JOURNAL OF MEMORY AND LANGUAGE 29, 316–335 (1990)

Monitoring the Lexicon with Normal and Compressed Speech: Frequency Effects and the Prelexical Code

EMMANUEL DUPOUX

Ecole Nationale Supérieure des Télécommunications and Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS/EHESS, France

AND

JACQUES MEHLER

Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS/EHESS, France

Previous reports suggest that initial phonemes are monitored on the basis of lexical information in monosyllabic words and on the basis of acoustic/phonetic information in multisyllabic words (Cutler, Mehler, Norris, & Segui, 1987). In Experiment 1, a frequency effect was found with item-initial phoneme monitoring for monosyllabic but not for bisyllabic words. In Experiments 2 and 3, we used speech time-compressed at a rate of 50% and failed to find a frequency effect for bisyllabic words, even though they were shorter than uncompressed monosyllables. In Experiment 4, we used a lexical decision task on the same items and found a frequency effect for both mono- and bisyllabic words. Results are interpreted on the basis of the dual code hypothesis. Implications for the nature of the prelexical code are discussed. © 1990 Academic Press, Inc.

The nature of the perceptual access code by which lexical entries are activated is still a matter of debate. The most common belief in spoken word recognition is that the speech signal is continuously transformed and processed and that even a few ms of speech stimuli can broadly activate the lexicon. This very diffuse lexical activity can be narrowed down as soon as new information is available, resulting in the isolation of a unique lexical candidate (see Marslen-Wilson & Tyler, 1980; McClelland & Elman, 1986). However, some psycholinguists believe that the signal is first parsed into large prelexical units, for instance, syllables, that are in turn the source of lexical look-up (see Massaro, 1972, 1975; Mehler, Dommergues, Frauenfelder, & Segui, 1981). In this view, lexical activation is essentially discontinuous, because nothing happens in the lexicon before a critical amount of prelexical information has been processed.

In this study, we will explore the nature of prelexical units in three experiments using word initial phoneme monitoring. This task seems adequate to explore the first stages of lexical activation because it has been shown to be sensitive both to acoustic/phonetic factors and to lexical factors (Cutler & Norris, 1979; Foss & Blank, 1980; Newman & Dell, 1978). Indeed, it is

We would like to thank Juan Segui and the members of the Laboratoire de Sciences Cognitives et Psycholinguistique for their help, comments, and fruitful discussion; Thierry Bommard for his technical help with speech-editing and compression routines; and Pierre Clergeot and Denis Acker for access to a large number of subjects and testing facilities at the Ecole Sainte Genevieve, Versailles. We would also like to thank Dr. D. A. Balota, Dr. N. F. Johnson, and Dr. David Birdsong for their helpful comments. This research was carried out with the help of CNET (Convention 837BD28 00790 9245 LAA/TSS/CMC), CNRS (ATP "Aspects Cognitifs et Neurobiologiques du Langage"); the European Science Foundation (TW 86/17); and the Ministére de la Recherche et de l'Enseigne ment Supérieur décision N°. 84C1390. Reprint requests should be sent to Emmanuel Dupoux, Laboratoire de Sciences Cognitives et Psycholinguistique, 54, Bd Raspail, 75270 Paris Cedex 06, France.

0749-596X/90 $3.00
Copyright © 1990 by Academic Press, Inc.
All rights of reproduction in any form reserved.

no longer possible to maintain that a single processing level is involved in phoneme monitoring. Morton and Long (1976) argued that phoneme monitoring responses were dependent upon the prior identification of the word-bearing item. This claim was undermined by a number of findings showing that under certain conditions, subjects could perform phoneme monitoring without being at all affected by lexical factors (Foss & Blank, 1980; Foss & Gernsbacher, 1983; Segui, Frauenfelder & Mehler, 1981). Moreover, it has been shown that phoneme detection is sensitive to many prelexical low-level factors. Response times are affected by vowel quality and vowel length in the target-bearing syllable (Diehl, Kluender, Foss, Parker, & Gernsbacher, 1987; Foss & Gernsbacher, 1983). RTs are also sensitive to the syllabic structure of the target bearing syllable (Cutler, Mehler, Norris, & Segui, 1987; Treiman, Salasoo, Slowiaczek, & Pisoni, 1982), and to the phonetic similarity between the word preceding the target and the target itself (Newman & Dell, 1978).

However, it would unfair to claim that phoneme monitoring is always sensitive to low-level factors. Contrary to Foss and Gernsbacher (1983), it seems that under certain conditions, phoneme detection can also be sensitive to lexical or postlexical factors. For one thing, lexical effects have been reported in sentences (Morton & Long, 1976; Segui, 1984). Moreover, in lists of words, Rubin, Turvey, and Van Gelder (1976) found a lexical superiority effect for monosyllabic items. That is, phonemes are detected faster in monosyllabic words than in similar, legal non-words. This effect has been successfully replicated by Cutler et al. (1987).

To sum up the above results, phonemes can be detected on the basis of two separate sources of information, lexical and prelexical. The lexicon contains a phonemic-like representation of words since subjects can pronounce and write them. Therefore, it is possible for subjects to detect a phoneme on the basis of the stored phonemic information that becomes available when a word is accessed. When a response is triggered after the lexical item has been accessed we say that the *lexical code* was used. However, subjects can also detect phonemes before lexical access has occurred, for instance when the target appears at the initial position of a very long word, or in a nonword. In that case, subjects are using information that is available at the *prelexical code,* most probably by performing an acoustic/phonetic conversion.

The *Dual Code* hypothesis (Cutler & Norris, 1979; Foss & Blank, 1980; Newman & Dell, 1978) acknowledges that both lexical and prelexical information can play a role, depending on the experimental conditions. Generally speaking, the lexical code is used when responding to short words (i.e., monosyllables) and not to polysyllabic words (Mehler, 1981). However, attentional factors can override this result. Homogenous lists of monosyllabic items can determine a shift towards the prelexical code (Cutler et al., 1987). In contrast, detecting a phoneme in an arbitrary position rather than in word-initial position favors reliance on the lexical code even with bisyllabic items (Frauenfelder & Segui, 1989; Marslen-Wilson, 1984). Moreover, Cutler and Norris (1979) outlined that if word identification is speeded by contextual focus, or by high predictability of the word in context, then the lexical code is likely to be used.

In this paper, we will use word initial phoneme monitoring which is sensitive to both codes. By manipulating the length of the target bearing item it is possible that latencies to short words can be influenced by lexical factors, whereas longer words are influenced mostly by prelexical factors. If a different code is used for short and long words it is possible to explore the nature of the prelexical unit.

Indeed, let us suppose that speech is segmented in terms of a rather coarse grained unit like the syllable, as was suggested for French by Mehler et al. (1981). Then one should expect to find a qualitative difference in the way monosyllabic and bisyllabic items are accessed. Indeed, if syllables are basic units, monosyllabic words might benefit from a direct access to the lexicon, whereas bisyllabic words might rely on further processing. In the latter case, phoneme monitoring responses could be issued before lexical access, on the basis of the prelexical information already available in the first syllable of the word.

Other researchers however (Klatt, 1989; Marslen-Wilson & Tyler, 1980; McClelland & Elman, 1986; among others) postulate that a finer grained unit is in operation during lexical access. Some go so far as to suggest that until the identification point is reached, the uptake of information is continuous. In their view, the relevant variable is the duration between word onset and identification point. Continuous models can also account for the effect of word length. In the case of a short word (such as a monosyllable), the identification point occurs very early, thus one should expect lexical involvement. In contrast, since bisyllabic words are usually longer, they are identified later, and the prelexical processing can be accomplished without any influence from the lexical level (Cutler & Norris, 1979).

