Published in *Journal of Memory and Language*, **39**, 331—346, 1998.

Where is the length effect?
A cross-linguistic study of speech production.*

A.-C. Bachoud-Lévi[1,2], E. Dupoux[1], L. Cohen[1,3], J. Mehler[1]

[1.] Laboratoire de Sciences Cognitives et Psycholinguistique, (EHESS-CNRS), Paris, France

[2.] Service de Neurologie, Hôpital Henri Mondor, Créteil, France

[3.] Service de Neurologie 1, Hôpital de la Salpétrière, Paris, France

---

* We wish to thank J. Morton for having allowed us to run our English experiments in his laboratory. We also thank B. Van Oyen, S. Peperkamp and S. Franck for their help and patience with various versions of the manuscript. This study was supported by grants from the Fondation pour la Recherche Médicale, from Assistance Publique-Hôpitaux de Paris and CNRS, an HFSP grant (1995-1997) and the Human Capital and Mobility Grant nb ERBCHRXCT 920031.

Correspondance should be sent to Anne-Catherine Bachoud-Lévi, Laboratoire de Sciences Cognitives et Psycholinguistique, Maison des Sciences de l'Homme, 54, Bd Raspail, 75006 Paris (email: bachoud@lscp.ehess.fr).

ABSTRACT

Many models of speech production assume that one cannot begin to articulate a word before all its segmental units are inserted into the articulatory plan. Moreover, some of these models assume that segments are serially inserted from left to right. As a consequence, latencies to name words should increase with word length. In a series of five experiments, however, we showed that the time to name a picture or retrieve a word associated with a symbol is not affected by the length of the word. Experiments 1 and 2 used French materials and participants, while Experiments 3, 4 and 5 were conducted with English materials and participants. These results are discussed in relation to current models of speech production, and previous reports of length effects are reevaluated in light of these findings. We conclude that if words are encoded serially, then articulation can start before an entire phonological word has been encoded.

Uttering words involves a series of processes that begins with the activation of concepts and results in overt articulation (Butterworth, 1980, 1989; Dell, 1981, 1986, 1988, 1992; Fay & Cutler, 1977; Garrett, 1975, 1992; Levelt, 1989; Shattuck-Hufnagel, 1979, 1992). Current models of speech production postulate that, when one wants to utter a word, its abstract representation is retrieved from the mental lexicon and is used to construct a detailed *phonological plan* to be executed by the articulatory system. The details of this process (which we refer to as *phonological encoding* are still being debated, but all models assume that some minimal portion of phonological plan has to be built before articulation can begin. What is the size of the minimal portion of the phonological plan?

Many authors have assumed that articulation can start when at least an entire word has been completely encoded (Levelt, 1989; Maclay & Osgood, 1959; Shattuck-Hufnagel, 1979; Sternberg, Monsell, Knoll & Wright, 1978, 1980). This assumption is supported in part by the observation that participants are slower in uttering bisyllabic rather than monosyllabic words (Eriksen, Pollock & Montague, 1970; Klapp, Anderson & Berrian, 1973). These effects of word length on naming latency supposedly arise because short words require less planning and hence can begin to be articulated more quickly than longer words. Indeed, most models assume that the phonological plan is built incrementally, with segmental units being specified one after the other (see Levelt, 1992; Meyer, 1990; 1991; Meyer & Schrieffers, 1991; Shattuck-Hufnagel, 1992 for strict seriality; see Costa, Sebastian and Pallier, submitted; Levelt & Wheeldon, 1994; Wheeldon & Levelt, 1995 for seriality with partial overlap). Only models with parallel insertion of all segments, as in Dell (1986), have problems accounting for an effect of word length on naming latencies.

The notion that the minimal planning unit is the word may seem counter-intuitive. Clearly, there are cases in which one can begin articulation before the entire word is programmed. For instance, Marslen-Wilson (1973) has shown that some people can repeat an utterance with a lag of around 250 msec. We can assume that people begin to articulate words before they know much more than the first syllable of, say, a quadrisyllabic utterance. However, shadowing experiments do not necessarily clarify what happens when people spontaneously generate speech.

What then is the empirical support for the length effect in the existing literature? Knowing whether the length effect is reliable is an essential empirical datum for models of speech production. The existence of a length effect would support the hypothesis that one cannot begin pronouncing a word until encoding has been completed. Furthermore, it would indirectly support the hypothesis that segments are serially encoded in the articulation plan. In contrast, the absence of a length effect would require that at least one of these two assumptions be abandoned.

Many studies have observed some effect of word length on reaction times. However, it is important to consider the precise task used. The relevant studies involved principally three tasks: word naming (Balota & Chumbley, 1984, 1985; Forster & Chambers, 1973; Frederiksen and Kroll, 1976; Jared & Seidenberg, 1990), list repetition (Sternberg et al., 1978, 1980), and picture or digit naming (Eriksen et al., 1970; Klapp et al., 1973). Since the studies by Morton and Patterson (1980) and Newcombe and Marshall (1980), it is accepted that naming of printed words may involve a lexical route - which involves phonological planning, as well as a nonlexical surface route - which partly short-circuits planning and directly converts graphemes to phonemes. Therefore, word naming cannot be assumed to unambiguously tap the phonological planning process. Some of the reported effects of length on word naming may be related either to the input phase of word reading or to the grapheme-phoneme transcription route, rather than to phonological encoding per se. A similar comment applies to the list repetition task, which involves both a memory retrieval process and phonological planning. Indeed, in the list repetition task, effects of length have been attributed to retrieving or triggering items from an articulatory buffer rather than to phonological planning per se (see Sternberg et al., 1978). Of the three tasks, digit or picture naming appears to be the most straightforward one to address the construction of the phonological plan. It involves visual processing, lexical access, and phonological planning. Provided confounding variables at the visual or lexical levels are controlled, it can be used to evaluate the time course of phonological planning.

Two naming studies report the existence of a length effect: Eriksen et al. (1970) and Klapp et al. (1973). Eriksen et al. (1970) used a digit naming paradigm while Klapp et al. (1973) used a picture naming task. Eriksen et al. (1970) found that numbers with long names had longer naming latencies than those with short names, while Klapp et al. (1973) found a similar length effect with pictures. In both of these experiments, length was measured in terms of number of syllables. Given that the first phonemes of the short versus long items were not matched in either of these studies, some low level property of the initial phonemes (e.g., ease of articulation or acoustic detectability) might have been confounded with the length variable. Both studies, however, used a delayed naming task as a control. In this task, no length effect was found, making it unlikely that the results in the immediate naming conditions were due to a low-level artifact.

