency characteristics of the N200 differ for the two contingency

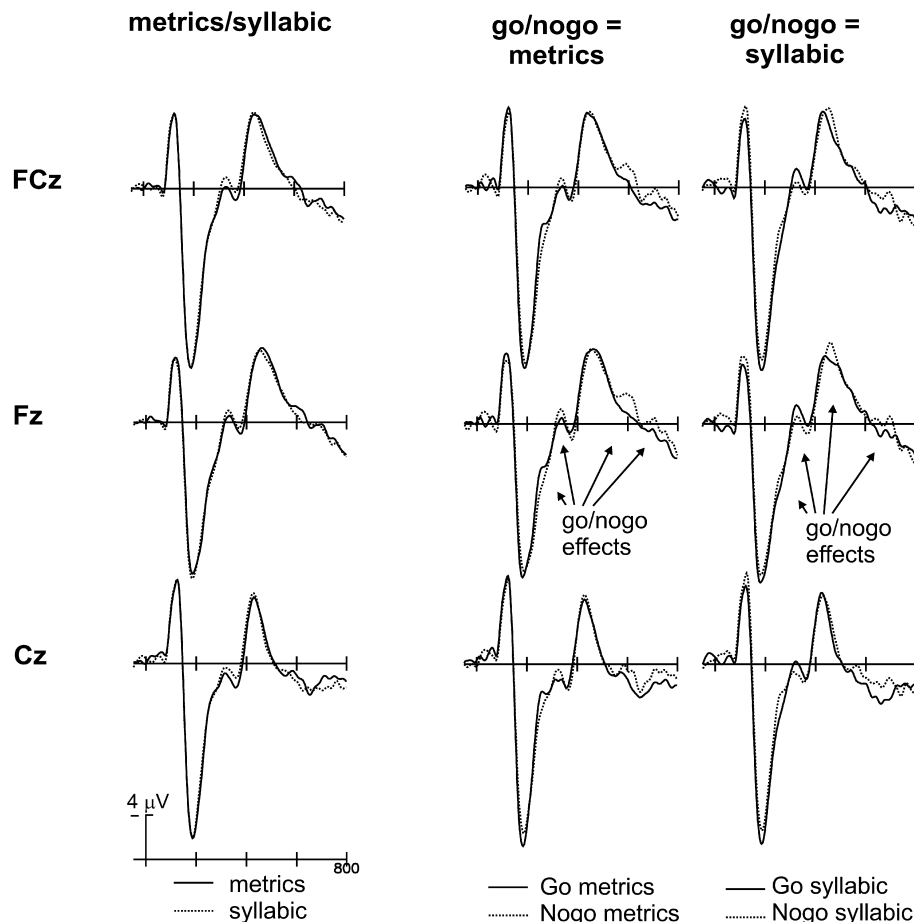


Fig. 4. Grand average ERPs for metrical and syllabic conditions collapsed across go and nogo trials (left column) and for go and nogo trials for metrical (middle column) and syllabic (right column) conditions. The ERPs were time-locked to picture onset. Both conditions are associated with two early and one late effect. Displayed are data from 16 participants (96 trials per participant per condition minus rejected trials) over three midline electrode sites. Negative voltage is plotted up.



conditions (i.e. key-press responses based on stress position vs. key-press responses based on syllable affiliation). Fig. 4 (middle and right column) shows grand average ERPs for both conditions, again for 16 participants at midline sites (Fz, FCz, and Cz). Both response contingency conditions show two early effects (see two leftmost arrows). With respect to the first early effect, go trials were more negative than nogo trials in the metrical condition (a ‘reversed N200’), whereas in the syllabic condition the pattern was reversed, i.e. nogo trials were more negative than go trials (a ‘classical N200’). In contrast, the morphology of the second early effect is such that in the metrical condition, nogo trials were more negative than go trials, but the reversed pattern was observed in the syllabic condition, i.e. go trials were more negative than nogo trials (again a ‘reversed N200’). Furthermore, there are late effects in each condition, but these effects occur much later than the N200 complex (see two rightmost arrows). Therefore, we will describe those effects separately from the two early effects. As can be seen in Fig. 4, the morphology of the two early effects looks very similar for the metrical and the syllabic condition, except for the switch in polarity. Fig. 5 shows the difference waves at frontal sites (top panel) and the corresponding topographic distribution of the difference wave in the time window for the second early effect. As can be seen, the nogo–go effect showed a reversal in polarity (negative for metrical and positive for syllabic conditions). The scalp distribution of the effect is similar in both conditions, showing a left frontal maximum.

