Psychological Review
1995, Vol. 102, No. 1, 146-184

0033-295X/95/$3.00

A Rose Is a REEZ: The Two-Cycles Model of Phonology Assembly in Reading English

Iris Berent and Charles A. Perfetti
University of Pittsburgh

The authors propose a model of phonological assembly that postulates a multilinear representation that segregates consonants and vowels in different planes. This representation determines the online process of assembly: Consonants and vowels are derived in 2 consecutive cycles that differ in their automaticity. The model's temporal properties resolve critical contradictions in the phonological processing literature. Its claims are further supported by a series of English-masking and English-priming experiments demonstrating that the contributions of consonants and vowels depend on target exposure duration and differ in their susceptibility to digit load. One methodological implication of the model is that regularity effects are not necessary evidence for assembly. This claim is supported by naming studies showing that vowel assembly requires long target durations, but short target durations permit consonant assembly despite null evidence for vowels.

Reading research clearly agrees that phonology plays a critical role in silent reading (e.g., Frost, 1991; Hanson & Fowler, 1987; Lukatela & Turvey, 1994; McCusker, Hillinger, & Bias, 1981; Patterson & Coltheart, 1987; Perfetti, 1994; Perfetti & McCutchen, 1982; Perfetti, Zhang, & Berent, 1992; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Van Orden, 1987; Van Orden, Johnston, & Hale, 1988; Van Orden, Pennington, & Stone, 1990). The consensus further points to the main function served by phonology in reading: Phonological representations help maintain information in working memory by securing reference (Perfetti & McCutchen, 1982) and coding the serial order information among text units (Seidenberg et al., 1984).

There is also considerable agreement about the possible sources of phonological codes during reading. English orthography, like any other linguistic system, has two principles of processing: **production** and **stipulation.** Production permits the decoding of new words and nonwords by the assembly of their phonological representation on the basis of subword orthographic units. For example, readers may assemble a pronunciation to the nonword *blar* by mapping each of its letters to a phoneme. Of course, the success of such a method in generating the pronunciation of many English words is limited. Consider, for instance, the word *cafe.* Because a final *e* normally does not map to any phoneme, the attempt to assemble a phonological representation to the word *cafe* will most likely fail to produce the correct pronunciation. For the pronunciation of such exception words, a second principle must be postulated—one that assigns a pronunciation by stipulation. The process that implements this principle must retrieve a stored representation that specifies the phonology of the word as a whole (e.g., Baron & Strawson, 1976; Besner, 1990; McCusker et al., 1981; Meyer & Ruddy, 1973; Meyer, Schvaneveldt, & Ruddy, 1974; Paap & Noel, 1991; Paap, Noel, & Johansen, 1992; Patterson & Coltheart, 1987; Patterson & Morton, 1985). To provide an adequate account for the decoding of regular and exception words, reading models must distinguish the relative contribution of stipulated versus productive information in decoding these two types of words.[1]

There is a considerable debate, however, regarding the relative contribution of each of these mechanisms to the identification of familiar words. The question concerns the contributions of the assembly mechanism in the recognition of familiar words. According to dual-route models, although the assembly of phonology is indispensable for the decoding of nonwords and new words, it is not necessary for familiar words, whose phonological forms can be obtained by retrieving their stored representation from the mental lexicon. Assembly in such cases is not only

---

Iris Berent and Charles A. Perfetti, Learning Research and Development Center and Department of Psychology, University of Pittsburgh.

Some of this research was carried out in partial fulfillment of Ph.D. requirements of Iris Berent. We wish to thank Ken Paap for his comments on a draft of this article. We are also grateful to Dan Everett, Ram Frost, Peter Gordon, and Jonathan Schooler for helpful discussions of this project. We owe special thanks to Barbara Frey, Tony Zuccolotto, and Walter Schneider for developing the Micro Experimental Lab code for these studies and to Mara Georgi for her technical assistance in the preparation of this article. This research was supported by National Science Foundation Grant SBR-9223125 to Charles A. Perfetti and by the Learning Research and Development Center at the University of Pittsburgh.

Correspondence concerning this article should be addressed to Iris Berent, 642 LRDC, University of Pittsburgh, 3939 O'Hara Street, Pittsburgh, Pennsylvania 15260. Electronic mail may be sent to BERENT@VMS.CIS.PITT.EDU.

[1] Reading models differ in how production and stipulation are implemented. Dual-route models attribute the productive and stipulated modes to two distinct mechanisms, whereas generalized activation models refer only to a quantitative distinction that is implemented by a single mechanism. In this discussion, we adopt dual-route terminology, attributing stipulation and production to two distinct mechanisms. This is a matter of expository convenience, however, because the choice of an architecture for the implementation of these two principles is not relevant to our main argument. Regardless of the precise representational choices made by a model, we believe that any theory of reading must implement at least a quantitative, if not a qualitative, distinction of these two principles of phonological generation in reading.

not necessary but is sometimes detrimental to recognition: For an exception word, such as *cafe,* the assembly process would yield an erroneous overregularized form (/keyf/ as rhyming with *safe*), instead of the correct stipulated output. Indeed, it has often been suggested that the recognition of written words is accomplished almost entirely by addressing a stored representation of a word. On this account, assembly is essentially a slow fall-back mechanism that has little influence on the identification of familiar words (e.g., Banks, Oka, & Shugarman, 1981; Baron, 1973; Bower, 1970; M. Coltheart, 1978; M. Coltheart, Davelaar, Jonasson, & Besner, 1977; M. Coltheart, Besner, Jonasson, & Davelaar, 1979; Martin, 1982; McCusker et al., 1981; Seidenberg, 1985; Seidenberg et al., 1984; Shulman, Hornak, & Sanders, 1978; Stanovich & Bauer, 1978). Its execution is viewed as a controlled process, attention demanding, and subject to strategic control (M. Coltheart, 1978; Paap & Noel, 1991).

These accounts, however, contrast with some recent evidence that points to a very early and automatic activation of assembled phonology in word recognition. Evidence for assembly is observed within the first 40 ms of word recognition in situations that do not require explicit pronunciation (Perfetti & Bell, 1991; Perfetti, Bell, & Delaney, 1988). Furthermore, assembly is detected even when the decoding of the word clearly interferes with the performance of the experimental task, as in the case of a Stroop task (Dennis & Newstead, 1981). In fact, Van Orden et al. (1990) argued that there is no convincing evidence for a direct visual access in word recognition and that the burden of proof is thus on visual-access theories (see also Carello, Turvey, & Lukatela, 1992; Ferrand & Grainger, 1992; Lukatela & Turvey, 1991, 1993, 1994).

The "What" Question

In contrast to the substantial research on *whether* assembled phonology plays a role in visual word recognition, very little attention has been given to the question of *what* assembled phonology is: What are the implications of the view that an assembled code is phonological? What is its nature and its internal structure?

Most previous research, we believe, has made some tacit assumptions regarding the nature of the assembled representation. One assumption, which we call the **linearity assumption,** views the assembled code as a linear string of phonemes that lacks any internal structure.[2] For example, the phonological representation for the word *sit* is simply viewed as three phonemes, s, I, and t, that are ordered consecutively left to right. This assumption, however, may be incompatible with modern theories of phonology, which view the assembled representation as composed of various levels of organization, or planes, each responsible for the organization of a different type of phonological information, such as features, syllable structure, and tone (Clements & Keyser, 1983; Durand, 1990; Kaye, 1989; Goldsmith, 1990; McCarthy, 1985, 1989). The linearity assumption is further challenged by the results of a recent comprehensive study by Treiman, Mullennix, Bijeljac-Babic, and Richmond-Welty (1993), demonstrating the effect of one such level of phonological organization, namely, syllable structure, on the assembly process and the encoding of grapheme-to-phoneme correspondences in the orthography.

This impoverished linear model assumption can be linked to a more fundamental tacit assumption, **the domain-independence assumption,** about the relationship between reading and specific language competence. The domain-independence assumption implicitly postulates that the assembly mechanisms, and word recognition in general, do not presuppose any linguistic knowledge. The assembly process, as well as the process leading to recognition, lexical access, is viewed as a mechanistic set of simple associations. Addressing the lexicon, on this assumption, is the formation of associations between a storage of graphemic units and a set of stored units encoding stipulated knowledge of the word as a whole, that is, graphemic, semantic, and phonological information. Similarly, assembly is viewed as a simple mechanistic set of associations between graphemes and phonemes. The output of assembly can be thus predicted on the basis of this limited set of associations, without recourse to readers' linguistic knowledge. Such an assumption has a clear implication regarding the architecture of the assembly system: Assembly is viewed as an autonomous mechanism that operates without interacting with a linguistic processor. Assembly, viewed this way, is not a linguistic skill in any interesting sense.

We believe that these assumptions of linearity and domain independence are incorrect. In what follows, we challenge both assumptions. In contrast to the linearity assumption, we argue that the phonological representation assembled in reading has an internal structure that is multilinear. We further claim that this structure determines the course of the on-line process of assembly. Finally, in contrast to the domain-independence assumption, we raise the possibility that the structure of the assembled representation may reflect the presence of linguistic constraints on the assembly process.

Our argument proceeds as follows: In the next section, we outline an alternative model of phonology assembly. We demonstrate the empirical adequacy of the model by showing its ability to account for the existing evidence regarding the role of assembled phonology in word recognition. We further show that the model is adequate also on grounds of computational plausibility. We then explore the model's explanatory adequacy (Chomsky, 1964), specifically, its ability to account for how reading is possible to acquire and compute. Finally, we discuss recent research that directly tests the predictions of the model.

The Two-Cycles Model of Phonology Assembly

The model we propose for the assembly of phonology in English visual word recognition, the two-cycles model, makes two interrelated claims. The first concerns the structure of the assembled representation and the second, its on-line processing. In accord with autosegmental theories of phonology (Archangeli, 1985; Clements & Keyser, 1983; Durand, 1990; Goldsmith, 1990; Kaye, 1989; McCarthy, 1985, 1989), the two-cycles model assumes that the assembled representation entails multiple levels of representation, or planes, defined by their phonological properties. Specifically, we suggest that the assem-

---

[2] Although considerable research effort has been given to the role of morphophonological and even phonological constituents in reading, these studies have been mostly concerned with the structure of the lexical representation, rather than specifically with that of the assembled code.

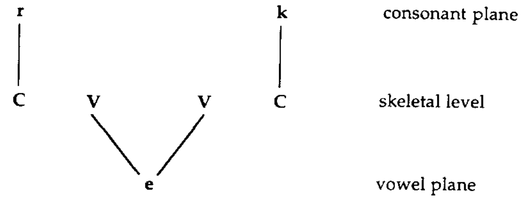

Figure 1. The structure of the phonological representation assembled for the word *rake* as proposed by the two-cycles model.

bled representation distinguishes between consonants and vowels by assigning them to two different planes. The coordination of these two planes is accomplished by a skeleton, an abstract set of timing units. A depiction of this representation is given in Figure 1.

In contrast to the assumption of linearity, the two-cycles model claims that the assembled representation is structured, at least as far as the representation of consonants and vowels is concerned. This claim is closely linked to a processing account for assembly: that the consonant and vowel planes are two distinct constituents in the assembled code. If mental processes are sensitive to such an internal constituency structure, then the distinction between consonants and vowels should emerge in the on-line processing of assembly as well.

According to the two-cycles model, the derivation of consonants and vowels is accomplished by two distinct mental processes that differ in their speed and automaticity. These processes constitute two cycles of assembly whose output is completed at two consecutive stages. The assembly of consonants is computed by the initial cycle through a relatively automatic process. Vowel information is added to the representation at the second cycle by a slower, controlled process. Thus, according to the model, given the output of the initial cycle, the assembled code specifies exclusively consonantal information. At that point, all words and nonwords sharing the same consonants (e.g., *same, some,* and *seem*) are essentially indistinguishable.

An important consequence of the two-cycles model concerns the effects of regularity. Most cases of irregularity in the correspondences between graphemes and phonemes arise from inconsistencies in the mapping of vowels. Thus, in the first cycle, the assembly process should be indifferent to the effect of orthographic irregularity: The phonological representation will be equally reliable for regular and exception words (e.g., *same* vs. *some*). It is only at the second cycle, when vowel information emerges, that the assembly of a phonological representation is completed and a distinction among words sharing the same consonantal information is created. Furthermore, because the assembly of exception words is expected to yield an erroneous, overregularized output, it is the second cycle in which regularity effects are expected to emerge.

The two-cycles model joins the predictions of other models that assume that the activation of assembled phonology is general (Perfetti & McCutchen, 1982; Perfetti et al., 1992). Specifically, it assumes that the output of the initial cycle generates a partial excitation of a word's entry regardless of its frequency. In addition, however, the model postulates a second route to recognition that directly addresses stipulated information. As in some modified dual-route models (Paap & Noel, 1991; Paap et al., 1992), it is this direct route that assures the correct pronunciation of exception words. Thus, the model generally joins dual-route models in assuming that the speed of the direct route and its contribution relative to the assembled route is determined by word frequency. A fast and direct lexical access for high-frequency words accounts for the confinement of regularity and homophony effects to low-frequency words. However, even for high-frequency words, our model assumes that there is always an initial activation based on the consonant cycle. Hence, the contribution of phonology is assumed to be general, and its detection depends on the sensitivity of the experimental task in question. The addressed route can only supplement the contribution of the assembly mechanism, but it cannot override the assembly process altogether.

The model's distinction between the processing of consonants and vowels in reading has several precedents in the English word-recognition literature. For instance, Brown and Besner (1987) suggested that the assembly of any consonant grapheme yields a single output, whereas the assembly of vowels results in a set of multiple outputs. The selection among these multiple candidates is guided by the lexicon. Carr and Pollatsek (1985), Perfetti and McCutchen (1982), and Shimron (1993) suggested that the reliability of consonantal information, consonants' informative value and their lack of temporal duration in the speech signal, might account for their preferred status in the phonological code and the assignment of distinct roles for consonants and vowels in the orthography. Specifically, Perfetti and McCutchen proposed that phonemic activation might initiate with the activation of consonants and spread to the activation of vowels. Their proposal, however, provides a general model of phonemic activation in reading rather than a specific account of the on-line course of assembly. A similar suggestion for a distinction in the time course of the assembly for consonants and vowels is provided by Allen (1979) in his computational model of speech synthesis from text. This proposal, however, does not seem to make an explicit psychological claim regarding the nature of the mental processes underlying assembly.

The two-cycles model differs from previous proposals in some important respects. First, the claims made by the model amount to a psychological account of the mental processes specifically responsible for assembly. Thus, although a distinction between consonants and vowels might emerge at various stages in the reading process, including the maintenance of verbal information in memory (e.g., Cowan, Lichty, & Grove, 1990) and its articulation (e.g., Fowler, 1980, 1983; Öhman, 1966; Perkell, 1969), the model makes a specific reference to the well-defined stage of assembly. Second, the model attributes the consonant–vowel distinction to the presence of two distinct mechanisms underlying their assembly, which differ not only in their onset but also in their computational characteristics. Finally, in the heart of the model is a claim regarding the structure of the assembled representation. The distinction between consonants and vowels, observed at the assembly process, is taken as evidence for the constituency structure of the assembled code. The similarity between the structure of the assembled code, as suggested by the two-cycles model, and that proposed by linguistic theories of autosegmental phonology has some important implications. It is at least possible that this structural resemblance is not merely coincidental. Instead, it might reflect the presence of linguistic constraints on the structure of the assembled code.

This hypothesis has some implications for the view of reading as a linguistic skill.

In the following sections, we develop three lines of argument in support of the model. We first demonstrate the empirical adequacy of the model by exposing some marked contradictions in the literature on English visual word recognition and showing that these contradictions can be neatly accounted for by the two-cycles model. We then argue that the hypothesis of two distinct mental processes underlying the computation of consonants and vowels is plausible from a computational point of view. Finally, we argue that to account for the invention of orthographies, for the acquisition of reading and for its processing, we must postulate some linguistic constraints on word recognition. The ability of the model to meet such constraints supports its explanatory adequacy.

The Empirical Adequacy of the Two-Cycles Model

Despite a large body of research on phonology in word recognition, we believe the question of the role of assembled phonology remains largely unanswered. A close examination of the existing evidence reveals some marked contradictions regarding the nature of the assembly process and its role in visual word recognition. These contradictions, however, seem to be systematically linked to the properties of the experimental method. In fact, it is useful to sort the experimental methods of visual word recognition according to the type of conclusions they provide regarding the nature of the assembly process. The sorting yields two general classes of experimental methods, which we designate, for now, by the neutral categories of **Type A** and **Type B**. Type A, including naming, lexical decision, semantic categorization, and proofreading tasks, using unmasked stimuli, all seem to agree on the view of assembly as a slow, controlled, and unreliable process, whose effect is limited to low-frequency words. In contrast, evidence coming from Type B procedures, obtained using masked stimuli, yields the opposite conclusions: Assembly emerges as a fast, automatic, and reliable process whose effect is general across word frequency.

Below, we review evidence supporting each of these contrasting profiles of assembly emerging from the two types of methods, addressing two fundamental questions: (a) How is it that assembly has such contradictory properties? (b) How do the particular facets of assembly observed for each type link to a particular property shared by its various methods? We show that both questions, which are inexplicable by existing accounts, can be handled by the two-cycles model. Our conclusion is that current models are inadequate precisely because they fail to provide an explicit account of the structure of the assembled code. Such an account is indispensable for gaining a deeper understanding of the nature of assembly as well as to adequately account for the data on phonological assembly in word recognition.

Defining a Set of Criteria for the Detection of Assembly

The issue at stake is not merely whether assembly plays a role in reading English: Such a role is well demonstrated by the fact that readers can pronounce new words and nonwords. Instead, it is the role of assembled phonology in the reading of familiar words, whose pronunciation can be achieved by the addressed route, that is open to debate. To determine the role of assembly in the recognition of familiar words, one needs a set of criteria to evaluate the contribution of assembled phonology relative to the addressed route. We focus on four such criteria: the speed of the assembly process, its generality, its reliability, and its automaticity.

Consider first the **speed** of the assembly process. To demonstrate that assembly contributes to the recognition of familiar words, it is not sufficient to show that an assembled code is computed during reading. Only if this computation occurs at a rate comparable with the computation of the addressed route can we conclude that assembly has a real effect on word recognition. However, because the speed with which a word can be identified varies with its frequency, the actual contribution of the assembled mechanism may depend on word frequency. Hence, the fact that the assembly mechanism contributes to the recognition of words of a relatively low frequency does not imply that the computation of this code is sufficiently fast and general to affect the recognition of high-frequency words as well. Thus, in evaluating the evidence for assembly, it is necessary to determine the **generality** of the contribution across different classes of word frequency. One of the strongest pieces of evidence for assembly is the difficulty in identifying words whose pronunciation cannot be predicted by the assembly mechanism (e.g., *cafe*). The **unreliability** of the assembly mechanism in such cases is thus considered one of its defining attributes. The assumption that assembly errs in computing a phonological representation for exception words is what underlies the use of such words as the paradigmatic case for detecting the contribution of assembly to recognition. Finally, to evaluate the contribution of assembled phonology to the recognition of familiar words, it is also important to determine the extent to which the assembly of phonology is subject to strategic control: Is assembly a relatively **automatic** process, likely to occur in normal reading, or is it merely an optional fall-back mechanism that can be activated or inhibited depending on the demands of the experimental task? As we show, the issues of the speed, generality, reliability, and automaticity of the assembly process are interrelated, and together they define the nature of the assembly process.

The Role of Assembled Phonology in Visual Word Recognition: The Story According to Type A Methods

In reviewing the evidence from Type A procedures, it is useful to further group these procedures according to the general type of rationale that underlies the experimental inference concerning assembly. Most studies can be classified according to three general types of reasoning: One assesses assembly on the basis of the effects of the predictability of the addressed form, a second seeks evidence for assembly by examining the effect of homophony, and a third infers the presence of assembly from the effects of encoding bias in priming.

Effects of the Predictability of the Addressed Form

A crucial feature distinguishing between assembled and addressed mechanisms is their ability to correctly predict the pronunciation of the word. Because the addressed mechanism stores the phonology of the word as a whole, its ability to predict the word's correct phonological representation is, by definition, perfect. Any effects of unpredictability of the word's pronunciation can result only from the operation of the assembled route. The predictability of a word's pronunciation is typically described by two concepts: regularity and consistency. **Regularity** is roughly the extent to which a word's pronunciation can be predicted by applying a set of grapheme-to-phoneme correspondence rules over its graphemic constituents. For instance, the word *come* is irregular, because a vowel followed by a consonant and a final *e* is normally mapped to its free alternate form, as in *home* (Venezky, 1970). **Consistency** refers to the extent to which the pronunciation of a word matches that of other words sharing the same orthographic pattern. For instance, *come* and *home* share the orthographic unit that corresponds to the rhyme, and they are thus considered "neighbors." However, because they disagree on the pronunciation of the vowel, they are considered inconsistent. Note that the inconsistency affects both words, regardless of their regularity. Thus, although the word *home* is regular, its predictability is diminished by the presence of the irregular neighbor *come.*

The two sources of predictability, consistency and regularity, are tightly linked, and it is often quite difficult to determine which of these factors is responsible for the unpredictability of the pronunciation (see also Norris & Brown, 1985; Patterson & Coltheart, 1987). Yet, the consequences of unpredictability are quite clear and well documented: Most models of word recognition assume that the unpredictability of a word's addressed form will result in **interference,** attributed to two distinct sources: the activation of competing phonological correspondences (e.g., Seidenberg & McClelland, 1989; Van Orden et al., 1990) and a mismatch between the output of the assembled and the addressed routes (Paap & Noel, 1991; Paap et al., 1992). Hence, evidence for assembled phonology is the demonstration of interference resulting from the unpredictability of the addressed form of the word, due to either its irregularity or its inconsistency.

Although there is much evidence that the unpredictability of the addressed form interferes with recognition, these effects are limited to low-frequency words. In the naming task, inconsistency impairs the recognition of low-frequency words only (Andrews, 1982; Jared, McRae, & Seidenberg, 1990; Rosson, 1985; Seidenberg et al., 1984). Similarly, regularity effects, observed using both the lexical decision (Parkin & Underwood, 1983; Seidenberg et al., 1984; Waters & Seidenberg, 1985) and the naming task (Andrews, 1982; Paap & Noel, 1991; Rosson, 1985; Seidenberg et al., 1984; Waters & Seidenberg, 1985; Taraban & McClelland, 1987), are limited to low-frequency words.

The fact that regularity and consistency effects seem confined to low-frequency words often has been taken as evidence that assembly is a slow process. On this view, the unpredictability of the addressed form is reflected in a competition between the outputs of the two routes, which is resolved at the photo finish of a horse race (Paap et al., 1992). This competition, however, emerges only when the addressed route is relatively slow, that is, for low-frequency words. In contrast, when the addressed route is fast, that is, for high-frequency words, the competition can be prevented because the addressed route can win the horse race well before the output of the assembly process reaches the photo finish. Thus, that the contribution of the assembled route affects recognition only when the addressed process is slow indicates that the assembled process is also slow.

This fact has been further taken to suggest that assembly is a controlled process, whose execution is resource demanding (Carr, Davidson, & Hawkins, 1978; M. Coltheart, 1978; McCusker et al., 1981). The view of assembly as a controlled process leads to some interesting predictions regarding the generality of predictability effects: If assembly is a slow and controlled process, then under circumstances that require extremely fast performance, disfavor assembly, or limit the availability of attention resources, the magnitude of predictability effects should decrease. Indeed, there have been various demonstrations that regularity effects are subject to strategic control. For example, requiring fast performance in the lexical decision task by a strict response deadline (Stanovich & Bauer, 1978) or by the exclusion of strange words from the experimental list (Waters & Seidenberg, 1985) results in the elimination of regularity effects.

Similarly, assembly is impaired under conditions that disfavor its use. In a naming task, presenting experimental materials in a blocked list of exclusively exception words results in a reduction in regularization errors for exception words (Monsell, Patterson, Graham, Hughes, & Milroy, 1992), a decrease in naming latency (Monsell et al., 1992, Paap & Noel, 1991), and a facilitation in a concurrent secondary task of tone detection (Paap & Noel, 1991). In the lexical-decision task, the inclusion of nonword foils that discourage assembly, pseudohomophones, results in the elimination of evidence for assembly such as interference from inconsistent neighbors and word-length effects (Pugh, Rexer, & Katz, 1994). As the reliance on assembly decreases, stronger lexical contribution is observed as larger frequency effects (Stone & Van Orden, 1993). Finally, the combination of the naming task with a concurrent secondary task of digit recall results in the emergence of frequency effects for regular words (Paap & Noel, 1991) and the elimination of regularity effects for low-frequency words (Bernstein & Carr, 1991; Paap & Noel, 1991). Similarly, the increase in digit load results in the elimination of nonword-length effects in the lexical-decision task (Pugh et al., 1994). The impairment of assembly under load has been attributed to the reduction in capacity available (Paap & Noel, 1991) and to crosstalk from distinct working-memory contents (Bernstein & Carr, 1991). Regardless of the exact mechanisms of impairment, this evidence shows that the assembly of phonology can be eliminated under blocking and load.

In summary, although the predictability of the addressed form affects the ease of recognition, these effects are limited to low-frequency words, and they are subject to strategic control. Such evidence supports the view of assembly as a slow, unreliable, and controlled process.

Homophony Effects in Word Recognition

A second form of experimental logic concerns the effects of homophony. If words are routinely recognized on the basis of their assembled phonology, then the activation of a candidate in the mental lexicon might occur on the basis of **any** input that is indistinguishable from the candidate on the phonological dimension. Thus, homophones (e.g., *hare*) and pseudohomophones (e.g., *bloo*) will have a greater probability of activating the lexical entry of a word candidate, compared with their graphemic controls. Consequently, readers might occasionally confuse *hare* and *bloo* with the intended word (*hair* and *blue*, respectively).