More importantly, however, both Eriksen et al. (1970) and Klapp et al. (1973) failed to control adequately for frequency or familiarity across lengths. In the Klapp et al. study, the monosyllabic words apparently were more familiar than the bisyllabic ones. Of the 14 pictures used in Klapp et al., 10 can be assessed on the Snodgrass and Vanderwart (1980) scale; their familiarity score is 3.7 for the monosyllables and 2.4 for the bisyllables (1-5 scale), a statistically significant difference ($t(8)=2.8$; $p<.02$). Hence, it is difficult to decide whether the 14.4 msec. effect is due to length, familiarity, or both. Eriksen et al. tested participants with two-digit numbers that were 2, 3 or 4 syllables long. The bisyllabic items were mostly "reference numerals" such as 15, 20, and 30, while the 3 and 4 syllable items included nonreference numerals such as 17, 28, and 37. They were also of greater magnitude than the bisyllabic items. Dehaene and Mehler (1992) have demonstrated a strong word frequency advantage for reference numerals as compared to neighboring nonreference ones across several languages. Moreover, they found a negative correlation between word frequency and magnitude for numerals. Hence, in the Eriksen et al. study, there is a confound among length, frequency, and magnitude, which makes it difficult to conclude that a true length effect was indeed observed.

Given these reservations concerning the results of Eriksen et al. and Klapp et al., we decided to explore once again whether long words have longer naming latencies than short ones. In Experiments 1 and 2 we tested the length effect with French words. In Experiments 3, 4 and 5, we used English words to compare our results more directly with those of Klapp et al. and Eriksen et al., as well as to assess the existence of the length effect cross-linguistically. This is an important issue because recent research (Cutler, Mehler, Norris & Segui, 1983, 1989) has shown that different languages, and in particular English and French, give rise to different processing routines. In our experiments we used picture naming as well as the symbol naming method developed by Levelt and Wheeldon (1994). The words in English and French were matched for frequency, syllabic structure and initial phoneme.

EXPERIMENT 1: PICTURE NAMING IN FRENCH

Experiment 1 was designed to investigate whether naming latencies increase with number of syllables in French mono- and bisyllabic words. A picture naming paradigm was used in order to ensure that no phonological information was provided to the participants. To name a picture, participants had to perceive and identify it and access the corresponding lexical entry containing the phonological and articulatory codes.

It is generally accepted that naming latencies increase as the frequency and the familiarity of the targets decrease (Connine, Mullenix, Shernoff & Yelen, 1990; Fraisse, 1964; Jescheniak & Levelt, 1994; Oldfield & Wingfield, 1965; Wingfield, 1968). Thus, frequency and familarity were systematically balanced across conditions. Moreover, each bisyllabic item had a monosyllabic counterpart with the same initial phoneme. Participants were told to name each picture as soon as they saw it.

Method

Materials. Thirty pictures of objects were selected from the set of line drawings in Snodgrass and Vanderwart (1980). In order to ensure that only the expected words were produced, the original corpus of 260 pictures was presented to 15 participants for naming. The experimental material was chosen from a set of 115 pictures on which all participants agreed completely. Fifteen pictures had monosyllabic names (median frequency: 22 per million according to Content, Mousty & Radeau, 1990), and 15 had bisyllabic names (median frequency: 31 per million), see Appendix. These items were arranged in pairs the members of which shared the same initial phoneme. The two classes of stimuli were also matched for word frequency (Content et al., 1990) and object familiarity (Snodgrass & Vanderwart, 1980). The location of stress followed its typical French distribution, that is, was on the last syllable in polysyllabic words.

Procedure. Pictures were digitized and presented on the plasma screen of a Toshiba T-5200 computer. A microphone was connected to an OROS AU-22 digital board that digitized the naming response (8 KHz, 16 bits) and ran a signal detection algorithm (using an adaptative threshold) to find word onset. Digitized responses were stored on a disk for subsequent scoring of errors.

Participants were instructed to name the picture as soon as it appeared on the screen. The target picture remained visible until a vocal response had been provided (with a time-out of 5 sec.). The response was followed by a 2 sec. blank screen before the next trial began. Naming latencies were measured from the onset of the picture. The set of 30 drawings was presented twice. The order of the pictures was randomized within each block of 30 pictures. Participants received five practice trials using different pictures.

Participants. Eighteen French students (10 men and 8 women), aged from 21 to 32 years of age, participated in this experiment. All were native speakers of French.

Results

Reaction times corresponding to erroneous responses, stuttering, self-correction and technical malfunction were excluded from the analyses. In all, these accounted for 4.9% of the responses.

Two ANOVAs were conducted on naming latencies and errors, one with participants and one with items as random variables. In the participant analysis, outliers were defined as latencies that differed by more than two standard deviations from the mean for each participant (1.6%). A similar definition was adopted in the item analysis (2.16%). Outliers were excluded from the analysis. The same procedure was used in all experiments. There was one within-participant factor (number of syllables). Importantly, there was no significant latency difference between mono- and bisyllabic words (respectively 565 msec. *vs* 561 msec.; $F1(1, 17) < 1$, $p = .39$; $F2(1, 14) = 1.7$, $p = .21$).

A post-hoc regression analysis assessed the effect of frequency, and familiarity on naming latencies. The regression analysis revealed that naming latencies decreased as familiarity and frequency increased (familiarity: $t(29) = -10.87$, $p < .02$, and frequency: $t(29) = -5.72$, $.05 < p < .1$ respectively).

There were significantly more errors with monosyllabic words than with bisyllabic words in the participants analysis but not in the items analysis (6% *vs* 4% errors, $F1(1, 17) = 5.6$, $p < .03$ and $F2(1, 14) = 1.3$, $p = .3$). Four pictures yielded more than 20% errors (namely "coq" (rooster), "gant" (glove), "chemise"(shirt), "mouche" (fly)). A new analysis was conducted excluding these pictures, but this did not alter the pattern of results.

Discussion

This experiment failed to reveal a length effect in French. The regression analysis revealed that naming latencies decreased as familiarity and frequency increased, in accordance with the findings of Balota and Chumbley (1985), Connine, Mullenix, Shernoff and Yelen (1990), Forster (1981), Fraisse (1964), and Monsell, Doyle and Haggard (1989).