The statistical comparison of the ERP difference waveforms (‘nogo minus go’) for both conditions at three midline electrodes (Fz, FCz, and Cz) supported the above description of the results based on visual inspection of the waveforms. For each participant, peak latencies and peak amplitudes (voltage value at the peak) of the two ERP components were measured between 200 and 400 ms at each of the three electrode sites for correct trials (96 trials minus errors). For the peak latencies as well as peak amplitudes, ANOVAs were carried out with Condition (metrical and syllabic) and Electrode Site (Fz, FCz, and Cz) as factors.

### 3.2.1. Peak latency of the first early component

When the go/nogo decision was contingent on metrical information, the mean peak latency of the early positive component was 255 ms (S.D. 30). In contrast, when the go/nogo decision was contingent on syllabic information, the mean peak latency of the early negative component was 269 ms (S.D. 27). The mean latency difference (across the three electrode sites) of the first early effects was 14 ms. With respect to the first early components, the main effect of peak latency was not significant for Condition ( $F(1,15)=1.80$ , n.s.) or for Electrode Site ( $F(2,14)=1.21$ , n.s.). Their interaction was not significant either ( $F(2,14)<1$ ).

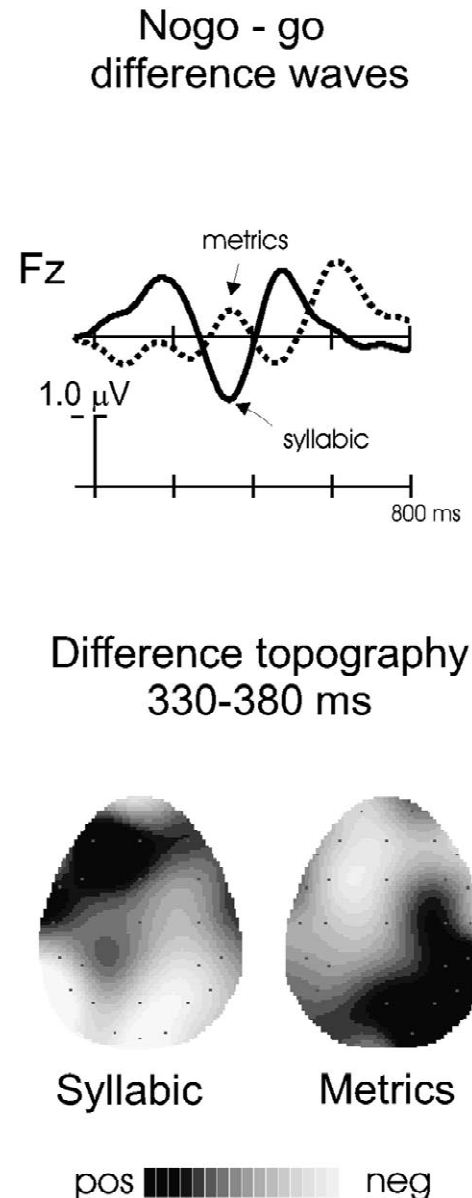


Fig. 5. Grand average difference waves nogo–go for metrical and syllabic conditions. The top panel displays the difference ERP waves at frontal site Fz. The difference waves are low pass filtered (5 Hz) for graphical display. The bottom panel shows the scalp distribution of the early nogo–go effect (left frontal positivity for syllabic and left frontal negativity for metrical condition in the time window 330–380 ms after picture onset).