In order to distinguish between these two types of models, we need a reliable diagnosis to assert whether one code or another has been used. The Dual Code Hypothesis has been mainly explored with experiments that manipulate the lexical status of target words (see Cutler et al., 1987). When latencies for words are faster than for nonwords, it is argued that responses (for words) are triggered from the lexical code. In contrast, when words and nonwords are responded to with comparable RTs it is generally claimed that all responses are elaborated at the prelexical code. However, the evidence is indirect, because it rests on the presupposition that responses from each one of the codes take a different amount of time. It would be preferable to compare responses derived from the same code. Moreover, the possibility remains that the presence of pseudo-words in the lists changes the subjects' normal behavior. In fact, subjects may stop using the lexical code because exclusive reliance on it might be a source of errors. If this is the case, using words and nonwords might induce subjects to shift their reliance from the lexical to the prelexical code. Our aim, therefore, is to explore whether both codes can be operational when only words are used. Thus, we will try to generalize previous results in an experiment in which frequency rather than lexical status is manipulated. Word frequency is one of the strongest indicators of lexical access and plays a prominent role in most word recognition models (Becker, 1976; Bradley & Forster, 1987; Forster, 1976, 1978, 1979;Glanzer & Ehrenreich, 1979; Morton, 1969, 1979, 1982; Norris, 1986). Thus, the word frequency effect can help to diagnose which stage of lexical processing underlies the phoneme monitoring response. We will assume that the lexical code has been used in order to detect a phoneme target if and only if latencies are correlated with word frequency.[1]

Experiment 1

Phoneme monitoring, as mentioned above, can be affected by lexical factors when the target is in initial position of short

[1] Of course, such an assumption may need some qualifications. Whereas it seems safe to assume that the presence of word frequency effects diagnoses lexical involvement (at some level of processing), it may not be very safe to assume that the absence of frequency effects means that the lexical codes have not been activated and used. In fact, some researchers argue that the first stages of lexical access may not be sensitive to word frequency at all (see McCann & Besner, 1987). However, in the following experiments we do not need to make strong claims about the frequency effect. A differential frequency effect on monosyllabic and bisyllabic words is all that is needed to assess that these two types of words are treated in a qualitatively different manner during speech processing. More precisely, the claim is that monosyllabic items contact lexical representations more directly than multisyllabic words.

monosyllabic items, whereas the prelexical code may be used with bisyllabic items (Mehler, 1981). Since word frequency is a property of lexical items, we can make a precise prediction concerning phoneme monitoring, namely, a frequency effect should be found for monosyllabic words, but not for bisyllabic ones. Experiment 1 tests this prediction by having subjects detect a phoneme that appears in item initial position. The items were presented in lists in which the syllable length and word frequency were crossed. Pairs of high and low frequency items had matched syllabic structure and shared the initial phoneme.

Method

Materials. High- and low-frequency French words (see the frequency of the materials in Table 1) were selected to construct 20 pairs of open-class items. Two words in a pair shared the same first phoneme, had matched or similar syllabic structure, and shared the maximum number of phonemes. On average, the first half of the two words were matched (e.g., DAME/DALLE TISSU/TYPHON). Ten pairs were monosyllabic (mean phonemic length 3.5) and 10 pairs were bisyllabic (4.9 phonemes). The 40 items were the *target words* whose initial phoneme was either /P/,/T/, /K/,/B/, or /D/; they are listed in the Appendix.

The target words were distributed into five lists, one for each target phoneme. For instance, DALLE, DAME, DOUZE, and DOUCHE were put into the /D/ list and distractor items were added in order to attain the ratio of one target to four distractors. The syllabic length of the word preceding the target word (the *critical word*), was not correlated with the target word's length or frequency. To avoid phonetic similarity effects, critical words did not begin with an occlusive and no distractor contained the target phoneme. Half the distractors were monosyllabic and half bisyllabic.

The lists were read by a native, French, female speaker at a regular rate of one word every 2 s. The stimuli were digitized at a sampling rate of 16kHz with an Oros system connected to a PDP-11/73 and stored on a Betamax video recorder. A speech editing system was used to place an inaudible marker at the onset of the stop consonant's burst. This inaudible time marker triggered a clock that was stopped by subject responses. Response times were collected on an Olivetti M24 computer with an error of less than 2 ms.

Subjects and procedure. Thirty right-handed students were tested (16 male and 14 female). Subjects were asked to push a button as rapidly as possible whenever they heard a previously specified phoneme target in initial position. Targets were specified auditorily with three town names: "/P/ as in Paris, Perpignan, Pau" (the other

TABLE 1
SUMMARY STATISTICS FOR ITEMS OF EXPERIMENTS 1 AND 2

Target word	Word length*	Vowel length*	Identification point	Frequency count**	Subjective frequency rating***
			Monosyllabic items		
Low frequency	530	157	3.5	3	3.06
High frequency	492	137	3.4	148	4.13
			Bisyllabic items		
Low frequency	610	90	4.1	3	2.92
High frequency	608	90	4.3	102	4.08

* In milliseconds; measured on a speech editor.

** Occurence per million, calculated from the spoken French table (Gougenheim et al., 1956). The numbers displayed are the geometric mean of the frequency of the words.

*** Subjective rating was obtained from a pool of 56 subjects who were asked to rate the frequency of the words on a 5-point scale.

specifications are listed in the Appendix). The experimental session lasted for 13 min.

Results

Reaction times above 1000 ms and below 100 ms were discarded; 98.8% of the raw data was within these two limits. Subjects did not respond on 2.7% of the items (including training items). Subjects made on the average 2.1 responses on distractor words (e.g., GARCON when the target was /P/), yielding a false-alarm rate of 0.8% per distractor per subject.

An analysis of variance with subjects as the random factor revealed a significant effect of the two main factors. There was a significant *frequency* effect: RTs were shorter for the high- than for the low-frequency targets (25 ms, $F1(1,29) = 12.9$, $p < .002$). This effect was also significant with an analysis by item ($F2(1,18) = 4.54$, $p < .05$; marginal $\min F'(1,31) = 3.36$, $.05 < p < .1$). There was also a *length* effect: monosyllables were responded to reliably faster than bisyllables (33 ms, $F1(1,29) = 12.1$, $p < .002$; $F2(1,18) = 5.18$, $p < .05$; marginal $\min F'(1,33) = 3.63$). The interaction between the two factors was only marginal ($F1(1,29) = 2.96$, $.05 < p < .1$; $F2 < 1$).

Reaction times for monosyllabic and bisyllabic items of high- and low-frequency are displayed in Fig. 1. The frequency effect is mostly due to monosyllabic items. Indeed, for the monosyllables alone, there was a very strong frequency effect (37 ms, $F1(1,29) = 15.9$, $p < .001$; $F2(1,9) = 9.86$, $p < .025$; $\min F'(1,21) = 6.10$, $p < .025$). For bisyllabic items, however, there was a nonsignificant trend in the direction of a frequency effect (14 ms, $F1(1,29) = 1.31$; $F2 < 1$)

Discussion

The results of this experiment have confirmed that the speed with which an item is responded to in a phoneme monitoring task is related to its frequency. This is true when the item is monosyllabic. However, for bisyllabic items, no significant frequency effect is found. This finding suggests that the length of the target bearing items is related to the frequency effects that can be measured in a phoneme monitoring task. Some investigators (see Marslen-Wilson, 1987)

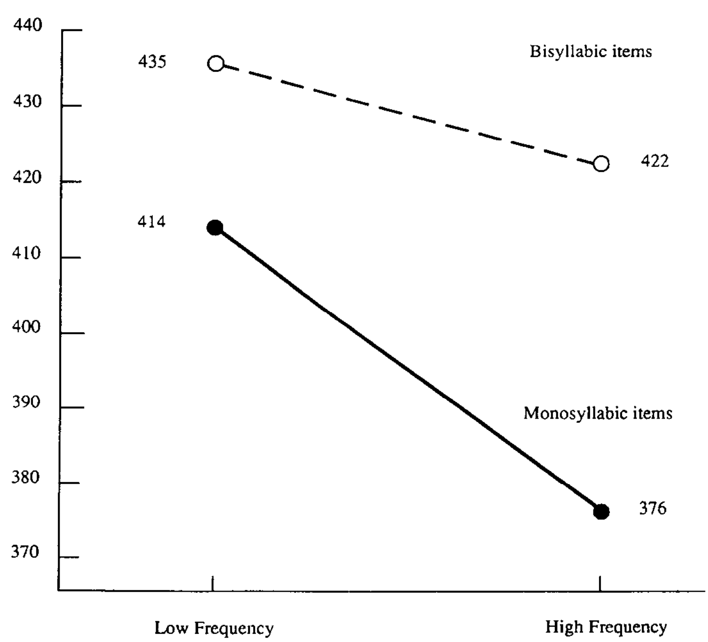

FIG. 1. Phoneme detection latencies for normal rate of speech.

have claimed that left to right properties of the speech signal play a crucial role in lexical processing, namely that the continuous input signal is accumulated until the *identification point* where a unique lexical candidate is isolated. By and large, the duration of a word is (negatively) correlated with frequency. It could be that the above results might be explained in terms of duration or identification point differences rather than word frequency. However, given the way in which our materials were constructed, this seems rather unlikely. Indeed, the identification point for monosyllabic items was 3.5 phonemes for low-frequency words, and 3.4 for frequent ones. For bisyllabic items it was 4.1 for low-frequency and 4.3 for high-frequency words.[2] There was no correlation between identification point and reaction time.