Homophony effects in the lexical-decision task. If readers access the mental lexicon on the basis of the word's assembled form, then they might face difficulties in classifying pseudohomophones as nonwords. Indeed, there are various studies demonstrating that the classification of pseudohomophones (e.g., *brane*) as nonwords is slower and more error prone, compared with nonwords that are not homophones (M. Coltheart et al., 1977; McCann, Besner, & Davelaar, 1988; McQuade, 1981, 1983; Meyer & Ruddy, 1973; Parkin & Ellingham, 1983; Pring, 1981; Rubinstein, Lewis, & Rubinstein, 1971; Underwood, Roberts, & Thomason, 1988). However, the implications of these findings are subject to an ongoing debate. First, the lack of a sufficient graphemic control in many of these studies raises the possibility that the ability of a homophone to activate the target candidate results from its graphemic, rather than its phonological, similarity to the target (Martin, 1982; Seidenberg & McClelland, 1989). Second, Coltheart and colleagues (M. Coltheart, 1978; M. Coltheart et al., 1977) suggested that the slow rejection of a pseudohomophone results from an increase in a response deadline due to an excitation of the logogen, rather than to the retrieval of any particular candidate, that is, the activation of any given word beyond a recognition threshold. Finally, even if the difficulty in rejecting the homophone were truly due to the activation of a candidate by the assembled code, it is still clear that these effects are highly flexible and subject to strategic control: Pseudohomophone effects in this task are affected by the proportion of pseudohomophones in the experimental list (McQuade, 1981), training (McQuade, 1983; Parkin & Ellingham, 1983), and the presence of word homophones (Dennis, Besner, & Davelaar, 1985; Underwood et al., 1988).

In summary, the evidence based on the pseudohomophone effect in lexical decision is generally weak and compatible with strategic control over phonological assembly. An adequate demonstration of the role of assembly must be established on genuine phonological effects, based on a positive erroneous classification of a homophone as a word within a framework that indicates the retrieval of a word candidate. These criteria are better met in tasks demonstrating semantic effects.

Homophony effects in semantic priming. A stronger test for the contribution of assembly is the semantic effects of homophony. If assembly results in the retrieval of a word candidate, then pseudohomophones should yield semantic effects that are greater than those elicited by their graphemic controls.

Using the lexical-decision task, Underwood and Thwaites (1982) demonstrated that the inhibitory effects caused by the priming of the word by an unattended semantic prime (e.g., *waste-rubbish*) are elicited also by a homophone of the semantic prime (*waist*), even when the assembly process is discouraged by the presence of pseudohomophone foils. In contrast, Fleming (1993) failed to obtain semantic priming effects by homophones. Unfortunately, the failure to provide an appropriate graphemic control prevents the attribution of both findings to phonological priming. However, visible real-word homophone primes produced null semantic priming effects relative to graphemic controls in the studies of Lesch and Pollatsek (1993) and Lukatela and Turvey (1994) using the naming task. These authors attributed the failure of visible word homophones to produce semantic priming to their suppression in a verification process.

Indeed, the assessment of homophony effects using pseudohomophones reveals significant semantic priming. In a series of naming studies, Lukatela and Turvey (1991, 1993, 1994) observed equal-magnitude priming effects of pseudohomophones (e.g., *hoap-despair*) and their target word counterparts (*hope-despair*). These effects held even in the presence of load created by a secondary task of digit recognition. Among Type A procedures, these results appear to provide the strongest support for the view of assembly as fast and automatic. Yet, a closer look at these data raises certain reservations. The heart of the debate is not whether assembly can occur but whether it occurs in circumstances in which the addressed route is relatively fast, that is, for high-frequency words or in circumstances that disfavor assembly. The demonstration of assembly in a task that requires the decoding of the pseudohomophones is thus a relatively weak test for the mandatory nature of assembly. Furthermore, the presence of pseudohomophones might result in an additional verification stage that might slow the recognition process and thus artificially increase the effects of assembly. Indeed, the presence of a verification mechanism is suggested by Lukatela and Turvey (1991, 1994) as an account for the null priming effects observed for visible word homophones and the facilitation of target naming by real words. Finally, the strongest test for the speed of assembly is its contribution to the recognition of high-frequency words. The evidence regarding this question is inconclusive. Although in their 1994 study (Experiment 6) Lukatela and Turvey observed equal facilitation by the appropriate primes and their pseudohomophones regardless of the frequency of the appropriate prime, their 1993 findings (Experiment 4) indicate that, for high-frequency primes in the low-load condition, the magnitude of the semantic priming effect for real words seems to be almost three times as big as that obtained for pseudohomophones (20 ms vs. 7 ms). This trend seems to suggest that the contribution of assembly to the recognition of high-frequency words may be rather limited, a conclusion that reemerges in other semantic tasks such as the semantic categorization and the proofreading tasks.

Homophony effects in the semantic categorization task. In this task, subjects determine whether a target word or nonword is a member of a given category (e.g., a "part of the human body"). The critical trials examine the categorization decisions for homophone (e.g., *hare*) and pseudohomophone (e.g., *brane*) foils. On the basis of a series of influential studies using the semantic categorization task (Van Orden, 1987; Van Orden et al., 1988), Van Orden et al. argued for an automatic, fast, and general contribution of assembled phonology in visual word recognition. In these experiments, subjects exhibited significantly more false-positive errors in classifying a homophone foil as a member of the category as compared with a graphemic control. Furthermore, the patterns of false-positive responses to homophones (e.g., *hare*) and pseudohomophones (e.g., *brane*) were comparable in both the rate of false positives and their latency. This finding supports the attribution of homophony effects to the assembled route. Note that for the word homophones, subjects could have used the foil's lexical orthographic form as additional information for avoiding miscategorization. Yet, the availability of an addressed representation for word homophones did not diminish the tendency of miscategorization.

This finding supports a strong version of the role of assembled phonology in recognition, suggesting that recognition is achieved exclusively from assembled phonology (Van Orden, 1987; Van Orden et al., 1988). Furthermore, on Van Orden's account, the contribution of assembled phonology is general: Assembly mediates the recognition of all words, regardless of their frequency. The ability of subjects to avoid miscategorization in most cases is due to a verification process that compares the spelling of the activated entry with that of the input. Hence, the probability of detecting a misspelling is determined by the familiarity with the spelling of the intended exemplar (e.g., *hair*), rather than by the familiarity with the foil (e.g., *hare*). Supporting this claim is Van Orden's finding (1987; Experiment 3) that false-positive errors are unaffected by the frequency of the foil but are greater for low-frequency exemplars compared with high-frequency exemplars.

Although Van Orden's results appear to provide strong support for a general, automatic, and fast activation of assembled phonology in visual word recognition, contrasting with the findings of most studies reviewed so far under "Type A" methods, his conclusions have come under various criticisms. V. Coltheart, Avons, Masterson, and Laxon (1991) argued that assembled phonology is subject to strategic control, finding that an increase in the proportion of homophones among the experimental stimuli from 10% to 25% eliminated the homophony effect for nonwords. (However, Jared & Seidenberg [1991] observed a significant homophony effect for pseudohomophones and low-frequency foils corresponding to low-frequency exemplars when the proportion of homophones was elevated to 83.3%.)

Less equivocal, and hence more serious, is a problem with the generality of assembly observed by Van Orden. Jared and Seidenberg (1991) argued that the observation of homophony effects for high-frequency words in Van Orden's experiments results from two sources of confound. One is the confound of the foil's frequency with that of the exemplar, due to the failure to cross these two factors. The second confound is the priming of the high-frequency exemplars by a top-down activation from their category's name. In support of this view, Jared and Seidenberg (1991) showed that crossing the frequencies of the foil and exemplar and using a higher level semantic category resulted in a cancellation of homophony effects for high-frequency foils altogether. Other experiments by Jared and Seidenberg appear to show that the spelling-verification process postulated by Van Orden is a strategic response to the inclusion of homophone foil trials. For example, in a categorization list without homophone foils, phonological effects were observed only for low-frequency targets. In contrast to the predictions of the verification model, an increase in familiarity with the foil's spelling resulted in an increase in false positives.

In summary, Jared and Seidenberg's (1991) results undermine Van Orden's (1987; Van Orden et al., 1988) claim for the generality of phonological mediation in word recognition. In accord with the conclusions emerging from the study of regularity effects, these data suggest that assembled phonology plays a rather limited role in word recognition, and its effect is observed only for low-frequency words.

Homophony effects in the proofreading task. Numerous studies report subjects' difficulty in detecting the misspelling of homophone foils (V. Coltheart et al., 1991; V. Coltheart, Avons, & Trollope, 1990; Daneman & Stainton, 1991; Doctor & Coltheart, 1980; Treiman, Freyd, & Baron, 1983) as well as pseudohomophones (V. Coltheart, Laxon, Rickard, & Elton, 1988; Doctor & Coltheart, 1980; Van Orden, 1991), compared with graphemic controls. Although these results suggest that assembled phonology contributes to reading, they do not lend themselves to clear conclusions on the role of assembly in the recognition of familiar words. Homophony effects observed for nonwords, whose decoding can be accomplished only by means of assembly, will not do as a demonstration of the role of assembly in real words, for which either the assembly or the addressed route is possible. What is required is an estimate of the relative contribution of these two routes resulting from the comparison of the speed of phonological activation from each route.

Such comparisons are reported by Treiman et al. (1983) and V. Coltheart et al. (1991). In both cases, the key data are the effect of regularity on the magnitude of homophony effects in the proofreading task. If assembly contributes to recognition, then the output of the assembly process will differ from that computed by the addressed route for exception words. Failures to detect misspellings in homophones should be more frequent when the homophonic pronunciation of the misspelling could be obtained from either of the two routes (regular words), than when only an addressed route is possible (exception words). Thus, misspelling of *fare* as *fair* should produce more failures than the misspelling of *nun* as *none,* because *none* is an exception word that allows only the addressed route. In fact, V. Coltheart et al. (1991) observed a higher failure-to-detect rate for regular than for exception homophones. Because the frequency of their stimuli was only moderate (93 per million), however, it is not clear whether the role of assembly could be generalized to high-frequency words.

Treiman et al. (1983) reasoned that if subjects apply the assembly of phonology in recognition, then they might tend to assign an overregularized form to an exception word. Hence, in cases where the target is an exception word, for example, *water,* whose homophony (*waiter*) can be generated only due to overregularization, subjects may be more likely to falsely accept the sentence "The water brought the food" than sentences with regular targets. Such a result was obtained by Treiman et al. (1983), although the difference was significant only in the subject analysis and the frequency of the words examined in the study was rather low. To test the generality of the contribution of assembled phonology, V. Coltheart et al. (1991) performed a similar manipulation and varied word frequency. The results replicated the findings of Treiman et al. (1983): False-positive errors were higher for exception compared with frequency-matched controls. However, the evidence for assembly was rather weak: The higher error rate for exception homophones was observed only among low-frequency words and was significant only in the subject analysis. The results of these two studies thus suggest that assembled phonology does play a role in word recognition, but its contribution is limited to low-frequency words.

Effects of Encoding Bias

A third logic for the study of phonological assembly concerns the effects of encoding bias. If word recognition is mediated by phonology assembly of grapheme to phoneme correspondences, then the priming of any given correspondence should facilitate the recognition of targets sharing the same correspondence (e.g., *bribe-tribe*). In contrast, when the prime and the target word share a set of graphemes whose correspondences are inconsistent (e.g., *couch-touch*), then the preactivation of the competing correspondence by the prime should inhibit the recognition of the target. The observation of such facilitation and inhibition effects in priming thus supports the view that assembled phonology mediates word recognition.

Although evidence for encoding bias has been reported in naming and lexical decision, the effects are not robust and are clearly subject to strategic control. Using a naming procedure, Bradshaw and Nettleton (1974) observed inhibition in target naming by an inconsistent phonological prime when subjects were required to name both target and prime. However, this inhibition disappeared when subjects were instructed to name only the target. The failure to observe inhibition when subjects are not required to pronounce the prime (Bradshaw & Nettleton, 1974; Peter, Lukatela, & Turvey, 1990) suggests that the source of the inhibition is in the articulatory response. The effects of inconsistency need to be examined in a task without an overt pronunciation.

Using the lexical-decision task, Meyer et al. (1974) were the first to report an interference effect with inconsistent prime and a persistent, but nonsignificant, facilitation by a consistent prime. Meyer et al. interpreted their results as demonstrating an encoding bias in the assembly of the target's phonology. However, both inhibition and facilitation effects seem to depend on the structure of the experimental list. Inhibition effects are eliminated in the presence of illegal nonwords (Hanson & Fowler, 1987[3]; Shulman et al., 1978), an increase in the stimulus onset asynchrony between the prime and the target (Hillinger, 1980), and the presence of pseudohomophones (Pugh et al., 1994). Similarly, the facilitation by a consistent prime is not robust: Several failures to obtain consistency effects are reported by Martin and Jensen (1988). In addition to the fragility of encoding bias effects, none of these studies manipulated the frequency of the target and prime, raising the question of generality.

Finally, there are questions about the source of these effects. Inconsistency effects may reflect a mismatch between different outputs of the addressed route, rather than a bias in the assembly process. This possibility is consistent with results suggesting that the pronunciation of nonwords can be biased to form an analogy with an exception word that shares the same body (e.g., pronouncing *nouch* in analogy to *touch*), by priming the nonword directly (e.g., priming *nouch* by *touch*; Burt & Humphreys, 1993; Kay & Marcel, 1981; Taraban & McClelland, 1987), or by a prime that is semantically related to the analogous word (e.g., by the prime *feel*; Rossen, 1983). The critical prediction distinguishing between the lexical account and the encoding bias hypothesis concerns the symmetry of the effect. The encoding bias hypothesis predicts that the effect of inconsistency should be asymmetric: Inconsistency should inhibit performance only when a regular word is followed by an exception word, but not vice versa. However, both Burt and Humphreys (1993) and Lupker and Colombo (1994) have demonstrated that target-naming latency can be affected by the prior presentation of a prime, regardless of the regularity of the prime. These effects are observed even when the prime and targets are separated by nine intervening trials (Burt & Humphreys, 1993), and the direction of these effects (facilitation vs. inhibition) depends on the frequency of the target and the prime's exposure duration (Lupker & Colombo, 1994). These results suggest that these priming effects may be at least partly due to lexical and strategic effects, rather than a short-term encoding bias during the stage of assembly.

In summary, in view of the uncertainty regarding their source, the effects of encoding bias may provide only weak evidence for the mediation of word recognition by assembly. At best, these effects appear to be highly flexible and subject to control.

Evidence for Assembly Using Type B Procedures

Whereas Type A methods suggest that assembly is a relatively unreliable and slow process whose effect is limited to low-frequency words and subject to strategic control, a second class of methods, termed Type B for now, suggests just the opposite conclusion.

Type B studies use masking and priming procedures. Importantly, in these methods, the stimuli are masked, and their exposure duration is brief. In the masking studies, a target word (e.g., *blue*) is presented for a brief moment and then replaced by a word or nonword mask and a visual pattern. The subject's task is target identification. The priming studies reverse the presentation order: The target is preceded by a briefly presented word or nonword prime. In studies where the target is further masked and its duration is brief, subjects are required to identify the target. In settings in which the target duration is unconstrained, subjects are typically asked to perform naming or lexical decision to the target. In masking, the brief exposure duration of the target and mask (typically, around 30 ms each) permits only a probabilistic identification. The mask impairs recognition of the target by interrupting recognition processes initiated by the target. However, to the extent that the information provided by the mask can reinstate such processes, the impairment can be reduced. The critical question is thus whether

---

[3] Hanson and Fowler's (1987) data do reflect a distinction between consistent and inconsistent primes even in the presence of illegal nonwords: Although both consistent and inconsistent primes led to facilitation, the magnitude of the facilitation was greater for consistent, as compared with inconsistent, words. A distinction between responses to inconsistent versus consistent primes was observed also by Treiman et al. (1983) in a forced-choice sentence completion task. Their data reflect inhibition in responding to sentences containing both consistent and inconsistent words, although the magnitude of the inhibition was greater for inconsistent words. Van Orden et al. (1990) suggested that the facilitation for inconsistent prime–target pairs in the presence of illegal words may reflect subjects' ability to adjust the level of cleaning-up of the phonological code depending on the nature of the nonwords. When the discrimination between words and nonwords can be determined on the basis of a partial phonological output, subjects may base their decision on a relatively noisy code that does not specify inconsistency and thus provides facilitation for both consistent and inconsistent primes. Although this suggestion neatly accounts for facilatory inconsistency effects in the presence of illegal nonwords, this explanation cannot account for the results of Treiman et al. (1983), for the failure of Hillinger (1980) to observe inhibition in the presence of legal nonwords, and for the failure to observe facilitation by consistent rhymes reported by Martin and Jensen (1988).

the reinstatement of the word's phonological properties will provide any additional benefit over the reinstatement of graphemic information. This question is examined by comparing the effect of two masks that are equated in terms of their graphemic similarity to the target but differ in the amount of phonemic overlap: a graphemic mask (e.g., *blar*) and a phonemic mask (e.g., *bloo*). If assembly mediates recognition, the reinstatement of the target by the phonemic mask should improve recognition compared with the graphemic mask. The priming logic is similar: The brief processing of the prime is interrupted by the appearance of the target. If assembly mediates recognition, the prior activation of the target's grapheme to phoneme correspondences by the prime should facilitate its processing.

Note that the rationale underlying Type B methods differs considerably from that guiding Type A procedures. Whereas Type A methods all examine the effect of phonological information on the outcomes of the identification process, Type B examines the contribution of phonological information before its final output by interrupting the identification process. Although this rationale of reinstating and preactivating phonological information is superficially similar to the encoding bias and preactivating rationale, the interruption of the identification process under Type B methods critically distinguishes the implications of these two sets of findings. In reviewing the studies based on the encoding bias rationale, we argued that the attribution of the observed phonological priming to the assembled, rather than the addressed, route remains uncertain. By contrast, several observations with the masking and priming methods secure the interpretation that phonological assembly is responsible for phonological effects. Specifically, the observation of masking and priming effects by nonwords, whose pronunciation requires a productive mechanism, the emergence of these effects in a setting that permits only a probabilistic identification of the target, the independence of such effects of word frequency, and the failure to observe phonemic effects in Chinese, an orthography that does not support the assembly of phonology (Perfetti & Zhang, 1991), all suggest that the source of the phonological code reinstated by the mask is in the assembled rather than in the addressed route.

Priming and masking studies provide strong evidence for an early contribution of assembly in both the shallow orthography of Serbo-Croatian (Lukatela & Turvey, 1990a, 1990b) and the deeper orthography of French (Ferrand & Grainger, 1992, 1993, 1994). In English, priming a target word by its homophone (Humphreys, Evett, & Taylor, 1982) or its pseudohomophone (Perfetti & Bell, 1991) increases its recognition accuracy compared with a graphemic control.[4] Similarly, Pollatsek, Lesch, Morris, and Rayner (1992) observed a facilitation in both naming and a silent reading task by a parafoveal preview of a homophone. Further support for the role of assembled phonology in the activation of a word's meaning emerges from studies of semantic priming. Lukatela and Turvey (1994, Experiments 8 and 9) observed significant facilitation in target naming (e.g., *frog*) by a pseudohomophone (e.g., *tode*) of its semantic associate (e.g., *toad*) compared with graphemic- and frequency-matched controls. This facilitation did not differ from that produced by the semantic associate. Similar evidence for priming effects by real-word homophone primes (e.g., *towed-frog*) that were equal in magnitude to those obtained using the appropriate prime (e.g., *toad*) were observed by Lesch and Pollatsek (1993) and Lukatela and Turvey (1994, Experiment 5). Interestingly, the advantage of real-word homophones compared with graphemic controls was observed only when the prime's exposure duration was brief (50–70 ms). The increase in the prime's exposure resulted in the elimination of the homophone advantage, while leaving the advantage of the appropriate prime intact (Lesch & Pollatsek, 1993; Lukatela & Turvey, 1994).

Homophony effects are found using the masking procedure as well. Both Naish (1980) and Perfetti et al. (Perfetti & Bell, 1991; Perfetti et al., 1988) reported that the deleterious effects of masking are reduced when a word is masked by a pseudohomophone compared with a graphemic control.[5] Thus, the priming and masking phonemic effects imply a very fast activation of assembled phonology in word recognition. These data suggest that phonemic information is activated within the first 30 ms of reading (Perfetti & Bell, 1991). The masking results further suggest that assembly is activated as quickly as the graphemic information.[6] Using this technique, no duration has been found so far that yields graphemic effects without phonemic effects. This observation converges with the results of semantic priming studies showing no advantage for the appropriate prime over its pseudohomophone or homophone when the prime duration is brief (Lesch & Pollatsek, 1993; Lukatela & Turvey, 1994). Thus, in complete contrast to the conclusion emerging from Type A procedures, these studies suggest that the activation of assembly is extremely **fast.** The speed of assembly and its occurrence in circumstances that do not require an articulatory response and do not allow conscious identification of the prime and mask are compatible with the view of assembly as an **automatic** process.

---

[4] In their 1982 study, Humphreys, Evett, and Taylor obtained phonemic effects only for real-word homophone primes but not for pseudohomophones. Consequently, they attributed the source of these phonemic effects to the addressed rather than the assembled route. This conclusion, however, is incompatible with numerous recent demonstrations of priming effects by pseudohomophones (e.g., Lukatela & Turvey, 1991, 1993, 1994; Perfetti & Bell, 1991). It is possible that this null finding is due to the fact that many of the pseudohomophone primes used by Humphreys et al. (1982) were orthographically ill-formed. Similarly, in an earlier study, Evett and Humphreys (1981) failed to obtain phonological priming by words sharing the same body but differing in their onsets. However, Humphreys et al. (1982) suggested that this null effect may be due to a lack of sufficient phonemic overlap between the target and prime. A Stroop-like interference due to a mismatch between the onsets of the target and prime was also suggested by Forster and Davis (1991).

[5] A study by Jacobson (1976) reported an inhibition in the recognition of a word masked by a homophone compared with an unrelated condition. However, some methodological difficulties in the experimental procedure (Naish, 1980) do not permit a clear interpretation of this effect. Most likely, this effect reflects priming due to the repeated consecutive presentations of the target, rather than a perceptual masking by the homophone (Naish, 1980).

[6] However, priming studies, in both English (Perfetti & Bell, 1991) as well as French (Ferrand & Grainger, 1992, 1993, 1994), suggest that the activation of graphemic information occurs before that of phonemic information. This finding, however, does not undermine the general claim that the activation of assembly is fast and general. In both languages, the assembly process is sufficiently fast to provide a significant contribution to recognition regardless of word frequency.

For the question of automaticity, ruling out strategic control is important in this paradigm as in others. Perfetti et al. (1988) ruled out the possibility that the homophone advantage reflects mask-based guessing by demonstrating that the homophone effect vanished in blank trials in which the mask was presented without its target. Similarly, no effects of homophony were observed when the target was re-paired with masks of other target words. However, the conclusion that the phonemic masking effect is automatic has been challenged by Brysbaert and Praet (1992), who concluded, from a study of Dutch word recognition, that the homophone effect depends on a high proportion of homophone masks. However, there is reason to doubt the validity of this challenge, both because of problems in their results and because of recent findings of homophone effects for a very low proportion of homophone masks.[7] Using an exposure of 45 ms for the target and mask, Perfetti and Georgi (1994) observed phonemic masking effects when the proportion of pseudohomophones was only 8.3%. Similarly, using lexical decision for primed targets that were presented under brief exposure, Ferrand and Grainger (1992) demonstrated a phonemic effect even when the proportion of pseudohomophone distracters was 25%. Moreover, we have obtained evidence for assembly during the performance of secondary tasks, such as digit recall and color naming—conditions in which phonological assembly and reading are clearly contrary to task demands. Such findings, which we describe in the final section of this article, contradict the view that phonemic effects arise from strategic activation. Instead, and in contrast to the conclusions emerging from Type A procedures, they support the view of assembly as relatively automatic.[8]

The masking and priming data suggest not only that phonological assembly is automatic but also that it is **general** and **reliable.** Manipulating the frequency and consistency of the target word, Perfetti and Bell (1991) observed no interaction of the type of mask or prime with either the frequency of the target or its consistency. Similarly, the magnitude of semantic priming effects by homophones (Lesch & Pollatsek, 1993; Lukatela & Turvey, 1994) and pseudohomophones (Lukatela & Turvey, 1994) was unaffected by the frequency of the homophone or the appropriate prime. This conclusion again contrasts with that from Type A procedures, which suggest that the contribution of assembly is limited to low-frequency words. Recall that the restriction of assembly to low-frequency words was taken as *prima facie* evidence for the view of assembly as a slow process. In contrast, the evidence from the masking and priming studies supports the view that the contribution of assembly is general, concordant with the conclusion that it is a fast process.

