Sequential activation processes in producing words and syllables: Evidence from picture naming

Julio Santiago

Departamento de Psicología Experimental y Fisiología del Comportamiento, Universidad de Granada, Spain

Donald G. MacKay

Psychology Department, University of California, Los Angeles

Alfonso Palma

Departamento de Psicología Experimental y Fisiología del Comportamiento, Universidad de Granada, Spain

Christine Rho

Psychology Department, University of California, Los Angeles

This study examines picture naming latencies for predicted effects of two word retrieval factors: onset complexity and number of syllables. In Experiment 1, naming latency was longer for depicted words with two syllables e.g., *demon*, than one syllable, e.g., *duck*, and longer for words beginning with consonant clusters, e.g., *drill*, than single consonants, e.g., *duck*. Experiment 2 replicated these results and showed that vowel nucleus and coda complexity did not interact with onset complexity, and did not affect naming latency. Moreover, effects of onset complexity and number of syllables were additive, and unrelated to word frequency, articulatory factors, or picture complexity. These results comport with evidence from speech errors and metalinguistic tasks and with predictions of the Node Structure theory of language production, but do not support production theories that do not predict special processing difficulty for words with complex onsets and multiple syllables.

INTRODUCTION

This paper addresses the question of how phonemes and syllables of a word are retrieved in proper order when naming pictures of common objects. Object naming is a basic task that requires a complex combination of perceptual and motor skills. To understand what units and processes underlie these skills, cognitive psychologists often break down the picture naming task into simpler aspects, e.g., perceptual processing of the picture, retrieval of the name and its phonology, and articulatory programming of the response. The present study develops and tests a theory of the second aspect of picture naming: how the name and its phonology are retrieved. We first describe the theory, and then review the relevant literature. However, to tap all and only the word retrieval aspects of picture naming, we had to solve some methodological problems, which we describe before the details of our experiments.

PHONOLOGICAL RETRIEVAL IN NODE STRUCTURE THEORY

This section applies Node Structure theory (NST; MacKay, 1981, 1987) to the task of picture naming. NST was developed originally to represent processes involved in normal, everyday production of words in sentences, and the present application assumes that these same processes play a role in retrieving single words when naming a picture. The present application allowed greater specification of the principles involved in the sequential activation of syllables and segments in words.

NST represents linguistic entities such as words and syllables via simple processing units known as nodes, and there are two interacting networks of nodes: a content network, and a sequence network. We first describe how these networks are organised, and then how they interact to control the serial order of behaviour.

Organisation of the Content Network

The content network consists of hierarchically organised nodes that represent the content of what is perceived or produced. In the case of words and their phonology, content nodes represent units such as syllables, syllable onsets, rhymes, vowel nuclei, codas, and individual segments or speech sounds. Thus, the lexical node for a multisyllabic word connects to its syllable nodes, each of which connects to nodes representing an onset and a rhyme. If the onset or rhyme nodes consist of more than one unit, then the onset node connects to its segment nodes, and the rhyme connects to the vowel nucleus and coda nodes, which can, in turn, connect to several segment nodes.

By way of illustration, Figure 1 shows the hierarchic structure of phonological content nodes for producing the word *truck*, and Figure 2 compares the structure of content nodes for producing words that begin with one versus two consonants (e.g., *duck* versus *drill*, and *demon* versus *dragon*), and contain one versus two syllables (e.g., *duck* versus *demon*, and *drill* versus *dragon*). Under NST, connections between phonological content nodes are all excitatory and two-way, i.e., they connect bottom-up as well as top-down, although only the top-down connections are relevant to present discussion.