### 3.2.2. Peak amplitudes of the first early component

Turning now to the mean peak amplitude analysis, the picture looks slightly different. Here, the main effect was significant for Condition ( $F(1,15)=27.17$ ,  $P<0.001$ ) but not for Electrode Site ( $F(2,14)=1.22$ , n.s.). The interaction between Condition and Electrode Site reached significance ( $F(2,14)=4.19$ ,  $P<0.05$ ), reflecting the fact that FCz was more positive than Fz in the metrical condition, whereas the pattern was reversed in the syllabic condition. When the go/nogo decision was contingent on metrical infor-

mation, the peak amplitude of the early positive component was 1.24  $\mu\text{V}$  (S.D. 1.13). In contrast, when the go/nogo decision was contingent on syllabic information, the mean peak amplitude of the early negative component was  $-1.02 \mu\text{V}$  (S.D. 1.31). The mean amplitude difference (across the three electrode sites) of the first early effects was 2.26  $\mu\text{V}$ .

### 3.2.3. Peak latency of the second early component

The second early component shows a similar pattern. The peak latency of the early negative component was 335 ms (S.D. 55) when the go/nogo decision was contingent on metrical information. However, when it was contingent on syllabic information, the peak latency of the early positive component was 329 ms (S.D. 23). The mean latency difference (across the three electrode sites) of the second early effects was merely 6 ms. The main effect of peak latency was not significant for Condition ( $F(1,15) < 1$ ) or for Electrode Site ( $F(2,14) < 1$ ), and their interaction was not significant either ( $F(2,14) = 1.02$ , n.s.).

### 3.2.4. Peak amplitude for the second early component

The main effect of peak amplitude was significant for Condition ( $F(1,15) = 27.05$ ,  $P < 0.001$ ), but not for Electrode Site ( $F(2,14) < 1$ ). The interaction between the two factors was not significant either ( $F(2,14) = 2.01$ , n.s.). When the go/nogo decision was contingent on metrical information, the peak amplitude of the early negative component was  $-0.92 \mu\text{V}$  (S.D. 1.11). In contrast, when the go/nogo decision was contingent on syllabic information, the mean peak amplitude of the early positive component was 0.86  $\mu\text{V}$  (S.D. 0.87). The mean amplitude difference (across the three electrode sites) of the second early effects was 1.78  $\mu\text{V}$ .

### 3.2.5. Late effect analysis

With respect to the later effects, we carried out serial *t*-test analyses at sites Fz, FPz, and Cz in the time window 400–800 ms after picture onset. We observed significant divergence of the difference waves from zero baseline ( $t > 1.64$ ;  $P < 0.05$ ) for the syllabic condition in two time windows, namely between 450 and 480 ms after trial onset at all three sites and between 690 and 710 ms after trial onset (for Cz even between 620 and 710 ms after trial onset). For the metrical condition, there is only one late effect, namely between 610 and 640 ms after picture onset (for Fz and FPz between 580 and 640 ms after picture onset).

## 4. Discussion

By applying high-temporal resolution ERP to tacit picture naming in a simple go/nogo N200 paradigm, we observed clear time course information of metrical and syllabic encoding. In the metrics=go condition, the ERP

waves of go and nogo responses diverge from each other in the time window of 250 to 350 ms, especially at frontal sites. The same holds for the syllabic=go condition. We were thus able to estimate for the first time on-line processing of metrical and syllabic encoding during tacit picture naming. The data also showed that there seems to be no difference between the metrical and the syllabic condition in terms of peak latency, i.e. in terms of information availability. The observed ERP effects within the 250–350 ms time windows can thus be interpreted as showing parallel processing of metrical and syllabic encoding.