However, the finding that the reliability of assembly does not vary as a function of the regularity and consistency of the target appears puzzling. Given inconsistent words that are also irregular, the assembly process should compute an overregularized rather than the correct phonological output. For instance, given the target *come,* the output of the assembly mechanism is expected to rhyme with *home.* Such an overregularized output differs from that provided by a homophone mask (*kumm*), because the homophone mimics the output of the addressed, rather than the assembled, route. Hence, the attempt to reinstate the output computed for an exception word by a homophone mask should result in a mismatch, analogous to the interference between the outputs of the assembled and addressed routes outlined previously in accounting for the Frequency × Regularity interaction. The same logic applied to masking and priming studies predicts interference between the homophone mask and the assembly process for exception target words, the effect of which should be to at least reduce the magnitude of the phonemic effect for exception words, especially low-frequency exception words. There was no evidence, however, of such reduction in the homophony effect for either exception or inconsistent words. Although strong inferences regarding the generality of the assembly process based on such null results are not legitimate, the result is quite persistent. Null interactions of consistency with mask types were obtained in

---

[7] In Brysbaert and Praet's (1992) third experiment, the homophone effect, observed under a high proportion of homophone masks (77.8%), vanished when the proportion of homophone masks was reduced from 77.8% to 5.6%. However, a closer examination of Brysbaert and Praet's methodology and results supports a rather different explanation for their data. First, it is not clear whether the failure to observe homophony effects results from the decrease in the proportion of homophones, because Brysbaert and Praet failed to show a significant homophony effect using the same materials even when the proportion of homophones was moderate (Experiment 2). Because the homophony effect was never properly demonstrated (perhaps due to the quality of their materials and specifically to the lack of sufficient graphemic similarity between the target and the mask), it is impossible to attribute its absence to the proportion of homophones. Second, even if the proportion manipulation was indeed effective, its effect may not have necessarily concerned the assembly of phonology. It should be noted that Brysbaert and Praet's design confounds the proportion of homophone masks with the proportion of control masks. Specifically, the decrease in the proportion of homophone masks was accompanied by an increase in the proportion of control masks. Because the control masks share neither graphemic nor phonemic information with the target, the masks create an extreme disruption in the processing of the target compared with both phonemic and graphemic masks, as well as a pattern mask (Perfetti & Bell, 1991). It is thus possible that the contingency of the effect of homophony on the proportion of masks may reflect a strategic response to the proportion of control masks rather than to that of homophone masks. In particular, subjects may adapt to the increase in the proportion of control masks by focusing their attention on the target and ignoring the mask, either completely or partially. Note that on this account, subjects' sensitivity to the proportion of masks reflects a strategic control over the allocation of attention to the mask stimulus, regardless of its content, rather than a control over the activation of phonology per se (see Johnston & McClelland, 1974, for a demonstration of a strategic control over the size of the attention window in a tachistoscopic recognition task). Thus, Brysbaert and Praet's criticism remains inconclusive.

[8] There are two other studies often cited as supporting strategic activation of assembly under tachistoscopic recognition. Carr et al. (1978) demonstrated that the advantage of pseudohomophones over nonwords is present only when subjects expect to form a discrimination between pseudohomophones. Hawkins, Reicher, and Peterson (1976) showed that the difficulty in discriminating between a pair of homophones (*sent-cent*) disappeared when the proportion of homophones was high. Common to these studies, however, is the examination of the effect of variables discouraging assembly on discrimination rather than recognition. Thus, it is possible that the effectiveness of the manipulation resides in the consideration of phonological information at a decision rather than at an encoding stage. This suggestion is further supported by the relatively high levels of recognition reported in those studies. (Recognition accuracy in the studies of Hawkins et al. and Carr et al. was 80% and 75%, respectively, whereas in Perfetti and Bell [1991], accuracy never exceeded 60% with the phonemic masks or primes.)

two different experiments (Perfetti & Bell, 1991, Experiments 1 and 3). A failure to detect the effect of target regularity was reported also by Humphreys et al. (1982). These findings are curious: If the advantage of the homophone mask results from its ability to accurately and completely reinstate the phonological representation of the target, then why is such an advantage still maintained when the properties reinstated by the mask do not match the code assembled for the target? This puzzle is but one in a series of contradictions emerging from the comparison of the conclusions as to the nature of assembly obtained by Type A and Type B methods.

The Puzzle

The comparison of evidence from Type A and Type B methods produces a puzzling contradiction: Although both types explore evidence for assembly, they give contrasting views on its nature. Assembly, according to Type A methods, is a slow and controlled process whose effect is limited and whose output is unreliable. Most critically, studies that examine the effects of regularity and homophony suggest that the effect of assembled phonology is limited to low-frequency words: Thus, it is low-frequency words that show effects of regularity and consistency (Andrews, 1982; Jared et al., 1990; Paap & Noel, 1991; Rosson, 1985; Seidenberg et al., 1984; Taraban & McClelland, 1987; Waters & Seidenberg, 1985); homophony inhibits semantic categorization (Jared & Seidenberg, 1991) and proofreading (V. Coltheart et al. 1991; Treiman et al., 1983) only for low-frequency words; and the effects of encoding bias are generally obtained for low-frequency words (Hanson & Fowler, 1987; Meyer et al., 1974). Furthermore, overregularization errors in reading exception words (Monsell et al., 1992; Treiman et al., 1983), interference in naming exception words (Seidenberg et al., 1984; Waters & Seidenberg, 1985), and the tendency in proofreading tasks to falsely accept exception words whose regularized form is a homophone of a real word (V. Coltheart et al., 1991; Treiman et al., 1983) all suggest that the output of the assembly route for exception words is clearly unreliable. Finally, the slowness of assembly is compatible with its view as a controlled, attention-demanding process that is subject to strategic control (Bernstein & Carr, 1991; McQuade, 1981, 1983; Monsell et al., 1992; Paap & Noel, 1991; Parkin & Ellingham, 1983; Pugh et al., 1994).

The view of assembly that emerges from Type B methods is quite different. The contribution of assembly is general for both low- and high-frequency words (Lesch & Pollatsek, 1993; Lukatela & Turvey, 1994; Perfetti & Bell, 1991). In the masking procedure, evidence for assembled phonology is observed at the earliest point in processing at which the contribution of graphemic information is detected. Such findings suggest that the assembly mechanism may be sufficiently fast to provide a general contribution in the recognition of high-frequency words as well. Even more surprising is the fact that there is no indication that the reliability of the assembled code is affected by the regularity and consistency of the word, that is, the contribution of the homophone is not reduced given inconsistent (Perfetti & Bell, 1991) or irregular (Humphreys et al., 1982) targets. The detection of assembly under conditions that do not permit identification of the mask or prime and target, the speed of assembly (within 30 ms), and its generality all support the view of the assembly as a relatively automatic process.

This state of affairs raises two questions: (a) How is it that assembly has such contradictory properties: both fast and slow, reliable and unreliable, automatic and strategic? (b) How is the particular property of assembly related to the properties of the experimental methods?

Resolving the Contradiction

The existence of these contradictory facets of assembly—its speed, generality, reliability, and automaticity—can be easily handled by a model that assumes the assembled code entails different constituents: The consonant constituent contains information that is mostly reliable in English, and its assembly is executed by a fast and automatic process. Because the assembly of the consonant constituent is fast, its effect can be detected even when the addressed mechanism is relatively fast, that is, for high-frequency words. In contrast, the assembly of vowels is accomplished by a slow and controlled process. Furthermore, in English, where vowels account for most of the irregularity in grapheme-to-phoneme correspondences, the reliability of the assembly of vowels is largely confined to regular words. However, because the assembly of vowels as a whole is relatively slow, its contribution to recognition may not be noticed in high-frequency words. Thus, the two facets reflected by Type A and Type B procedures closely match the two components of assembly predicted by the two-cycles model.

This account, however, still leaves the puzzle of how these distinct facets are related to the experimental methods producing contrasting results. We suggest that the Type A and Type B methods contrast on one essential property, namely, the amount of processing permitted by each type of method, which, in turn, is determined by the exposure duration of the orthographic input (the target and, when present, the prime or mask). Type A methods, including naming, lexical decision, semantic categorization, and proofreading, all share the feature that the amount of processing they permit for the orthographic input is practically unconstrained. Although some of these procedures require a relatively fast response, the method of presentation permits a complete processing of the orthographic input and essentially perfect identification. By contrast, in Type B methods, the orthographic input is masked, and its processing is highly constrained: The exposure duration is determined by the experimenter to allow only a partial processing of the orthographic input.

This difference in the amount of processing permitted becomes crucial within a model that predicts qualitatively different outputs at different times in processing. The two-cycles model indeed predicts that given a brief target exposure, as in the priming and masking techniques, the limitation in the amount of processing will increase the chances of tapping exclusively into the first cycle. At sufficiently brief durations, the representation assembled for the orthographic input specifies only consonantal information, which is highly reliable and derived quickly and automatically. In contrast, given a long exposure duration, as in Type A methods, the assembly is likely to yield the output of the second cycle as well, which includes unreliable vowel information. Because the process of vowel assembly is slow, regularity effects are unnoticed for high-frequency words. Furthermore, given circumstances that disfavor assembly or have high attention demands, the assembly of vowels is expected to be impaired, thereby leading to the elimination of regularity effects in low-frequency words as well, just as reported by the studies of Bernstein and Carr, 1991; Monsell et al., 1992; Paap and Noel, 1991; Pugh et al., 1994; Seidenberg et al., 1984; and Waters and Seidenberg, 1985. Similarly, the inhibition of vowel assembly is expected to reduce the difficulty in the classification of pseudohomophones as nonwords (Dennis et al., 1985; McQuade, 1981, 1983; Parkin & Elling-ham, 1983; Underwood et al., 1988) and the interference by an inconsistent prime (Hanson & Fowler, 1987; Hillinger, 1980; Pugh et al., 1994; Shulman et al., 1978).

In summary, the two-cycles model can provide a principled account of the distinct facets of assembly. It can also account for why experimental methods show different facets of assembly depending on the amount of processing they permit for the orthographic input. By assuming that consonants and vowels are computed in different cycles, the model accounts for the contradictory evidence regarding the role of assembled phonology in visual word recognition and provides some methodological insights concerning how experimental methods have determined research conclusions about the nature of the assembly mechanism.

The Computational Adequacy of the Two-Cycles Model

In addition to its ability to account for empirical results, the model's two-cycles assumption is motivated by an analysis of the mapping functions of consonants compared with vowels. It has been widely recognized that the mapping of consonant graphemes to their phonemes in the English orthography is far more consistent and context independent compared with that of the mapping of vowels (Allen, 1979; Brown & Besner, 1987; Carr & Pollatsek, 1985; Henderson, 1985; Perfetti & McCutchen, 1982). Given that consistent mapping functions are more likely to be performed by a fast and automatic process (Shiffrin, 1988; Shiffrin & Schneider, 1977), the derivation of consonants and vowels may develop into computational processes that differ in their speed and automaticity. Thus, the distinction between the computational processes underlying the assembly of consonants and vowels is independently motivated in view of the consistency of their mapping. Indeed, a faster assembly of consonants is implied by several recent computational models of word recognition (e.g., Kawamoto, 1993; Plaut & McClelland, 1993; Van Orden & Goldinger, in press).

In fact, a consideration of the mapping function of consonant and vowel graphemes suggests that the segregation of input into these two classes might be necessary even at a much earlier stage, namely, at the stage of constructing a graphemic representation. Many of the grapheme–phoneme correspondences in English are sensitive to whether the letters in their graphemic environment are consonants or vowels. For instance, the mapping of the vowel unit in *make* can be predicted on the generalization that a vowel followed by a consonant and a final *e* are mapped to the vowel's free alternate (Venezky, 1970). To achieve the mapping of the *a* grapheme unit it is necessary to determine whether the graphemes in its environment are consonants or vowels.[9] Accordingly, it is conceivable that the assembly process is preceded by a process that sorts graphemes into consonants and vowels. Because the mapping of vowels is more sensitive to the status of its environment as consonants versus vowels, it is plausible that the assembly of vowels is suspended pending the results of an earlier process that sorts graphemes into types and immediately maps the consonant letters onto their phonemes. A segregation of consonants and vowels at the graphemic level is suggested by Caramazza and Miceli (1990) and supported by breakdown patterns due to brain injuries (Caramazza & Miceli, 1990; Cubelli, 1991).

The Explanatory Adequacy of the Two-Cycles Model

So far we have provided two lines of reasoning in support of the model: Its empirical adequacy and its computational plausibility. We have argued that a distinction between consonant and vowel assembly is necessary to account for the word-recognition literature. A processing distinction between consonants and vowels is expected on purely computational grounds, and it is likely to emerge in several existing computational models of word recognition that do not make an explicit structural distinction between the representations of consonants and vowels. However, given two computational models that are empirically comparable in their ability to account for the processing characteristics of a cognitive system, a preference should be given to a model that can explain a wider range of theoretical issues, such as the acquisition of that system and its relation to other systems. Accordingly, we now turn to a third argument, one concerning the explanatory adequacy of the model. Although this argument is speculative, we believe that it potentially has some far-reaching implications for our understanding of reading in general as well as for the adequacy of the two-cycles model. The essence of the argument is this: The identification of some purely linguistic constraints on the assembly process supports the role of linguistic competence as an essential component of reading ability (Perfetti, 1985, 1994). And, given that such general constraints are independently motivated, then the ability of the two-cycles model to meet these constraints supports its explanatory adequacy. The development of the argument is as follows: We first suggest that, to account for the acquisition of reading and for the ability of readers to assign surface phonetic forms to written words, it is necessary to postulate some linguistic constraints on the assembled representations. On the basis of these constraints, we then formulate some specific criteria for the adequacy of reading models in general. Finally, we show that the two-cycles model meets these criteria and is thus preferred on the grounds of its explanatory adequacy.

Against the Domain-Independence Assumption: A Prelude

The claim that reading is subject to linguistic constraints is clearly contrary to a tacit assumption typically made by word-recognition models, namely, the assumption of domain inde-

---

[9] Of course, it is possible to avoid the labeling of graphemes as consonants and vowels if correspondence rules are represented for each grapheme separately. But such a system would fail to show significant generalization about orthographic systems. Moreover, given the enormous increase in the number of stored correspondences required to implement this suggestion, it seems highly undesirable on the grounds of cognitive economy.

pendence. As a prelude to the main argument, we want to suggest why the domain-independence assumption must be rejected: The assembly process cannot be reduced to a simple set of associations between graphemes and phonemes, because the output of the assembly process cannot be predicted without reference to phonological knowledge.

There are various demonstrations that some principles guiding grapheme-to-phoneme correspondences are essentially phonological principles. For instance, Venezky (1970) noted phonotactic influences on the mapping of graphemes to phonemes, such as the influence of particular consonants on the mapping of vowels (e.g., compare the vowels in *bint* vs. *birm* and *scam* vs. *swamp*). Similarly, some of the regularities in the mapping of the graphemes can be described in terms of the assimilation of voicing from preceding consonants (compare *haps* vs. *hags*). Note, however, that such phonotactic constraints do not need to be postulated explicitly, because their effects completely overlap orthographic principles. That is, the fact that the correspondence of graphemes varies depending on the phonotactic properties of their environment can be simply encoded in the assembly mechanism as context-sensitive principles that are orthographic, rather than phonological, in nature. However, we try to show that such a solution will not do, because some of the constraints affecting grapheme-to-phoneme correspondences cannot be reduced to principles that are sensitive only to the identity of other orthographic units. To predict the correspondences chosen by a native speaker, what is needed, instead, is an explicit consideration of phonological principles. We illustrate this point by considering the effect of stress on the mapping function.

The complexity of the principles governing stress assignment in English led written-word-recognition research to the assumption that the assignment of stress is achieved by stipulation. On this view, stress is retrieved along with the word's addressed form the mental lexicon, rather than generated on-line by the assembly process.[10] Whereas readers may well have a stored representation of the word's metrical structure, it is also clear that readers are equipped with a productive mechanism that permits stress assignment, because they are certainly able to assign stress when their lexicon cannot provide them with such a prestored representation, that is, in reading multisyllabic nonwords. Our claim is that although the stress assigned by any native speaker to a given string cannot always be fully predicted by phonological principles, such principles constrain the range of metrical representations that can be generated by a native speaker. For instance, given the nonword *palany,* readers may assign the main stress to either the first or possibly the second syllable, but not to the final syllable. More important, the assignment of stress itself constrains the mapping of the graphemes to phonemes. For example, the mapping of vowels is subject to vowel reduction of unstressed syllables. In pronouncing the nonword *palany,* one's mapping of the vowels is conditioned by stress: If a reader chooses to stress the first syllable, then he or she will also reduce the second vowel to a schwa, resulting in /páləni/. Inversely, a reader who stresses the second syllable will map the first vowel into a schwa (/pəláni/). Thus, although native readers may pronounce the nonword *palany* in several ways, they will not produce certain outputs including /pálani/, /paláni/, and /palaní/. Thus, it is impossible to account for the decoding of multisyllabic nonword strings without taking into consideration the metrical structure assigned to the string. The reader's phonological knowledge determines stress assignment, which, in turn, imposes constraints on the output of the assembly process.

This example demonstrates an important interdependence between the mapping of the vowel and the metrical structure of the word, which can be accounted for only by the consideration of the reader's phonological knowledge. This suggests the assumption of domain independence is incorrect. In fact, to account for the invention of orthographies, the acquisition of reading, and the computation of phonetic surface forms in reading, it may be necessary to postulate some linguistic constraints on reading. Below, we consider two such types of constraints: on the type of units encoded by the orthography and on the structure of assembled representations.

Linguistic Constraints on the Type of Units Encoded by the Orthography

The argument we develop here closely follows the proposal of Mattingly (1992). An essential property of all orthographies is their productivity, which is a consequence of compositionality or the existence of spelling principles. Thus, to invent orthographies and to acquire reading, readers must become aware of such spelling principles. An examination of the spelling principles across orthographies, however, reveals that essentially any orthographic unit corresponds to only certain linguistic units, either a syllable or a phoneme. Human beings appear not to invent some arbitrary correspondences between acoustic units and graphemic signs. Instead, spelling principles are discovered on the basis of readers' linguistic knowledge. Thus, readers' linguistic knowledge of their native language and its morphophonological properties determines the type of units encoded by the orthography.

This idea can be reformulated in the form of the following criterion for the adequacy of reading models: A model of reading is adequate if and only if the types of phonological units it assumes are also differentiated by the grammar of the language in question. Thus, to determine whether the two-cycles model is adequate, we want to know whether the model's distinction between consonants and vowels is also given by the grammar of English. This criterion of adequacy is rather weak, requiring merely the following: If the grammar distinguishes between two types of units, and these units are also represented by the orthography, then they must be assigned distinct representations by the model of reading. This criterion does not constrain how reading models must capture the distinction. A distinction between consonants and vowels, for example, might be captured at the level of the orthography by distinguishing the graphemic properties of consonant and vowel letters without a further dis-

---

[10] However, for the possibility of stress assignment by the assembly mechanism, see Colombo and Tabossi (1992) and Monsell, Doyle, and Haggard (1989). The involvement of such a productive mechanism in the assignment of stress to words is further supported by Miceli and Caramazza's (1993) report of an Italian patient exhibiting overregularization errors in the assignment of stress to words whose stress pattern is unpredictable.

tinction that is specific to the mental processes and outputs of the assembly process.[11]

We thus consider an alternative account, on which the constraints imposed by speakers' phonological knowledge are much stronger: These constraints strictly determine the structure of the assembled representation, rather than merely suggesting the type of distinctions that the orthography may encompass. Furthermore, the construction of that particular representation is indispensable for the execution of the mental processes underlying lexical access, rather than merely for providing guidance to the evolution of the orthography.

Linguistic Constraints on the Structure of Assembled Representations

Readers' ability to read nonwords suggests that they are able to assign nonwords a surface phonetic form, thus applying phonological processes to the output of the assembly mechanism.[12] However, if we assume that phonological processes operate within a module (Fodor, 1983), a computational system that is domain specific, then the output of assembly must be written in a computational language that can be read by the module. In other words, the assembled output must qualify as a well-formed phonological representation: one subject to linguistic constraints of well-formedness. These well-formedness conditions, however, are not specific to the outputs of assembly. Within a phonological module, any of its inputs must conform to these well-formedness conditions. If phonological representations assembled in reading and those assigned to spoken language are subject to similar well-formedness conditions, then we should expect them to share some significant structural similarities. In fact, these representations must be **isomorphic.** This expectation may be formulated by the following principle:

> *The isomorphism condition:* The phonological representations computed in reading, assembled or addressed, are ismorphic to those assigned to spoken language, and their structure is determined by well-formedness conditions given by the grammar of the language.

The isomorphism condition imposes a very strong constraint on the adequacy of phonological and reading theories. If this principle is correct, then reading and phonological theories may have a mutual adequacy relationship: A theory of assembly is adequate if and only if the structure of the representation it predicts is isomorphic to that motivated by linguistic evidence. Similarly, evidence on the structure of the phonological code computed during word recognition may have clear relevance to the adequacy of phonological theories. Furthermore, to the extent that any universal constraints affect phonological representations, they should be reflected in both phonological theory and reading theory. To evaluate any model of reading, one needs to discover the well-formedness conditions imposed by the grammar of the language. In particular, to determine the adequacy of the two-cycles model, one needs to determine whether a distinction between consonants and vowels, predicted by the model, is also given by the grammar of English. This requires a brief overview of the structure of phonological representations as proposed by modern autosegmental theories of phonology.

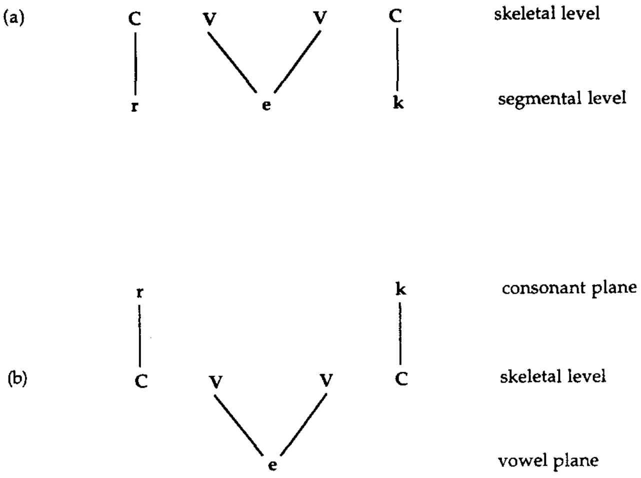

Figure 2. Two alternative representations for the word *rake.* Both representations entail a distinction between consonants and vowels at the skeletal level. However, representation (b) also segregates consonants and vowels at the segmental level by assigning them to different planes.

The Distinction Between Consonants and Vowels in Autosegmental Theories of Phonology

Autosegmental phonology views phonological representations as multidimensional organizations in which different types of phonological information, such as tone, syllable structure, and phonological features, are each arranged on a different level (Archangeli, 1985; Clements & Keyser, 1983; Durand, 1990; Goldsmith, 1990; Kaye, 1989; McCarthy, 1985, 1989). The representation can be thought of as an open book, in which the different pages each display a certain type of phonological information. The coordination of the different planes is achieved by a skeleton: A set of abstract timing units that mediates the planes, analogous, in the book metaphor, to the book's binding (Durand, 1990).

Within such a representation, two forms of differentiation between the representation of consonants and vowels have been proposed (see Figure 2): One is a distinction at the **skeletal** level. On this proposal, the skeleton has two distinct sets of timing units, one for consonants and the other for vowels (Clements & Keyser, 1983). A second proposal provides differentiation of consonants and vowels at the **segmental** level. On this proposal, consonant and vowel segments are segregated into two separate planes (McCarthy, 1979). Note that the resulting representation is, in fact, isomorphic to that suggested by the two-cycles model. Given such a representation, phonological processes are free to operate on adjacent consonants, ignoring the intermediate vow-

---

[11] As a matter of fact, there are orthographies that make a strictly orthographic distinction in the visual representation of consonants and vowels in the script (e.g., Hebrew and Arabic). Thus, such a hypothesis is clearly plausible. The claim here is that the linguistic constraints might have even a stronger form that applies to the structure of assembled phonology rather than merely to the structure of the orthography.

[12] As far as the addressed form is concerned, there is evidence suggesting that the activation of a surface phonetic representation during reading may be automatic (Frost, 1991; Frost, Repp, & Katz, 1988).

els, and vice versa. The existence of phonological processes that operate in such a fashion, such as the obligatory contour principle in Arabic (McCarthy, 1979), supports this suggestion.

Evaluating the Explanatory Adequacy of the Two-Cycles Model

This brief overview of the structure of phonological representations suggests that some form of differentiation between consonants and vowels may be required by the grammar of English. Thus, the two-cycles model is clearly consistent with the requirement that reading models must preserve the distinction between phonological units made by the grammar of the language, provided that these units are represented by the orthography. As to the stronger claim of isomorphism, we may conclude that a segregation between consonant and vowel planes is consistent with the existence of such a segregation in the grammars of various languages (e.g., Arabic and Yawelmani), although a segregation has not been specifically proposed for English. Even for English, however, a segregation of vowels and consonants is compatible with current autosegmental accounts with the assumption of plane conflation: A segregation early in a phonological derivation is erased by a process of plane conflation, making it invisible to subsequent phonological processes. Indeed, a segregation of consonants and vowels is strongly implicated by evidence from the production and perception of speech. Several models of speech perception and production argue for a nonlinear representation in which consonants and vowels form distinct constituents. Specifically, the production of consonants and vowels is attributed to two distinct systems whose output is coproduced simultaneously (Fowler, 1980, 1983; Öhman, 1966; Perkell, 1969). Vowels are produced as a continuous single stream on which the consonant plane is superimposed. Hearer/speakers' knowledge of the coproduction of consonants and vowels further determines the perception of speech. On this account, speakers apply a vector analysis that separates consonants and vowels into two qualitatively distinct constituents. It is this qualitative distinction that accounts for the illusion of discrete segments inferred from the overlapping interdependent events in the acoustic signal (Diehl, Kluender, Foss, Parker, & Gernsbacher, 1987; Fowler, 1983; Fowler & Smith, 1986). These empirical data can be taken as supporting the segregation of consonants and vowels in the linguistic representation.

The segregation assumption has far-reaching implications not only for the adequacy of the two-cycles model but also for the nature of reading in general. First, if consonants and vowels are segregated in the representations of both assembled phonology and spoken language, there would be evidence for isomorphism between the structure of these representations that supports the explanatory adequacy of the model. Second, if assembled representations conform to linguistic conditions of well-formedness, this implies a strong interaction between the assembly mechanism and the linguistic processor. Specifically, linguistic constraints would define the structure of assembled representations, which in turn determines the on-line assembly of phonology. On this view, the linguistic processor determines the on-line functioning of the assembly system. Such an assembly process would represent a very strong form of interaction between the linguistic processor and the assembly mechanism. Thus, in contrast to the domain-independence assumption, assembly would be clearly affected by the readers' linguistic competence, giving specific meaning to the claim that reading is partly a linguistic skill.