Organisation of the Sequence Network

The sequence network consists of nodes that control the sequencing of behaviour by imposing serial order on the activity of content nodes. These sequence nodes ensure that content nodes are activated in proper order and that only one content node at a time becomes activated (see Houghton & Hartley, 1995; Lashley, 1951; MacKay, 1981, 1987). To do this, sequence nodes represent abstract sequential classes by virtue of how they connect and interact with content nodes and with other sequence nodes. Each sequence node has two-way excitatory connections with a set of content nodes that constitute its "sequential domain". In general, a domain is a set of content nodes that all have the same sequential properties or privileges of occurrence in relation to other domains. As applied to syllable structure, a domain corresponds to the content units that can occupy a given sequential position in syllables.

By way of illustration, Figure 1 shows the sequence nodes for producing the word *truck* in NST, with sequence nodes identified via a capitalised label such as ONSET, and content nodes identified via a two part lower case label such as *tr*(onset). The first part of this label (e.g., *tr*) identifies the content, and the second part identifies the sequential domain in parentheses, e.g., (onset). Thus, the sequence node ONSET (see Figure 1) has the domain (onset) and includes all content nodes
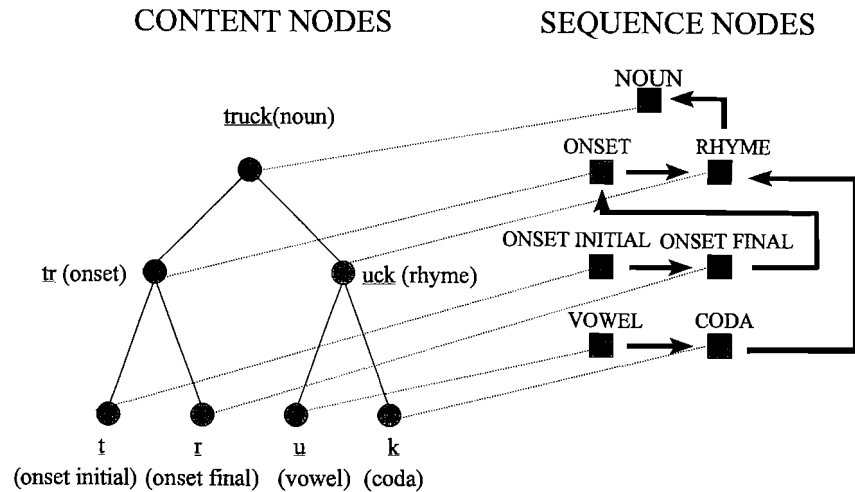
CONTENT NODES   SEQUENCE NODES

Figure 1. Content and sequence nodes for producing the word *truck* in NST. Thin lines represent two-way excitatory connections. Thick lines represent uni-directional inhibitory connections.

representing consonantal units that can precede the vowel in English syllables. This onset domain includes both syllable-initial singletons, e.g., *t*(onset) and *s*(onset), and syllable-initial clusters such as *st*(onset) and *str*(onset).

NST assumes that sequence nodes connect to one another via undirectional links that are inhibitory in nature, and Figure 1 indicates these undirectional inhibitory connections via thick arrows. These inhibitory connections fall into two categories: lateral versus terminal inhibitory connections. Lateral inhibition between sequence nodes represents a sequential rule, e.g., the rule that onsets always precede rhymes in syllables, and is illustrated via thick horizontal arrows in Figure 1, e.g., the one linking ONSET to RHYME. Terminal inhibition signals when one sequential process ends and the next begins, and is indicated via thick vertical arrows in Figure 1, e.g., the one linking CODA to RHYME. The functions and detailed operation of these two types of inhibitory connections are discussed in subsequent sections.

How Content and Sequence Nodes Interact

The interaction between content and sequence nodes is central to the distinction between node priming and node activation in NST. Priming and activation are theoretical processes that originated with Lashley

(1917) and are similar to but also distinct from the concepts of spreading activation and selection in more recent theories. Node priming is a process that prepares a node for possible activation. Priming summates both spatially and temporally, is not self-sustaining, is a parallel and automatic process, and is nonsequential. By contrast, activation is all-or-none rather than graded, taking the form of a sudden increase in activity of a content node. Unlike priming, activation is followed by a brief period during which the content node is self-inhibited. Activation is also discrete and sequential: nodes within the phonological system are activated one after the other rather than in parallel. Finally, activation is not an automatic process, but requires application of an activating mechanism, i.e., sequence nodes.