Furthermore, we observed significant mean amplitude differences (switch in polarity of the N200 effects) within each condition in the 250–350 ms time window. This N200 amplitude pattern is reversed between metrical and syllabic conditions, a finding we did not expect. In the metrical condition (button press contingent on metrical information), the waveform for go-trials is more negative than the waveform for nogo-trials between 250 and 300 ms after picture onset ('reversed N200'). This is the first early effect. The second early effect in the metrical condition (between 300 and 350 ms after picture onset) is such that nogo-waveforms are more negative than go-waveforms. In the syllabic condition (button press contingent on syllabic information), nogo-waveforms are more negative than go-waveforms between 250 and 300 ms after picture onset. Between 300 and 350 ms, the waveform for go-trials is more negative than the waveform for nogo-trials (again, a 'reversed N200'). Although these effects are small, serial *t*-tests showed that they are statistically robust and the peak amplitude statistics revealed significant differences between the metrical and the syllabic condition.

We can only speculate here on the nature of this switch of polarity. Positive N200 effects have been reported in the literature before. In single cell recording in monkeys the surface negative N200 is usually positive in subcortical structures, indicating polarity switches based on differences of measurement methods and locations [13]. Furthermore, in humans, surface positive N200 have been reported for go/nogo tasks in earlier research (e.g., Ref. [20]). Kiefer et al. [20] suggested that overlap of the N200 complex with the P300 component was responsible for their N200 becoming positive in polarity. The P300 is often related to task difficulty: the more difficult a task is to perform for the participants, the larger the amplitude of the P300, and consequently the more positive the whole signal. Our data, however, do not display a clear P300 at all, and, as tested in the pretest, the tasks are comparable in terms of difficulty, so this account is difficult to apply to the current data set. Most importantly, for the purpose of interpretation of the N200 as being related to response inhibition (and information availability), the authors of Ref. [20] interpreted their positive N200 in the same way as the negative N200.

Sensitivity of the N200 amplitudes to task difficulties

has been reported in monkeys [13] as well as for humans in priming tasks [22]. It has also been related to false alarm rates [9]. But, again, based on our pretest data, we cannot explain the amplitude shift in this way since we did not observe differences in task difficulties or error proportions. Others showed that polarity changes or switches detected on the surface might be related to the involvement of a different set of underlying neural generators. This has been suggested, for example, for the mismatch negativity (MMN) polarity reversals [2], as well as for P200 polarity reversals related to face recognition [14]. It remains an open question whether or not the same holds for the nogo–go N200 in humans. The topography of the positive and negative N200 effect in our study (see Fig. 5) implies that we are dealing with similar neural structures that are responsible for the observed activation. However, the limited spatial resolution of our data does not allow for detailed dipole localization. As far as we are aware, no one has yet compared positive and negative N200 and their potential difference in underlying neural structures and functions in a systematic way. This issue clearly merits further research. Most important for the present study, however, is that we take the observed *difference* between go and nogo conditions and its maximum as an upper limit of information availability for metrical and syllabic encoding regardless of the polarity switch.

In sum, based on the early-observed ERP effects we estimated the time course of information availability for metrical and syntactic encoding to take place in the time window of 250–350 ms after picture onset. The peak latency analysis did not show a difference between the two conditions. Based on these data we conclude, therefore, that both processes occur more or less at the same time—at least our ERP measures (based on peak latency analysis of the nogo–go difference waves) could not differentiate between the two components. Interestingly, the time window between 250 and 350 ms falls exactly within the time window that Indefrey and Levelt [16,17] propose for lexical phonological code retrieval and phonological encoding in speech production. These authors assume that lexical selection (all processes included in the upper gray box in Fig. 1) is accomplished within 275 ms from picture onset. Phonological form encoding is estimated to take place between 275 and 400 ms after picture onset [16]. This includes phonological code retrieval and syllabification, and it fits our results here. Therefore, we would like to suggest that the two early effects we observed in both conditions reflect production components proper, one for metrical encoding, and one for syllabic encoding.