Before concluding that a length effect does not arise in naming with French participants, we have to address the following potential shortcoming: Naming latencies not only reflect phonological encoding but also earlier visual and lexical processes. Could the observed results be due to the fact that, for some unknown reason, such nonphonological processes take longer for monosyllabic than for bisyllabic items? To answer this question, we ran a control experiment using a task that requires the same visual and lexical processes as above, but no phonological planning. The difference in reaction times between mono- and bisyllabic items in this control task can be used to assess the visual/lexical contribution to the effects observed with the naming task. We presented sixteen participants from the same populations as those of Experiment 1 with a written word for 500 msec., followed by a picture after 1.5 sec. They had to press on one of two buttons to indicate whether the picture matched or did not match the word. We found no significant latency difference between mono- and bisyllabic words (514 msec. *vs* 520 msec. respectively; all Fs <1). Nor was the number of errors significantly different (4.2% for monosyllables, 3.3 % for bisyllables, Fs < 1). This control experiment suggests that the observed absence of length effect is not due to putative differences in early visual or lexical processes between mono- and bisyllabic items.

We must also consider why, in the present experiment, the frequency effect was only marginal while the familiarity effect was significant. In some studies, familiarity has been shown to be a stronger predictor of naming latencies than frequency (Fraisse, 1964; but see Jescheniak & Levelt, 1994). More importantly, the materials we used were in a narrow frequency range, which might have prevented the emergence of a frequency effect. Experiment 2 was designed to confirm the absence of length effect in the presence of an explicit frequency manipulation. In order to achieve this, we used a different technique that allows greater flexibility in the selection of stimuli.

EXPERIMENT 2: NAMING OF ARBITRARY SYMBOLS IN FRENCH

In the present experiment, we investigated the existence of a length effect using a symbol naming paradigm similar to the one developed by Levelt and Wheeldon (1994). In this paradigm, participants are required to learn an arbitrary association between a symbol and a word. During the test phase, a symbol is displayed on the screen and participants have to pronounce the corresponding word as quickly as possible. This procedure has the advantage that, since symbols are rotated across subjects, differences in visual processing across items of different length cannot interfere. Moreover, the symbol naming procedure allows flexibility in the selection and control of stimuli.

As in Experiment 1, we explored naming latencies for French monosyllabic and bisyllabic words. This time, the materials were also partitioned into low and high frequency words.

Method

Materials. We constructed 12 quadruplets containing a pair of mono- and a pair of bisyllabic words (see Appendix). All members of the quadruplets shared the initial phoneme and structure of the first syllable (e.g., CV or CVC for example). In each pair, one word had a low frequency (less than 5 per million according to Content et al. (1990)) and the other a high frequency (more than 100 per million). For example, the monosyllabic low frequency word "pope" (pope) was matched with the monosyllabic high frequency word "peur" (fear), and also with the bisyllabic low frequency word "pactole" (treasure) and the bisyllabic high frequency word "personne" (person). The stress pattern followed the French distribution. The 12 quadruplets were divided into three sets of 4 quadruplets each. Each set of 4 quadruplets was rearranged to compose four groups of 4 words each. Each group contained one member of each quadruplet and only one member of each condition for length and frequency. No semantically or phonologically close words belonged to the same group. Three such rearrangements were prepared for each set. In this way, 9 rearrangements were obtained for the entire set of stimuli. Nine additional rearragements were made by changing the order of the groups of the 9 previously obtained rearrangements. This procedure resulted in 18 experimental lists consisting of 4 groups of 4 words each.

Four groups of 4 symbol pairs were also constructed (see Table 1). For each experimental list, we established a random mapping between group of symbols and group of words.

Table 1. Groups of symbols used in Experiment 2.

Vocabulary			
A	B	C	D
{ {	– –	++	; ;
° °	££	@@	<<

//	^^	**	))
==	!!	~~	&&

Design and procedure. Each participant was assigned to one list. The experiment was divided into four blocks for each group of 4 words. Each experimental session was subdivided into a learning phase, a training phase and a speeded naming phase. As soon as one block was over, another one began with the next group of words.

At the onset of the learning phase, participants saw the instructions displayed on the computer screen. They were instructed to learn each of the four written words and their associated symbols. When participants felt they had mastered the associations, they pressed a key to stop the learning phase and start the training phase. In this second phase, symbols were presented one-by-one on the screen. Participants were told to pronounce the corresponding word loudly and clearly and then press a key to verify whether their response was correct. When participants made no errors, they were told to start the next phase by pressing another key. The instructions for the speeded naming phase were then displayed and the four words and their corresponding symbols were presented again on the screen. In this phase, a symbol was presented on the screen and remained visible until a vocal response was detected by the computer. If the participant gave no response, the symbol disappeared after five seconds. The next trial was initiated after a 2 second delay. In each naming phase, each symbols was tested 5 times. The order was random with the constraint that the same symbol did not occur in consecutive trials. The first time a word was uttered was considered a practice trial and its reaction time was not recorded.

Overall, 64 responses and reaction times (4 x 4 x 4) were recorded for each participant.

Participants. The participants were the same as in Experiment 1. Experiments 1 and 2 were performed in separate sessions and Experiment 1 preceded Experiment 2.

Results

Three participants who either failed to learn the associations properly or made more than 15 % errors were replaced. The mean reaction times are shown in Table 2. A total of 7.5% of the responses (omissions, false starts, breaths, etc) were excluded from the analysis. 4.3% of the responses were excluded as outliers from the participants analyses and 5% from the items analyses. Two ANOVAs were conducted on naming latencies; one with participants and one with items as random variables. There were one between-participant factor (list) and two within-participant factors (number of syllables, frequency of words). There was no significant latency difference between mono and bisyllabic words ($F1(1, 15) = 2.24$, $p = .15$ and $F2(1, 36) = 0.75$, $p = .40$). High frequency words were produced significantly faster than low frequency ones ($F1(1, 15) = 12$, $p < .004$; $F2(1, 36) = 5.2$, $p < .03$). There was no interaction between length and frequency ($F1$ and $F2 < 2$, $p > 0.1$)

Table 2. Mean RT (in msec.), percent error, standard error (in msec.) for Experiment 2.

	Word length								
	Monosyllables			Bisyllables			Average		
	RT	Err	SE	RT	Err	SE	RT	Err	SE
High Frequency	640	4.3%	18	634	3.7%	13	637	4.0%	11
Low Frequency	688	8.7%	23	671	11.8%	61	679	10.3%	14
Average	664	6.5%	15	652	8.6%	10			

Analyses of errors yielded the same pattern of results as the analyses of the reaction times. Mono- and bisyllabic words did not differ in terms of number of errors ($F1(1, 15) < 1$, $p = .7$; $F2(1, 36) < 1$, $p = .6$). High frequency words were pronounced with significantly fewer errors than low frequency words ($F1(1, 15) = 6.4$, $p < .03$; $F2(1, 36) = 13.5$, $p < .001$).