When activated, a sequence node multiplies the priming level of its entire domain of content nodes until the most-primed content node in the domain reaches threshold and becomes activated. Thus, activating the sequence node ONSET necessarily precedes and causes activation of the most highly primed content node in its domain, e.g., *tr*(onset) in Figure 1. However, sequence nodes are themselves activated under this same, general, "most-primed-wins" principle.

How Sequence Nodes Interact with One Another

Sequence nodes interact with one another in two ways. One way is through lateral inhibitory connections. Two sequence nodes often receive comparable degrees of excitatory priming at the same time, and when this happens, one of the two sequence nodes inhibits the other via their lateral inhibitory link, a process known as a sequential decision. For example, if ONSET and RHYME simultaneously receive excitatory priming, ONSET inhibits RHYME by virtue of its lateral inhibitory link (see Figure 1). This enables ONSET to accrue greater priming than any other phonological sequence node and to become activated. Such sequential decisions or lateral inhibitory interactions between sequence nodes are time consuming, relative to the time it takes a sequence node to activate a content node.

Sequential nodes also interact via terminal inhibitory connections. Terminal connections link a terminal sequence node to its parent node. A terminal sequence node represents the last domain in a content node hierarchy, and its parent node dominates that particular hierarchy. For example, RHYME represents the last domain in the content node hierarchy for the word *truck*, and NOUN is the parent node dominating that hierarchy (see Figure 1). Activation of RHYME therefore inhibits NOUN, so that the next word can be activated. Similarly, ONSET FINAL represents the last domain in the content node hierarchy for the word onset in *truck*, and ONSET is the parent node dominating that hierarchy

(see Figure 1). Activation of ONSET FINAL therefore inhibits ONSET, so that RHYME can be released from inhibition and activated as the next (i.e., most primed) sequence node.

Such interactions between sequence nodes ensure that activation of content nodes proceeds in a top-down, left-to-right tree-traversal manner.[1] To illustrate this tree-traversal activation process in detail for the picture naming task, consider the hierarchically organised content nodes in Figure 1 for producing the word *truck*. Presenting a picture of a truck strongly primes the lexical node for *truck*(noun), which primes its sequence node NOUN. Because no other sequence node is primed, NOUN can be quickly activated, which multiplies the priming of all nodes in the (noun) domain. Because *truck*(noun) has just been primed, it will therefore be activated, and transmit strong, first-order priming to the content nodes *tr*(word onset), and *uk*(word rhyme), which in turn will prime their sequence nodes, ONSET and RHYME. Due to its lateral inhibitory connection with RHYME (see Figure 1), ONSET inhibits RHYME, and becomes activated as the most primed node in the sequence network. ONSET therefore multiplies the priming of every content node in its domain, activating the most primed one. Because *tr*(onset) has just been primed, it will therefore be activated, causing transmission of strong, first-order priming to the content nodes *t*(onset initial) and *r*(onset final), which in turn simultaneously prime ONSET INITIAL and ONSET FINAL. Then the lateral inhibition process repeats, enabling ONSET INITIAL to activate its most primed content node *t*(onset initial), the first terminal node in the phonological system. Activation of *t*(onset initial) primes its articulatory nodes, enabling the articulatory processes that constitute onset of the overt response.