The later effects were not expected, and they occur relatively late to be part of proper speech production components. Especially, they are too late to reflect phonological encoding, because that should be finished around 400 ms after picture onset [16]. Because the effects are statistically sound, we would like to propose an interpretation here. Let us first have a look at the metrical

condition. Here, there is one clear effect between 610 and 640 ms after picture onset (as indicated by serial *t*-tests;  $P < 0.05$ , at all three target frontal sites). We would like to cautiously suggest that internal self-monitoring caused this effect. Before participants can decide whether or not to press the button in the metrical condition, they must first generate the target picture name and then self-monitor the stress pattern of the target. Wheeldon and Levelt [53] argued that phonological information is not directly available to the speaker, but rather after phonological word formation only, i.e. when a fully prosodified phonological representation of a word is generated. It is assumed that speakers build a phonological word and then put it into an articulatory buffer [15,24] so that they can monitor it for certain (phonological) information, such as stress position or syllable boundaries. It is possible that this late ERP component we observed in the metrical condition reflects the self-monitoring process for stress position. Note that we have behavioral evidence [42] demonstrating that the monitoring of metrical information such as stress is possible.

Turning to the syllabic condition, there is also a late component that serial *t*-tests revealed to be significant ( $P < 0.05$ ) between 690 and 710 ms after picture onset (at all three frontal target sites), i.e. a bit later than in the metrical condition. We would like to speculate that this also reflects a self-monitoring component, i.e. monitoring where the syllable boundary in a particular word is. Again, we have behavioral data [18] showing that speakers are able to self-monitor syllable boundaries at the phonological word level. However, there is another relatively late effect in the syllabic condition, occurring between 450 and 480 ms (as revealed by serial *t*-tests;  $P < 0.05$ ). So far, we do not have a clear interpretation as to the nature of this effect. Its timing is also too late to belong to phonological encoding proper, but probably it is too early to be considered as a reflection of internal monitoring.

In summary, we used a specific ERP component to investigate the relative time course of two different processes within phonological encoding in language production. Specifically, we employed the N200 effect (related to response inhibition) to investigate on-line picture naming. Brain waves did not show a difference of peak latencies for two relatively early effects, which were detected in the signals. However, peak amplitudes were significantly different for both effects, probably due to a change in polarity. Although our data do not show any differences with respect to the relative time course of syllabic and metrical processing, both early effects fall exactly within the time window identified in the literature for phonological encoding on the basis of independent data. Furthermore, there were two late effects in the data, one in the metrical and one in the syllabic condition, which we speculated to be related to internal self-monitoring before response execution. However, the monitoring aspect and the reversal in polarity merit further research.

## Acknowledgements

Niels O. Schiller is supported by the Royal Dutch Academy of Arts and Sciences (KNAW), M. Bles and B.M. Jansma by the Dutch Science Foundation (NWO). B.M. Jansma published under the name B.M. Schmitt before 2003. The authors thank Iemke Horemans (University of Maastricht) for her help during the analysis. The research reported in this paper benefited from discussions at the Eighth Annual Meeting of the Cognitive Neuroscience Society in New York (March, 2001) and at the conference for the Neurological Basis of Language in Groningen (July, 2001).