However, the identification point measured in phonemes may not be very relevant for speech. The total duration of the words might perhaps provide a better estimate of the psychologically relevant identification duration. The temporal length of targets in Experiment 1 was measured using a speech editor and is shown in Table 1. The difference in length between monosyllabic and bisyllabic items was quite large (98 ms, significant $F(1,18) = 4.7, p < .05$), and the difference in duration between high- and low-frequency words was rather small; frequent monosyllables were slightly shorter than infrequent ones (38 ms, not significant). A correlation of the duration of monosyllabic items with the RTs was not significant ($r = -0.11$); furthermore, the slope was negative, thus indicating, if anything, that the length difference should go against the observed frequency effects.[3] An explanation of the above results in terms of identification point thus seems unsuitable.

[2] These data were obtained for each target item, by counting the number of initial phonemes that correspond to a unique candidate in the French dictionary "Petit Robert."

[3] Unexpectedly, a trend of a correlation of latencies with duration was found for bisyllables ($r = 0.4$, $F(1,18) = 3.23, .05 < p < .1$).

In short, a global frequency effect for a reduced set of French words was uncovered; the effect was reliable for monosyllabic but not for bisyllabic items. This result is compatible with the Dual Code Hypothesis, namely, the lexical code is on the whole used for monosyllables, and the acoustical/phonetical code for bisyllables. It is the length of the target-bearing item that determines whether response latencies will be correlated with the frequency of an item.

As before, we may ask whether it is the syllabic structure or the duration that is critical here. Informally, we may assume that the subjects' motor response takes about 100 ms. Since latencies were on the order of 400 ms, responses should on the average be initiated 300 ms after the onset of the word. The duration of monosyllables is 500 ms, and their acoustic identification point is close to the middle of the signal (i.e., 250 ms); thus subjects have enough time to access the lexical code by the time they start their response. Bisyllables are longer (600 ms), and a rough estimate of the identification point can be placed at 300 ms. Thus, subjects are not always able to use the lexical code.

No reliable frequency effect was found for bisyllabic words. However, if the above estimations are correct, it is possible to predict a differential tendency depending on the speed of subjects. Faster subjects should not show a frequency effect for bisyllabic items since they always respond before the lexical information becomes available, whereas slower subjects may show such a tendency. We thus divided subjects into two equal groups according to their latencies: the 15 fastest subjects had an average latency of 353 ms and the 15 slowest an average of 483 ms. Both groups showed a frequency effect for monosyllables. Fast subjects gave no sign of a frequency effect for bisyllabic items (0 ms), but slow subjects showed a trend (28 ms, $F1(1,14) = 3.39, .05 < p < .1$).

Qualitatively, the above length effects are compatible with an explanation in terms

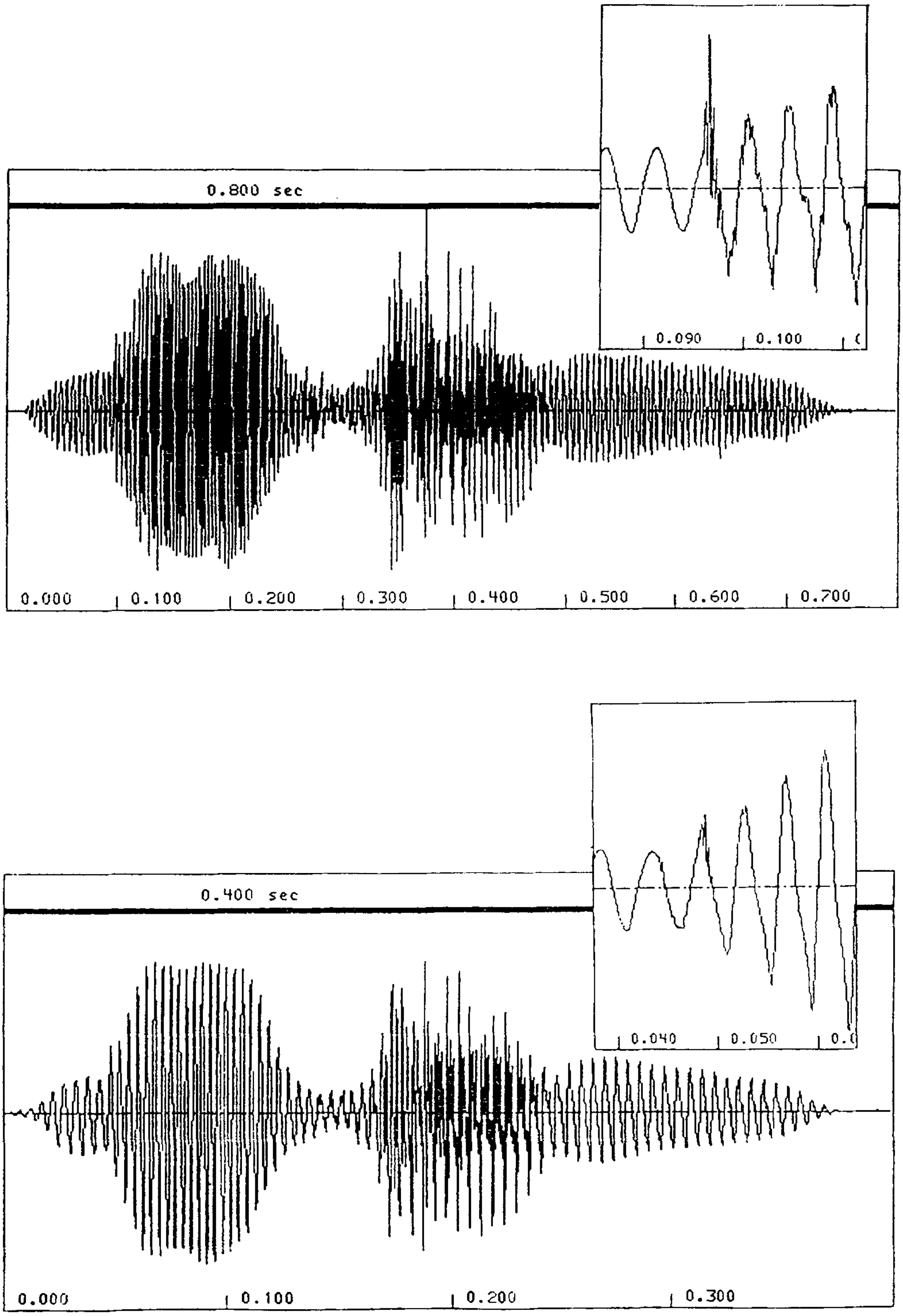

FIG. 2. Speech waveform of an utterance of the word "DOUZAINE." The upper part is natural speech and the lower part is compressed by a 50% factor. The time scale of the compressed signal is twice the time scale of the uncompressed one. The little window is at the same time scale and shows a detail at the burst of the initial /D/.

of identification point differences. They are also compatible with another interpretation, namely, that the number of syllables in the target determines the code that will be used. At first sight, this *structural model* seems a mere rephrasing of the durational model in terms of a less fine grained signal analysis. But predictions are not the same: number of syllables is a structural parameter and does not depend on rate of speech, whereas the time taken to attain an identification point is a function of speech rate. Experiment 2 uses compressed speech to evaluate the two hypotheses.

EXPERIMENT 2

Modifying the rate of speech makes it possible to alter the durational parameters, leaving the structural ones unchanged. The present experiment was designed to evaluate whether a structural or a durational hypothesis is best at predicting which code will be used. Subjects listened to the same stimuli as in the previous experiment, but at rates that were twice as fast.

Speech compression allows a modification of the rate of speech, only minimally altering spectral characteristics such as formants and fundamental frequency. Our compression algorithm was run on a PDP-11/73; it was originally developed by the CNET, Lannion, France (Charpentier & Stella, 1986) and was later adapted at our laboratory. The algorithm operates on digitized speech sampled at 16kHz. When the speech segment is voiced, the program inserts pitch-synchrous temporal marks. For nonvoiced segments, the mark is randomly placed. In a second pass, depending on the compression rate, adjacent periods of the signal are superimposed. More specifically, they are averaged and weighted on a Hamming window. The result, a very smooth, high quality stimulus, is illustrated in Fig. 2. It can be seen that the overall energy envelope is very well preserved by speech compression. However, rapid transitions such as stop consonant bursts are somewhat eroded by the averaging technique. Still, at a 50% compression rate the intellibility of the phonemes is unimpaired.[4]

In the past, some studies have explored intelligibility as a function of compression rate (Chodorow, 1979). However, their compression algorithms were far more primitive than ours (earlier, compression was done by cutting out pieces of tape and splicing them by hand), but they yielded consistent results. Garvey (1953) reported that individual words remain intelligible (more than 90%) at a compression rate of less than 50%. Above 50%, intelligibility drops rapidly. More recently, the same threshold was found for the intelligibility of connected speech (deHann, 1977, 1982).