Testing the Two-Cycles Model

The two-cycles model makes two central claims. The first concerns the structure of the assembled representation and its on-line derivation. The second claim concerns the automaticity of the assembly process in general and the relative automaticity of the consonant versus the vowel cycle in particular. An additional important implication of the model is methodological: The model predicts that experimental methods that differ in target exposure duration should reveal different facets of the assembled code. In this section, we provide empirical support for these claims by reviewing some studies that directly address each of these issues. A detailed description of some of these studies can be found in Berent (1993).

The Structure of Assembled Phonology and Its On-Line Derivation: Experiment 1

The central distinguishing feature of the two-cycles model is its focus on the structure of assembled phonology and its temporal derivation. The model claims that the assembled code is computed in two consecutive cycles, a first cycle that generates consonants and a second cycle that generates vowels. This claim leads to two empirical tests: **What** are the constituents of the assembled representation? and **when** in the time-course of assembly do specific constituents emerge? The unique ability of the masking technique to directly assess both the content of the representation and its onset makes it particularly suitable for testing these two claims.

Consider first the issue of the phonological constituents of the representation. The masking procedure allows a test of these constituents by manipulating the phonological overlap between the target and the mask. Target identification accuracy under masking depends on the extent to which the mask can reinstate information that is used in the processing of the target. By comparing recognition accuracy under masks equated on their graphemic similarity to the target, but differing in the phonemic contents they reinstate, we may infer the content computed in assembling the target's phonological code.

The critical prediction of the two-cycles model concerns the consonant and vowel constituents of this code. The test requires that a target be masked by two kinds of nonword masks equated on their graphemic similarity to the target: A consonant-preserving mask and a vowel-preserving mask. For example, the word *rake* is masked either by RIKK, which preserves the consonants but not the vowels of *rake,* and RAIB, which preserves the vowels but not the consonants of *rake.* A third kind of mask is a homophone (e.g., RAIK), which preserves both consonantal and vowel information.[13] (As the example shows, it is nearly

---

[13] In view of the reports of specific articulatory interference between the target and masks due to a mismatch in the onset position (Forster & Davis, 1991), we decided to avoid altering the graphemes occupying the word's initial position in the homophone, consonant-preserving, and vowel-preserving masks, because the reinstatement of a target by a mask of a different onset might generate an additional source of interference that might be confounded with its phonological effect. For instance, the alteration of the onset in the vowel-preserving mask (*rake*-NAKE)

impossible to completely equate the graphemic similarity of the homophone to the target with that of the other two masks.) Finally, as a baseline condition, a fourth kind of mask is included: a control mask. This mask preserves neither the consonants nor the vowels of the target (e.g., BLIN).

These four types of masks yield precise predictions that can test the two-cycles model, if target and mask exposure variation is also introduced. According to the two-cycles model, the content of the representation (the *what* question) can be formulated only in reference to a particular point in time along the assembly process (the *when* question). Relatively brief exposure durations should reflect the outcome of the first cycle, whereas the contents of the second cycle will be observed under longer exposure durations. The predictions of the model regarding each of the cycles are summarized in the following paragraphs.

Evidence for the first cycle. The critical evidence for the first cycle is in the comparison of the consonant-preserving (RIKK) and vowel-preserving (RAIB) masks: If the representation at the first cycle contains primarily consonantal information, then at brief exposure durations, the masking of *rake* by its consonants should produce higher accuracy than a mask that preserves the vowels. A strong version of this claim states that, at the first cycle, the code contains exclusively consonantal information. This strong version would predict that the reinstatement of vowels by a homophone mask will not increase recognition accuracy compared with a consonant-preserving mask. Note, however, that the greater graphemic similarity of the homophone mask to the target biases the design against this strong prediction. Finally, because vowel information, which is the major contributor to irregularity in English, is absent at the first cycle, and because the assembly process occurs before lexical access, the effect of the mask should be general with respect to the properties of the target word; that is, the effect of the mask should not interact with regularity and frequency.

Evidence for the second cycle. Longer exposure durations are expected to tap into the second cycle, in which vowels are generated. At the second cycle, the contribution of the vowel-preserving mask to the recognition of the target should be at least equivalent to that of the consonant-preserving mask. Hence, the advantage of the consonant-preserving mask is expected to disappear. At longer exposure durations, we also expect the effect of the mask to interact with the regularity of the target. Given a regular word, such as *rake,* the homophone, RAIK, is expected to reinstate both the consonants of the target and its vowels. Thus, the homophone should have an advantage over all other masks. However, the state of affairs for exception words is different. For exception words, the assembly process is expected to compute an overregularized output, instead of the word's exact form. The reinstatement of this output by a homophone mask should result in a mismatch between the phonological representation assembled for the target and the homophone mask. Hence, for exception words, the homophone mask may interfere with the vowels assembled for the target. Consequently, the homophone advantage should be greater for regular compared with exception words.

The experimental materials contained 64 target words that represented a 2 × 2 combination of Regularity × Frequency (see Appendix A). Each of these targets was paired with the four types of masks described above and counterbalanced across subjects. The stimuli were presented on a computer screen,[14] and the order of trials was random. In other essential aspects, the procedure followed that used by Perfetti et al., 1988, and Perfetti and Bell, 1991. A trial contained three events: A target word, presented in lowercase, which was immediately followed by a nonword mask, presented in uppercase, and a pattern mask. Subjects were asked to write down exactly what they had perceived.

In the experimental session, the target and mask were exposed for one of four duration combinations, 15:30, 30:30, 45:30, and 45:60 ms, respectively. The first two durations were intended to capture the first cycle, whereas the latter two conditions were designed to reflect the computation of the second cycle. Exposure duration was a between-subjects factor. A cutoff procedure was used to eliminate floor and ceiling performance, with a criterion of 10–90% accuracy. The results are based on 24 subjects who met the criterion in each condition.

Results and Discussion

To test predictions at the first cycle, an analysis of variance (ANOVA; 4 Mask × 2 Duration × 2 Frequency × 2 Regularity) of identification accuracy at the two brief durations was carried out over both subjects and items. Both analyses yielded significant main effects of mask, $F(3, 138) = 64.18$, $p < .0001$; $F(3, 180) = 44.31$, $p < .0001$, for subjects and items, respectively; and an interaction of Mask × Duration, $F(3, 138) = 6.22$; $p < .001$; $F(3, 180) = 5.25$, $p < .005$, for subjects and items, respectively. The effect of mask type in each of the brief exposure durations is plotted in Figure 3.

A series of planned comparisons in each of the two short durations tested the specific predictions of the model regarding the effect of mask type at the first cycle. Because assembly in the first cycle computes consonantal information, the reinstatement of consonantal information is expected to improve recognition compared with the reinstatement of vowels. Consistent with this prediction, there was a significantly greater accuracy for the consonant-preserving mask compared with the vowel-preserving mask at each of the short durations. The consonant mask advantage was nearly identical at each duration: 6.25% at the 15:30 duration, $t(138) = 2.24$, $p < .05$; $t(180) = 2.06$, $p < .05$, for the subjects and items analyses, respectively, and 6% at the 30:30 duration, $t(138) = 2.15$, $p < .05$; $t(180) = 1.97$, $p < .06$, for the subjects and items analysis, respectively. Importantly, the mask factor did not interact with either target frequency or regularity. This result is compatible with the claim that the information specifying regularity, namely, vowels, is not present at the assembled code at the first cycle. Actually, the results obtained at the 30:30 duration support the strong ver-

---

might have generated such an articulatory interference and thus biased the design in favor of the consonant-preserving mask. Whereas the preservation of the target's onset prevented a confound of the consonant advantage in the first cycle, this decision resulted in the limitation of our conclusions. Further research is needed to examine the generality of the results reported here to onset positions.

[14] The precision in the exposure duration of the target and mask was achieved by locking the electron gun to the top of the screen at the beginning of each trial and the use of exposure durations that correspond to full refresh cycles.

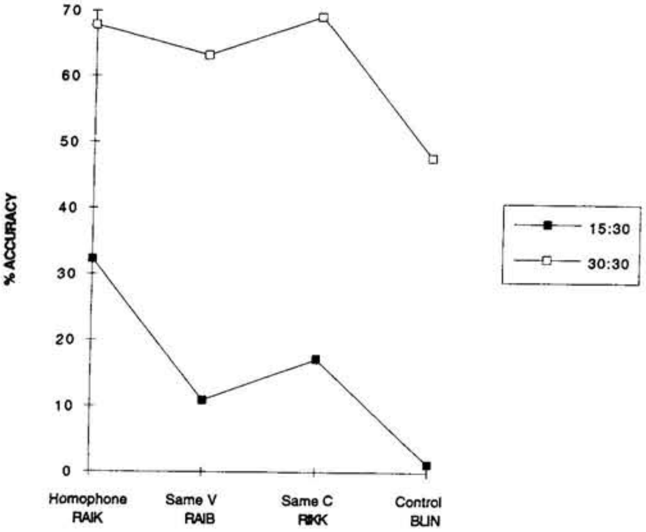

Figure 3. Recognition accuracy in each of the brief exposure durations as a function of mask type. (Same V = a vowel-preserving mask; same C = a consonant-preserving mask.)

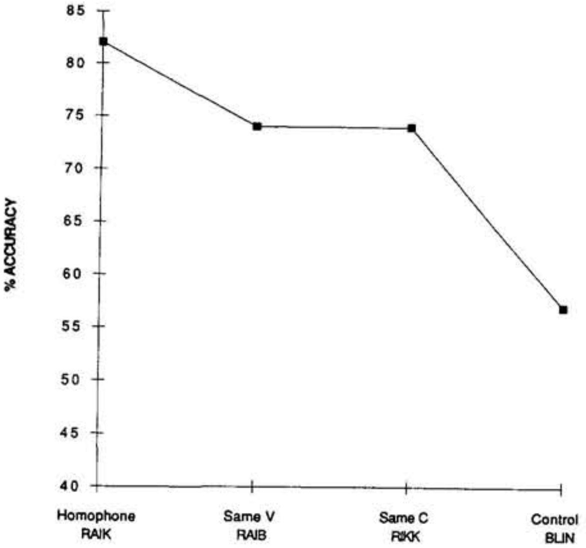

Figure 4. Recognition accuracy at a long exposure duration (target and mask are presented for 45 ms and 60 ms, respectively) as a function of mask type. (Same V = a vowel-preserving mask; same C = a consonant-preserving mask.)

sion of the model that assumes the code assembled at the first cycle specifies exclusively consonantal information. At the 30:30 duration, the reinstatement of the vowels by the homophone provided no additional benefit over the reinstatement of consonants by the consonant-preserving mask ($\Delta = 0.5\%$), $t(138) = 0.17$; $t(180) = 0.2$, *ns,* for subjects and items. However, at the shorter duration of 15:30, there was a significant advantage for the homophone compared with the consonant-preserving mask ($\Delta = 15.1\%$), $t(138) = 5.42$, $p < .0001$; $t(180) = 4.97$, $p < .0001$, for the subjects and items analyses, respectively. We attribute this advantage to the greater graphemic similarity of the homophone to the target word, rather than to the reinstatement of vowels per se. We return to this point later in the discussion.

The two longer durations, the 45:30 and 45:60 conditions, both provide evidence for the generation of vowels at the second cycle. However, because most of the subjects in the 45:30 condition were at ceiling (mean accuracy reached 79%, corresponding to 50.66 points), an additional cutoff of 50 points was applied to this duration. The remaining subjects in this condition ($n = 9$) were analyzed separately from the 45:60 duration. The ANOVA for the 45:60 duration (4 Mask × 2 Frequency × 2 Regularity) revealed a significant main effect of mask type, $F(3, 69) = 16.89$, $p < .0001$; $F(3, 180) = 21.62$, $p < .0001$, for the subject and item analyses, respectively. Recognition accuracy for the four mask types is depicted in Figure 4. The specific predictions of the model were further tested by planned comparisons. Evidence for the generation of vowels was observed in the form of a significant advantage of the homophone over the mean of the consonant and the vowel-preserving masks ($\Delta = 8.25\%$), $t(69) = 2.68$, $p < .01$; $t(180) = 2.93$, $p < .005$, for the subjects and items analyses, respectively. As expected, the advantage of the consonant-preserving mask over the vowel-preserving mask disappeared ($\Delta = 0.5\%$), $t(69) = 0.16$; $t(180) = 0.4$, *ns,* for subjects and items.

In contrast to the model's predictions, the advantage of the homophone did not vary as a function of the word's regularity. However, evidence for the effect of regularity on the homophone advantage is suggested by the data obtained at the 45:30 duration. Although the small number of subjects warrants caution, the homophone advantage was affected by the regularity of the target (Figure 5). For regular words, there was a marginally significant advantage for the homophone compared with the mean of the consonant and vowel-preserving masks ($\Delta = 11.12\%$), $t(24) = 1.42$, $p < .09$, one-tailed. In contrast, for exception words, no such advantage was observed ($\Delta = -1.4\%$), $t(24) = 0.18$.

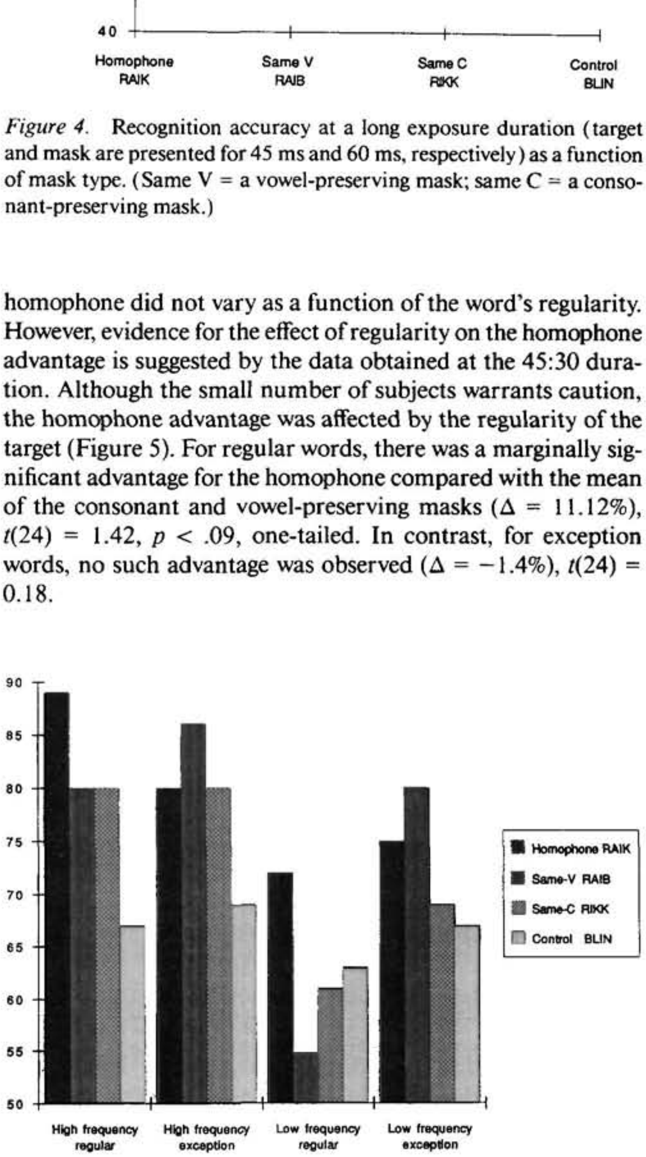

Figure 5. Recognition accuracy at a long exposure duration (target and mask are presented for 45 ms and 30 ms, respectively) as a function of mask and target type. (Same V = a vowel-preserving mask; same C = a consonant-preserving mask.)

The results of this experiment demonstrate that the content of the assembled representation varies with processing time, in accord with the predictions of the two-cycles model. At brief durations, evidence for the generation of consonants was obtained in the form of a significant advantage of the consonant-preserving mask compared with the vowel-preserving mask. In contrast, at longer exposure durations, the generation of vowels was reflected in a significant advantage for the homophone. The results of the 45:30 duration further suggest that the homophone advantage depends on the word's regularity. The sensitivity of the homophone advantage to the predictability of the vowel component indicates that this advantage is specifically related to the emergence of vowels at the second cycle.

One point that remains to be explained is the observation of a homophone advantage at the 15:30 duration. One possible explanation is that although the first cycle computes mostly consonantal information, some vowel information is present as well. However, this explanation raises several difficulties. In general, any account proposed for the homophone advantage of the 15:30 duration must be able to provide an explanation for the global pattern of homophone effect across durations. This pattern is clearly nonmonotonic: The homophone advantage emerges at the 15:30 duration, disappears at the 30:30 duration, and reappears at the longer durations. Accounting for this complex pattern by attributing it to a single source, the activation of vowels, fails to explain the sudden drop in activation at the 30:30 duration and its reemergence at longer durations. A further difficulty for this explanation is the failure to observe any interactions of mask type with regularity at the shorter durations.

We believe that the data can be better accounted for by assuming that the homophone advantage results from two distinct sources: graphemic activation and the phonemic activation of vowels. Recall that the homophone is slightly biased in its graphemic similarity to the target. The advantage of the homophone at the very brief duration may reflect graphemic activation, rather than the assembly of vowels. The early contribution of graphemic activation in visual word recognition is supported by priming studies in French (Ferrand & Grainger, 1992, 1993, 1994), which find a short-lived graphemic activation that emerges before the activation of phonemic information and decays with the emergence of the phonemic effect. Similarly, in their English priming studies, Perfetti and Bell (1991) also observed early graphemic activation before the emergence of the phonemic effect. In the present study, the detection of a homophone advantage at the very brief duration, and its cancellation with the increase in exposure duration, is compatible with the graphemic explanation. Although the early role of graphemic information in recognition clearly merits some further study, the goal here concerns the role of assembled phonological information in word recognition. The contribution of the results is in demonstrating that the consonant and vowel constituents in the assembled representation are assembled in two consecutive stages.

The Mandatory Nature of Assembly: Experiment 2

The two-cycles model further claims that the assembly of phonology is mandatory and relatively automatic. The fact that evidence for the two cycles is observed under conditions that do not require an articulatory response or a conscious identification of the mask is compatible with this view. However, an alternative view emphasizes the strategic nature of assembly, even within the masking procedure (Brysbaert & Praet, 1992). On this view, evidence for assembly obtained with the masking technique may have little relevance to a natural reading setting. To counter this claim, it is useful to obtain evidence for the two cycles of assembly in a situation in which reading, in general, and the assembly of phonology, in particular, are clearly contrary to task demands.

Experiment 2 seeks evidence for two cycles of assembly by using a hybrid of the Stroop task and the masking technique. Evidence for the two cycles of assembly obtained under these circumstances would strongly suggest that the generation of phonology is automatic and mandatory. Furthermore, the demonstration of phonological effects in the Stroop task using the masking procedure validates the use of the masking technique as a means for assessing the structure of assembled phonology.

In this experiment, subjects are presented with words printed in color. The subject's task is to name the color and ignore the content of the words. The words are color names (*blue, green, gray,* and *red*), and there are three conditions of compatibility between the content of the word and its color: A compatible condition, an incompatible condition, and a control, in which the word is substituted by a row of Xs. To the extent that the reading of the word is automatic relative to color naming, subjects will not be able to ignore the content of the word, and color naming will be affected by the congruence between the word's meaning and its color. Each word, however, is masked by one of four masks: A homophone, a consonant-preserving, a vowel-preserving, and a control mask (see Appendix B). And, of course, the exposure duration of the target and mask is manipulated. At any given exposure duration, the relation between the target and the mask should determine the quality of the representation assembled for the word. According to the two-cycles model, under short exposure durations, the reinstatement of the consonants should be more advantageous than the reinstatement of the vowels. In contrast, later on, the reinstatement of vowel information should improve recognition.

The predictions of the two-cycles model concern the quality of the representation assembled for the word. If assembly mediates word recognition, then the reinstatement of phonological information assembled at the current cycle should improve the quality of the representation available for the word. Note, however, that the quality of the word's representation can be detected only in the performance of the color-naming task. To infer the presence of an assembled representation, it is thus necessary to understand the effect of the quality of the word's representation on color naming. If the Stroop effect results from a response competition between the different representations available for the word and the color, then the better the quality of the word's representation, the stronger the competition, and the worse the performance on color naming. Hence, in the incompatible condition, the quality of the word's representation should be inversely related to performance on the color-naming task, and it should depend on both mask type and exposure duration. The rate of color-naming errors in the incompatible conditions should thus mirror the two cycles of assembly: In the brief duration, errors should be greater when consonants are reinstated, whereas at the longer duration, the rate of errors should increase because of the reinstatement of the vowels. Exposure durations for the target and mask at the brief

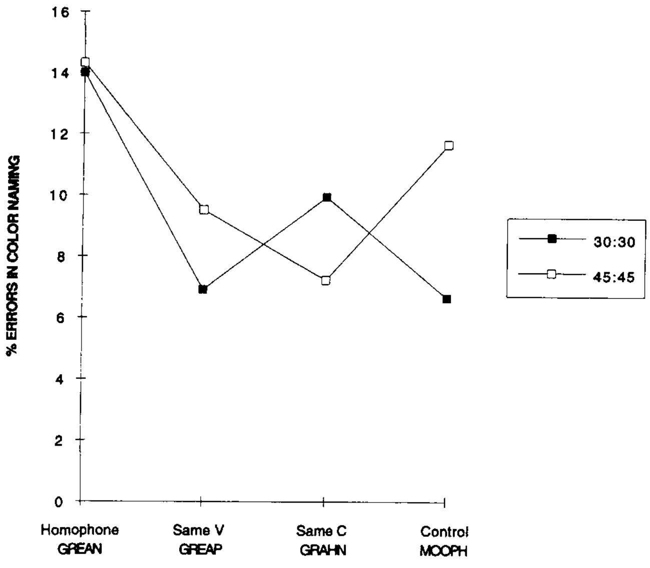

Figure 6. Color-naming accuracy (% error) in the incompatible condition as a function of mask type and exposure duration. (Same V = a vowel-preserving mask; same C = a consonant-preserving mask.)

exposure condition were each 30 ms, whereas at the longer exposure duration, the target and mask were each presented for 45 ms. Twenty-five University of Pittsburgh native English speakers served in each exposure-duration condition.

Results and Discussion

As expected, subjects' performance on the color-naming task was affected by the compatibility between the word's meaning and the color. Overall ANOVA (3 Compatibility × 2 Duration) produced significant main effects of compatibility for both accuracy, $F(2, 96) = 67.16, p < .0001$, and reaction time, $F(2, 96) = 31.9, p < .0001$. Tukey post hoc tests showed that responses in the incompatible conditions were both slower and less accurate than in the control condition ($\Delta = 37$ ms, $p < .01$; $\Delta = 8\%$, $p < .01$). Conversely, responses in the compatible conditions were faster than in the control conditions ($\Delta = 29.14, p < .05$). This replication of the Stroop effect demonstrates the effectiveness of the experimental manipulation combining the Stroop and the masking techniques.

However, the main question is the sensitivity of color naming to the quality of the representation assembled for the word as predicted by the two-cycles hypothesis. At the first cycle, the reinstatement of consonantal information should improve the quality of the word's representation compared with the vowel-preserving mask, resulting in an increase in the error rate for words with a consonant-preserving mask in the incompatible condition. In contrast, at the longer exposure duration, the reinstatement of vowel information should have an advantage over consonantal information, resulting in a higher error rate with both vowel-preserving and homophone masks compared with the consonant-preserving mask. Figure 6 plots the rate of errors as a function of mask type for each exposure duration.

The overall ANOVA (2 Compatibility × 2 Duration × 4 Mask Type) showed a significant interaction of Mask Type × Exposure Duration × Compatibility, $F(3, 144) = 2.82, p < .05$. The model's predictions were tested by planned comparisons. In accord with the predictions for the first cycle, at the brief duration, there was a significant increase of errors with the consonant compared with the vowel-preserving mask ($\Delta = 3.04\%$), $t(144) = 2.13, p < .05$. This state of affairs was practically reversed at the longer duration: There was a marginally significant decrease in accuracy with the vowel-preserving mask compared with the consonant-preserving mask ($\Delta = 2.3\%$), $t(144) = -1.59, p < .06$, one-tailed. In addition, the longer duration showed a significant increase in error rate for the homophone compared with the consonant-preserving mask ($\Delta = 7.08\%$), $t(144) = 4.94, p < .0001$. Although a disadvantage of the homophone compared with the consonant-preserving mask, possibly due to its greater graphemic similarity to the target, was observed also at the first cycle ($\Delta = 4.04\%$), $t = 2.83, p < .006$, the magnitude of this disadvantage increased at the longer duration. This observation is compatible with the hypothesized increase in the activation of vowels at the second cycle.

In summary, performance on the color-naming task is remarkably sensitive to the phonological properties of the mask. The effect of the mask, moreover, varies dramatically as a function of exposure duration. At brief exposure durations, consonantal information is assembled, whereas longer exposure durations give rise to the processing of vowels. These data provide converging evidence supporting the two-cycles model.

Furthermore, the observation of such indirect phonological effects in the Stroop situation has important implications for the issue of the automaticity of the assembly of phonology and the methodological validity of the masking technique. Because word reading is contrary to the demands of the Stroop task, the phonemic effects reflect the mandatory and relatively automatic nature of the assembly process. That such phonemic effects are detected with masking is incompatible with the claim that the phonemic masking effect reflects a guessing strategy (Brysbaert & Praet, 1992). Instead, the discovery of masking effects under Stroop conditions validates the use of the masking technique as a tool for detecting automatic prelexical activation and further establishes its remarkable sensitivity to both the specific phonological contents and the temporal course of assembly.