## Appendix A

Targets with initial stress		Targets with final stress	
CV	CVC	CV	CVC
bezem ('broom')	banjo ('banjo')	banaan ('banana')	balkon ('balcony')
boter ('butter')	borstel ('brush')	beha ('bra')	biljart ('pool')
hamer ('hammer')	bunker ('bunker')	bureau ('desk')	bonbon ('candy')
jager ('hunter')	cactus ('cactus')	citroen ('lemon')	dolfijn ('dolphin')
kabel ('cable')	cirkel ('circle')	fabriek ('factory')	garnaal ('shrimp')
kano ('canoe')	dokter ('doctor')	gebit ('dentures')	gordijn ('curtain')
kegel ('bowling pin')	gondel ('gondola')	geweer ('rifle')	harpoen ('harpoon')
ketel ('kettle')	halter ('weight')	giraf ('giraffe')	kalkoen ('turkey')
koning ('king')	herder ('shepherd')	gitaar ('guitar')	karkas ('skeleton')
lepel ('spoon')	kansel ('pulpit')	kameel ('camel')	kasteel ('castle')
molen ('wind mill')	lifter ('hitch hiker')	kanon ('canon')	kompas ('compass')
motor ('motor bike')	masker ('mask')	karaf ('pitcher')	lantaarn ('lantern')
nagel ('finger nail')	panter ('panther')	konijn ('rabbit')	magneet ('magnet')
navel ('navel')	parfum ('perfume')	libel ('dragonfly')	pastoor ('priest')
ratel ('rattle')	pinguin ('penguin')	loket ('counter')	penseel ('brush')
robot ('robot')	pleister ('band aid')	matras ('mattress')	pincet ('tweezers')
sleutel ('key')	scalpel ('scalpel')	meloen ('melon')	pistool ('gun')
spijker ('nail')	stempel ('stamp')	piraat ('pirate')	pompoen ('pumpkin')
tafel ('table')	tempel ('temple')	piloot ('pilot')	portret ('portrait')
tijger ('tiger')	tractor ('tractor')	raket ('rocket')	sandaal ('sandal')
toren ('tower')	varken ('pig')	rivier ('river')	soldaat ('soldier')
vlieger ('kite')	vlinder ('butterfly')	sigaar ('cigar')	tampon ('tampon')
vogel ('bird')	wortel ('carrot')	tomaat ('tomato')	trompet ('trumpet')
zebra ('zebra')	zuster ('nurse')	toneel ('stage')	vampier ('vampire')