Next, ONSET FINAL is activated, causing activation of its most primed content node, *r*(onset final), and inhibition of ONSET due to the terminal

---

[1] Although many theories assume tree traversal activation processes, NST is the only theory to propose detailed node structures for achieving this top-down, left-to-right pattern. This tree traversal pattern was proposed originally as a way to solve the problem of serial order in higher order planning and memory retrieval processes (Miller, Galanter, & Pribram, 1960; Simon, 1972; Greeno & Simon, 1974; Restle, 1970; Volpert, 1982; Yngve, 1960), and subsequently has been postulated for the perception of structured patterns (Povel, 1981) and the planning and production of sequences of limb movements and key presses (see Collard & Povel, 1982; Gordon & Meyer, 1987; Jordan & Rosenbaum, 1989; Pew & Rosenbaum, 1988; Rosenbaum, 1990, 1991; Rosenbaum, Gordon, Stillings, & Feinstein, 1987; Rosenbaum, Inhoff, & Gordon, 1984; Rosenbaum, Kenny, & Derr, 1983; Rosenbaum, Weber, Hazelett, & Hindorff, 1986). Even for phonological retrieval, the top-down, left-to-right tree traversal idea is not new (Gordon & Meyer, 1987; MacKay, 1970, 1972, 1974; Rosenbaum, 1985; Rosenbaum et al., 1986, 1987).

inhibitory connection between ONSET FINAL and ONSET. This terminal inhibition will release the lateral inhibition on RHYME, and enable RHYME to become activated as the most primed sequence node, causing activation of uk(rhyme). Similar processes enable sequential activation of u(vowel nucleus) and k(coda) (see Figure 1).[2]

NST PREDICTIONS FOR VOCAL LATENCY IN STANDARD PICTURE NAMING

The tree-traversal activation process in NST generates several novel predictions that are tested in the present experiments. These predictions are based on how long it takes to activate the first terminal node, i.e., left-most, bottom-most node in the hierarchy of phonological content nodes for articulating the word. Because inhibitory interactions between sequence nodes are relatively protracted, the time it takes to retrieve (activate) a terminal content node will increase as a direct function of how many sequential decisions or inhibitory interactions between sequence nodes are required before the terminal sequence node is activated. Under NST, vocal production follows on the heels of activating a terminal node, so that the main factor influencing vocal latency under NST is how many sequential decisions must be taken in activating nodes in the left branch of the tree, prior to articulating the first segment in the word. Factors that affect the time to activate subsequent nodes should affect the duration of the vocal response, but not its latency under NST.

To illustrate this general point in more detail, consider the node structures for words with one versus two syllables, say, *duck* versus *demon* in Figure 2. As the Roman numerals in Figure 2 indicate, only one sequential decision is required to activate the first segment of a monosyllabic word with a simple onset, e.g., *duck*, whereas two sequential decisions are required to activate the first segment of a bisyllabic word with a simple onset, e.g., *demon*. Consequently, activating the first segment should take longer for bisyllabic than monosyllabic words.

Consider now the structures of content nodes for words with one versus two onset consonants, say, *duck* versus *drill* in Figure 2. Only one sequential decision is required for activating the first segment of a monosyllabic word with simple onset, e.g., *duck*, whereas two sequential decisions are required for activating a monosyllabic word with a complex onset, e.g., *drill*. Consequently, activating the first segment should take

---

[2] To simplify this illustration, we ignore the case of errors, where a content node receiving top-down priming happens not to be the most-primed node in its domain when the activating mechanism (sequence node) is applied, so that the wrong content node is activated under the most-primed-wins principle.