What are the effects of speech compression on the activation time course of the lexical and the prelexical codes? The predictions of a durational based model are as follows: since compressed bisyllables are shorter than uncompressed monosyllables, the identification point is attained faster, making it possible for the lexical route to win the race. Thus, speech compression should speed up the activation of the lexical code. In contrast, compression does not necessarily result in a considerable processing advantage for the prelexical code. Indeed, the phonetic information relevant to the target phoneme is concentrated around the first few milliseconds of the words. Thus variations in overall duration of the word should play a minor role with respect to phonemic extraction. Moreover, for a compression that deletes nearly 50% of the information, the extraction of an acoustic/phonetic code is likely to be disrupted. A durational based model still predicts frequency effects with monosyllabic items, but one should also expect a high incidence of lexical factors when responding to compressed bisyllabic items.

In contrast, the structural model predicts the same pattern of RTs as in the previous experiment, namely, a frequency effect for

[4] A pilot experiment with a group of 10 subjects for 20 initial stop consonant words showed a 100% identification performance.

monosyllables because the processing units remain the same. Bisyllabic items are still responded to on the basis of the prelexical code. Since the analyzer relies on a large segment of speech (like the first syllable), phoneme extraction may not necessarily be disrupted even by massive compression.

Method

Materials. A speech compression routine was run on the materials used in Experiment 1 with a compression rate of 50%.

Procedure and subjects. Thirty right-handed students (23 males, 7 females) were tested. The same experimental procedure was used as in Experiment 1, except that subjects were told that they would hear compressed speech.

Results

As in Experiment 1, reaction times above 1000 ms and below 100 ms were discarded; 4.2% of the responses were eliminated. However, two experimental items provoked more than 20% of the misses and had very long RTs: "parfum" (20%, 716 ms) and "terrasse" (30%, 660 ms). These were removed from the analysis, as well as their paired frequent word ("pardon" and "terrible").

The mean RT was 430 ms (18 ms longer than in Experiment 1). This small increase in RT (about 3%) is consistent with the excellent quality of the compression. An analysis of variance, with subjects as random factor, revealed a significant frequency effect: RTs were shorter for the high- than for the low-frequency targets (16 ms, $F1(1,29) = 6.46, p < .02$). This effect failed to reach significance with an analysis by item ($F2(1,16) = 1.79, p > .1$) The syllabic length did not yield a significant effect (−10 ms, $F1(1,29) = 1.25$, ns); the interaction between the two factors was marginal ($F1(1,29) = 3.34, .05 < p < .1$). Reaction times for monosyllabic and bisyllabic items of high and low frequency are displayed in Fig. 3. For the monosyllables alone, there was a marginally reliable frequency effect (30 ms, $F1(1,29) = 11.7, p < .002$; marginal $F2(1,9) = 4.25, .05 < p < .1$). No frequency effect was found for bisyllables (3 ms, $F1 < 1$ and $F2 < 1$).

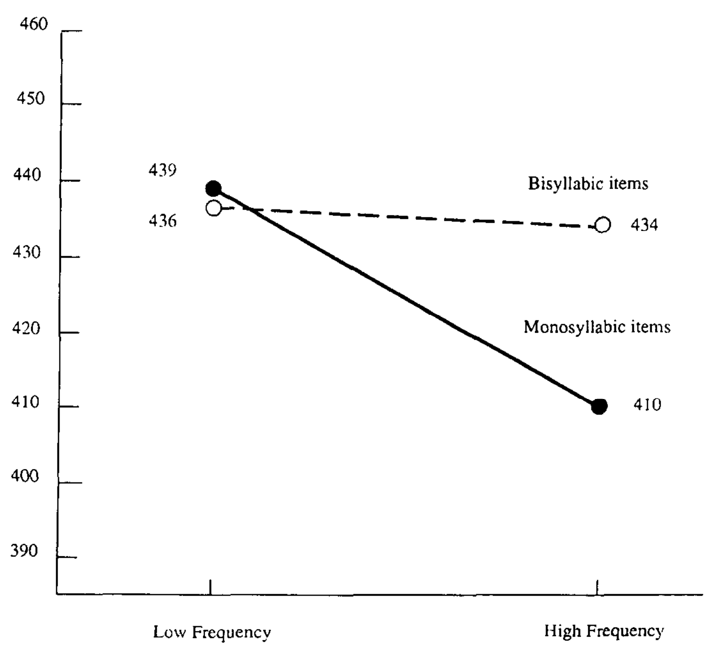

FIG. 3. Phoneme detection latencies for compressed rate of speech.

Speech compression did not disrupt phoneme monitoring. Furthermore, misses and false alarm rates remained very low, though increased by this transformation and taking aside the two discarded items. Tentatively, a signal detection analysis was performed to determine the effect of speech compression on subjects' responses. It was found, following McCarthy and Davison (1981)[5] that discriminability decreased from 67.9 for the normal rate to 43.3 for compressed speech. Response bias, on the other hand, remained quite stable ($\beta = 0.527$ for normal rate, $\beta = 0.517$ for rapid rate)[6] This suggests that speech compression mostly affects the discriminability of speech sounds. Thus, if anything, compression should have disrupted the phonetic code.

A good correlation was found between RTs for normal speech and RTs for compressed speech ($r = 0.70$, $t(35) = 5.72$, $p < .001$). Although speech compression changes a number of acoustic parameters, the determinants of subjects' responses seem to remain unchanged.

Speech compression only increased global reaction times by 15 ms. The results of Experiments 1 and 2 were merged, and a new analysis of variance was run with three factors, frequency, length, and compression. The first two factors are within-subject while compression is a between-subject factor. The two within-subject factors were tested and were significant only by subject (frequency: $F1(1,58) = 16.9$, $p < .001$; $F2(1,16) = 3.2$, ns; length: $F1(1,58) = 10.9$, $p < .005$; $F2(1,16) = 2.3$, ns). The analysis by effect of compression was only significant by item ($F(1,58) = 0.5$, ns; $F2(1,16) = 6.02$, $p < .05$). Notice that the interaction between length and frequency was significant by subject ($F1(1,58) = 7.1$, $p < .01$; $F2(1,16) = 1.8$, ns). This interaction is compatible with the hypothesis that monosyllabic and bisyllabic items yield responses that reflect differentially the two codes, regardless of speech rate. No other significant interaction was found. We want to stress the fact that globally, monosyllabic items are significantly affected by frequency while bisyllabic items are not.

Discussion

The compression algorithm we used, with a rate of 50%, leaves intelligibility nearly intact. It provokes a slight increase in RTs, and a small decrease in performance. However, the performance level is still very good, given that speech compression removes 50% of the original information present in the signal. Furthermore, the pattern of response is very similar to that obtained for normal speech. The basic processes underlying phoneme detection thus seem to be relatively unaffected by speech compression.

Indeed, a reliable frequency effect for monosyllables was established with compressed and normal speech. However, speech compression did not cause a frequency effect to emerge for bisyllables. Moreover, the interaction between frequency and syllabic length was significant by subjects in a global analysis taking into account Experiments 1 and 2. The structural hypothesis would thus appear to explain better the results of Experiments 1 and 2 than the durational hypothesis: it is the number of syllables in the target-bearing item that determines the code used in phoneme detection.

The durational hypothesis should predict the emergence of a frequency effect both for fast and slow subjects. Indeed, even fast subjects should have time to access the lex-

[5] Adapted from McCarthy and Davison (1981) for phoneme monitoring.

[6] A value of d and β was computed for each subject. A nonparametric test (Mann–Whitney) showed that the difference in discriminability between compressed and uncompressed speech was significant ($Z = 3.15$, $p < .001$, one-tailed), but the difference in bias failed to reach significance ($Z = 0.4$, $p = .32$, one-tailed)

$$d = \left[\frac{\text{Hit} \cdot \text{Correct Rejection}}{\text{Miss} \cdot \text{False Alarm}}\right]^{1/2}$$

$$\beta = \left[\frac{\text{Hit} \cdot \text{False Alarm}}{\text{Miss} \cdot \text{Correct Rejection}}\right]^{1/2}$$

ical representation of compressed bisyllables before they begin their response.[7] In Experiment 2, the 15 fast subjects and 15 slow subjects showed latencies that were very similar to those in Experiment 1. Namely, both classes of subjects showed a reliable frequency effect for monosyllables, and none for bisyllabic items.[8] The structural hypothesis can thus accommodate our results without much difficulty. With bisyllables, most subjects rely on the first syllable to detect a phoneme regardless of speech rate.