## References

- [1] R.H. Baayen, R. Piepenbrock, L. Gulikers, The CELEX lexical database, Linguistic Data Consortium, University of Pennsylvania, Philadelphia, 1995 (CD-ROM).
- [2] T. Baldeweg, J.D. Williams, J.H. Gruzeliier, Differential changes in frontal and sub-temporal components of mismatch negativity, *Int. J. Psychophysiol.* 33 (1999) 143–148.
- [3] G. Booij, *The Phonology of Dutch*, Clarendon Press, Oxford, 1995.
- [4] J. Cholin, N.O. Schiller, W.J.M. Levelt, The role of the syllable at the interface of phonology and phonetics in speech production, *J. Mem. Language* (in press).
- [5] A. Crompton, Syllables and segments in speech production, *Linguistics* 19 (1981) 663–716.
- [6] M.F. Damian, R.C. Martin, Semantic and phonological codes interact in single word production, *J. Exp. Psychol. Learn. Mem. Cogn.* 25 (1999) 345–361.
- [7] G.S. Dell, A spreading-activation theory of retrieval in sentence production, *Psychol. Rev.* 93 (1986) 283–321.
- [8] G.S. Dell, The retrieval of phonological forms in production: tests of predictions from a connectionist model, *J. Mem. Language* 27 (1988) 124–142.
- [9] M. Falkenstein, J. Hoormann, J. Hohnsbein, ERP components in go/nogo tasks and their relation to inhibition, *Acta Psychol.* 101 (1999) 267–291.
- [10] L. Ferrand, J. Segui, J. Grainger, Masked priming of word and picture naming: the role of syllabic units, *J. Mem. Language* 35 (1996) 708–723.
- [11] L. Ferrand, J. Segui, G.W. Humphreys, The syllable's role in word naming, *Mem. Cogn.* 35 (1997) 458–470.
- [12] O. Fujimura, J.B. Lovins, Syllables as concatenative phonetic units, in: A. Bell, J.B. Hooper (Eds.), *Syllables and Segments*, North-Holland, Amsterdam, 1978, pp. 107–120.
- [13] H. Gemba, K. Sasaki, Potential related to no-go reaction of go/no-go hand movement task with color discrimination in human, *Neurosci. Lett.* 101 (1989) 262–268.
- [14] N. George, J. Evans, N. Fiori, J. Davidoff, B. Renault, Brain events related to normal and moderately scrambled faces, *Cogn. Brain Res.* 4 (1996) 65–76.
- [15] R.J. Hartsuiker, H.H.J. Kolk, Error monitoring in speech production: a computational test of the perceptual loop theory, *Cogn. Psychol.* 42 (2001) 113–157.
- [16] P. Indefrey, W.J.M. Levelt, The neural correlates of language production, in: M. Gazzaniga (Ed.), *The New Cognitive Neurosciences*, MIT Press, Cambridge, MA, 2000, pp. 845–865.
- [17] P. Indefrey, W.J.M. Levelt, The spatial and temporal signatures of word production components, *Cognition* (in press).
- [18] B.M. Jansma, N.O. Schiller, Monitoring syllable boundaries during speech production, *Brain Language* (in press).
- [19] E. Jodo, Y. Kayama, Relation of a negative ERP component to response inhibition in a go/nogo task, *Electroencephalogr. Clin. Neurophysiol.* 82 (1992) 477–482.
- [20] M. Kiefer, F. Marzinsik, M. Weisbrod, M. Scherg, M. Spitzer, The time course of brain activation during response inhibition: evidence from event-related potentials in a go/no go task, *Neuroreport* 9 (1998) 765–770.
- [21] A. Kok, Effects of degradation of visual stimuli on components of the event-related potential (ERP) in go/nogo reaction tasks, *Biol. Psychol.* 23 (1986) 21–38.
- [22] B. Kopp, U. Mattler, R. Goerty, F. Rist, N2, P3 and the lateralized readiness potential in a nogo task involving selective response priming, *Electroencephalogr. Clin. Neurophysiol.* 99 (1996) 19–27.
- [24] W.J.M. Levelt, *Speaking. From Intention to Articulation*, MIT Press, Cambridge, MA, 1989, pp. 566.
- [25] W.J.M. Levelt, Spoken word production: a theory of lexical access, *Proc. Natl. Acad. Sci.* 98 (2001) 13464–13471.
- [26] W.J.M. Levelt, N.O. Schiller, Is the syllable frame stored?, *Behav. Brain Sci.* 21 (1998) 520.
- [27] W.J.M. Levelt, A. Roelofs, A.S. Meyer, A theory of lexical access in speech production, *Behav. Brain Sci.* 22 (1999) 1–75.
- [28] W.J.M. Levelt, L. Wheeldon, Do speakers have access to a mental syllabary?, *Cognition* 50 (1994) 239–269.
- [29] A.S. Meyer, The time course of phonological encoding in language production: the encoding of successive syllables of a word, *J. Mem. Language* 29 (1990) 524–545.
- [30] A.S. Meyer, The time course of phonological encoding in language production: phonological encoding inside a syllable, *J. Mem. Language* 30 (1991) 69–89.
- [31] A. Pfefferbaum, J.M. Ford, B.J. Weller, B.S. Kopell, ERPs to response production and inhibition, *Electroencephalogr. Clin. Neurophysiol.* 60 (1985) 423–434.