The Relative Automaticity of the Two Cycles: Experiment 3

The conclusion that assembly is a mandatory process, however, is challenged by the results of studies demonstrating that the assembly of phonology is attention demanding and subject to strategic control. Particularly strong is the evidence of Paap and Noel (1991, Experiment 1), who assessed the attention demands of assembly in a dual-task experiment with word naming and digit recognition as concurrent tasks. The critical manipulation concerned the size of the digit set to be recognized. When the digit set contained a single digit, resulting in a negligible increase in load, evidence for assembly was observed in the form of a regularity effect among low-frequency words: Low-frequency exception words were named slower than regular words. When the processing load was increased to five digits, regularity effects disappeared. In fact, the naming of low-frequency exception words was facilitated. Because the increase in processing load impaired a phenomenon directly linked to assembly, the regularity effect, the elimination of the regularity effect with increased load was interpreted as indicating the impairment of assembly.[15] This finding was taken to support the view of assembly as a controlled attention-demanding process. However, this view of the assembly mechanism seems incompatible with the observation of a fast and mandatory activation of assembly under procedures of brief exposure durations and the demonstration of assembly in circumstances in which its computation is clearly contrary to task demands, such as in the semantic categorization, proofreading, and Stroop tasks.

The two-cycles model attempts to resolve the contradiction by assuming that the assembly process entails two distinct processes that differ in their attention demands. The assembly of consonants is relatively automatic. The control over assembly is limited to the assembly of vowels in the second cycle. Because the model postulates an automatic stage of assembly in which the phonological input is underspecified with respect to vowels, such regularity effects are not a necessary condition for the detection of assembly. What is required, instead, is a direct estimate of the content of the representation in terms of both consonants and vowels as a function of exposure duration.

The masking technique provides a unique opportunity to assess these aspects. To examine the phonological content of the assembly code as well as the relative demands of its two components, we can combine the masking technique with the concurrent secondary task of digit recall. In Experiment 3, subjects were first presented with a memory set of one or five digits to be recalled, followed by a target word, masked by one of the four masks described in Experiment 1. Subjects were asked to write down the word and then to recall the digit set backwards. (Pilot work demonstrated that this increase in the demands of the secondary task was necessary to assure a sufficient memory load.)

The predictions of the two-cycles model are as follows. In general, the attention demands of the assembly process depend on the current cycle of assembly. The first cycle computes consonantal information, and its computation is relatively automatic. Thus, the advantage of the consonant-preserving mask over the vowel-preserving mask should be maintained regardless of load. In contrast, the assembly of vowels under longer exposure durations is more susceptible to load. The effect of load in the second cycle is expected to emerge in two ways. First, under low load, the assembled representation should specify vowel information, whereas under high load, the generation of vowels may be impaired. This may result in a greater advantage of the vowel-preserving mask under the low- compared with the high-load condition. Second, because the assembly of vowels at the second cycle is particularly susceptible to load, we may expect an increase in the overall effect of load at longer exposure durations. A final prediction of the model concerns the relative automaticity of the two tasks. Although the assembly of vowels is resource demanding, the assembly process as a whole is more automatic than digit recall. Thus, attention required for the word-recognition task should be provided at the expense of the digit recall task. A mask that reinstates incompatible phonological information is expected to create interference that draws resources from digit memory. Thus, digit recall should mirror the effect of the mask on word recognition.

Twelve University of Pittsburgh native English speakers were randomly assigned to each of the four conditions created by the combination of load (one or five digits) and exposure duration. The brief exposure-duration target and mask was 30 ms, and the longer duration was 45 ms. The materials were generally the same targets and masks described in Experiment 1.[16] The counterbalancing of the mask type and the word characteristics factors across subjects and materials was the same as in Experiment 1. The order of trials within a list was random.

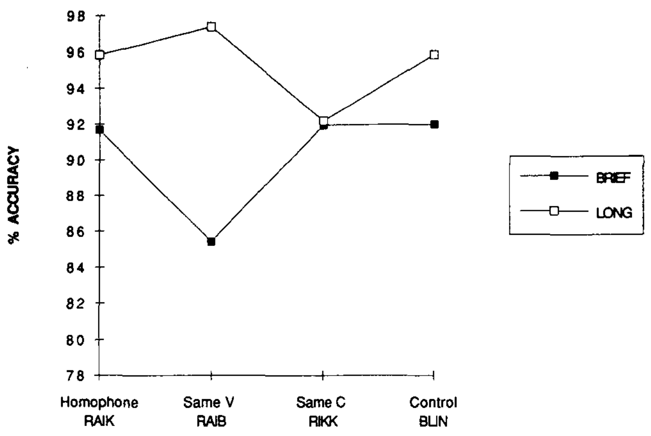

Figure 7. Digit recall accuracy as a function of mask type and exposure duration at the low-load condition. (Same V = a vowel-preserving mask; same C = a consonant-preserving mask.)

Results and Discussion

The predictions for word identification accuracy were tested in a five-way ANOVA (2 Duration × 2 Load × 4 Mask × 2 Regularity × 2 Frequency). The significance of the main effects of duration, $F(1, 44) = 73.3, p < .0001$; load, $F(1, 44) = 25.11$, $p < .0001$; and mask, $F(3, 132) = 13.25, p < .0001$, demonstrates that the primary manipulations were effective.[17] However, there was no indication that the effect of mask type varied as a function of exposure duration and load: The three-way interaction of Duration × Mask × Load and the higher level interactions involving these three factors did not reach significance.

In contrast, accuracy on the digit recall task showed a remarkable sensitivity to the content of the assembled code as predicted by the two-cycles hypothesis. Subjects' accuracy on the digit recall task as a function of mask type and exposure duration is plotted for each of the load conditions in Figures 7 and 8.

The predictions of the model for digit recall were tested by a set of planned comparisons. According to the two-cycles model, the assembly of consonantal information should be unaffected

---

[15] In accounting for the results of their Experiment 1, Paap and Noel (1991) did not claim that the assembly process was eliminated under load but that it was slowed to the extent that its contribution failed to affect the horse race before the photo finish.

[16] Based on subjects' ratings of the quality of our materials, a few of the masks used in Experiment 1 were replaced. These items are provided in Appendix C.

[17] The mean target recognition accuracy for each of these main effects are as follows: Duration (long: 80%, brief: 48.25%); Load (high: 54%, low: 73.25%); Mask type (homophone: 70%, vowel-preserving: 65%, consonant-preserving: 66%, control: 35%).

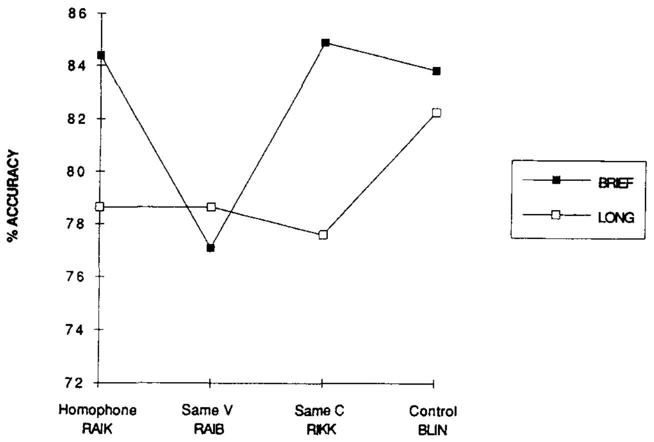

Figure 8. Digit recall as a function of mask type and exposure duration at the high-load condition. (Same V = a vowel-preserving mask; same C = a consonant-preserving mask.)

by load. Indeed, at the shorter exposure duration, there was a significant advantage of the consonant-preserving mask compared with the vowel-preserving mask at each of the load conditions ($\Delta = 7.81\%$), $t(132) = 2.59$, $p < .006$, and ($\Delta = 6.5\%$), $t(132) = 2.11$, $p < .05$, for the high- and low-load conditions, respectively. Furthermore, in accord with the strong hypothesis that the initial cycle computes exclusively consonantal information, there was no evidence for an advantage of the homophone over the vowel-preserving mask in any of the load conditions ($\Delta = 0.52\%$), $t(132) < 1$; ($\Delta = 0.25\%$), $t(132) < 1$, for the high- and low-load conditions, respectively. These findings demonstrate that the assembly of consonantal information is maintained regardless of load. In contrast, at the longer exposure duration, evidence for the assembly of vowels appeared only in the low-load condition. Specifically, with low load, there was a significant advantage of the vowel-preserving mask compared with the consonant-preserving mask ($\Delta = 5.21\%$), $t(132) = 1.69$, $p < .05$, one-tailed. In contrast, there is no such evidence for vowel assembly under the high-load condition ($\Delta = 1.04\%$), $t(132) < 1$. This finding supports the model's claim that it is the assembly of vowels that is eliminated under load, consistent with the assumption that vowel assembly is a resource-demanding process.

To assess the attention demands of the assembly of vowels at the second cycle, the overall effect of load at the longer exposure duration was compared with that observed under the brief exposure duration (Figure 9). At the second cycle, the effect of load is twice the size of that observed under the first cycle (16.01% vs. 7.7%, respectively, $\Delta = 8.3\%$), $t(44) = 2.19$, $p < .05$.

In summary, the data of digit recall support the predictions of the two-cycles hypothesis. The first cycle shows evidence for the assembly of consonantal information alone. In contrast to previous findings by Bernstein and Carr (1991), Paap and Noel (1991), and Pugh et al. (1994), the present results suggest that the contribution of the assembly process is not eliminated under load. Evidence for consonant assembly, in the form of an advantage for the consonant-preserving mask over the vowel-preserving mask, is observed at each of the load levels, indicating that the computation of consonants is accomplished by a relatively automatic process. Moreover, the present study may accommodate previous reports of the elimination of regularity effects under load by pointing to a well-localized effect of load on the assembly process. It is the assembly of vowels, the component accounting for most cases of irregularity in the English orthography, that is affected by load. Evidence for vowel assembly, observed in an advantage for the vowel-preserving mask over the consonant-preserving mask in the low-load condition, is eliminated under high load. The attention demands of vowel assembly are also evident in the form of an overall increase in load effect at the longer exposure duration. At present, our findings are restricted to digit recall accuracy and require replication and extension by some more direct measures of word identification. These results nevertheless converge with our previous masking studies in suggesting that the assembly of consonants and vowels is computed in two consecutive stages. The contrast between the susceptibility of the assembly of consonants and vowels to attention demands provides initial support for the view that these constituents of the assembled code are generated by two processes that differ in their computational characteristics.

The Basis of the Two Cycles of Assembly—The Phonological Versus Graphemic Hypotheses: Experiment 4

The previous sections have demonstrated that consonants and vowels differ in their contribution to word recognition and that their relative contribution varies along the time dimension. The two-cycles model attributes the distinction in the contribution of consonant and vowel information to the structure of the assembled code: Consonants and vowels are two distinct constituents. This structural distinction determines the order in which they are assembled: Consonantal information facilitates recognition before vowel information because consonant phonemes are assembled before vowel phonemes. However, an alternative explanation might attribute the early consonant ad-

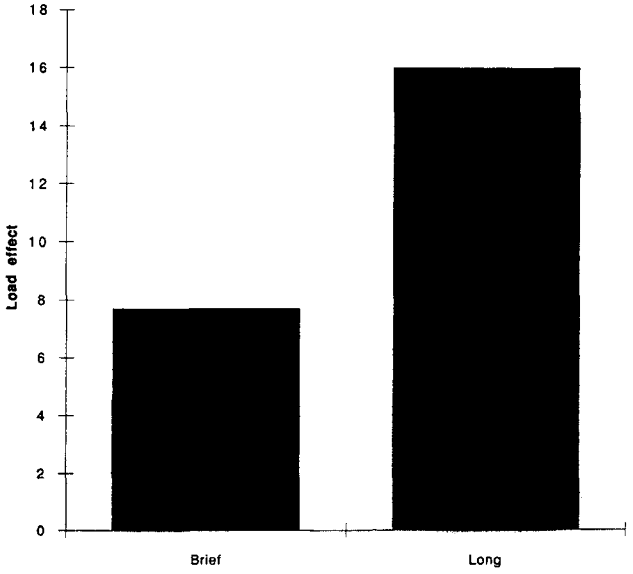

Figure 9. The effect of load (the difference between recognition accuracy under the low- and high-load conditions) on digit recall in each exposure duration.

vantage to graphemic factors. Note that the phonological status of consonant and vowel phonemes is often confounded with their graphemic complexity. For instance, in the word *rake,* the vowel grapheme consists of two letters, whereas the consonant *k* consists of only one. Indeed, in our materials, vowel graphemes are generally constructed of more letters than consonant graphemes. It is thus possible that the early advantage of consonants is due to their graphemic simplicity.

According to a **graphemic hypothesis,** the two cycles reflect primarily the order in which graphemes are parsed: Vowel phonemes are later to emerge because the parsing of vowel graphemes from their letter components is often more complex compared with that of consonant phonemes. This graphemic hypothesis predicts that the early advantage of consonants should emerge only when the graphemic complexity of vowels is greater than that of consonants (e.g., *rake*). Given words in which consonants and vowels are equated in their graphemic complexity (e.g., *pest*), there should be no advantage for the consonant-preserving mask over the vowel-preserving mask. In contrast, if the order of assembly is due to a structural distinction in the assembled code, then the early consonantal advantage should emerge regardless of graphemic complexity. The **phonological hypothesis** thus predicts that the early advantage of the consonant-preserving mask will emerge even for targets such as *pest,* in which the complexity of the vowel is not greater than that of the consonant.

To contrast these two accounts for the two cycles, Experiment 4 compared the contribution of the consonant- and vowel-preserving masks for two sets of monosyllabic targets that differed in the complexity of the vowel spelling. The **complex-vowel** group included 48 words in which the vowel grapheme consisted of multiple letters (e.g., *lace, stain,* and *deep*). For these materials, the graphemic complexity of the vowel graphemes was greater than that of consonants. Hence, an early advantage of the consonant-preserving mask is predicted by both the graphemic and the phonological hypotheses. The critical test for these two hypotheses lies in the case where the graphemic complexity of the vowels is equated to that of consonants. For this end, a second **simple-vowel** group was created. This group included 48 words in which the vowel grapheme consisted of one letter (e.g., *tank, big,* and *went*). For such targets, the phonological hypothesis predicts an early advantage for the reinstatement of consonantal information, whereas no such advantage is expected by the graphemic hypothesis. To examine these predictions, each target was masked by either a consonant-preserving mask, a vowel-preserving mask, or a control mask (see Appendix D). The consonant- and the vowel-preserving masks were matched for their graphemic similarity to the target. For instance, the simple-vowel target *tank* was masked by either TAFK, TINK, or MELF. For the complex-vowel target *lace,* the masks were LABE, LOCE, and MUFT. The exposure duration of target and mask was set to one of three durations: 15:30, 30:30, or 45:60 ms. The two brief durations were chosen to tap into the first cycle, whereas the longer duration was intended to provide evidence for the second cycle. The simple and complex target groups included 2 × 2 combinations of Frequency × Regularity. Within each of the Frequency × Regularity combinations, the simple and complex targets were matched for the number of letters and frequency. To control for the visual salience of consonant and vowel letters, the targets and masks were presented bounded by number signs (#) at the 30:30 and 45:60 durations. Because this procedure resulted in a floor effect at the 15:30 duration, the number signs were omitted at the shortest duration, and a cutoff procedure of 5% was applied. The results are based on 30 subjects who met the cutoff procedure in each duration.

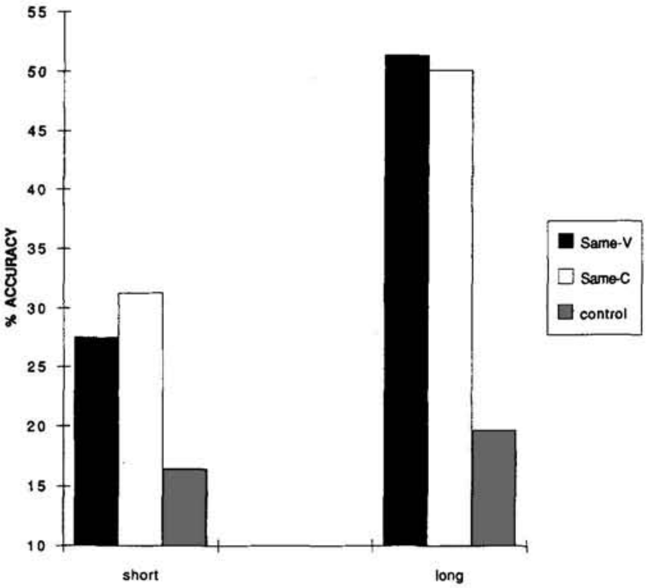

Figure 10. Recognition accuracy at the brief and long durations as a function of mask type. (Same V = a vowel-preserving mask; same C = a consonant-preserving mask.)

Results and Discussion

To examine evidence for the two cycles, subjects' accuracy data in the two brief durations and at the long duration were submitted to separate ANOVAs by subject and by item (2 Complexity × 3 Mask × 2 Frequency × 2 Regularity × 2 Duration and 2 Complexity × 3 Mask × 2 Frequency × 2 Regularity, for the brief and long durations, respectively). These ANOVAs revealed a significant main effect of mask (at the brief durations, $F[2, 116] = 67.91, p < .0001$, $F[2, 176] = 41.83, p < .0001$, by subjects and items; at the long duration, $F[2, 58] = 108.82, p < .0001$, $F[2, 176] = 94.39, p < .0001$, by subjects and items) and no interaction of Mask × Duration. At the short durations, the consonant-preserving mask had a significant advantage of 3.75% over the vowel-preserving mask, $t(116) = 2.81, p < .01$, $t(176) = 2.22, p < .05$, by subjects and items, whereas at the long duration, associated with the second cycle, this advantage disappeared ($\Delta = -0.6\%, t < 1$, by subjects and items) and overall recognition accuracy increased ($\Delta = 15.5\%$; Figure 10). These results replicate the evidence for two cycles obtained in Experiment 1.

Our main interest, however, was in the effect of graphemic complexity on the consonantal advantage. The contrast between the graphemic and phonological hypotheses concerns the advantage of consonantal information for targets in which the graphemic complexity of consonants and vowels is equivalent, namely, for simple-vowel targets. Supporting the phonological hypothesis, the consonant advantage at the first cycle was highly significant for the simple-vowel targets ($\Delta = 6.98\%$), $t(116) = 4.96, p < .001$; $t(176) = 2.967, p < .005$; by subjects and items (Figure 11). In fact, this group of targets accounted for the entire consonant advantage at the short durations, resulting in a significant Mask × Complexity interaction in the ANOVA by subjects, $F(2, 116) = 5.54, p < .01$.[18]

The emergence of consonant advantage for the simple-vowel targets suggests that our previous findings of an early advantage for the consonant-preserving mask are not due to the greater graphemic complexity of vowels. To provide further support for this claim, we performed a post hoc analysis of the data from Experiment 1 by dividing the monosyllabic targets into simple- and complex-vowel groups. This division resulted in 29 targets in the complex-vowel group and 15 targets in the simple-vowel group. To examine the consonant advantage in the first cycle for these two groups, we compared the effect of the consonant-preserving, vowel-preserving, and the control mask at each of the two brief durations used in Experiment 1. The item ANOVA (3 Mask × 2 Duration × 2 Complexity) revealed a significant main effect of mask, $F(2, 84) = 19.16, p < .0001$, and no interactions with the complexity factor. Across the two levels of complexity, the consonant-preserving mask had a significant advantage of 5.68% over the vowel-preserving mask, $t(84) = 1.86, p < .05$, one-tailed. In contrast to the graphemic hypothesis, however, the simple-vowel group accounted for most of the consonant advantage: For simple-vowel targets, the consonant-preserving mask had a significant advantage of 8.9% over the vowel-preserving mask, $t(84) = 1.71, p < .05$, one-tailed, whereas for the complex-vowel targets, this advantage reached only 4.03%, $t(84) = 1.08$, *ns* (Figure 12). This result is especially striking in view of the fact that complex-vowel targets outnumbered the simple-vowel ones.

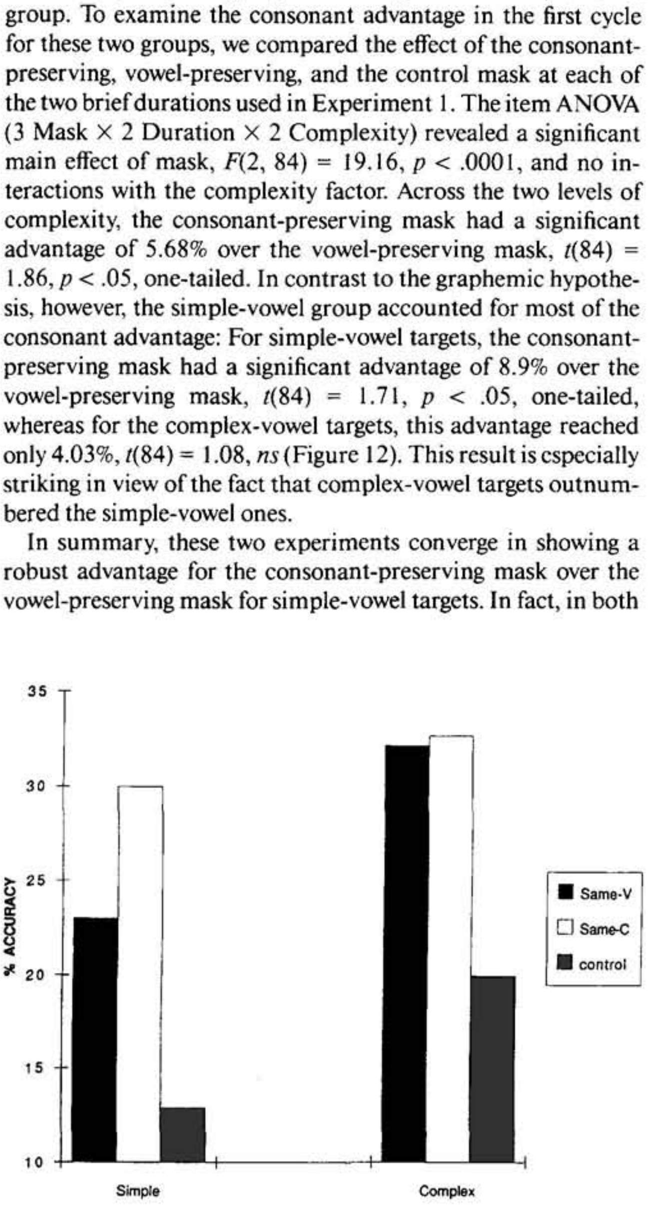

Figure 11. Recognition accuracy at the brief duration for simple- and complex-vowel targets as a function of mask type. (Same V = a vowel-preserving mask; same C = a consonant-preserving mask.)

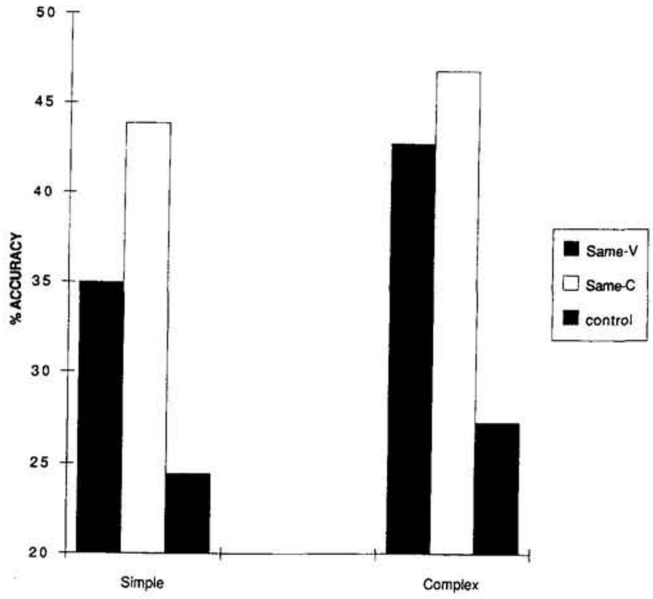

Figure 12. Recognition accuracy at the brief duration in Experiment 1 for simple- and complex-vowel targets as a function of mask type. (Same V = a vowel-preserving mask; same C = a consonant-preserving mask.)

In summary, these two experiments converge in showing a robust advantage for the consonant-preserving mask over the vowel-preserving mask for simple-vowel targets. In fact, in both experiments, the consonantal advantage obtained for simple-vowel targets was greater than for the complex-vowel group. This unexpected finding requires future research. This result, however, is not due to a visual confound, because it was obtained even when the visual salience of the consonant and vowel letters was controlled for by presenting the stimuli bounded with number signs in Experiment 4,[19] nor can it be attributed

---

[18] The ANOVAs performed at the brief durations also revealed significant interactions of Mask × Frequency, $F1(2, 116) = 5.54, p < .01$, $F2(2, 76) = 4.74, p < .01$, and a Mask × Regularity interaction, significant only by subjects, $F(2, 116) = 3.09, p < .05$). The interpretation of these interactions by a Tukey honestly significant difference test suggested that consonant advantage was significant only for low-frequency words and only for regular words. These interactions are due to a consonant advantage for complex vowel targets emerging only for low-frequency regular words. Because as a whole, the complex-vowel targets failed to show a consonant advantage, this finding seems to reflect mainly item variability rather than some principled constraints on the generality of the first cycle. Indeed, for simple-vowel targets, the group that accounts for the entire consonantal advantage, the consonantal advantage emerged in each of the Frequency × Regularity combinations, and its size for high-frequency words ($\Delta = 9.4\%$) was in fact greater than that for low-frequency words ($\Delta = 4.58\%$).