However, a couple of objections prevent any strong conclusion. Despite the fact that frequency effects are robust when the target is carried by a monosyllabic word, but nonexistent when the target is carried by a bisyllabic word, in neither experiment is there a significant interaction between word frequency and syllabic length. Even when the two experiments are combined in an omnibus analysis, the requisite interaction is significant only across subjects. Moreover, two pairs of bisyllabic items out of 10 were excluded from the analysis due to high error rates. In fact the high error rate was caused by two low-frequency items which showed very long reaction times. We cannot rule out the possibility that at least some bisyllabic items could show a frequency effect, when presented under compression. Thus we wished to enlarge our database, by running a new experiment using a larger set of bisyllabic items.

EXPERIMENT 3

Experiment 3 was basically a replication of Experiment 2. The focus of this experiment was the behavior of bisyllabic items under speech compression. We chose to favor the bisyllabic items by constructing 15 new bisyllabic pairs and only 9 pairs of monosyllabic pairs.

Moreover, the instructions were modified in order to favor the use of the lexical code. Indeed, Dupoux and Mehler (in preparation) have found that in lists with words and pseudo words, subjects are very sensitive to attentional factors and to the exact wording of the task. If they are just asked to perform phoneme monitoring, they can do it on the basis of a pure prelexical strategy (even on monosyllabic items). However, when the subjects are instructed to pay attention to the meaning of the words, they are likely to switch back to a lexical strategy.

It could be that in the compressed experiment, subjects were processing the items as nonwords, because they were distorted by the speech compression routine. Thus compression could bias subjects to use the prelexical code (despite a robust frequency effect found in monosyllables). At any rate, a modification in the instruction set that biases subjects to use the lexical code can be used to make our point stronger. Our prediction is that even in a compressed situation, even with a bias to use the lexical code, no frequency effect should be found in bisyllabic items.

Method

Materials. Fifteen pairs of high- and low-frequency bisyllabic items (4.9 mean phonemic length) were chosen as in Experiment 1. Two monosyllabic pairs of Experiments 1 and 2 (DALLE/DAME and POULE/POUF) plus seven new monosyllabic pairs (3.5 mean phonetic length) were also included.

Five lists were constructed, corresponding to the targets /P/, /T/, /D/, /B/, and /P/ again. Distractor items were added in order to attain the ratio of one target to four distractors. The syllabic length of the word preceding the target word (the *critical word*), was not correlated with the target

[7] Given the 100-ms estimation of the motor response, rapid subjects probably heard about 90% of the bisyllables, and passed their supposed identification point by 120 ms.

[8] The trend found in Experiment 1 was still observed here since fast subjects had a quasi null frequency effect for bisyllabic items (−7 ms), and slow subjects showed a trend of 14 ms.

word's length or frequency. To avoid phonetic similarity effects, critical words did not begin with an occlusive and no distractor contained the target phoneme. Half the distractors were monosyllabic and half bisyllabic. As in Experiments 1 and 2, the distractors were randomized in the lists so that there was no systematic association or meaning correlation among the items.

The lists were read by a female native speaker of French at a regular rate of one word every 2 s. Stimuli were recorded and temporally marked with the same procedure as in Experiments 1 and 2. The five digitized lists were applied to the compression algorithm described in Experiment 2 with the same rate of 50%.

Subjects and procedure. Forty right-handed students (27 males, 13 females) were tested. The same experimental procedure was used as in Experiment 2, except that subjects were asked to focus their attention on the meaning of the words. Parallel to their task of detection, they had to pay attention to associations and semantic relations between successive items in the list. They were told that a series of questions would be asked after the experiment to test their attention to the meaning of the items.

Results

As in Experiment 1, reaction times above 1000 ms and below 100 ms were discarded. The mean error rate was 3.8%, and the mean RT was 472 ms (only 42 ms longer than in Experiment 2).

An analysis of variance showed no significant effect of frequency nor length (all the F ratios were less than 1). However a reliable interaction between frequency and length was found ($F1(1,39) = 4.85$, $p < .035$; $F2(1,22) = 4.52$, $p < .05$; but $\text{Min}F'(1,55) = 2.34$, $p > .1$). Reaction times for monosyllabic and bisyllabic items of high and low frequency are displayed in Fig. 4. For the monosyllables alone, there was a small but reliable frequency effect (19 ms, $F1(1,39) = 5.17$, $p < .03$; $F2(1,8) = 5.25, p = .051$; $\text{Min}F'(1,28) = 2.60, p > .1$). No frequency effect was found for the bi-

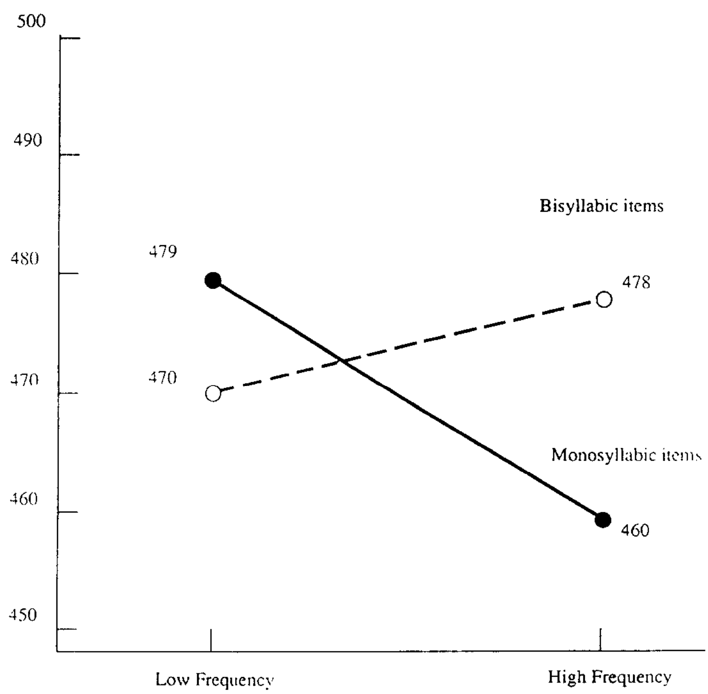

FIG. 4. Phoneme detection latencies for compressed rate of speech.

syllabic items (−8 ms, $F1(1,39) = 1.22$, $p > .1$; $F2 < 1$).

Discussion

This experiment fully replicates Experiment 2. As expected, a frequency effect was found for the nine pairs of monosyllables. The important result is the complete absence of a frequency effect for the 15 new pairs of bisyllabic items. The trend is even in the opposite direction, which gives rise to a significant interaction between frequency and length. This result confirms the hypothesis that even in a compressed situation, monosyllables and bisyllables are treated in a qualitatively different manner during lexical processing.

However, before drawing any strong conclusions, a last control is called for. The absence of a frequency effect for bisyllables, regardless of rate, could be due to a poor choice of the items themselves. The French frequency table (Gougenheim, Michea, Rivenc, & Savvageot, 1956) may be outdated and it is entirely possible that the estimates are unreliable. A different possibility would be that bisyllabic items are less sensitive to frequency effect, due to the time course of lexical activation. For instance, Bradley and Forster (1987) reported that frequency effects might be smaller in polysyllabic items than in monosyllabic items. Thus the absence of frequency effects for bisyllabic items in phoneme monitoring could be an artifact of our material, or could reflect some specificities of lexical organization in speech. The next experiment tests these hypotheses using a lexical decision task.

Experiment 4

The purpose of this experiment was twofold. First, lexical decision provides a means of assessing the validity of the materials used in Experiments 1, 2, and 3. Second, it allows a comparison of the parameters that affect the early and late components of lexical processing. One way to estimate whether the items used in the preceding experiments are good representatives of high- and low-frequency classes is to try and determine whether a frequency effect is observed with lexical decision. The lexical decision method necessarily involves lexical access and probably some postaccess mechanisms. Since the words used in Experiments 1, 2, and 3 have matched identification points and initial phonetic properties, lexical decision should allow us to determine whether the material used in our phoneme monitoring experiments can generate a reliable frequency effect.