DUCK
/dʌk/

/d/   I   /ʌk/

/ʌ/   /k/

R start

DEMON
/dimɔn/

/di/   I   /mɔn/

II

/d/   /i/   /m/   /ɔn/

/ɔ/   /n/

R start

DRILL
/drɪl/

/dr/   I   /ɪl/

II

/d/   /r/   /ɪ/   /l/

R start

DRAGON
/drægən/

/dræ/   I   /gən/

/dr/   II

III

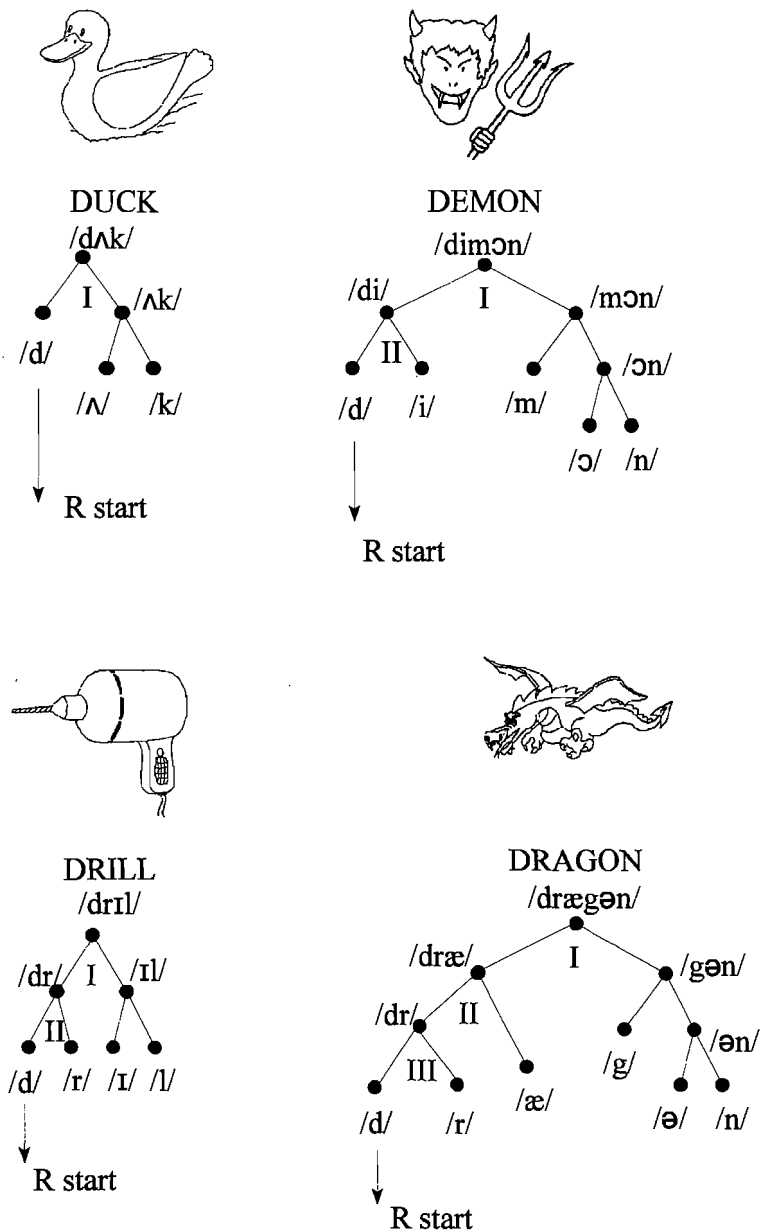/d/   /r/   /æ/   /g/   /ən/

/ə/   /n/

R start

Figure 2. Node structures for example words with one versus two onset consonants (e.g., *duck* versus *drill*), and one versus two syllables (e.g., *duck* versus *demon*). Roman numerals indicate how many sequential decisions precede activation of the left-most, bottom-most content node.

8

longer for words that contain complex onsets than simple or singleton onsets.

If sequential decisions require comparable amounts of time at all levels in the content network, then NST generates five predictions for vocal latencies in picture naming tasks: First, there should be an effect of syllable length such that vocal latency is longer for bisyllabic than monosyllabic words. Second, there should be an effect of onset complexity such that vocal latency is longer for words starting with onset clusters than with singletons. Third, effects of word and onset complexity should be independent and therefore additive. Fourth, complex vowel nuclei and complex codas should not affect vocal latency because sequential decisions for these units follow activation of the first segment in the word, which determines onset of the response. Fifth, there should be no interactions between onset, vowel nucleus and coda complexity. Experiment 1 tested Predictions 1–3, and Experiment 2 replicated the results of Experiment 1 and tested Predictions 4 and 5.