- [32] A. Rodriguez-Fornells, B.M. Schmitt, M. Kutas, T.F. Münte, Electrophysiological estimates of the time course of semantic and phonological encoding during listening and naming, *Neuropsychologia* 40 (2002) 778–787.
- [33] A. Roelofs, The WEAVER model of word-form encoding in speech production, *Cognition* 64 (1997) 249–284.
- [34] A. Roelofs, A.S. Meyer, Metrical structure in planning the production of spoken words, *J. Exp. Psychol. Learn. Mem. Cogn.* 24 (1998) 922–939.
- [35] K. Sasaki, H. Gemba, Prefrontal cortex in the organization and control of voluntary movement, in: T. Ono, L.R. Squire, M.E. Raichle, D.I. Perret, M. Fukuda (Eds.), *Brain Mechanisms of Perception and Memory: From Neuro To Behavior*, Oxford University Press, New York, 1993, pp. 473–496.
- [36] K. Sasaki, H. Gemba, A. Nambu, R. Matsuzaki, No-go activity in the frontal association cortex of human subjects, *Neurosci. Res.* 18 (1993) 249–252.
- [37] N.O. Schiller, The effect of visually masked syllable primes on the naming latencies of words and pictures, *J. Mem. Language* 39 (1998) 484–507.
- [38] N.O. Schiller, Masked syllable priming of English nouns, *Brain Language* 68 (1999) 300–305.
- [39] N.O. Schiller, Single word production in English: the role of subsyllabic units during phonological encoding, *J. Exp. Psychol. Learn. Mem. Cogn.* 26 (2000) 512–528.
- [40] N.O. Schiller, A. Costa, A. Colomé, Phonological encoding of single words: in search of the lost syllable, in: C. Gussenhoven, N. Warner (Eds.), *Papers in Laboratory Phonology 7*, Mouton de Gruyter, Berlin, 2002, pp. 35–59.
- [41] N.O. Schiller, P. Fikkert, C.C. Levelt, Stress priming in picture naming: an SOA study, *Brain Language* (in press).
- [42] N.O. Schiller, B.M. Jansma, J. Peters, W.J.M. Levelt, Monitoring metrical stress in polysyllabic words (submitted for publication).
- [43] N.O. Schiller, A.S. Meyer, W.J.M. Levelt, The syllabic structure of spoken words: evidence from the syllabification of intervocalic consonants, *Language Speech* 40 (1997) 103–140.
- [44] B.M. Schmitt, T.F. Münte, M. Kutas, Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming, *Psychophysiology* 37 (2000) 473–484.
- [45] B.M. Schmitt, A. Rodriguez-Fornells, M. Kutas, T.F. Münte, Electrophysiological estimates of semantic and syntactic information access during tacit picture naming and listening to words, *Neurosci. Res.* 41 (2001) 293–298.
- [46] B.M. Schmitt, K. Schiltz, W. Zaake, M. Kutas, T.F. Münte, An electrophysiological analysis of the time course of conceptual and syntactic encoding during tacit picture naming, *J. Cogn. Neurosci.* 13 (2001) 510–522.
- [47] H. Schriefers, A.S. Meyer, W.J.M. Levelt, Exploring the time course of lexical access in language-production: picture–word interference studies, *J. Mem. Language* 29 (1990) 86–102.
- [48] R. Simson, H.G. Vaughan, W. Ritter, The scalp topography of potentials in auditory and visual go/nogo tasks, *Electroencephalogr. Clin. Neurophysiol.* 43 (1977) 864–875.
- [49] S. Thorpe, D. Fize, C. Marlot, Speed of processing in the human visual system, *Nature* 381 (1996) 520–522.
- [50] M. van Turenout, P. Hagoort, C.M. Brown, Electrophysiological evidence on the time course of semantic and phonological processes in speech production, *J. Exp. Psychol. Learn. Mem. Cogn.* 23 (1997) 787–806.
- [51] M. van Turenout, P. Hagoort, C.M. Brown, Brain activity during speaking: from syntax to phonology in 40 milliseconds, *Science* 280 (1998) 572–574.
- [52] J. Waals, An experimental view of the Dutch syllable, Ph.D. dissertation (Netherlands Graduate School of Linguistics; 18), Holland Academic Graphics, The Hague, 1999, 156 pp.
- [53] L. Wheeldon, W.J.M. Levelt, Monitoring the time course of phonological encoding, *J. Mem. Language* 34 (1995) 311–334.