Discussion

As in Experiment 1, naming latencies and errors were not affected by the length of the target words. We did find a frequency effect as in previous studies (Balota & Chumbley, 1984, 1985; Connine et al., 1990; Forster, 1981; Jescheniak & Levelt, 1994). Frequency effects are an indication that the paradigm is sensitive to word form retrieval. Taken together with the results of Experiment 1, Experiment 2 shows that naming is influenced by word frequency/familiarity but not by word length.

Our results contrast with previous findings (Eriksen et al., 1970; Klapp et al., 1973). However, both Eriksen and Klapp's experiments were run in English, whereas our study involves only French materials. Could it be that phonological planning in French speakers is different from that of English speakers? Recent investigations suggest that basic speech perception strategies often differ across languages (Cutler et al.,1983, 1986; Cutler & Norris, 1988; Mehler, Dommergues, Frauenfelder & Segui, 1981; Morais, Cary, Alegria & Bertelson,. 1986; Otake, Hatano, Cutler & Mehler, 1993; Pallier, Sebastian-Galles, Felguera, Christophe & Mehler, 1993; Sebastian-Galles, Dupoux, Segui & Mehler, 1992; Segui, Dupoux & Mehler, 1990). For instance, French participants are sensitive to syllables (Mehler et al., 1981; Segui, Dupoux & Mehler, 1990 for a review), Japanese participants to morae (Otake et al., 1993), and English participants to the rhythmicity resulting from the alternation of full and reduced vowels (Cutler et al., 1986; Cutler & Norris, 1988). A difference between French and English production routines, therefore, cannot be discarded on a-priori grounds. However, before speculating further about potential processing differences between the two languages, it would seem wise to check for the existence of a length effect in English using the same procedure as we used in French. This was the aim of the next three experiments.

EXPERIMENT 3: PICTURE NAMING IN ENGLISH

This experiment was carried out to assess the existence of a length effect in a picture naming task in English. The design of this experiment is similar in most respects to that of Experiment 1. The location of stress followed its typical English distribution, that is, most words were stressed in the first syllable. If speech processing in French is similar to that of English, no differences should be found between Experiments 1 and 3. In contrast, if a length effect is observed in English, phonological processing differences between English and French may be the cause.

Method

Materials and procedure. As in Experiment 1, thirty pictures were selected with corresponding words that were mono- (median frequency: 18 according to Kuçera and Francis (1967)) and bisyllabic (median frequency: 17.5). Most of the bisyllabic items had stress on the first syllable (see Appendix). The frequency and familiarity of the monosyllabic words matched those of the bisyllabic words. The procedure and the equipment were identical to those employed in Experiment 1.

Participants. Eighteen students from London University (11 men and 7 women), aged from 19 to 22 years, participated in this experiment. All were native speakers of British English.

Results

A total of 6.7% responses were excluded from the analysis. There were 2% unexpected productions and 3.7% false starts and technical problems. 5.3% of the responses in the participants analysis and 5.6% in the items analysis were rejected as outliers. Two ANOVAs were conducted on naming latencies and errors; one with participants and one with items as random variables. There was one within-participant factor (number of syllables). There was no significant latency difference between mono- and bisyllabic words (601 msec. *vs* 602 msec. respectively; $F1(1, 17) < 1$, $p = .79$; $F2(1, 14) < 1$, $p = .87$). Analyses of errors yielded the same results. That is, there was no significant difference between mono- and bisyllabic words (7.2% *vs* 6% errors, $F1(1, 17) = 1$, $p = .3$ and $F2(1, 14)=1.4$, $p = .3$).

A post-hoc multiple regression analysis was conducted to evaluate the effects of familiarity and lexical frequency. A familiarity effect appeared (familiarity: $t(29) = -2.65$, $p < .02$) while there was no frequency effect ($p > .1$).

Table 3. Mean RT (in msec.), percent error, standard error (in msec.) for Experiment 4.

	Word length								
	Monosyllables			Bisyllables			Average		
	RT	Err	SE	RT	Err	SE	RT	Err	SE
High Frequency	639	5.0%	16	656	7.5%	16	647	6.2%	11
Low Frequency	687	4.3%	27	677	5.6%	22	682	4.9%	17
Average	663	4.6%	16	667	6.5%	14			

Discussion

This experiment used English materials, and no length effect was found. This outcome is quite similar to that of Experiment 1 that used French materials and identical procedure. Up to this point, there is hence no evidence of a processing difference between French and English. However, our results are at odds with previously obtained length effects (Eriksen et al., 1970; Klapp et al. 1973). Given the importance of the length effect as an indicator of the phonological planning process, we decided to look for this effect again, using a procedure similar to that used in Experiment 2.

EXPERIMENT 4: NAMING OF ARBITRARY SYMBOLS IN ENGLISH

Experiment 4 was designed to assess the influence of frequency and length on naming latencies. To avoid the limitations imposed by pictures, Experiment 4 used naming of arbitrary symbols. Naming latencies of mono- and bisyllabic words of high or low frequency were compared. In this experiment (as in the corresponding French experiment), stress location was kept constant. In contrast with French, however, the stress was always located on the first syllable, which is the most typical stress pattern in English (Cutler & Carter, 1987). Experiment 4 was similar to Experiment 2, with minor differences as indicated below.

Method

Materials and procedure. Twelve quadruplets containing one pair of mono- and one pair of bisyllabic English words were constructed. The members of each quadruplet shared both the initial phoneme and the structure of the first syllable. In each pair, one word had a low frequency of occurence (less than 5 per million according to Kuçera and Francis, 1967) and the other a high frequency of occurence (more than 100 per million), for example: "sash"/"sun", and "suction"/"section". Stress was always on the first syllable. The 18 lists were constructed in the same way as in Experiment 2 (see Appendix).

The procedure and the equipment were identical to those employed in Experiment 2. As in French, the first production of each word was considered as a practice trial and was not recorded.

Participants. The participants were the same as in Experiment 3. Experiment 3 and 4 were performed in the same session, with a ten minute pause between experiments. Experiment 3 preceded Experiment 4.

Results

One participant who failed to learn the associations properly was replaced. Mean reaction times across length are presented in Table 3. A total of 5.5% responses was excluded from the analyses (1.2% unexpected productions and 3.3% false starts and technical problems, while 4.3% of the data in the participant analysis and 5.4 % in the item analysis were rejected as outliers.