DATA RELATED TO PHONOLOGICAL RETRIEVAL IN NST

We do not reiterate MacKay's (1987) comprehensive review of phono-logical retrieval data here. Instead we concentrate on phonological retrieval research carried out since 1987, focusing on the aspects of NST that might seem controversial within recent literature, especially the appropriateness of picture naming for testing the separation of content and sequence networks, the existence of syllables, and the hierarchic organisation of syllables in NST.

The Separation of Content versus Sequence Networks in NST

The separation of content versus sequence networks in Figure 1 is currently a controversial issue. The main evidence for the separation comes from a regularity that has been observed in a wide range of speech error data: Segments involved in phonological substitution errors almost invariably occupy identical positions within their syllables (Boomer & Laver, 1968; Fromkin, 1971; MacKay, 1970, 1987 (pp. 120–121)). Although details of this "syllable position regularity" may differ somewhat between languages (compare, for example, Shattuck-Hufnagel, 1987, 1992, for English, versus Berg, 1991, and García-Albea, del Viso, & Igoa, 1989, for Spanish), it is also central to language production theories that have adopted a "slots-and-fillers" approach (Dell, 1986; Fromkin, 1971; Garrett,

1975; Levelt, 1989; Shattuck-Hufnagel & Klatt, 1979; Stemberger, 1985, 1990; Stemberger & Treiman, 1986).

Nevertheless, some investigators have argued that other factors, such as featural similarity (Shattuck-Hufnagel, 1986; Shattuck-Hufnagel & Klatt, 1979), context similarity (Dell, Juliano, & Govindjee, 1993), and sonority (e.g., Hooper, 1972), and other processes such as the Onset Principle (Selkirk, 1982) can explain the syllable position regularity in phonological speech errors without assuming separate representations for content versus sequential form (Dell et al., 1993; Levelt, 1992; Roelofs, 1997a). Levelt and colleagues (Levelt, 1992; see also Levelt, Roelofs, & Meyer, in press; Meyer, 1997; and Roelofs, 1997a, b) have also argued against the concept of pre-stored syllable structures because connected speech often exhibits resyllabification across word boundaries (Schiller, Meyer, Baayen, & Levelt, 1996), suggesting that syllable structure may be created on the fly rather than stored with each word in the lexicon and then changed depending on sentence context. However, this issue is far from settled (compare Roelofs & Meyer, 1998, versus Costa & Sebastián, 1998; Ferrand & Segui, 1998; Meijer, 1994, 1996; Romani, 1992; Sevald & Dell, 1994; Sevald, Dell, & Cole, 1995). By examining predictions derived from how sequence nodes function in NST, the present research will therefore bear on this controversy concerning the separation of sequence versus content mechanisms.

The Hierarchic Organisation of Syllables in NST

Although most of the units in NST are generally accepted constructs, the hierarchic organisation of content nodes is currently a controversial issue. NST divides syllables into onset and rhyme, and the rhyme into vowel nucleus and coda, where the vowel nucleus consists of a simple vowel and a glide or liquid, and the coda consists of all consonants following the vowel nucleus. These units have been given a variety of names, but are now widely accepted in both linguistics (Fudge, 1969, 1987; Harris, 1983; Selkirk, 1982) and psycholinguistics, with supporting evidence from speech errors (MacKay, 1972; Shattuck-Hufnagel, 1986; Stemberger, 1983, 1984, 1985; Treiman, 1995; Treiman & Danis, 1988), metalinguistic tasks (Treiman, 1983, 1984, 1986; see Treiman, 1989, for a review), distributional analyses (Kessler & Treiman, 1997) and in the case of rhyme units, on-line response times in production tasks (Meyer, 1991; Yaniv, Meyer, Gordon, Huff, & Sevald, 1990; but see Sevald & Dell, 1994). In the case of complex onsets, Stemberger & Treiman (1986) showed that onset-initial units are less vulnerable to errors than onset-final units, suggesting that these domains of content nodes constitute independent structural positions, as in NST. In the case of complex codas, consonants in a coda cluster often