[19] Recall that in Experiment 4, the number signs were omitted from the shortest duration of 15:30 ms because of a floor effect. Hence, in this duration, the location of the consonant for the simple-vowel targets is more visible than that for complex-vowel targets. This state of affairs raises the possibility that the consonantal advantage for simple-vowel targets at the brief durations is driven solely by their greater visual salience at the shortest duration. However, the consonantal advantage obtained for simple-vowel targets at the 30:30 duration was significant ($\Delta = 5.6\%$), $t(116) = 2.85$; $t(176) = 2.05, p < .05$; by subjects and items) and similar in its magnitude to that obtained at the 15:30 duration ($\Delta = 8.25\%$). The interaction of Mask × Duration × Complexity was not significant, $F(2, 116) = 1.72$. Further evidence against the attribution

to some item-specific effects, because it emerged with two distinct sets of materials. However, these data do suggest a rather clear answer to our primary question of interest, namely, the contrast between the graphemic and phonemic accounts for the two cycles. The finding of an early advantage for the consonant-preserving mask for simple-vowel targets, in which the graphemic complexity of consonant and vowel graphemes is controlled, contradicts the graphemic hypothesis that the consonantal advantage is merely due to the greater graphemic complexity of vowels. Instead, these data support the phonological hypothesis that the early advantage of consonantal information reflects the order of assembly. This finding is compatible with the two-cycles model's claim that consonants and vowels are distinct structural components in the assembled representation, and it is this phonological, rather than graphemic, distinction that determines the order of assembly.

A Methodological Conclusion: Experiments 5–7

One of the most important implications of the two-cycles model is methodological. According to the model, the content of the assembled representation varies with processing time. Accordingly, tapping into the assembled code at distinct points along the assembly process should yield qualitatively different conclusions regarding the content of the assembled code. Specifically, we suggest that experimental methods used in investigating the assembly process differ in the amount of processing they allow for the target, and consequently, they tap into distinct stages of the assembly process.

We have shown that Type A methods, those characterized by an unconstrained exposure duration for the target, and Type B methods, for which the target's exposure duration is limited because of masking, produce contradictory accounts about the nature of the assembly process. Within existing atemporal models of word recognition, these contradictions remain unexplained. In contrast, an account for these contradictions easily follows from the two-cycles model's assumption of two temporally distinct stages of assembly.

So far we have demonstrated qualitative changes in the content of the assembled code by varying the exposure duration within one single paradigm, the masking technique. In the following experiments, we attempt to demonstrate that similar changes can be attributed to the choice of experimental method. Our previous analysis of Type A and Type B methods leads to a prediction: For a given set of materials, the results emerging from the naming technique, a member of Type A methods, should differ substantially from the results observed using masking under a brief exposure duration, a Type B procedure. Furthermore, differences between experimental procedures will be directly related to the target's exposure duration.

Our demonstration focuses on the assembly of vowels, evidenced in regularity effects, which are a primary source of disagreement between the results of Type A and Type B methods. Regularity effects in Type A methods are robust; indeed, they are considered the hallmark of assembly. In contrast, the effect of mask in Type B methods is generally unaffected by the regularity of the target. On our account, the absence of regularity effects in Type B methods when exposure duration is brief is due to the absence of vowels in the first cycle. We thus make three interrelated claims: (a) The emergence of regularity effects in Type A methods requires long exposure durations; (b) the contingency of regularity effects on long durations is due to the late emergence of vowels; (c) brief exposure durations that yield null effects of vowel assembly permit the assembly of consonantal information. We illustrate these claims in three sets of experiments. We first demonstrate the contingency of regularity effects on the amount of processing by manipulating target exposure duration using the naming technique. We then proceed to assess the contents of the assembled code at different exposure durations by examining the effect of priming on target naming. We demonstrate that vowel priming and regularity effects emerge under the same conditions. However, circumstances leading to null regularity and vowel-priming effects can nevertheless produce positive evidence for the assembly of consonantal information.

Regularity Effects Are Contingent on the Target's Exposure Duration: Experiments 5a–b

The first step in our inquiry was to demonstrate that regularity effects are contingent on the amount of processing allowed for the target. To illustrate this point, we performed two simple manipulations. First, we subjected the same materials used in Experiment 1, a masking experiment, to a naming procedure. We refer to this condition as **standard naming.** In this experiment the target was clearly visible (900 ms). Second, to secure the interpretation that it is the exposure duration that is specifically critical for the emergence of regularity effects, we then reduced the target's exposure duration to 60 ms (**brief naming**). To control the effective exposure duration of the target, targets were preceded and followed by a pattern mask (XXX) in each of these conditions. Eighteen subjects, asked to name the word as fast as possible, participated in each of these conditions. Response latencies slower than 1,000 ms received negative feedback in the form of a computer message accompanied by a short tone informing the subject that his or her response was too slow. These slow responses (0.9% and 5.6% of the total correct responses in the standard- and brief-naming conditions, respectively[20]) were excluded from the analyses of naming time.

---

of the consonant advantage to visual confounds comes from examining the effect of the position of the masked consonant on the magnitude of the consonant advantage among simple-vowel targets. If the consonant advantage is merely due to the greater visual salience of consonants at external positions in the word (e.g., *t* in the word *sit*), then for words in which the masked consonant occupies internal positions (e.g., *s* in *blast*), no such advantage should emerge. However, for targets such as *blast* (27% of the simple-vowel targets), the consonant advantage reached 13.8% and 17.7% in each of the two brief durations, and its magnitude clearly exceeded the effect observed for simple-vowel targets as a whole in each of these durations (8.32% and 5.63%, respectively).

[20] Supporting the claim that the effect of exposure duration on the emergence of regularity effect is *not due to the exclusion of slow responses*, the distribution of slow correct responses across the two levels of regularity was similar. Specifically, the percentage of slow responses for low-frequency regular and exception words was 0.09% and 0.73% at the standard-naming condition and 1.52% and 2.39% at the brief-naming condition.

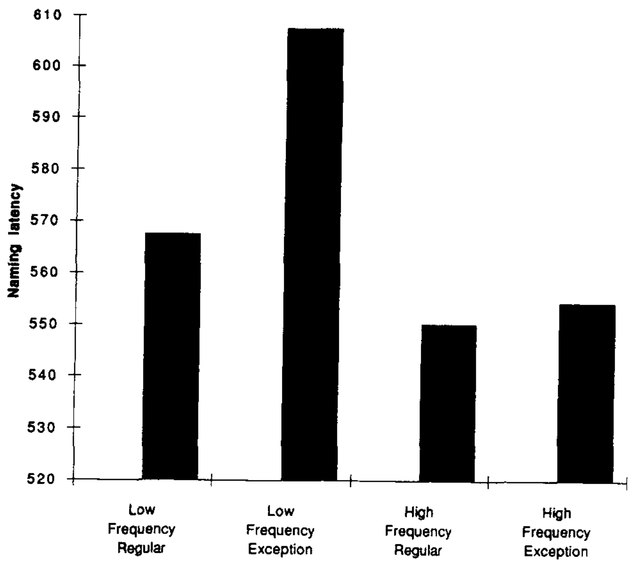

Figure 13. Naming latency in Experiment 5a (standard naming) for each of the Frequency × Regularity target combinations.

Results and Discussion

The results of the standard-naming condition showed strong regularity effects for low-frequency words (see Figure 13). Low-frequency exception words were named significantly slower (Δ = 40.33 ms), $t(17) = 6.02, p < .001$; $t(60) = 2.91, p < .006$, for the subject and item analyses, respectively, and less accurately (Δ = 7.7%), $t(17) = 6.5, p < .001$; $t(60) = 2.05, p < .05$, for the subject and item analyses, respectively, than regular words. The interaction of Frequency × Regularity in the overall ANOVA (2 Frequency × 2 Regularity) on response latencies was highly significant by subjects, $F(1, 17) = 14.56, p < .005$, and marginally significant by items, $F(1, 60) = 3.3, p < .08$. This result agrees with many other results using the standard-naming task (Andrews, 1982; Paap & Noel, 1991; Rosson, 1985; Seidenberg et al., 1984; Taraban & McClelland, 1987; Waters & Seidenberg, 1985). For accuracy, this interaction was significant by subjects, $F(1, 17) = 25.54, p < .0005$, but not items, $F(1, 60) = 2.52, p < .12$.

These results demonstrate that the choice of the experimental method clearly affects the magnitude of regularity effects. Using the masking procedure, no evidence for an interaction of the phonemic masking effect with regularity was obtained in Experiment 1. In contrast, the same materials in standard naming showed strong regularity effects for low-frequency words.

The 60-ms (brief-naming) condition further tests the assumption that it is the exposure duration of the target that specifically accounts for the emergence of the Regularity × Frequency interaction. Although this exposure duration would normally yield a ceiling effect in the masking procedure, the demand for a fast and confident naming response increases the chance that subjects might be forced to skip the assembly of vowels and access the word's entry on the basis of the output of the initial consonantal cycle. Such a strategy could maximize both subjects' speed and their confidence. The elimination of vowel assembly could save a slow and resource-demanding process. In addition, it could increase identification certainty by permitting a verification of the lexical candidate against the visual representation of the target before its masking. Such a strategy, however, is expected to result in the elimination of regularity effects. Indeed, under brief naming, regularity effects disappeared in both naming latency and accuracy analyses (see Figure 14). The naming of low-frequency exception words was not significantly slower than regular words in the subject analysis (Δ = 11.48 ms), $t(17) = 1.18$, *ns*, and was only marginally slower by items, $t(59) = 1.784, p < .08$. Similarly, no significant differences were observed in accuracy for low-frequency regular versus exception words (Δ = 3.77%), $t(17) = 1.79$, *ns;* $t(60) = 0.55$, for subjects and items, respectively. As expected, however, the level of accuracy in the brief-naming (81%) condition- was lower than in the standard-naming condition (97%).

In sum, these results reflect a clear relation between the exposure duration of the target and the emergence of regularity effects: Regularity effects are obtained for low-frequency words given a long exposure duration; the decrease in exposure duration results in the elimination of regularity effects.

Vowel Assembly Requires Long Exposure Durations: Experiments 6a–b

The standard- and brief-naming results demonstrate that regularity effects are contingent on long exposure durations. According to the two-cycles model, this contingency is due to the later emergence of vowel information in the assembled code. Regularity effects are due to a mismatch between the contents of the phonological representation assembled for exception words and their addressed form. However, because the content of the mismatch, mainly vowel, information is late to occur, regularity effects are expected only at longer exposure durations. In the following experiments, we provide more direct evidence for the claim that the contingency of regularity effects on

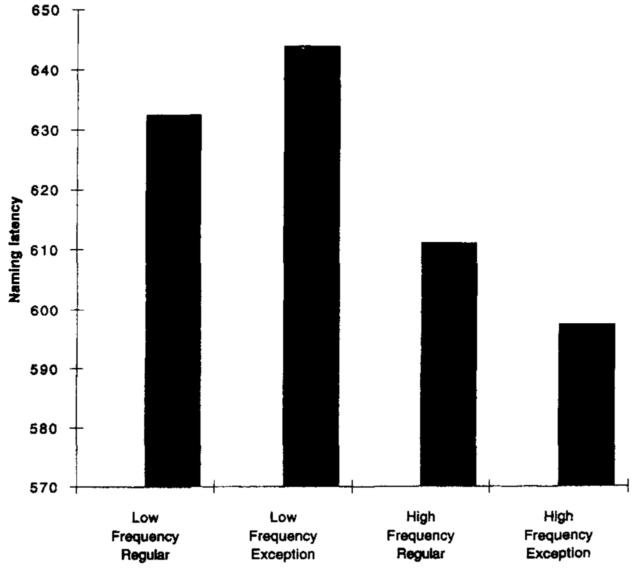

Figure 14. Naming latency in Experiment 5b (brief naming) for each of the Frequency × Regularity target combinations.

exposure duration is due to the late emergence of the second vowel cycle by directly assessing the contents of the assembled code at distinct points along the time dimension. We show that the emergence of vowel information and regularity effects is limited by several factors—exposure duration, frequency, and predictability—that apply to both phenomena in a similar fashion.

To assess the presence of vowel information in the assembled code, we embedded the standard- and brief-naming techniques in a priming procedure.[21] We compared the effect of three types of primes on target-naming latency: A homophone (e.g., RAIK), a consonant-preserving prime (RUKK), and a control prime (BLIN, see Appendix E). The use of only three primes permitted us to equate the homophone and the consonant-preserving prime in their graphemic similarity to the target. Hence, any additional benefit of the homophone compared with the consonant-preserving prime could be attributed to the contribution of vowel information. To examine the dependence of both the contents of the representation and regularity effect on the amount of processing allowed, we further manipulated the exposure duration of the target and prime. At the long-priming condition, the prime and target were presented for 60 and 150 ms, respectively, whereas at the brief condition, their exposure duration was set to 30 and 90 ms, respectively. Because our main interest in this experiment was in the effect of exposure duration on target-naming latency, we attempted to prevent a significant decrease in naming accuracy by using relatively long exposure durations for the target in both conditions. These relatively long exposure durations are expected to permit some vowel assembly for the target in each of the two durations, although the probability of reaching the second cycle, and hence, the strength of evidence for vowel assembly, should increase at the longer exposure duration.

Our primary interest concerns the effect of exposure duration on the contents of the representation assembled for the prime. Because vowel-priming and regularity effects are viewed as two manifestations of the same phenomenon, namely, the second vowel cycle, a common set of constraints is expected to affect them both. These constraints are exposure duration, predictability, and frequency. Because vowel information is late occurring, the priming of the target by its vowel is expected only in long-exposure durations. Hence, in the long-exposure duration, naming latency for the target should be facilitated when preceded by the homophone compared with the consonant-preserving prime. In contrast, the early termination of prime processing at the brief duration is expected to prevent the assembly of vowel information. Thus, at the brief duration, the homophone should provide no additional facilitation compared with the consonant-preserving prime. A second factor that constrains vowel-priming effects is the predictability of the target's addressed form, or its regularity. In general, the prime should facilitate target recognition to the extent that its grapheme-to-phoneme correspondences match those assembled for the target. However, for exception words (e.g., *come*), the phonological representation assembled for the target should yield an overregularized output, rhyming with *home*. Because this overregularized form mismatches the phonological representation assembled for the prime (e.g., KUMM), the homophone is expected to provide no additional facilitation compared with the consonant-preserving prime. Hence, the homophone advantage should be limited to regular words. Finally, vowel-priming effects should be further modulated by the target's frequency. To the extent that vowel assembly is a slow process compared with the addressed route, the homophone advantage should be limited to low-frequency words. High-frequency words are expected to be addressed directly by the consonantal cycle and therefore override both vowel-priming and regularity effects. We thus predict that in long exposure durations, low-frequency regular words will exhibit a homophone advantage compared with the consonant-preserving prime. In short durations, the advantage of the homophone is expected to disappear. However, supporting the notion that this null effect is not simply due to a failure to process the prime, a disadvantage for the control prime should be manifested.

The procedure in these priming experiments was the same as in the standard- and brief-naming studies except for the inclusion of a prime. In this and all subsequent experiments, naming latency was measured from the prime's onset. Eighteen subjects participated in each of these conditions. Response latencies slower than 1,500 ms (1% of the total correct responses in each of the two durations) received negative feedback from the computer and were excluded from the analyses of naming time.

Results and Discussion

Subjects' latency and accuracy in target naming in each of the two durations were analyzed in separate ANOVAs by subjects and items (3 Mask × 2 Frequency × 2 Regularity). The primary question concerned the contingency of vowel information on exposure duration. We hypothesized that the recognition of low-frequency regular words is facilitated by the availability of vowel information. This question was assessed by planned comparisons comparing the effect of homophone and the consonant-preserving prime on the naming of low-frequency regular words in each duration. Subjects' naming accuracy was high (89% and 87.5% at the long and brief conditions, respectively) and showed no significant differences between the homophone and consonant-preserving primes in any of the two exposure durations. In contrast, the latency data support the claim that vowel assembly requires long exposure duration. As expected, at the long duration, the homophone produced significant facilitation in the naming latency of low-frequency regular words compared with the consonant-preserving prime ($\Delta = 53$ ms), $t(34) = 2.14, p < .05$; $t(54) = 1.73, p < .05$, one-tailed, by subjects and items (Figure 15). In contrast, no significant facilitation was observed for low-frequency exception words ($\Delta = 10$ ms, $t < 1$, by subjects and items). However, this advantage of the homophone for low-frequency regular words depended on its exposure duration. At the brief duration, the advantage of the homophone for low-frequency regular words disappeared ($\Delta = -6$ ms, $t < 1$, by subjects and items, Figure 16). Similarly, no facilitation was observed for exception words ($\Delta = -18$ ms), $t(34) < 1$; $t(112) = 1.62$, *ns*, by subjects and items. Supporting the assumption that the elimination of vowel assembly at the brief duration is not simply due to a floor effect, the control

---

[21] To counterbalance the prime and target properties factors, four of the targets used in Experiment 5 (*hotter, finale, past,* and *woman*) were excluded from Experiments 6a–b.

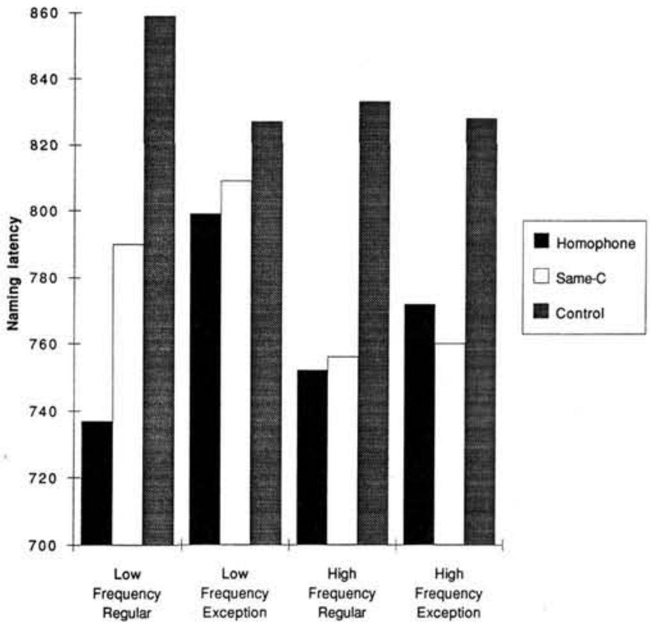

Figure 15. Naming latency in Experiment 6a (long duration) for each of the Frequency × Regularity target combinations as a function of prime type. (Same C = a consonant-preserving prime.)

prime impaired target naming, resulting in a significant main effect of prime type for both latency, $F(2, 34) = 24.3, p < .0001$, $F(2, 112) = 15.19, p < .0001$, for subjects and items, and accuracy, $F(2, 34) = 5.57, p < .01$, $F(2, 112) = 5.38, p < .01$, by subjects and items. A Tukey honestly significant difference test comparing the control prime with the other two primes found a significant disadvantage in both latency ($p < .01$, by subjects and items) and accuracy ($p < .05$, by subjects and items).

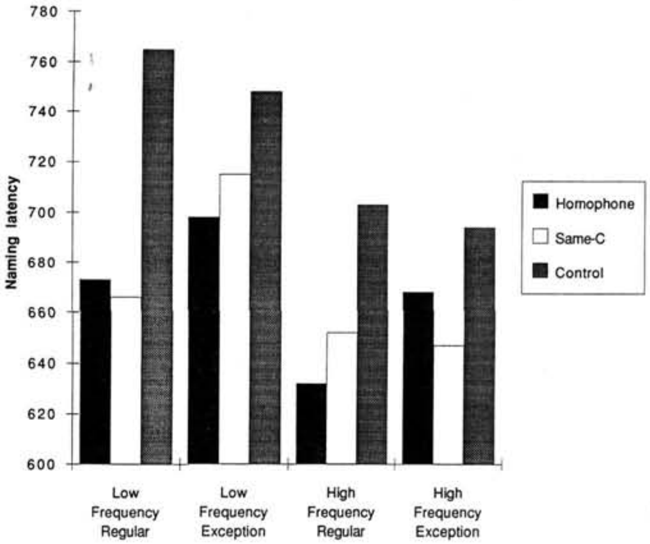

Figure 16. Naming latency in Experiment 6b (brief duration) for each of the Frequency × Regularity target combinations as a function of prime type. (Same C = a consonant-preserving prime.)

Hence, the priming of the target by its vowel phonemes can facilitate the assembly of low-frequency words provided this phonological information matches that assembled for the target, that is, for regular words, and that their exposure duration is relatively long.

Vowel-priming effects are thus contingent on the same set of constraints affecting regularity effects, namely, exposure duration, frequency, and the predictability of the addressed form. This conclusion is compatible with the view that our previous findings, demonstrating a dependency between regularity effect and exposure duration, are due to the specific effect of exposure duration on vowel assembly. Further support for this view is provided by the pattern of regularity effects in the present experiment as well. At the long exposure duration, there was evidence for vowel assembly in the form of a Frequency × Regularity interaction in the accuracy data that was significant by subjects, $F(1, 17) = 5.1, p < .05$. Although naming latency for low-frequency regular words was faster than exception words, this trend failed to reach significance in the present experiment. The reasons for this null effect are unclear. Strong evidence for Frequency × Regularity effects using longer exposure durations or unmasked targets was obtained in Experiment 5a and Experiments 7a–b. As expected, however, the short duration did not reveal any significant interaction of Frequency × Regularity in either latency or accuracy, although there was a suggestion of a marginally significant regularity effect in the latency analysis by subject, $F(1, 17) = 4.02, p < .07$.

In summary, vowel-priming and regularity effects emerge under limited, albeit similar, circumstances: under long exposure duration and for low-frequency regular words. These results support the view that these are two manifestations of the same phenomenon, namely, the second cycle of assembly. Thus, the contingency of regularity effect on long exposure durations is due to the assembly of vowel information being late occurring.

Brief Exposure Durations Permit Consonant Assembly: Experiments 7a–b

The results of Experiment 6 provide a rather clear picture: Vowel assembly is contingent on long exposure durations. Long exposure durations produce significant regularity and vowel-priming effects, whereas brief exposure durations result in their elimination. Within linear atemporal models of assembly, such evidence for the elimination of vowel assembly is often interpreted as evidence for the elimination of assembly as a whole. However, the postulation of distinct consonant and vowel cycles by the two-cycles model predicts that consonant assembly proceeds in the face of the truncation of the later vowel assembly. To support this point, Experiments 7a–b demonstrate that brief exposure durations, which resulted in null evidence for vowel assembly in our previous primed naming experiments, can produce significant evidence for consonantal assembly. In contrast, the increase in exposure duration using the same materials results in the emergence of evidence of vowel assembly for low-frequency regular words.

To assess the assembly of consonantal information, the targets used in Experiment 5 were primed by the four types of primes corresponding to the four types of masks used in Experiment 1[22]: a homophone (RAIK), a vowel-preserving (RAIB), a consonant-preserving (RIKK), and a control prime (BLIN). These primes were presented for 15 ms. They were preceded by a pattern mask (XXX) presented for 500 ms and followed by an unmasked target, presented for 60 ms. (The elimination of the final pattern mask was designed to prevent an increase in errors in target naming.) At this duration, we expect that the assembly of a phonological representation for the prime will yield only consonantal information. In contrast to the second cycle, consonant assembly is viewed as fast, general, and reliable. Hence, the consonant-preserving prime should facilitate target naming compared with the vowel-preserving prime, and this facilitation should be general, regardless of the target's frequency and regularity. In addition, the availability of vowel information in the homophone should provide no additional benefit as compared with the consonant-preserving prime. To demonstrate that these materials can result in vowel assembly with longer exposure duration, Experiment 7b examined the effect of these four types of primes in a slightly longer prime duration (30 ms) and a different pattern mask (###).[23] At this longer duration, we thus expect a disappearance of the consonant-prime advantage and the emergence of a significant vowel-priming effect for low-frequency regular words in the form of an advantage for the homophone compared with the consonant-preserving prime. To encourage fast responding, response latencies slower than 800 ms received a negative feedback in the form of a computer message and a short tone. Twenty-four subjects were assigned to each of the two exposure durations.

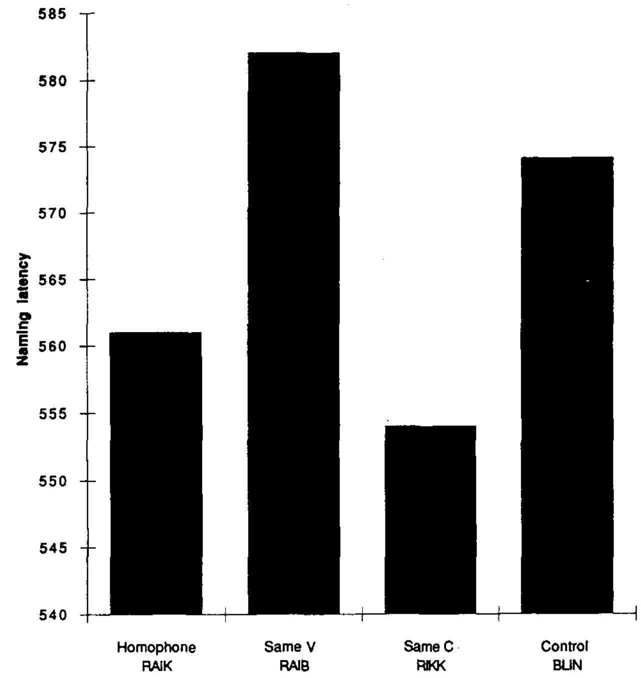

Figure 17. Naming latency in Experiment 7a (brief duration) as a function of prime type. (Same V = a vowel-preserving prime; same C = a consonant-preserving prime.)