Method

Materials. The materials of Experiments 1 and 2 and Experiment 3 were merged together. The monosyllabic items consisted of the 10 pairs of monosyllables of Experiments 1 and 2 plus the 7 pairs of Experiment 3. Similarly, the 10 original bisyllabic pairs of Experiments 1 and 2 plus the 15 new pairs of Experiment 3 were included. Legal nonwords were constructed by exchanging the initial phonemes of two of the monosyllabic items or the initial syllable of two of the bisyllabic items. One hundred filler words and nonwords with a nonocclusive initial phoneme were added. The list was split into four blocks of equal size that satisfied the following requirements: each member of a high/low frequency pair (e.g., DAME/DALLE) occurred in a different block, (e.g., DAME in Block 1 and DALLE in Block 3). The order was globally counterbalanced with frequency (another pair DOUZE/DOUCHE had a distribution opposite to, say, DAME/DALLE). Subjects were presented with the blocks in either of two possible orders: 1,2,3,4 or 4,3,2,1. In this way one group of subjects heard "DALLE" and "DOUZE" first, whereas the other group heard "DAME" and "DOUCHE" first.

Procedure and subjects. The experimental lists were read by a female native speaker of French at a regular rate of one word every 2 s. As before, stimuli were sampled at 16kHz; test words were edited and marked at the onset. These inaudible time marks triggered a clock that was stopped by subject responses. Response times were collected by an Olivetti M24 with an error of less than 2 ms.

Twenty-five right-handed students, 21 male and 4 female, were tested. Subjects were split into two groups. Each group heard the four blocks in either of two orders (1,2,3,4 or 4,3,2,1). Subjects were asked to push a button on their right when they heard a French word, and on their left when they heard a nonword. Both speed and accuracy were emphasized. The whole session lasted for 11 min.

Results

For data analysis, reaction times above 1800 ms and below 100 ms were discarded. The mean error rate was 5.9% for words and 6.1% for nonwords, that is, higher than for phoneme detection. Errors appeared to be concentrated in a small number of items. Words with more than a 20% error rate were discarded from further analysis: three low-frequency monosyllables [BRU (44%), POUF (28%), TANK (20%)], and one high-frequency monosyllable [BREF (20%)]; four low-frequency bisyllables: [PETON (60%), CREDO (40%), CANTATE (36%), BURIN (24%)]. The other word belonging to the pair was accordingly discarded.

A total of 12 monosyllables and 21 bisyllables were subsequently analyzed. An analysis of variance with subjects as a random factor revealed a strong frequency effect (47 ms, $F1(1,23) = 45.7$, $p < .001$). This effect was also significant with an analysis by item ($F2(1,31) = 7.87$, $p < .01$; significant $\min F'(1,41) = 6.71$, $p < .02$). The length effect was not significant ($F1(1,23) = 2.86$, $.05 < p < .1$; $F2 < 1$). The group factor did not approach significance ($F1 < 1$). The interaction between length and frequency was not significant (both F ratios < 1), nor was any other interaction.

Figure 5 shows separate reaction times for high- and low-frequency monosyllables and bisyllables. For the monosyllables, there was a strong trend towards a frequency effect (48 ms, $F1(1,23) = 13.29$, $p < .002$; marginal $F2(1,11) = 3.64$, $.05 < p < .1$). For bisyllables, the frequency effect is robust (47 ms, $F1(1,23) = 23.66$, $p < .001$; $F2(1,20) = 4.38$, $p < .05$; $\min F'(1,27) = 3.70$, $.05 < p < .1$).

Errors and latencies were consistent since low-frequency monosyllables elicited a 9% error rate and high-frequency ones only a 4% error rate (this difference was significant: Wilcoxon test $Z = 3.5$, $p < .001$). Low-frequency bisyllables have a 4% error rate, whereas frequent ones have error rates of only 1% ($Z = 4.76$, $p < .001$). Notice that monosyllables showed a higher error rate for both frequency categories. This is an indication that short words are easier to confuse with nonwords than longer words.

However, since 16 words were excluded from the analysis because of high error rate, a *subjective frequency rating* experiment was run with 56 subjects on the material used in Experiments 1, 2, and 3. The results confirmed that the values given in Gougenheim's table for both monosyllables and bisyllables are reliable estimates of word frequency. As shown in Table 1 and Table 2, high- and low-frequency items differed significantly from each other.[9]

Discussion

Lexical decision elicits more errors than phoneme monitoring. Subjects complained that some words were problematic. Several reasons were invoked; i.e., some were not typically French ("TANK"), others were rated as idiosyncratic ("PETON") ("POUF"), and yet others were unusual ("CANTATE," "CREDO," etc.). Subjects are not always clear about what counts as a French word in a lexical deci-

[9] The frequency factor was very strong for items of both Experiments 1 and 2 (monosyllables: $\mathrm{Min}F'(1,10) = 17.87$, $p < .005$; bisyllables: $\mathrm{Min}F'(1,10) = 14.7$, $p < .005$) and items of Experiment 3 (monosyllables: $\mathrm{Min}F'(1,9) = 8.93$, $p < .02$; bisyllables: $\mathrm{Min}F'(1,17) = 33.97$, $p < .001$).

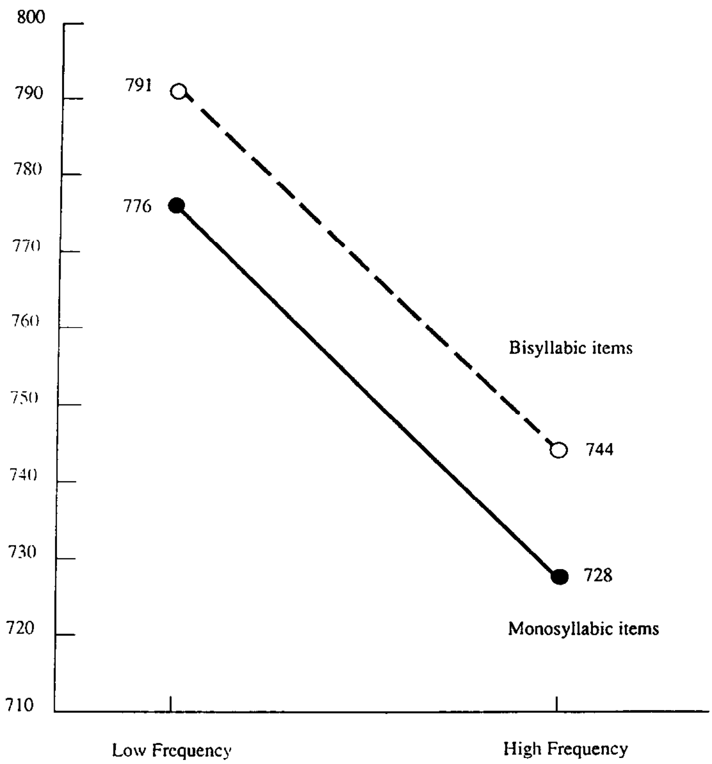

FIG. 5. Lexical decision latencies for high- and low-frequency, mono- and bisyllabic words

sion task. They seem to rely on some notion like "prototypicality".[10] However, that notion correlates with frequency of use, since seven out of the eight words that were mistakenly classified as nonwords were low-frequency items.

For words that are consistently judged as real words, the results are straightforward. A frequency effect is observed for bisyllabic and also for monosyllabic items. This finding attests that frequency effects in the auditory modality can be observed using a lexical decision task. This result, together with those of a subjective rating task, showed that the items of Experiments 1, 2, and 3 were reasonably chosen for word frequency.

Thus it is possible to compare the performance of subjects on the two types of tasks. In monosyllables, a word-frequency effect is found both with phoneme monitoring and lexical decision. However, in bisyllables, a word-frequency effect is found only with lexical decision. No frequency effect is found with the phoneme detection task, even when bisyllabic words are compressed so as to be shorter than uncompressed monosyllables. To assess this result statistically, we computed the mean reaction time for each monosyllabic and bisyllabic item in the two tasks—that is, phoneme monitoring with compressed speech (Experiments 2 and 3) and lexical decision (Experiment 4). As expected, there was a reasonably good correlation between phoneme monitoring and lexical decision latencies for the monosyllabic items ($r = .482$, $t(34) = 3.21$, $p < .01$). In contrast, there was no significant correlation

[10] In fact, it appeared that nouns are better prototypes than adjective or verbs. "BREF," a high-frequency adjective provoked more than 20% errors. Moreover, among the 33 high-frequency words, 9 were not nouns (adjectives, verbs, etc.). Interestingly, the reaction time for these words was significantly lower than the high-frequency nouns (56 ms, $F(1,31) = 5.62$, $p < .03$). This category effect was powerful enough to wipe out the frequency effect in the nine pairs (−31 ms ns; significant interaction between frequency and category: $F(1,31) = 7.88$, $p < .001$).