Two ANOVAs were conducted on naming latencies and errors; one with participants and one with items as random variables. There were one between-participant factor (list) and two within-participant factors (number of syllables, frequency of words). There was no significant latency difference between mono- and bisyllabic words ($F1(1, 15) = .02$, $p = .9$ and $F2(1, 36) = 0.25$, $p = .62$). High frequency words were produced significantly faster than low frequency words (36 msec. difference, $F1(1, 15) = 5.18$, $p < .04$ and $F2(1, 36) = 3.83$, $p = .058$). There was no interaction between frequency and length ($F1$ and $F2 < 1$, $p > 0.1$).

Finally, there were no differences in error rates either between mono- and bisyllabic words or between high and low frequency words, and there was no interaction between length and frequency ($F1(1, 15)$ and $F2(1, 36) < 1$).

Discussion

In this experiment, we used English materials and found results very similar to those of Experiment 2 with French materials. In both cases we failed to find a length effect. The frequency effect found in both experiments indicates that the paradigm was sensitive to phonological encoding (Jescheniak & Levelt, 1994).

In Experiment 3, stress location was not controlled but the first syllable of the majority of the items was stressed, following the typical stress pattern of English. In Experiment 4, however, the location of stress was fixed on the first syllable. The experiments by Klapp et al. (1973) and Eriksen et al. (1970) were not designed to study the effect of stress location, and they generally contained words with stress on the first syllable. Although differences due to stress between these studies and ours seem unlikely, we wanted to carry out one last experiment where stress location was systematically varied. It is possible that if stress is kept constant throughout the experiment, participants may use a strategy that takes into account this special property of the list. As a result, potential effects of length could have been missed.

This final experiment also controlled for another potential difference between our study and previous ones. Although Klapp et al. used mono- and bisyllabic items as we did, Eriksen et al. used words containing 2, 3, and 4 syllables. Hence, length effects may be more apparent when longer words are used, and for this reason we used mono-, bi-, and trisyllabic words in our final experiment.

EXPERIMENT 5: NAMING OF ARBITRARY SYMBOLS AND ROLE OF STRESS IN ENGLISH

The aim of this experiment was to test the generality of the findings of Experiments 2 and 3, by using more varied materials. We used mono-, bi-, and trisyllabic English words in order to increase length differences, and hence the likelihood of finding a length effect. In order to avoid potential strategies due to a fixed stress pattern throughout the experiment, half of the polysyllabic items had stress on the first syllable, and half on the second.

Method

Material. We constructed 8 sextuplets containing a pair of mono-, a pair of bi-, and a pair of trisyllabic words each (median frequency: 30, 38, 34, respectively; see Appendix). All members of a sextuplet had the same initial phoneme and the same first syllable structure (e.g., CV or CVC). In one member of each bisyllabic and trisyllabic pair, the first syllable was stressed, while in the other, stress was on the second syllable. Frequency was matched across number of syllables and across stress location. The eight sextuplets were divided into four sets of four triplets each. Each set of triplets contained two triplets carrying stress on the first syllable, and two triplets with stress on the second. Each set was rearranged to make a list consisting of four groups of words. Each group contained a mono-, a bi-, and a trisyllabic word. Three rearrangements were made for each set. Two words never occurred together in two different groups. There were no semantically or phonologically close words in the same group. This resulted in 12 lists for the entire set of stimuli. Twelve more lists were created by changing the order of the groups of the 12 previously obtained lists. In this way, a total of 24 experimental lists of 4 groups of 3 words each were constructed.

Procedure. Each of the 24 participants was assigned to a list. The procedure and the design were identical in most respects to that of Experiments 2 and 4. In contrast with Experiment 2 and 4, though, the four experimental sessions involved groups of three words rather than four. Moreover, the number of trials was smaller (60 per participant). As before, the first presentation of the word was considered a practice trial and the answer was not recorded. The equipment was identical to that used in the previous experiments.

Participants. Twenty-four English students from London University (15 men and 9 women), aged from 18 to 41 years, participated in this experiment. All were native speakers of English. None had participated in the other experiments.

Results

The mean reaction times are displayed in Table 4. A total of 3.7% responses were excluded from the analyses (1.5% errors and 2.2% false starts and technical problems). Furthermore, 2.5% of the data in the participant analysis and 4.6% in the item analysis were excluded as outliers. Two ANOVAs were conducted on naming latencies; one with the participants and one with the items as random variables. There was one between-participant factor (list) and two within-participant factors (number of syllables, stress location). There were no significant latency difference either between mono-, bi-, and trisyllabic words ($F1(2, 40) < 1$, $p = .40$, $F2(2, 24) < 1$, $p = 0.98$), or between words carrying stress on the first or on the second syllable ($F1(1, 20) < 1$, $p = .74$, $F2(1, 24) < 1$, $p = .52$). There was also no interaction between stress position and length (F1 and F2 < 2, $p > .1$).

As the monosyllabic words cannot be considered to carry stress on the second syllable in this experiment, four restricted ANOVAS were performed. In the first, reaction times for mono-, bi-, and trisyllabic

Table 4. Mean RT (in msec.), percent error, standard error (in msec.) for Experiment 5.

	Word length								
	Monosyllables			Bisyllables			Trisyllables		
	RT	Err	SE	RT	Err	SE	RT	Err	SE
First syllable stress	603	3.7%	21	645	3.7%	31	629	2.5%	19
Second syllable stress	627	6.2%	22	623	3.7%	25	619	3.7%	24
Average	615	4.8%	15	634	3.7%	20	624	3.1%	15

targets with stress on the first syllable were compared. The second analysis was restricted to the latencies for mono- and bisyllabic words. In the third, latencies of bisyllabic and trisyllabic words with stress on the first or second syllable were analyzed. Finally, in the last analysis, mono- and trisyllabic words were compared. None of these ANOVAs yielded an effect of either length or stress (F1 and F2 < 2, p > .1).

There were no significant effect of number of syllables or stress on the errors; no interactions between these factors emerged either in the combined or in the restricted analyses (F1 and F2 < 1, p > .1).

Discussion

In this experiment, we found no significant effect of length although we used a wider range of length and variations in word stress. However, our results show a numerical trend for subjects to name monosyllables with shorter latencies than polysyllabic items. This trend does not approach statistical significance but it is interesting and merits future studies. Possibly, adding variability in both stress and length promotes such a trend, although not enough to make it significant (Colombo, 1992). It is also possible that this trend is simply due to error of measurement. Note that the trend does not appear to be stronger for one pattern of stress than for another. In fact, we found that the latencies for the triplets of items carrying stress on the second syllable were not significantly different from those of the triplets carrying stress on the first syllable.