Results and Discussion

Subjects' latency and accuracy in the brief and long durations were submitted to separate ANOVAs (4 Mask × 2 Frequency × 2 Regularity) by subjects and by items. The overall level of accuracy was 92% and 93% at the brief and long durations, respectively. In accord with our previous findings, the presentation of the unmasked target resulted in a significant regularity effect for low-frequency words in each of the two durations. In the brief duration, the Frequency × Regularity interaction was found significant by subjects and items in the latency data, $F(1, 23) = 35.77$, $p < .0001$, $F(1, 60) = 5.23$, $p < .05$, for subjects and items, and by items in the accuracy data, $F(1, 23) = 21.55$, $p < .0005$. At the long duration, the interaction was significant in the accuracy and latency data by subjects, $F(1, 23) = 23.49$, $p < .001$, $F(1, 23) = 10.9$, $p < .005$, for latency and accuracy, and marginally significant in the accuracy data by items, $F(1, 60) = 3.4$, $p < .08$. These results fit our previous conclusions that long or unmasked presentation of the target is necessary for vowel assembly.

However, evidence for assembly is obtained even under brief exposure durations when consonantal information is considered. The latency analyses at the brief exposure duration produced a significant main effect of prime type, $F(3, 69) = 3.09$, $p < .05$, $F(3, 180) = 2.58$, $p < .06$, by subjects and items, and no interaction of the prime type with either frequency or regularity. As expected, the consonant-preserving-prime produced a significant facilitation in target-naming latency compared with the vowel-preserving prime ($\Delta = 28$ ms), $t(69) = 2.79$, $p < .01$; $t(180) = 2.74$, $p < .01$, by subjects and items (Figure 17). This finding reaffirms the claim that brief exposure durations result in the assembly of consonantal information that is general in regard to the target's properties. Furthermore, the introduction of vowel information by the homophone prime did not facilitate recognition compared with the consonant-preserving prime ($\Delta = 7$ ms), $t(69) < 1$, $t(180) = 1.16$, *ns*. These results support the model's claim that such a brief exposure duration permits the assembly of exclusively consonantal information.[24]

In contrast, the increase in exposure duration is expected to give rise to the second cycle of vowel assembly. Thus, at the longer duration, the priming of the target by its vowels should yield facilitation when this vowel information matches that assembled for the target, that is, for regular words. Because vowel

---

[22] A few of these primes were replaced on the basis of subjects' rating of the materials (see Appendix C).

[23] Pilot work suggests that this pattern mask increases the functional exposure of the prime compared with the xxx mask.

[24] The consonant-preserving prime resulted in lower naming accuracy than either the homophone and vowel-preserving primes ($\Delta = 3\%$) or the control prime ($\Delta = 1\%$). However, the main effect of prime type was significant only by items, $F2(3, 180) = 2.67$, $p < .05$, and not by subjects, $F1(3, 69) = 1.73$. The data provide no indication of a speed–accuracy trade-off. In fact, there was a significant negative correlation between naming latency and accuracy under either the consonant- or the vowel-preserving prime. This inhibition by the consonant-preserving prime can be attributed to five targets, three of them are multisyllabic words, which also exhibit slow and variable performance on both latency and accuracy. The exclusion of these targets from the accuracy analysis by items resulted in the elimination of the inhibition by the consonant prime in the accuracy data ($\Delta = 1.39\%$, $t < 1$) while maintaining the facilitation in the naming latency ($\Delta = 26.28$ ms), $t(165) = 2.2$.

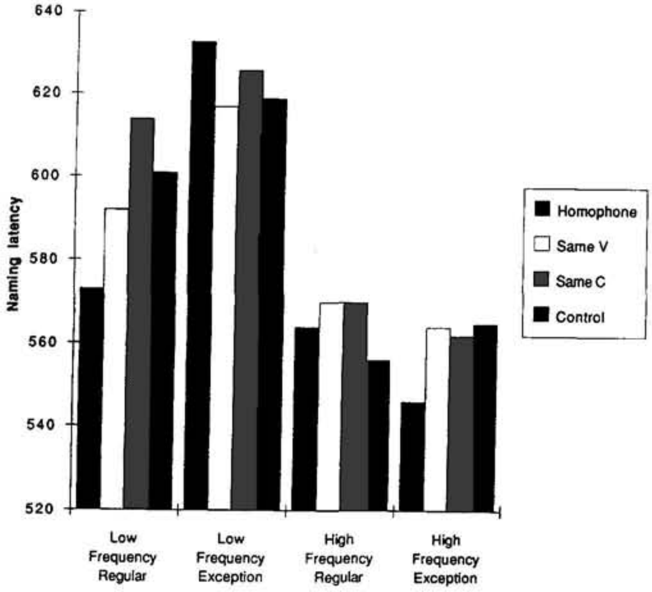

Figure 18. Naming latency in Experiment 7b (long duration) for each of the Frequency × Regularity target combinations as a function of prime type. (Same V = a vowel-preserving prime; same C = a consonant-preserving prime.)

assembly is slow compared with the addressed route, the facilitation should emerge only for low-frequency words. Thus, vowel-priming effects should be limited to low-frequency regular words. The accuracy data showed no significant difference between the homophone and the consonant- and vowel-preserving primes. Strong support for the model's prediction is provided in the naming latency data. For low-frequency regular words, the availability of vowel information by the homophone prime resulted in a significant facilitation compared with the consonant-preserving prime ($\Delta = 41$ ms), $t(69) = 2.37$, $p < .05$, $t(180) = 1.97$, $p < .05$, by subjects and items (Figure 18). (For comparison, at the brief duration, low-frequency words exhibited no facilitation by the homophone, $\Delta = -3$ ms, $t < 1$. In contrast, low-frequency exception words exhibit a nonsignificant inhibition by the homophone compared with the consonant-preserving prime, $\Delta = 6$ ms, $t < 1$, by subjects and items.) In summary, these experiments clearly demonstrate that for a single set of materials, the amount of processing time allowed for the prime determines the contents of the assembled code. The longer exposure duration yields evidence for vowel assembly in terms of a homophone advantage for low-frequency regular words. The decrease in exposure duration results in null effects of vowel priming, that is, the elimination of the homophone advantage. Yet, this short duration does provide positive evidence for the assembly of consonantal information that is general in respect to the target's properties. These results support the model's claim that the assembly of the first consonantal cycle proceeds in the face of a truncation of the later vowel cycle.

General Discussion of Experiments 5–7

We initiated this article with a puzzle: We identified two types of methods that reveal contradictory evidence regarding the role of assembled phonology in word recognition and its nature. We suggested a solution to these contradictions by making two interrelated claims: (a) that there is a methodological difference between these types of methods in terms of exposure duration and (b) that there are systematic changes in the contents of the assembled code that critically depend on this methodological factor, namely, exposure duration. These differences in exposure duration thus determine the extent to which each of these two types of methods is sensitive to vowel information, a later emerging constituent in the assembled code.

Experiments 5–7 provide support for these two claims. Experiment 5 demonstrated that the type of experimental method used determines the emergence of regularity effects by showing that the same set of materials, which, under brief exposure durations using the masking procedure showed no interaction of phonemic masking effects with regularity, produce a significant Frequency × Regularity interaction using the naming technique. To support the claim that it is the target's exposure duration that accounts for this contrast, we further showed that the reduction of exposure duration results in the elimination of the Frequency × Regularity interaction. We then provided evidence for the claim that the sensitivity of regularity effects to the methodological variable of exposure duration is due to temporal changes in the contents of the assembled code, that is, the late emergence of vowel information. In Experiment 6, we demonstrated that vowel-priming effects are late occurring and they are constrained by the same factors affecting the emergence of regularity effects, namely, exposure duration, frequency, and the predictability of the addressed form. Experiment 7 further showed that brief exposure durations that typically produce null regularity effects do permit phonological assembly of consonants.

The conclusions we draw from this demonstration are as follows. First, the choice of experimental method determines whether evidence for assembly will be observed. The critical variable is the target's exposure duration. The detection of vowel assembly in the form of regularity and vowel-priming effects requires relatively long exposure duration. Second, this methodological conclusion reaffirms the two-cycles model's predictions regarding the time course of assembly. The content of the assembled representation changes along the time dimension, and the relatively late onset of the regularity and vowel-priming effects is compatible with the view of vowels as derived at a later stage of assembly. Finally, these results suggest some caution in using regularity effects as the sole indicator for assembly. Regularity effects are observed only when the target's exposure duration is relatively long. Yet, brief exposure durations do permit the assembly of consonants. Thus, although regularity effects are a sufficient condition for the detection of assembly, they are not necessary. Hence, the failure to detect regularity effects cannot be taken as evidence for a direct access, achieved without any mediation of a (partial) assembled representation. In general, the results of these studies suggest that an accurate interpretation of data on assembled phonology in word recognition requires a close consideration of the timing properties of the experimental method.

Conclusions and Extensions

We have proposed a model of phonology assembly in English visual word recognition whose central claim is that the assembled code is structured in a multilinear fashion: Consonants and vowels are two distinct constituents of the assembled representation. The structure of the assembled representation determines the computational properties of the assembly process: The consonant and vowel constituents are derived by two mechanisms that differ in their onset and automaticity. Consonants are derived in the initial cycle by a relatively fast and automatic process, whereas vowels are added to the code at the second cycle by a controlled process.

The two-cycles model is supported by empirical evidence as well as principled arguments. In a series of experiments, we have shown that the relative contribution of consonant versus vowel information varies with time. Early in processing there is an advantage for the priming or reinstatement of consonant information over vowel information that emerges regardless of the target's frequency and regularity. In contrast, in later processing stages, the priming or reinstatement of vowel information facilitates the recognition of low-frequency regular words compared with consonantal information. The distinct time course of consonant and vowel assembly cannot be attributed to differences in their graphemic complexity, because evidence for the two cycles is obtained even when the graphemic complexity of consonants and vowels is equivalent. The emergence of this pattern of results in the performance of secondary tasks, as well as reading, suggests that, as a whole, the assembly of phonology is a relatively automatic process that is mandatory in recognition. However, the consonant and vowel components do differ in their attention demands. The assembly of consonants proceeds regardless of load. In contrast, the increase in attention demands eliminates vowel assembly. Similarly, at the second cycle, the overall effect of load increases compared with the initial cycle.

In addition to having direct empirical support, the two-cycles model is able to account for some contradictions about the nature of phonological assembly in the English word-recognition literature. These contradictions, unaccounted for by existing models of word recognition, are resolved by the two-cycles model, which assumes that the assembled code entails two distinct components that contrast in their computational properties. The model's claims regarding the temporal onset of these two components also carry some important methodological implications. In general, experimental methods in which the exposure duration of the target is unconstrained suggest that assembly is a slow and controlled process. In contrast, experimental methods that constrain the target's exposure duration reveal assembly as a fast and automatic process. Clearly, the properties of the experimental method dramatically affect the facet of assembly it reflects, and these properties are critically determined by the exposure duration of the target. This methodological claim is supported by a series of studies examining the effect of exposure duration on the emergence of regularity and vowel-priming effects in the naming technique.

Although the two-cycles model is an account of phonological assembly in skilled reading, its predictions are highly compatible with the evidence on the acquisition of reading as well. According to the model, it is the assembly of vowels in English that is particularly demanding in attention. Indeed, beginning and poor English readers face special difficulty in the assembly of vowels (Bryson & Werker, 1989; Fowler, Liberman, & Shankweiler, 1977; R. Shankweiler & Liberman, 1972). The pattern of errors for consonants and vowels differs not only quantitatively but also qualitatively: Whereas consonant errors are primarily due to phonemic similarity (Fowler et al., 1977; Werker, Bryson, & Wassenberg, 1989), the rate of vowel errors is related to the complexity of the graphemes' mapping (Fowler et al., 1977; R. Shankweiler & Liberman, 1972). Specifically, vowel errors result from the failure to apply context-sensitive rules in the mapping of vowel graphemes (Bryson & Werker, 1989; Fowler, Shankweiler, & Liberman, 1979; Zinna, Liberman, & Shankweiler, 1986). The sensitivity of vowel errors to the complexity of grapheme-to-phoneme mapping conforms to the model's claim about the attention demands of vowel assembly. An impairment of the reader's processing resources should have a particularly devastating effect on the assembly of vowels. Although a detailed discussion of the etiology of reading disability is beyond the scope of this article, we note that this prediction at least fits the existing reports of a verbal working-memory deficit among poor readers (Brady, 1991; I. Y. Liberman, 1989; I. Y. Liberman, Mann, Shankweiler, & Werfelman, 1982; Perfetti, 1994; Perfetti & Goldman, 1976; D. Shankweiler, 1989).

Why are consonants and vowels assembled in two cycles? One possibility is that the differences in the speed and automaticity of assembling consonants and vowels merely reflects the differences in the consistency of grapheme to phoneme mapping in the English orthography. The later onset and greater attention demands of vowel assembly simply follow from the inconsistency of the mapping of vowel graphemes onto their phonemes. On this view, the two cycles of assembly are merely a consequence of the differences in the consistency of mapping in the input, rather than a reflection of the linguistic constituency structure of the assembled code. The distinction between consonants and vowels is thus limited in its scope to English word recognition and bears no linguistic significance. In fact, the differences in the computational properties of consonants and vowels may be attributed to quantitative differences in the operation of a single mental mechanism, without the need to postulate distinct systems for the assembly of consonants and vowels.

Although we recognize that the different computational properties of consonant and vowel assembly may stem from the specific properties of the English orthography, this state of affairs does not preclude a deeper processing explanation. The computational properties of consonant and vowel graphemes in English are confounded with their phonological characteristics. English vowels are less distinct phonologically than consonants (A. M. Liberman, 1970; A. M. Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). The fact that the English orthography tolerates a greater inconsistency in the mapping of vowels may thus reflect this phonological factor (R. Shankweiler & Liberman, 1972). Furthermore, the variations in the mapping of English vowels are not random but often correspond to the phonological deep structure of the language (Chomsky, 1970). In fact, the depth of the English orthography might itself be due to the tendency of the writing system to reflect the deep phonological structure. Thus, the distinction between consonants and vowels, and the particular choices made by the orthography in the representation of vowels, may reflect the presence of linguistic constraints on word recognition.

We have hypothesized two such possible constraints on reading. One constrains the kind of linguistic units represented by the orthography. The second concerns the structure of the assembled code, requiring an isomorphism between the structure of phonological representations assembled in reading and those assigned to spoken language. On this view, a distinction between the representation of consonants and vowels in the assembled code, and the presence of two mental systems underlying their computation, might reflect a linguistic constraint of Universal Grammar. Thus the potential for the segregation of consonant and vowel assembly may be universal, whereas the particular computational characteristics of these two mechanisms in English may result from the differential degrees of the consistency of their mapping in the English orthography.

The suggestion of two distinct systems for the assembly of consonants and vowels is supported by the observation of separate mechanisms for the processing of consonants and vowels in other mental processes specialized in linguistic input. Most notable is the existence of distinct specialized systems for the production of consonants and vowels in speech production (Fowler, 1980, 1983; Öhman, 1966; Perkell, 1969). A segregation of the auditory input into two distinct but interrelated components of consonants and vowels is manifested in the perception of speech as well (Diehl et al., 1987; Fowler, 1983; Fowler & Smith, 1986).

An additional support for the segregation of the consonant and vowel plains in Universal Grammar comes from children's phonology. Macken (1979, 1992) has claimed that children's early hypotheses regarding the grammar of their native language are formulated in terms of separate whole-word templates for consonants and vowels. Children's early speech "errors" thus reflect an attempt to fit the adult's output into their predetermined templates. For instance, many of the speech errors of Si, a Spanish-speaking girl during the age of 19 months–25 months could be accounted for by the imposition of a "labial + dental" template. For instance, Si produced the words *pelota sopa* and *manzana* as *pata pota* and *mana,* respectively. Similar evidence for the segregation of consonants and vowels and the use of a "labial-alveolar" template are observed by Studdert-Kennedy and Whitney Goodell (1992) in the speech of an English-speaking girl.

The possibility of two systems for the processing of consonants and vowels is also strongly implicated by an apparently unrelated phenomenon: spelling errors. Cubelli (1991) reported two cases of Italian patients exhibiting a specialized deficit in the spelling of vowels. In one case (C. F.), the patient completely omitted all vowels but correctly wrote consonant letters. A second patient (C. W.) had a milder deficit in the spelling of vowels: Although most of his spelling errors concerned vowels, some errors in consonant letters were reported as well. Interestingly, however, substitution errors, which account for 83% of C. W.'s errors, respected the distinction between the graphemic status of the letters as consonant or vowel: C. W. tended to substitute a consonant letter with another consonant letter and a vowel letter with another vowel letter. A similar pattern of substitution errors was reported by Caramazza and Miceli (1990) regarding the case of L. B. Caramazza and Miceli accounted for this pattern by postulating a nonlinear orthographic representation that specifies the graphemic identity, quantity, status (consonant or vowel), and the location of graphosyllabic boundaries on distinct tiers. Note, however, that the proposed representation is strictly orthographic rather than phonological in nature. Supporting this claim is the observation that the errors made by both L. B. and C. W. violated phonological constraints and failed to respect the unity of graphemic clusters corresponding to a single phoneme.

The functional distinction between the processing of consonant and vowel input in four distinct phenomena—speech production and perception, children's phonology, the structure of the graphemic representation, and the assembly process—may not be coincidental. Instead, it may reflect a general common constraint on the representation of linguistic input. The evidence for distinct consonant- and vowel-processing mechanisms in speech production, language acquisition, and spelling is at least compatible with the two-cycles model's suggestion that the assembly of consonants and vowels is achieved by two different systems. We speculated earlier that the phonological assembly process might be preceded by an early stage that sorts the graphemes into these two types, thereby preparing the input for each of the subsequent cycles of assembly. For the moment, however, the evidence supports the conclusion that there are distinct stages of assembly defined by their phonological content. Regardless of the precise implementation of the distinction between consonant and vowel processing, or its causes, it is clear that the maintenance of such a distinction is necessary for the empirical adequacy of English word-recognition models.

Finally, it is difficult to conclude a discussion of phonological processes in word recognition without addressing the controversy regarding dual- versus single-route models. The issue at stake entails two distinct questions. One concerns the generality of phonology assembly in word recognition. A second question regards the architecture of the system. Our discussion has taken a rather neutral stance regarding the architecture of the system. Our results, however, do suggest a mandatory activation of the consonantal cycle that is fast and general across word frequency. Perhaps the most important contribution of the present discussion does not reside in the answers it provides as to the role of assembly in word recognition but, instead, in the reformulation of the question. The question of **whether** assembled phonology plays a role in recognition cannot be adequately addressed without an explicit and detailed account of **what** assembled phonology is. We have shown that the failure to consider the nature of the assembled code and its multilinear structure results in contradictions as to the role of assembly in recognition. We conclude that the examination of the structure of the assembled representation as a linguistic entity is indispensable for understanding the role of assembled phonology in word recognition, in particular, and the nature of reading as a linguistic skill, in general.

References

Allen, J. (1979). Speech synthesis from text. In J. C. Simon (Ed.), *Spoken language generation and understanding* (pp. 383–396). Dordecht, The Netherlands: D. Reidel.

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory & Cognition, 10,* 565–575.

Archangeli, D. (1985). Yokuts harmony: Evidence for coplanar representation in nonlinear phonology. *Linguistic Inquiry, 16,* 335–372.

Banks, W. P., Oka, E., & Shugarman, S. (1981). Recoding of printed words to internal speech: Does recoding come before lexical access? In O. J. L. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology* (pp. 137–170). Hillsdale, NJ: Erlbaum.

Baron, J. (1973). Phonemic stage not necessary for reading. *Quarterly Journal of Experimental Psychology, 25,* 241-246.

Baron, J., & Strawson, C. (1976). Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 386-393.

Berent, I. (1993). *The role of assembled phonology in English visual word recognition: Towards a nonlinear model of phonology assembly.* Unpublished doctoral dissertation, University of Pittsburgh.

Bernstein, S. E., & Carr, T. H. (1991, November). *Dual-route theories in reading: When does a concurrent task facilitate pronunciation?* Paper presented at the meeting of the Psychonomic Society, San Francisco.

Besner, D. (1990). Does the reading system need a lexicon? In D. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 1-25). Hillsdale, NJ: Erlbaum.

Bower, T. G. (1970). Reading by eye. Phonology and reading. In H. Levin & J. O. Williams (Eds.), *Basic studies in reading* (pp. 134-146). New York: Basic Books.

Bradshaw, J. L., & Nettleton, N. C. (1974). Articulatory interference and the mown-down heterophone effect. *Journal of Experimental Psychology, 102,* 88-94.

Brady, S. A. (1991). The role of working memory in reading disability. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 129-151). Hillsdale, NJ: Erlbaum.

Brown, P., & Besner, D. (1987). The assembly of phonology in oral reading: A new model. In M. Coltheart (Ed.), *Attention and performance XII. The psychology of reading* (pp. 471-489). Hillsdale, NJ: Erlbaum.

Brysbaert, M., & Praet, C. (1992). Reading isolated words: No evidence for automatic incorporation of the phonetic code. *Psychological Research, 54,* 91-102.

Bryson, S. E., & Werker, J. F. (1989). Towards understanding the problem in severely disabled readers. Part I: Vowel errors. *Applied Psycholinguistics, 10,* 1-12.

Burt, J. S., & Humphreys, M. S. (1993). Delayed priming and pronunciation of inconsistent words and pseudowords. *Journal of Memory and Language, 32,* 743-765.

Caramazza, A., & Miceli, G. (1990). The structure of graphemic representations. *Cognition, 37,* 243-297.

Carello, C., Turvey, M. T., & Lukatela, G. (1992). Can theories of word recognition remain stubbornly nonphonological? In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 211-226). Amsterdam: North-Holland.

Carr, T. H., Davidson, B. J., & Hawkins, H. L. (1978). Perceptual flexibility in word recognition: Strategies affect orthographic computation but not lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 674-690.

Carr, T. H., & Pollatsek, A. (1985). Recognizing printed words: A look at current models. In D. Besner, T. G. Waller, & G. E. Mackinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 5, pp. 1-82). Orlando, FL: Academic Press.

Chomsky, N. (1964). Current issues in linguistic theory. In J. Fodor & J. J. Katz (Eds.), *The structure of language. Reading in the philosophy of language* (pp. 50-118). Englewood Cliffs, NJ: Prentice Hall.

Chomsky, N. (1970). Phonology and reading. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading* (pp. 3-18). New York: Basic Books.

Clements, G. N., & Keyser, S. J. (1983). *CV phonology: LI Monography Series* (No. 9). Cambridge, MA: MIT Press.

Colombo, L., & Tabossi, P. (1992). Strategies and stress assignment: Evidence from a shallow orthography. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 319-399). Amsterdam: North-Holland.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151-216). London: Academic Press.

Coltheart, M., Besner, D., Jonasson, T. J., & Davelaar, E. (1979). Phonological encoding in the lexical decision task. *Quarterly Journal of Experimental Psychology, 31,* 489-507.

Coltheart, M., Davelaar, E., Jonasson, T. V., & Besner, D. (1977). Access to the internal lexicon. In D. Stanislav (Ed.), *Attention and performance* (Vol. 6, pp. 532-555). Hillsdale, NJ: Erlbaum.

Coltheart, V., Avons, S. E., Masterson, J., & Laxon, V. J. (1991). The role of assembled phonology in reading comprehension. *Memory & Cognition, 19,* 387-400.

Coltheart, V., Avons, S. E., & Trollope, J. (1990). Articulatory suppression and phonological codes in reading for meaning. *Quarterly Journal of Experimental Psychology, 42A,* 375-399.

Coltheart, V., Laxon, V., Rickard, M., & Elton, C. (1988). Phonological recoding in reading for meaning by adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 387-397.

Cowan, N., Lichty, W., & Grove, T. R. (1990). Properties of memory for unattended spoken syllables. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 258-269.

Cubelli, R. (1991). A selective deficit for writing vowels in acquired dysgraphia. *Nature, 353,* 258-260.

Daneman, M., & Stainton, M. (1991). Phonological recoding in silent reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 618-632.

Dennis, I., Besner, D., & Davelaar, E. (1985). Phonology in visual word recognition: Their is more two this than meats the I. In D. Besner, T. G. Waller, & G. E. Mackinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 5, pp. 167-197). Orlando, FL: Academic Press.

Dennis, I., & Newstead, S. E. (1981). Is phonological recoding under strategic control? *Memory & Cognition, 9,* 472-477.

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.

Doctor, E. A., & Coltheart, M. (1980). Children's use of phonological encoding when reading for meaning. *Memory & Cognition, 8,* 195-209.

Durand, J. (1990). *Generative and nonlinear phonology.* London: Longman.

Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology, 33A,* 325-350.

Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked nonword priming. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 45A,* 353-372.

Ferrand, L., & Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society, 31,* 119-122.

Ferrand, L., & Grainger, J. (1994). Effects of orthography are independent of phonology in masked form priming. *Quarterly Journal of Experimental Psychology, 47A,* 365-382.

Fleming, K. (1993). Phonologically mediated priming in spoken and printed recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 272-284.

Fodor, J. A. (1983). *The modularity of mind.* Cambridge, MA: MIT Press.

Forster, K. I., & Davis, C. (1991). The density constraint on form-priming in the naming task: Interference effects from a masked prime. *Journal of Memory and Language, 30,* 1-25.

Fowler, C. A. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics, 8,* 113-133.

Fowler, C. A. (1983). Converging sources of evidence on spoken and perceived rhythms of speech: Cyclic production of vowels in monosyllabic stress feet. *Journal of Experimental Psychology: General, 112,* 386–412.

Fowler, C. A., Liberman, I. Y., & Shankweiler, D. (1977). On interpreting the error pattern in beginning reading. *Language and Speech, 20,* 162–173.

Fowler, C. A., Shankweiler, D., & Liberman, I. Y. (1979). Apprehending spelling patterns of vowels: A developmental study. *Language and Speech, 22,* 243–252.

Fowler, C. A., & Smith, M. R. (1986). Speech perception as "vector analysis": An approach to the problems of invariance and segmentation. In J. S. Perkell & H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 123–139). Hillsdale, NJ: Erlbaum.

Frost, R. (1991). Phonetic recoding of print and its effect on the detection of concurrent speech in amplitude-modulated noise. *Cognition, 39,* 195–214.