TABLE 2
SUMMARY STATISTICS FOR ITEMS OF EXPERIMENT 3

Target word	Word length*	Vowel length*	Identification point	Frequency count**	Subjective frequency rating***
			Monosyllabic items		
Low frequency	484	133	2.9	6	2.73
High frequency	453	147	3.1	227	3.87
			Bisyllabic items		
Low frequency	559	84	3.6	5	2.90
High frequency	588	84	3.7	374	4.34

* In milliseconds; measured on a speech editor.

** Occurence per million, calculated from the spoken French table (Gougenheim et al., 1956). The numbers displayed are the geometric mean of the frequency of the words.

*** Subjective rating was obtained from a pool of 56 subjects who were asked to rate the frequency of the words on a 5-point scale.

between lexical decision and phoneme monitoring for the bisyllabic items (r = .048, $t(48) = 0.34$, $p > .1$).

GENERAL DISCUSSION

A frequency effect was found with phoneme detection for a small sample of French monosyllables. With those words, the effect was quite strong. However, the results do not allow the claim that the frequency effect obtains for all monosyllabic French words. There are several reasons that prevent such a generalization. The sample tested was not very large. In fact, it was already quite difficult to find 17 pairs of properly controlled monosyllables of high and low frequency, given the relative scarceness of monosyllabic words in French. Indeed, more than 85% of lexical entries in French are polysyllabic (Gauvain, 1986). Furthermore, the monosyllabic pairs were less well controlled than the bisyllabic pairs. Indeed, some of the pairs were not exactly matched with respect to syllabic complexity: e.g., "POIRE/PEUR," "TACT/TEXTE," "TRACE/TRAIN," "PIAULE/PIED," and "TAXE/TARD." The syllabic structure of "poire" (/pwar/) incorporates a diphthong. The syllabic structure of "TEXTE" is not even clearly monosyllabic (see below).

Lastly, it proved impossible to control for the vowel quality in 7 out of the 17 monosyllabic pairs ("BRUME/BREF," "TEINTE/TYPE," "BRAS/BRU," plus four of the pairs already mentioned). This is quite unfortunate since vowel quality, and more specifically, vowel length has been shown to correlate well with phoneme detection latencies (Diehl et al., 1987; Foss & Gernsbacher, 1983). In fact, it appeared that mean vowel duration, measured with a speech editor, was less well matched in monosyllabic words than in bisyllabic ones. In Experiments 1 and 2 (see Table 1), vowel duration was longer for low-frequency monosyllables than for high-frequency ones (20 ms, but nonsignificant $F(1,9) = 1.59$). However, in Experiment 3, the duration difference was in the other direction (−10 ms, see Table 2). This mismatch could explain that a higher frequency effect was found in Experiments 1 and 2 (37 ms and 30 ms) than in Experiment 3 (19 ms). Unfortunately, given the properties of French it might prove very difficult to design an experiment that incorporates a larger number of well-controlled pairs.

Another problem relates to the syllabic structure of words such as "TEXTE." It is unclear whether they can be considered monosyllabic or not. To avoid this issue, in the experiment by Cutler et al. (1987), only very simple monosyllables were tested (CVs, CVCs, and CCVs). In summary, although a frequency effect with a very limited set of monosyllabic items was established it remains for future research to assess whether the frequency effect can be generalized to all monosyllabic words.

Most French lexical entries are bisyllabic or longer. For these items, a strong frequency effect was found using lexical decision. However, the phoneme monitoring task failed to uncover a frequency effect for bisyllabic words. Why does one of the tasks result in a frequency effect and not the other? For the Dual Code hypothesis, an absence of frequency effects in phoneme monitoring is taken to imply that the word has been responded to on the basis of a prelexical code, while lexical activation is still taking place. Moreover, no frequency effect emerges when these items are compressed by 50%. This suggests that regardless of duration, bisyllabic items yield a response on the basis of the prelexical code without being influenced by lexical factors.

This result is difficult to interpret within the framework of some of the more powerful models of auditory word recognition currently available, (Elman & McClelland, 1984; Marslen-Wilson & Welsh, 1978; Marslen-Wilson & Tyler, 1980). Many models presuppose a rather small perceptual unit (e.g., distinctive features in TRACE II; McClelland & Elman, 1986). These units provide a gradual increment of evidence for lexical items in a maximally efficient process. For such models, the crucial parameter is the amount of perceptual information necessary for the selection of a unique lexical candidate. Models that presuppose a small perceptual unit can be formulated in either an autonomous (such as the dual-code model), or an interactive framework. But both models make the same prediction with respect to compressed speech, namely, speech compression should reduce the time needed to identify bisyllables to roughly that needed for uncompressed monosyllables. Thus, under compression, one would expect the lexicon to play a greater role on phoneme detection latencies. In an autonomous framework, one would predict that compression should speed up lexical access, and make the lexical code available sooner. Likewise, in interactive models (e.g., TRACE II) since rapid speech certainly allows the lexical level to send more feedback to the phonetic level, lexical variables such as word frequency should also play a greater role. However, these predictions are not reflected in the data.

One way interactive models could accommodate the observed results would be to reduce the role of word frequency during the lexical access stage. For instance, it could be argued that word frequency plays a more important role for short than for long words because the former have many more competitors. Long words have fewer neighbors and therefore should be more dependent on temporal parameters such as isolation point. This argument could account for the absence of a frequency effect with normal and compressed bisyllabic items. This view, entirely compatible with Savin's (1963) is nonetheless at odds with the recent proposal made by Marslen-Wilson, namely that frequency plays an important role in the cohort model (Marslen-Wilson, 1987).

However, for the processing of French, an alternative view exists, which suggests that a discrete segmentation unit, the syllable, is used during early processing. On this view (see Bertoncini & Mehler, 1981; Mehler, 1981; Segui, 1984), the speech stream is segmented into syllable-like chunks before the lexicon is accessed. In all likelihood, these chunks serve to decompose the speech stream into phoentic units. This view predicts that monosyllables are not at first segmented into smaller units. The syllable directly accesses the lexicon and is simultaneously analyzed into phonetic components. In a race between these two processes, lexical activation is bound to be completed before phonetic analysis, and the monitoring response is triggered from the lexical code (which, in all likelihood, is sensitive to frequency). For bisyllables, the first syllable is used for cohort reduction but the next syllable is usually necessary to single out a unique candidate. By the time the last syllable becomes available, the analysis into phonetic components has won the race, and a response is triggered from the phonetic code (except, maybe, for slow subjects who have a tendency to wait for confirmation from the lexical code). In this account, the time course is, within reasonable limits, independent of speech rate.

It should be noted that the structural hypothesis does not necessarily entail a syllabic hypothesis. The critical parameter responsible for the code used in phoneme detection may as well be the number of phonemes, the number of diphones, or any other structural parameter. Indeed, another way of stating the structural hypothesis would be to postulate that the prelexical processor performs a *speech rate normalization*. This view is compatible with a number of speech recognition systems which perform some kind of *dynamic time warping* on the signal. The output of such a transformation is a spectral template where irrelevant durational information has been factored out. Such a normalization would account for the fact that reaction times are virtually unaffected by speech compression. Likewise, normalization may also explain the excellent correlation between latencies in the two conditions. This proposal is related to the empirical finding that very early processes such as phonetic categorization are sensitive to speech rate (Miller, 1981).

In brief, lexical decision and phoneme detection can be used to measure quite different stages of lexical processing. As expected, lexical decision reflects mostly lexical and postlexical processing. This task is thus highly sensitive to a lexical variable such as word frequency. In contrast, phoneme detection is in general less sensitive to the lexical stage. Phoneme detection is sensitive to lexical variables only for the relatively small class of monosyllabic words, whereas it taps the prelexical stage for polysyllabic ones. Phoneme detection with compressed speech yielded novel results. Compressed bisyllabic items were not affected by word frequency. This result may be explained in two ways: (i) by denying that frequency plays a crucial role during the generation and matching of lexical candidates, or (ii) by arguing that during the *prelexical stage* listeners segment the speech stream into syllable-like units. In this second case, the time course of lexical access is affected by structural parameters such as the number of syllables rather than by temporal parameters.

The present results should be verified for English as well as for other languages. French is a syllable-based language of the oxytonic family, while English is a stress-based language. Considerable processing differences between the two languages have been reported by Cutler, Mehler, Norris, and Segui (1986). Consequently, until such data becomes available, the implications of our results should be limited to French.

APPENDIX

Materials for Experiments 1 and 2

Specification of the Target Phoneme

/P/ as in Paris, Perpignan, Pau.
/T/ as in Tarascon, Troie, Toulouse.
/K/ as in Carcassonne, Creusot, Clermont.
/B/ as in Bayeux, Bordeaux, Bruxelles.
/D/ as in Dijon, Dreux, Douai.