GENERAL DISCUSSION

We reported the results from five naming experiments. Experiments 1 and 2 reported the behavior of French participants naming French words, while Experiments 3, 4 and 5 reported results of English participants naming English words. In all experiments, word length was the dependent variable, while frequency and/or familiarity was controlled. Moreover, for English, Experiment 5 explored the role of stress variability. In none of the five experiments did we find that naming latencies or errors were significantly related to the number of syllables in the target. Neither did we find that stress location affects naming latencies for English items. Two procedures were used. In Experiment 1 and 3, we used picture naming, while in Experiments 2, 4 and 5, we used the symbol naming task of Levelt and Wheeldon (1994).

How can we account for the absence of a length effect in our data? How likely is it that we missed a true but small length effect? Klapp et al. (1973) found a 14 msec. difference between mono- and bisyllabic items. Eriksen et al. (1970) did not directly compare mono- and bisyllabic items, and measured naming times for words containing 2, 3 and 4 syllables. They found a 17 msec. difference in 2 versus 3 syllable words[1]. On the basis of these results, let us assume that the expected value of a potential length effect is of the order of 15 msec. How can we assess the likelihood that we missed a length effect of 15 msec. or more?

To assess the robustness of our experiments, we performed two kinds of analyses. In the first analysis, the length effect for each participant was computed separately for each experiment. Within each experiment, we computed the average and variability of the length effect, from which confidence intervals were derived (two standard deviations of the mean). The first two experiments in French allow us to discard a potential 15 msec. effect (in fact, they allow

[1] Note that they found only a (presumably nonsignificant) 5 msec. effect for 3 versus 4 syllables.

us to discard a length effect of 11 msec. or more, at $p < .05$). The data of the final three experiments are somewhat noisier and do not allow us to reject a 15 msec. effect (although the third experiment allowed us to reject a 17 msec. effect). The second analysis rests on the distribution of mean length effects across experiments. If the underlying effect was 15 msec., we should have found a distribution of effects centered around 15 msec. However, we found effects distributed around 1 msec. (-4, -11, 1, 3, 14). A non-parametric rank test for location confirmed that our sample had a median significantly deviant from 15 msec. ($Z = 1.99$, $p < .05$)[2]. Consequently, we feel confident in rejecting the existence of a length effect of a size comparable to what was reported by Klapp et al. and Eriksen et al. (Our own estimate of the putative length effect is 1.3 msec., plus or minus 9.7 msec. We obtained this estimate by computing the mean length effect across participants in our five experiments, and computing a $p<.05$ confidence interval.).

How can we account for the discrepancy between our 1.3 msec. and the reported 14 msec. and 17 msec. length effects found in Klapp et al. and Eriksen et al.? As mentioned in the introduction, these two studies had some shortcomings. In particular, the Klapp et al. (1976) experiment controlled for frequency but not for familiarity, while the Eriksen et al. experiment did not control for frequency. From our results, we estimate that large differences in frequency can give rise to a 35-40 msec. advantage for frequent words. It is plausible that smaller differences in frequency or familiarity may have resulted in the observed 14 msec. and 17 msec. of Eriksen et al. and Klapp et al.

[2] For Experiment 5, we used the mean difference between mono and polysyllabic items.

As mentioned in the introduction, some studies using other paradigms have reported that the length of an utterance to be produced influences production latencies. Most notably, Sternberg et al. (1978, 1980) used a list repetition paradigm in which participants had to retain a list composed of several items and produce it whenever a response signal appeared on the screen. The latency between the response signal and the onset of the utterance was linearly related to the number of items in the list (about 10 msec. for each additional item). However, such a "length effect" is very different from the one we have discussed. First, participants had about four seconds to study the list and prepare to respond, a situation similar to delayed naming. Under these conditions, people can presumably prepare a large part of the plan before overt articulation, and hence latencies may only partially reflect the building up of the plan. Indeed, Sternberg et al. interpret latencies as reflecting the time needed to transfer a previously constructed motor program to the articulators (motor preparation). Second, the effects found with this paradigm are quite different from the effects found in immediate naming. Sternberg et al. argue that the critical variable is not the number of syllables, but rather the number of words. Indeed, the slope of the latency function was not influenced by the addition of unstressed syllables to each word (e.g. bay-rum-cow *vs* baby-rumble-coward)[3]. These results are difficult to compare with results obtained in immediate naming. Indeed, in the Klapp et al. and Eriksen et al. naming studies, only one word was tested, and length was manipulated by the addition of (mostly

[3] In fact, the addition of unstressed syllables between words (bay-and-rum-and-cow) did not change the slope either. The authors concluded that the relevant unit was the stress group (or metrical foot), and not the word.

unstressed) syllables. In our experiments, we were interested in the time it takes to build the phonological plan, and we tried to keep constant the variables that affect other processes. In particular, we neutralized the factors that affect motor preparation by keeping constant the number of words, as well as the number of stressed syllables to be pronounced. Thus, the absence of length effect in our studies is not contradictory with the results of Sternberg et al. (1978; 1980).

Let us grant that word length does not affect naming latencies. Such a result obliges us to reevaluate some basic assumptions made by models of speech production. The absence of a length effect can be accounted for in two ways. The first assumes that the phonological plan is stored as a whole in the lexicon, or that the phonological plan is built in a massively parallel fashion. Following such a view, it takes the same amount of time to compute the plans for short or long words. The second possible interpretation is that plans are built in a sequential fashion, but that speakers can start uttering a word before the phonological plan has been completed. As we mentioned in the introduction, there is one model that is incompatible with the absence of a length effect: i.e., the one which claims that encoding is sequential *and* that people cannot begin speaking until they have finished computing the plan for the whole word. Thus, our results force us to make the following exclusive choice: Either the segments of a word are encoded in parallel, or articulation can be initiated before the entire phonological plan for the word is available.