Frost, R., Repp, B. H., & Katz, L. (1988). Can speech perception be influenced by simultaneous presentation of print? *Journal of Memory and Language, 27,* 741–755.

Goldsmith, J. A. (1990). *Autosegmental & metrical phonology.* Cambridge, England: Basil Blackwell.

Hanson, V. L., & Fowler, C. A. (1987). Phonological coding in word reading: Evidence from hearing and deaf readers. *Memory & Cognition, 15,* 199–207.

Hawkins, H. L., Reicher, G. M., & Peterson, L. (1976). Flexible coding in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 380–385.

Henderson, L. (1985). Issues in the modeling of pronunciation assembly in normal reading. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 459–508). Hillsdale, NJ: Erlbaum.

Hillinger, M. L. (1980). Priming effects with phonemically similar words: The encoding-bias hypothesis reconsidered. *Memory & Cognition, 8,* 115–123.

Humphreys, G. W., Evett, L. J., & Taylor, D. E. (1982). Automatic phonological priming in visual word recognition. *Memory & Cognition, 10,* 576–590.

Jacobson, J. Z. (1976). Visual masking by homonyms. *Canadian Journal of Psychology, 30,* 174–177.

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29,* 687–715.

Jared, D., & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General, 120,* 358–394.

Johnston, J. C., & McClelland, J. L. (1974). Perception of letters in words: Seek not and ye shall find. *Science, 184,* 1192–1194.

Kawamoto, A. (1993). Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account. *Journal of Memory and Language, 32,* 474–516.

Kay, J., & Marcel, A. (1981). One process, not two, in reading aloud: Lexical analogies do the work of non-lexical rules. *Quarterly Journal of Experimental Psychology, 33A,* 397–413.

Kaye, J. (1989). *Phonology: A cognitive view.* Hillsdale, NJ: Erlbaum.

Lesch, M., & Pollatsek, A. (1993). Automatic access of semantic information by phonological codes in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 285–294.

Liberman, A. M. (1970). The grammars of speech and language. *Cognitive Psychology, 1,* 301–323.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74,* 431–461.

Liberman, I. Y. (1989). Phonology and beginning reading revisited. In C. von Euler (Ed.), *Werner-Gren international symposium series: Brain and reading* (pp. 207–220). Hampshire, England: MacMillan.

Liberman, I. Y., Mann, V. A., Shankweiler, D., & Werfelman, M. (1982). Children's memory for recurring linguistic and nonlinguistic materials in relation to reading ability. *Cortex, 18,* 367–375.

Lukatela, G., & Turvey, M. T. (1990a). Automatic and pre-lexical computation of phonology in visual word identification. *European Journal of Cognitive Psychology, 2,* 325–343.

Lukatela, G., & Turvey, M. T. (1990b). Phonemic similarity effects and prelexical phonology. *Memory & Cognition, 18,* 128–152.

Lukatela, G., & Turvey, M. T. (1991). Phonological access of the lexicon: Evidence from associative priming with pseudohomophones. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 951–966.

Lukatela, G., & Turvey, M. T. (1993). Similar attentional, frequency and associative effects for pseudohomophones and words. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 166–178.

Lukatela, G., & Turvey, M. T. (1994). Visual lexical access is initially phonological: 1. Evidence from associate priming by words, homophones, and pseudohomophones. *Journal of Experimental Psychology: General, 123,* 107–128.

Lupker, S. J., & Colombo, L. (1994). Inhibitory effects in form priming: Evaluating a phonological competition explanation. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 437–452.

Macken, M. (1979). Developmental reorganization of phonology: A hierarch of basic units of acquisition. *Lingua, 49,* 11–49.

Macken, M. (1992). Where's phonology? In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development. Models, research, implications* (pp. 255–282). Timonium, MD: York Press.

Martin, R. C. (1982). The pseudohomophone effect: The role of visual similarity in non-word decisions. *Quarterly Journal of Experimental Psychology, 34A,* 395–409.

Martin, R. C., & Jensen, C. R. (1988). Phonological priming in the lexical decision task: A failure to replicate. *Memory & Cognition, 16,* 505–521.

Mattingly, I. G. (1992). Phonological awareness and orthographic form. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology and meaning* (pp. 11–26). Amsterdam: North Holland.

McCann, R. S., Besner, D., & Davelaar, E. (1988). Word recognition and identification: Do word-frequency effects reflect lexical access? *Journal of Experimental Psychology: Human Perception and Performance, 14,* 693–706.

McCarthy, J. J. (1985). *Formal problems in Semitic phonology and morphology.* New York: Garland.

McCarthy, J. J. (1989). Linear order in phonological representation. *Linguistic Inquiry, 20,* 71–99.

McCusker, L. X., Hillinger, M. L., & Bias, R. G. (1981). Phonological recoding in reading. *Psychological Bulletin, 89,* 217–245.

McQuade, D. V. (1981). Variable reliance on phonological information in visual word recognition. *Language and Speech, 24,* 99–109.

McQuade, D. V. (1983). Pre-lexical phonological recoding becomes automatic with stimulus repetitions. *Language and Speech, 26,* 393–405.

Meyer, D. E., & Ruddy, M. G. (1973, November). *Lexical-memory retrieval based on graphemic and phonemic representations of printed words.* Paper presented at the meeting of the Psychonomic Society, St. Louis, MO.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1974). Functions of graphemic and phonemic codes in visual-word recognition. *Memory & Cognition, 2,* 309–321.

Miceli, G., & Caramazza, A. (1993). The assignment of word stress in oral reading: Evidence from a case of acquired dyslexia. *Cognitive Neuropsychology, 10,* 273–296.

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, 119,* 43–71.

Monsell, S., Patterson, K., Graham, A., Hughes, C. H., & Milroy, R.

(1992). Lexical and sublexical translation of spelling to sound: Strategic anticipation of lexical status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 451–467.

Naish, P. (1980). The effects of graphemic and phonemic similarity between targets and masks in a backward visual masking paradigm. *Quarterly Journal of Psychology, 32,* 57–68.

Norris, D., & Brown, G. (1985). Race models and analogy theories: A dead heat? Reply to Seidenberg. *Cognition, 20,* 155–168.

Öhman, S. E. G. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America, 39,* 151–168.

Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research, 53,* 13–24.

Paap, K. R., Noel, R. W., & Johansen, L. S. (1992). Dual-route models of print to sound: Red herrings and real horses. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology and meaning* (pp. 293–318). Amsterdam: North Holland.

Parkin, A. J., & Ellingham, R. (1983). Phonological recoding in lexical decision: The influence of pseudohomophones. *Language and Speech, 26,* 81–90.

Parkin, A. J., & Underwood, G. (1983). Orthographic vs. phonological irregularity in lexical decision. *Memory & Cognition, 11,* 351–355.

Patterson, K., & Coltheart, V. (1987). Phonological processes in reading. A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 421–447). Hillsdale, NJ: Erlbaum.

Patterson, K. E., & Morton, J. (1985). From orthography to phonology: An attempt at an old interpretation. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 335-364). Hillsdale, NJ: Erlbaum.

Perfetti, C. A. (1985). *Reading ability.* New York: Oxford University Press.

Perfetti, C. A. (1994). Psycholinguistics and reading ability. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 849–894). Orlando, FL: Academic Press.

Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language, 30,* 473–485.

Perfetti, C. A., Bell, L. C., & Delaney, S. M. (1988). Automatic (prelexical) phonetic activation in silent reading: Evidence from backward masking. *Journal of Memory and Language, 32,* 57–68.

Perfetti, C. A., & Georgi, M. C. (1994). *Phonemic masking effects are independent of phonemic mask properties.* Manuscript in preparation.

Perfetti, C. A., & Goldman, S. R. (1976). Discourse memory and reading comprehension skill. *Journal of Verbal Learning and Verbal Behavior, 14,* 33–42.

Perfetti, C. A., & McCutchen, D. (1982). Speech processes in reading. In N. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 7, pp. 237–269). New York: Academic Press.

Perfetti, C. A., & Zhang, S. (1991). Phonological processes in reading Chinese words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 633–643.

Perfetti, C. A., Zhang, S., & Berent, I. (1992). Reading in English and Chinese: Evidence for a "universal" phonological principle. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology and meaning* (pp. 227–248). Amsterdam: North-Holland.

Perkell, J. S. (1969). *Physiology of speech production: Results and implications of a quantitative cineradiographic study.* Cambridge, MA: MIT Press.

Peter, M., Lukatela, G., & Turvey, M. T. (1990). Phonological priming: Failure to replicate in the rapid naming task. *Bulletin of the Psychonomic Society, 28,* 389–392.

Plaut, D., & McClelland, J. (1993). Generalization with componential attractors: Word and nonword reading in an attractor network. In *Proceedings of the 15th Annual Conference of the Cognitive Science Society* (pp. 824–829). Hillsdale, NJ: Erlbaum.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 148–162.

Pring, L. (1981). Phonological codes and functional spelling units: Reality and implications. *Perception and Psychophysics, 30,* 573–578.

Pugh, K. R., Rexer, K., & Katz, L. (1994). Evidence of flexible coding in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 807–825.

Rosson, M. B. (1983). From SOFA to LOUCH: Lexical contributions to pseudoword pronunciation. *Memory & Cognition, 11,* 152–160.

Rosson, M. B. (1985). The interaction of pronunciation rules and lexical representations in reading aloud. *Memory & Cognition, 13,* 90–99.

Rubinstein, H., Lewis, S. S., & Rubinstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior, 10,* 645–657.

Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition, 19,* 1–30.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23,* 383–404.

Shankweiler, R. (1989). How problems of comprehension are related to difficulties in decoding. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle, IARLD Research Monograph Series* (pp. 36–68). Ann Arbor: University of Michigan Press.

Shankweiler, D., & Liberman, I. Y. (1972). Misreading: A search for causes. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye* (pp. 293–317). Cambridge, MA: MIT Press.

Shiffrin, R. M. (1988). Attention. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Steven's handbook of experimental psychology. Volume 2: Learning and cognition* (pp. 739–811). New York: Wiley.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–189.

Shimron, J. (1993). The role of vowels in reading: A review of studies of English and Hebrew. *Psychological Bulletin, 114,* 52–67.

Shulman, H. G., Hornak, R., & Sanders, E. (1978). The effects of graphemic, phonetic, and semantic relationships on access to lexical structures. *Memory & Cognition, 6,* 115–123.

Stanovich, K. E., & Bauer, D. W. (1978). Experiments on the spelling-to-sound regularity effect in word recognition. *Memory & Cognition, 6,* 410–415.

Stone, G. O., & Van Orden, G. C. (1993). Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 744–774.

Studdert-Kennedy, M., & Whitney Goodell, E. (1992). Gestures, features, and segments in early child speech. In *Haskins Laboratories Status Report on Speech Research* (pp. 89–102). New Haven, CT: Haskins Laboratories.

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language, 26,* 608–631.

Treiman, R., Freyd, J. J., & Baron, J. (1983). Phonological recoding and use of spelling-sound rules in reading of sentences. *Journal of Verbal Learning and Verbal Behavior, 22,* 682–700.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. (1993, November). *Spelling-sound relations in English and their* *effects on reading.* Paper presented at the meeting of the Psychonomic Society, Washington, DC.

Underwood, G., Roberts, M., & Thomason, H. (1988). Strategical invariance in lexical access: The reappearance of the pseudohomophone effect. *Canadian Journal of Psychology, 42,* 24–34.

Underwood, G., & Thwaites, S. (1982). Automatic phonological coding of unattended printed words. *Memory & Cognition, 10,* 434–442.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound and reading. *Memory & Cognition, 15,* 181–190.

Van Orden, G. C. (1991). Phonologic mediation is fundamental to reading. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 77–103). Hillsdale, NJ: Erlbaum.

Van Orden, G. C., & Goldinger, S. D. (in press). The interdependence of form and function in cognitive systems explains phonologic mediation in reading. In P. Tallal (Ed.), *The exceptional brain: Neural and cognitive mechanisms underlying speech, language and reading.*

Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 371–386.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97,* 488–522.

Venezky, R. L. (1970). *The structure of English orthography.* The Hague, The Netherlands: Mouton.

Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: Time-course and decision criteria. *Memory & Cognition, 13,* 557–572.

Werker, J. F., Bryson, S. E., & Wassenberg, K. (1989). Towards understanding the problem in severely disabled readers. Part II: Consonant errors. *Applied Psycholinguistics, 10,* 13–30.

Zinna, D. R., Liberman, I. Y., & Shankweiler, D. (1986). Children's sensitivity to factors influencing vowel reading. *Reading Research Quarterly, 21,* 465–480.

Appendix A

Materials Used in Experiment 1

Target	Homophone	Same Vowel	Same Consonant	Control
Low-Frequency Regular-Consistent Words				
rake	RAIK	RAIB	RIKK	BLIN
pest	PEHST	PEHDT	PEEST	DANKO
grill	GRYLL	GHYLL	GREEL	FETAN
bake	BAIK	BAIM	BIKK	DINN
glaze	GLAIZ	GRAIZ	GLEEZ	CRILP
cane	CAIN	CAID	CUNN	SIKOO
cable	KAIBLE	KAITLE	KIBBLE	GITONT
dental	DEHNTAL	DEHKTAL	DONNTAL	JORBOO
hotter	HOHTER	HOHBER	HAITER	PAICKA
modular	MAHDULAR	MAHKULAR	MEEDULAR	BIKENCHY
lazy	LAIZY	LAIMY	LAWZY	FOOND
pillow	PILLAW	PIDDAW	POOLAW	HUNGAD
locus	LOKUS	LOBUS	LACUS	BAPAR
rumor	ROOMOR	ROODOR	REEMOR	KOSHEE
bite	BIGHT	BIGHS	BIETT	MORDA
mellow	MEHLOW	MEBOW	MOOLOW	HANTHY
Low-Frequency Exception-Inconsistent Words				
wolf	WULFF	WULK#	WILLF	MASHO
ghost	GOAST	GOAPT	GAIST	YIEND
allege	ALLEDGE	ABBEDGE	ALOUDGE	BROONKO
shoe	SHOO	SPOO	SHAY	MAIG
liter	LITTER	LIBBER	LUTTER	DONBAN
tomb	TUME	TUPE	TEIM	DEEK
shove	SHUV#	SHUM#	SHEEV	CRANT
lemon	LEMMON	LEGGON	LEHMIN	PAIKAT
bury	BEHRY	BEHBY	BEERY	DANGO
recipe	RESEPEE	REMEPEE	ROSIPEE	BAKLOUN
elite	ELEET	ELEEM	ELOUT	FROOG
cafe	KAFFAY	KANNAY	KIFFAY	PURSCH
climb	CLIME	CLIPE	CLUME	TREST
glove	GLUV#	GLUK#	GLEAV	CHIB#
finale	FINNALY	FITTALY	FONNALY	BOMOODY
treble	TREBBLE	TREMMLE	TRIBBLE	POMICHY
High-Frequency Regular-Consistent Words				
late	LAIT	LAIB	LAUT	SIRP
name	NAIM	NAIB	NIMM	CRUT
not	NOTT	NOFF	NUTE	SERM
face	PHACE	PHABE	PHICE	TONDY
red	REHD	REHK	RADD	PAIF
past	PAHST	PAHLT	PAIST	WROOG
fine	FYNE	FYKE	FENN	SOAT
side	SYDE	SYME	SUDD	BROT

case	KAIS	KAIB	KESS	GOTH
rate	RAIT	RAIM	RITT	BUSK
met	METT	MEBB	MOTT	HOGA
got	GOTT	GOPH	GOOT	FERD
make	MAIK	MAIB	MUK#	PHIN
same	SAIM	SAIB	SIMM	GURN
letter	LEHTER	LEHBER	LOUTER	NADBAM
wife	WYFE	WYCE	WAF#	PLAS
	High-Frequency Exception-Inconsistent Words			
gone	GAUN	GAUP	GAHN	PRID
move	MOOV	MOOK	MAVV	RACH
whom	HOOM	HOOG	HAIM	PREE
cost	KAWST	KAWSH	KEEST	LAMIP
some	SUMM	SUGH	SEIM	PARB
police	POLEECE	POREECE	POOLYCE	BAMEDDY
give	GIV#	GIP#	GAV#	COWN
done	DUNN	DUPH	DIEN	TRIB
none	NUNN	NUPH	NEAN	DAIK
love	LUVV	LUMM	LEHV	PUCH
both	BOETH	BOECH	BOUTH	DRAIF
were	WIR#	WIRP	WOR#	GANK
woman	WUMAN	WUPAN	WUMIN	PELLED
come	KUMM	KULL	KEEM	YARK
whose	WHOOZ	WHOOP	WHESE	CRAGH
work	WOORK	WOORB	WARRK	KALAT

Appendix B

Materials Used in Experiment 2

Target	Homophone	Same Vowel	Same Consonant	Control
blue	BLOO	BREW	BLAY	NOOM
green	GREAN	GREAP	GRAHN	MOOPH
red	REHD	REHZ	RAHD	KIFF
gray	GREY	GLAY	GROY	WEEM

Appendix C

Corrections to the Stimuli Used in Experiment 1 Based on Subjects' Rating of the Quality of the Mask

Target	Homophone	Same Vowel	Same Consonant	Control
	Low-Frequency Regular-Consistent Words			
grill	GRYLL	GHYLL	GROUL	FETAN
pillow	PILLOH	PIMMOH	POOLAH	HUNGAD
bite	BIGHT	BIGHS	BAHTT	MORDA
	Low-Frequency Exception-Inconsistent Words			
wolf	WULFF	WULS#	WILLF	MASHO
liter	LEETER	LEEBER	LOUTER	DONBAN
lemon	LEHRY	BEHNY	BOORY	DANGO
	High-Frequency Regular-Consistent Words			
past	PAHST	PAHKT	POOST	WROOG
	High-Frequency Exception-Inconsistent Words			
gone	GAUN	GAUK	GUEN	PRID
some	SUMM	SUTT	SEIM	PARB
woman	WUMAN	WUMAR	WYMIN	PELED
work	WURK#	WURB#	WEERK	KALAT

(*Appendixes continue on next page*)

Appendix D

Materials Used in Experiment 4

Target	Same Vowel	Same Consonant	Control
	Complex Vowels		
	Low-Frequency Regular		
glaze	GLAME	GLEZE	PROFT
spike	SPIME	SPUKE	CRORT
globe	GLOCE	GLIBE	TRAIZ
wipe	WIKE	WUPE	DORF
lace	LABE	LOCE	MUFT
peel	PEDE	POEL	DAIK
stain	STAIP	STAUN	PROOK
skate	SKAME	SKOTE	PLOUM
pale	PABE	PULE	DRIS
spine	SPIFE	SPENE	GLOUF
wired	WIMED	WARED	MONST
deed	DEES	DUED	CAUB
	Low-Frequency Exception		
glove	GLUPH	GLAIV	PIRCH
shove	SHUM#	SHEEV	MANCK
gauge	GAFE#	GOGE#	POORD
weird	WEERM	WAIRD	FOUTH
lure	LOOB	LOIR	KAID
pour	PONE	PEIR	DLIM
seize	SEFE#	SOUZ#	BRAUG
sieve	SIG#	SAV#	BAV#
prove	PROOM	PRAIV	TLIMT
steak	STAPE	STUKE	DRIPH
wool	WHUM	WELE	FAIG
deaf	DET#	DIF#	DIB#
	High-Frequency Regular		
note	NOLE	NUTE	FIRD
came	CADE	CUME	BUND
date	DAPE	DITE	LIPH
deep	DEEG	DEIP	MAIG
game	GABE	GOME	ROLD
side	SIPE	SODE	GLIR
more	MAWG	MEIR#	DUND#
times	TIKES	TUEMS	POPHD
while	WHIBE	WHULE	JOONG
brain	BRAIT	BROON	FLOOZ
like	LISE	LOKE	NAIG
light	LIPE#	LEET#	CROUS
	High-Frequency Exception		
none	NUCK	NAIN	GUEL
come	CUPH	CAUM	PRIZ
done	DUSH	DAIN	KISP
door	DOBE	DUER	MILP
give	GHIT	GOOV	NOBS
some	SUTH	SOUM	GLIF
move	MOOP	MAIV	GRAG
touch	TUTH#	TICH#	NAIF
would	WOOR#	WOED#	NAISK
break	BRAFE	BRUK#	GLOPH
lose	LOOG	LEEZ	HARV
love	LUN#	LAIV	MAIC
	Simple Vowels		
	Low-Frequency Regular		
blast	BLANT	BLIST	GREMP
tank	TAFK	TINK	MELF
whit	WHIF	WOOT	GAUP
width	WIMTH	WADTH	KERCK

weld	WEKD	WALD	MONF
wig	WIB	WUG	LEP
dig	DIF	DAG	KOR
whiff	WHIB	WHAF	NOUS
grass	GRASH	GRESS	DEETH
blank	BLASK	BLENK	REESP
pill	PIM#	PEL#	KENF
gram	GRAZ	GREM	PLOK

Low-Frequency Exception

bush	BUHR#	BEESH	TORG
tomb	TOOG	TEMM	BLAG
wand	WAHD	WUND	METH
warn	WORB#	WEERN	NOOCH
wasp	WOSCK	WELSP	BLITT
wolf	WULK	WELF	GAHN
doll	DAWB	DAIL	BUBE
wads	WOBS	WIDS	PILP
gross	GROBE	GRAUS	BIBTH
bull	BUHR	BEEL	SHOR
pint	PRITE	POONT	MEELD
ghost	GLOAS	GAIST	RUEMP

High-Frequency Regular

big	BICK	BAIG	MALD
class	CLAF#	CLES#	MOOX#
held	HESD	HILD	GRAN
well	WEG#	WUL#	TYND
west	WEKT	WIST	FOMF
man	MAK	MUN	SIZ
wet	WES	WOT	LOL
past	PASH	PIST	GEND
went	WENS	WINT	MISF
path	PAGH	PETH	LIRD
wish	WIPH	WESH	NEPE
when	WHEV	WOON	GLIG

High-Frequency Exception

both	BOTE	BITH	GERP
cost	CAWNT	KEEST	GWELL
whom	HOOV	HAIM	NERK
work	WERTT	WEERK	FENFT
word	WURP#	WAIRD	BLET#
most	MOANT	MEEST	FAIRN
warm	WORB	WIRM	GLOP
put	POOK	PAIT	GORD
war	WAUL	WOOR	MOUD
pull	PUHR	PELL	KORK
watch	WOTH#	WETCH	DEDGE
what	WUM#	WIT#	PUD#

Appendix E

Materials Used in Experiment 6

Low-Frequency Regular Consistent Words

Target	Homophone	Same Consonant	Control
rake	RAIK	RUKK	BLIN
pest	PEHST	PEEST	DANKO
grill	GRYLL	GRALL	FETAN
bake	BAIK	BAHK	DINN
glaze	GLAIZ	GLEZZ	CRILP
cane	KAIN	KAUN	SIKOO
cable	KAYBLE	KEBBLE	GITONT
dental	DEHNTAL	DONNTAL	JORBOO
hotter	HAWTER	HETTER	PAICKA

(*Appendixes continue on next page*)

modular	MAUDULAR	MEEDULAR	BIKENCHY
lazy	LAIZY	LAWZY	FOOND
pillow	PILLOH	PALLOW	HUNGAD
locus	LOKUS	LACUS	BAPAR
rumor	ROOMOR	REEMOR	KOSHEE
bite	BIGHT	BOTTE	MORDA
mellow	MEHLOW	MEELOW	HANTHY

Low-Frequency Exception Inconsistent Words

wolf	WULFF	WILLF	MASHO
ghost	GOAST	GAIST	YIEND
allege	ALLEDGE	ALLIDGE	BROONKO
shoe	SHEW	SHEE	MAIG
liter	LEETER	LOUTER	DONBAN
tomb	TUME	TEMM	DEEK
shove	SHUVV	SHIVV	CRANT
lemon	LEMMON	LEEMON	PAIKAT
bury	BEHRY	BOORY	DANGO
recipe	RESEPEE	ROSIPEE	BAKLOUN
elite	ELEET	ELOUT	FROOG
cafe	KAFFAY	KEEFAY	PURSCH
climb	CLYME	CLUME	TREST
glove	GLUVV	GLUVE	CHIB#
finale	FYNNALY	FOONALY	BOMOODY
treble	TREBBLE	TRIBBLE	POMICHY

High-Frequency Regular Consistent Words

late	LAIT	LAWT	SIRP
name	NAIM	NEMM	CRUT
not	NOHT	NOUT	SERM
face	PHACE	PHICE	TONDY
red	REHD	RUDD	PAIF
past	PAHST	PAIST	WROOG
fine	FYNE	FUNE	SOAT
side	SYDE	SUDE	BROT
case	KACE	KUSE	GOTH
rate	RAIT	RITT	BUSK
met	MEHT	MOTE	HOGA
got	GOTT	GUTE	FERD
make	MAIK	MUKE	PHIN
same	SAIM	SIMM	GURN
letter	LEHTER	LOTTER	NADBAM
wife	WYFE	WAFE	PLAS

High-Frequency Exception Inconsistent Words

gone	GAWN	GUEN	PRID
move	MOOV	MAVV	RACH
whom	HOOM	HEEM	PREE
cost	KAWST	KEEST	LAMIP
some	SUMM	SIME	PARB
police	POLEECE	POOLICE	BAMEDDY
give	GIV#	GAV#	COWN
done	DUNN	DENN	TRIB
none	NUNN	NINN	DAIK
love	LUVV	LOUV	PUCH
both	BOETH	BOUTH	DRAIF
were	WIR#	WOR#	GANK
woman	WUMAN	WEMAN	ELED
come	KUMM	KEMM	YARK
whose	WHOOZ	WHEEZ	CRAGH
work	WURK#	WARK#	ALAT

Received June 18, 1993
Revision received May 30, 1994
Accepted July 12, 1994 ■