Low-Frequency Monosyllables: DALLE, DOUCHE, BLAME, BRUME, TRACE, TEINTE, TACT, CASTE, POUF, POIRE.

High-Frequency Monosyllables: DAME, DOUZE, BLAGUE, BREF, TRAIN, TYPE, TEXTE, CARTE, POULE, PEUR.

Low-Frequency Bisyllables: BILAN, BOUDOIR, TYPHON, TERRASSE, CASIER, COMPOTE, CREDO, CANTATE, PARFUM, POTAGE.

High-Frequency Bisyllables: BILLET, BOUTEILLE, TISSU, TERRIBLE, CAMION, CONFIANCE, CREDIT, CAMPAGNE, PARDON, POLICE.

Materials for Experiment 3

Low-Frequency Monosyllables: POUF,

DALLE, DIGUE, BRU, BRAIZE, PIAULE, TANK, TAXE, TRESSE.

High-Frequency Monosyllables: POULE, DAME, DIX, BRAS, BREF, PIED, TENTE, TARD, TREIZE.

Low-Frequency Bisyllables: POUBELLE, PARURE, DEBAT, DOUZAINE, DORTOIR, DOCTRINE, BOULON, BADEAU, BURIN, PENICHE, PANIQUE, PROFIL, PETON, TERREAU, TRACHEE.

High-Frequency Bisyllables: POUSSIERE, PAREIL, DEJA, DOULEUR, DORMIR, DOCTEUR, BOUQUIN, BATEAU, BUREAU, PENIBLE, PAROLE, PROVINCE, PETIT, TERRAIN, TRAVAIL.

References

Becker, C. A. (1976). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 556–566.

Bertoncini, J., & Mehler, J. (1981). Syllables as units in infant perception. *Infant Behavior and Development*, 4, 247–260.

Bradley, D. C., & Forster, K. I. (1987). A reader's view of listening. *Cognition*, 25, 103–134.

Charpentier, F. J., & Stella, M. G. (1986). Diphone synthesis using an overlap-add technique for speech waveform concatenation. *IEEE ICASSP*, Tokyo. New York: IEEE.

Chodorow, M. S. (1979). Time compressed speech and the study of lexical and syntactic processing. In W. E. Cooper & E. T. C. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garret*. Hillsdale, NJ: Erlbaum.

Cutler, A., & Norris, D. (1979). Monitoring sentence comprehension. In W. E. Cooper & E. T. C. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*. Hillsdale, NJ: Erlbaum.

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

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology*, 19, 141–177.

deHann, H. J. (1977). A speech rate intelligibility threshold for connected speech and time-compressed connected speech. *Perception and Psychophysics*, 22, 366–372.

deHann, H. J. (1982). The relationship of estimated comprehensibility to the rate of connected speech. *Perception and Psychophysics*, 32, 27–31.

Dell, G. S., & Newman, J. E. (1980). Detecting phonemes in fluent speech. *Journal of Verbal Learning and Verbal Behavior*, 19, 608–623.

Diehl, R. L., Kluender, K. R., Foss, D. J., Parker, E. M., & Gernsbacher, M. A. (1987). Vowels as islands of reliability. *Journal of Memory and Language*, 26, 564–573.

Dupoux, E., & Mehler, J. Attentional effects in phoneme monitoring: The dual code revisited. In preparation.

Elman, J. L., & McClelland, J. L. (1984). Speech perception as a cognitive process: The interactive activation model. In J. Lass (Ed.), *Speech and Language: Advances in basic research and practice* (Vol. 10). New York: Academic Press.

Forster, K. I. (1976). Accessing the mental lexicon. In E. C. Walker & R. J. Wales (Eds.), *New approach to language mechanisms*. Amsterdam: North Holland.

Forster, K. I. (1978). Accessing the mental lexicon. In E. C. Walker (Ed.), *Explorations in the biology of languages*. Assoks: Harvester Press.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. Walker (Eds.), *Sentence processing: Psycholinguistics studies presented to Merrill Garrett*. Hillsdale, NJ: Erlbaum.

Foss, D. J., & Blank, M. A. (1980). Identifying the speech codes. *Cognitive Psychology*, 12, 1–31.

Foss, D. J., & Gernsbacher, M. A. (1983). Cracking the dual code: Toward a unitary model of phoneme identification. *Journal of Verbal Learning and Verbal Behavior*, 22, 609–632.

Frauenfelder, U., & Segui, J. (1989). Phoneme monitoring and lexical processing: evidence for associative context effects. *Memory and Cognition*, 17, 134–140.

Garvey, W. D. (1953). The intelligibility of speeded speech. *Journal of Experimental Psychology*, 54, 102–108.

Gauvain, J. L. (1986). A syllable-based isolated word recognition experiment. *Proceedings IEEE ICASSP*, Tokyo. New York: IEEE.

Glanzer, M., & Ehrenreich, S. L. (1979). Structure and search of the mental lexicon. *Journal of Verbal Learning and Verbal Behavior*, 18, 381–398.

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

Klatt, D. (1989). Review of selected models in speech perception. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge: MIT Press.

Marslen-Wilson, W. (1984). Functions and process in spoken word recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X:*

Control of language process. Hillsdale, NJ: Erlbaum.

MARSLEN-WILSON, W. D. (1987). Functional parallelism in spoken word recognition, *Cognition*, **71**, 71–102.

MARSLEN-WILSON, W., & TYLER, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, **8**, 1–71.

MARSLEN-WILSON, W., & WELSH, A. (1978). Processing interaction and lexical access during word recognition in continuous speech. *Cognitive Psychology*, **10**, 29–63.

MASSARO, D. W. (1972). Perceptual images, processing time and preperceptual units in auditory perception. *Psychological Review*, **79**, 124–125.

MASSARO, D. W. (1975). Backward recognition masking. *Journal of the Acoustical Society of America*, **58**, 1059–1065.

MCCANN, R. S., & BESNER, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance*, **13**(1), 14–24.

MCCARTHY, D., & DAVISON, M. (1981). Towards a behavioral theory of bias in signal detection. *Perception and Psychophysics*, **29**(4), 371–382.

MCCLELLAND, J. L., & ELMAN, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1–86.

MEHLER, J. (1981). The role of syllables in speech processing: Infant and adult data. *Phil. Trans. R. Soc. Lond. B*, **295**, 333–352.

MEHLER, J., DOMMERGUES, J. Y., FRAUENFELDER, U., & SEGUI, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, **20**, 298–305.

MILLER, J. L. (1981). Effects of speaking rate on segmental distinctions. In P. D. Eimas & J. L. Miller (Eds.), *Perspectives on the study of speech* (pp. 39–74) Hillsdale, NJ: Erlbaum.

MORTON, J. (1969). The interaction of information in word recognition. *Psychological Review*, **76**, 165–178.

MORTON, J. (1979). Some experiments of facilitation in word and picture recognition and their relevance for the evolution of a theoretical position. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *The processing of visible language, I*. New York: Plenum Press.

MORTON, J. (1982). Desintegrating the lexicon. In J. Mehler, E. T. C. Walker, & M. Garrett (Eds.), *Perspectives in mental representation*. Hillsdale, NJ: Erlbaum.

MORTON, J., & LONG, J. (1976). Effect of word transitional probability on phoneme identification. *Journal of Verbal Learning & Verbal Behavior*, **15**, 43–51.

NEWMAN, J. E., & DELL, G. S. (1978). The phonological nature of phoneme monitoring: A critique of some ambiguity studies. *Journal of Verbal Learning and Verbal Behavior*, **17**, 359–374.

NORRIS, D. G. (1986). Word recognition: Context effects without priming. *Cognition*, **22**, 93–136.

RUBIN, P., TURVEY, M. T., & VAN GELDER, P. (1976). Initial phonemes are detected faster in spoken words than in nonwords. *Perception and Psychophysics*, **19**, 394–398.

SAVIN, H. (1963). Word-frequency effects and errors in the perception of speech. *Journal of the Acoustical Society of America*, **35**, 200–206.

SEGUI, J. (1984). The syllable: A basic perceptual unit in speech perception? In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X, Control of language processes*. Hillsdale, NJ: Erlbaum.

SEGUI, J., FRAUENFELDER, U., & MEHLER, J. (1981). Phoneme monitoring & lexical access. *British Journal of Psychology*, **72**, 471–477.

TREIMAN, R., SALASOO, A., SLOWIACZEK, L. M., & PISONI, D. B. (1982). *Effects of syllable structure on adults' phoneme monitoring performance*. (Progress Report No. 8) Bloomington: Indiana University, Speech Research Laboratory.

(Received: February 28, 1989)

(Revision received: August 16, 1989)