Evidence in favor of serial encoding has been accumulating fast. Meyer (1990, 1991), using an implicit priming paradigm based on a paired-associate learning task, showed that when the first segment of a word is predictable, naming latencies are shorter. The more predictable the segments from left to right, the shorter the naming latencies. In contrast, predictable segments starting from the right do not result in a decrease in naming latencies. This suggests that both phonemes and syllables are encoded from left to right. Sevald and Dell (1994) also found results that led them to reject a parallel encoding model: the time to repeat a sequence of items is shorter when the items share the coda than when they share the onset of the syllable. The sequence "pick tick pick tick" for example, requires less time to repeat than "pick pin pick pin". Such differences would not appear in a purely parallel model. Using a picture-word interference paradigm, Meyer and Schriefers (1991) found phonological facilitation when the probe was related to the onset of the name of the picture and was presented 150 msec. before or after the picture. However, they could only find facilitation for the end of the picture name if the probe was presented simultaneously or 150 msec. after the picture. The authors also concluded that the phonological encoding of a word proceeds incrementally. Finally, using phoneme monitoring in a covert translation task, Wheeldon and Levelt (1995) found that earlier phonemes in a word are available before later phonemes. In Experiment 1, they found that the difference in phoneme monitoring time in the first versus the second syllable was 123 msec. (and 72 msec. in a replication with articulatory suppression). In Experiment 3, they found a difference of 111 msec. for onset phonemes and 69 msec. for coda phonemes. Costa, Sebastian and Pallier (submitted), using a similar task applied to picture naming, showed that word onsets were available before the other phonemes of the word, with a difference in latencies comparable to that reported by Wheeldon and Levelt (1995).

On the basis of the above experiments, one can estimate that the speed of serial encoding is in the range of 70-150 msec. per syllable (which is faster than the average articulation time for syllables in continuous speech: 150-200 msec.). Even if we halve the low estimate of serial encoding to 30 msec., our experiments should have detected such a difference. Thus, we have to ask why naming times are not affected by word length in any of our experiments. We suggest that people (at least in our experiments) do not need to complete the plan for the whole word before they start speaking. We are aware, however, that several researchers claim to have found empirical support for the notion that subjects cannot start articulating before the whole word plan has been completed. For instance, Levelt and Wheeldon (1994) found that naming latencies are related to the syllabic frequency of the *second* syllable of a word (but not the first). They claim that their finding implies that articulation requires full encoding of all the syllables in a word. How could the frequency of the last syllable have an impact on naming if subjects do not wait until the second syllable becomes available? Similarly, Meyer and Schriefers (1991) found that, apparently, articulation does not start until the second syllable of a word is planned. How can we reconcile these discrepant results? We can offer at best some speculations. Many processing models accept that encoding arises simultaneously at different levels of a hierarchically organized structure, including phonemes, syllables, phonological words, and so on. It is possible that, depending on task demands, articulation can be initiated using different levels of the hierarchy. Presumably, the larger the utterance to be produced, the more advance planning is necessary, and hence the higher the level of the hierarchy that will be concerned. Likewise, different pressures on speed or accuracy may cause people to use a planning strategy based on a smaller or larger structural level. The notion that planning is flexible is supported by recent work by Schriefers and Teruel (submitted) and Huitema (1994). Of course, the extent and conditions of this flexibility have strong methodological implications on the use of naming latencies as an index of planning time (see Lupker, Brown & Colombo, 1997). More experiments are needed to explore these issues.

REFERENCES

Balota, D. A. & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, **10**, 340-357.

Balota, D. A. & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production. *Journal of Memory and Language*, **24**, 89-106.

Butterworth, B. (1980). Some constraints on models of language production. In B. Butterworth (Ed.), *Language Production: Vol. I. Speech and Talk* (pp. 423-459). San Diego, CA: Academic Press.

Butterworth, B., (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 108-135). Cambridge, MA: MIT Press.

Colombo, L. (1992). Lexical stress effect and its interaction with frequency in word prononciation. *Journal of Experimental Psychology: Human Perception and Performance*, **18**, 987-1003.

Connine, C. M., Mullenix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **16**, 1084-1096.

Content, A., Mousty, P., & Radeau, M. (1990). Brulex, une base de données lexicales informatisée pour le français écrit et parlé. *L'Année Psychologique*, **90**, 551-566.

Costa, A., Sebastian, N., & Pallier, C. The time course of phonological encoding: sequential and parallel processing. (submitted)

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access.

Journal of Experimental Psychology: Human Perception and Performance, **14**, 113-121.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1983). A language-specific comprehension strategy. *Nature*, **304**, 5922: 159-160.

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

Cutler, A. & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, **2**, 133-142.

Dehaene, S. & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, **43**, 1, 1-29.

Dell, G. S. & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, **20**, 611-629.

Dell, G. S. & O'Seaghdha, P.G. (1992). Stages of lexical access in language production. *Cognition*, **42**, 287-314.

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, **93**, 283-321.

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, **27**, 124-142.

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

Forster, K. I. & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, **12**, 627-635.

Forster K. I. (1981). Frequency blocking and lexical access: One mental lexicon or two? *Journal of Verbal Learning and Verbal Behavior*, **20**, 190-203.

Fraisse, P. (1964) Le temps de réaction verbale: I. - Dénomination et lecture. L'*Année Psychologique*, **64**, 21-46.

Frederiksen, J. R. & Kroll J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 361-379.

Gougenheim, G., Rivenc, P., Michea, R., & Sauvageot, A. (1956). *L'élaboration du français élémentaire*, Paris: Didier

Huitema, J.S. (1994). Referential Planning and Phonological Retrieval in Sentence Production, *Cognitive Science Technical Report*, Beckman Institute: University of Illinois.

Jared, D. & Seidenberg, M. S. (1990). Naming multisyllabic words. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 92-105.

Jescheniak, J. D. & Levelt, W. J. M. (1994). Word frequency effects in speech production: retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **20**, 4, 824-843.

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

Kuçera, H. & Francis, W.N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

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

Levelt, W. J. M. (1989). In Aravin Joshi, (Ed.). *Speaking: From Intention to Articulation*. Cambridge, Mass: The MIT Press.

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

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

Marslen-Wilson, W. D. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, **244**, 522-523.

Mehler, J., Dommergues, J., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, **10**, 298-305.

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language*, **29**, 524-545.

Meyer, A. S. (1991). The time course of phonological encoding in language production:

Phonological encoding inside a syllable. *Journal of Memory and Language*, **30**, 69-89.

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitaion in picture-word interference experiments: eefects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **6**, 1146-1160.

Monsell, S., Doyle, M-C., Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental psychology: General*, **118**, 43-71.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phonemes arises spontaneously. *Cognition*, **24**, 45-64.

Morton, J. & Patterson, K. (1980). A new attempt at an interpretation, or, an attempt at a new interpretation. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia*, London: Routledge & Kegan Paul.

Newcombe, F., & Marshall, J. C. (1980). Transcoding and lexical stabilisation in deep dyslexia. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia*, London: Routledge& Kegan Paul.

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

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory an Language*, **32**, 258-278.

Pallier, C., Sebastian-Gallés, N., Felguera, T., Christophe, A., & Mehler, J. (1993). Attentional allocation within the syllabic structure of spoken words. *Journal of Memory and Language*, **32**, 373-389.

Sebastian-Galles, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in catalan and spanish. *Journal of Memory and Language*, **31**, 18-32.

Segui, J., Dupoux, E. & Mehler, J. (1990). The role of the syllable in speech segmentation, phoneme identification, and lexical access. In G. Altman (Ed.), *Cognitive models of speech processing*, (pp. 263-280). Cambridge, MA: MIT Press

Sevald, C. A. & Dell, G. S. (1994). The sequential cueing effect in speech production. *Cognition*, **53**, 91-127.

Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial ordering mechanism in sentence production. In *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*, (pp. 295-342). Hillsdale, New Jersey: Lawrence Erlbaum Associates (Eds).

Schriefers, H., & Teruel, (submitted). Phonological facilitation in the production of two-word utterances.

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

Snodgrass J. G. & Vanderwart M., (1980). A standardized set of 260 pictures: norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, **6**, 174-215.

Sternberg S., Monsell S., Knoll R. L., & Wright C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning*, (pp. 117-152). New York: Academic Press.

Sternberg S., Wright C. E., Knoll R. L., & Monsell S. (1980). Motor programs in rapid speech: additional evidence. In R. A. Cole (Ed.), *Perception and production of fluent speech*, (pp. 469-505). Hillsdale, NJ: Lawrence Erlbaum.

Thorndike, E. L. & Lorge, I. (1944). *The teacher's word book of 30.000 words.* New York: Columbia University, Teachers College Press.

Wheeldon, L.R. & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language*, **34**, 311-334.

Wingfield A. (1968). Effects of frequency on identification and naming of objects. *American Journal of Psychology*, **81**, 226-234.

APPENDIX: MATERIALS FOR EXPERIMENTS

Experiment 1

Monosyllables		Bisyllables	
chaise	*(chair)*	chapeau	*(hat)*
chat	*(cat)*	chemise	*(shirt)*
coq	*(cockerel)*	canon	*(canon)*
cygne	*(swan)*	serpent	*(snake)*
feuille	*(leaf)*	fourchette	*(fork)*
gant	*(glove)*	guitare	*(guitar)*
jupe	*(skirt)*	girafe	*(giraffe)*
lit	*(bed)*	lapin	*(rabbit)*
pied	*(foot)*	piano	*(piano)*
pipe	*(pipe)*	pinceau	*(paintbrush)*
poire	*(pear)*	poisson	*(fish)*
mouche	*(fly)*	maison	*(house)*
scie	*(saw)*	sifflet	*(whistle)*
singe	*(monkey)*	soleil	*(sun)*
tasse	*(cup)*	tortue	*(tortoise)*

Note: The mean log word frequency and the mean familiarity score were 1.47 and 3.54 respectively for the monosyllables and 1.33 and 3.19 respectively for the bisyllables.

Experiment 2

Monosyllables		**Bisyllables**	
Low Freq	**High Freq**	**Low Freq**	**High Freq**
	Set 1		
cosse	cause	cardan	question
(hull)	*(cause)*	*(universal joint)*	*(question)*
pôpe	peur	pactole	personne
(pope)	*(fear)*	*(booty)*	*(person)*
faon	fond	félon	figure
(fawn)	*(bottom)*	*(felon)*	*(figure)*
jante	genre	gerçure	jardin
(felloe)	*(gender)*	*(crack)*	*(garden)*
	Set 2		
as	âge	acné	argent
(ace)	*(age)*	*(acne)*	*(money)*
dague	doute	dolmen	docteur
(dagger)	*(doubt)*	*(dolmen)*	*(doctor)*
sauge	soeur	sulfure	soldat
(sage)	*(sister)*	*(sulfide)*	*(soldier)*
rôt	rue	robot	regard
(roast)	*(street)*	*(robot)*	*(gaze)*
	Set 3		
pal	page	pastis	partie
(pale)	*(page)*	*(pastis)*	*(piece)*
soude	sol	servage	service
(soda)	*(ground)*	*(serfdom)*	*(service)*
vamp	vague	vodka	vertu
(vamp)	*(wave)*	*(vodka)*	*(valour)*
gigue	jambe	jersey	journal
(jig)	*(leg)*	*(jersey)*	*(newspaper)*

Note: The mean log word frequency were 1.57 and 2.29, respectively, for the monosyllabic low and high frequency groups and 1.58 and 2.28, respectively, for the bisyllabic low and high frequency groups.

Experiment 3

Monosyllables	Bisyllables
axe	ashtray
bear	barrel
bell	basket
belt	button
cat	camel
comb	candle
cow	cannon
cup	carrot
flag	flower
goat	guitar
lamp	ladder
pear	pencil
pig	penguin
pipe	piano
sock	cigar

Note: The mean log word frequency and the mean familiarity score were 1.17 and 3.2, respectively, for the monosyllables and 0.99 and 2.9, respectively, for the bisyllables.

Experiment 4

Monosyllables		Bisyllables	
Low Freq.	**High Freq.**	**Low Freq.**	**High Freq.**
Set1			
beak	book	bismuth	business
dune	death	dolmen	distance
mane	man	mistress	member
soy	sea	savor	season
Set2			
bough	boy	beaver	basis
pock	peace	panther	picture
nun	name	nimbus	number
sash	sun	suction	section
Set3			
cud	cause	cactus	congress
dam	date	dolphin	doctor
lair	line	lancet	language
chess	chief	charger	children

The mean log word frequency were 0.22 and 2.32, respectively, for the monosyllabic low and high frequency groups and 0.08 and 2.25, respectively, for the bisyllabic low and high frequency groups.

Experiment 5

Stress First			Stress Second		
Monosyll.	**Bisyll.**	**Trisyll.**	**Monosyll.**	**Bisyll.**	**Trisyll.**
Set 1					
gay	guidance	governor	bay	balloon	behavior
piece	pencil	parliament	moon	monsoon	mosquito
Set 2					
tile	target	terminus	dew	degree	division
ray	razor	relative	guy	guitar	gorilla
Set 3					
bee	barrel	battery	tool	technique	temptation
cat	capture	confidence	rye	relief	reduction
Set 4					
meat	mention	mystery	pot	pursuit	perception
doe	detail	dominance	cup	cartoon	component

Note: The mean log word frequency were 1.30, 1.53 and 1.35, respectively,for the monosyllables, bisyllables and trisyllables belonging to the group with stress first and 1.36, 1.29 and 1.24, respectively, for the group with stress second.