behave as independent units in errors, indicating that coda clusters are not indivisible, but consist of independent consonants. Regarding vowel nuclei, the error induction experiment of MacKay (1978) demonstrated that the simple vowel in a diphthong can participate in errors independently of its subsequent glide. However, for errors in everyday speech, components of a diphthong rarely act as independent units (Shattuck-Hufnagel, 1986; and Stemberger, 1984; but see Stemberger's, 1984, analysis of German and Swedish errors), as if complex vowel nuclei are much more cohesive units than coda and onset clusters (see also Meijer, 1994, 1996).

However, alternative accounts of these data have been proposed (Hayes, 1989; Levelt, 1992; Pierrehumbert & Nair, 1995), and several current theories of phonological retrieval pose a flatter onset-vowel nucleus-coda structure, e.g., where units following the onset are not grouped as a rhyme consisting of vowel nucleus and coda (Dell, 1986; Eikemeyer & Schade, 1991; Hartley & Houghton, 1996). In short, further research is needed to convincingly demonstrate the psychological reality and hierarchic organisation of onset, vowel, and coda clusters. By examining predicted effects of onset complexity, vowel complexity, and coda complexity on response times, the present research will therefore bring a new source of data to bear on these syllable structure controversies.

The Existence of Syllables in NST

Under NST, syllable nodes represent fundamental phonological units in the hierarchy of content nodes for producing a word. However, available evidence on the syllable as a unit of language processing is mixed, being stronger in speech perception and reading tasks (Carreiras, Alvarez, & de Vega, 1993; Jared & Seidenberg, 1990; Mehler, Dommergues, Frauen-felder, & Segui, 1981), and especially weak in the case of language production. Perhaps the best evidence comes from the tip-of-the-tongue (TOT) phenomenon, where speakers are unable to retrieve the full phonology of a familiar word, but can correctly guess how many syllables it contains with greater than chance probability (Burke, MacKay, Worthley, & Wade, 1991). However, sophisticated guessing hypotheses (Brown, 1991) have undermined this TOT evidence.

Output preparation effects also indicate mixed support for syllables as output units. Meyer (1991) reported a first-syllable preparation effect for Dutch speakers that was independent of how many segments that syllable contained. However, subsequent studies in Dutch (Roelofs, personal communication, 1997) and Spanish (Costa & Sebastián, 1996) failed to replicate this result, reporting only an effect of the number of prepared segments. Similarly, Romani (1992) found a preparation effect when prime and target shared number of syllables without sharing segmental content.

However, this effect may be limited to cases where prime and target share initial segments (Roelofs & Meyer, 1998), and to words with irregular stress patterns (Meyer, Roelofs & Schiller, in preparation, cited in Levelt et al., in press), at least in Dutch.

A number-of-syllables-effect on vocal latency was reported by Stern-berg, Monsell, Knoll, and Wright (1978) in a delayed naming task, where participants know what word(s) they must produce well in advance of the reaction signal. However, Klapp, Anderson, and Berrian (1973) did not find a number-of-syllables-effect in delayed naming, but only in standard picture naming tasks where participants do not know in advance what name to produce. Further complicating the picture, Bachoud-Levy, Dupoux, Cohen, and Mehler (1998) did not replicate Klapp et al.'s results in five experiments involving standard picture and symbol naming in French or English.[3]

Using masked primes and picture naming, Ferrand, Segui, and Grainger (1996) found a syllable-match effect in French. Primes such as *pa* speeded latencies for words like *pa.lace*, where it corresponds to the first syllable, compared to words like *pal.mier*, where it does not. The opposite happened for the prime *pal*. However, Schiller (1998) failed to replicate this effect using otherwise similar procedures in Dutch. Baumann (1995) and Romani (1992) did not find a syllable-match effect using other variants of priming tasks in production, even though all three studies found a general effect of segmental overlap. This contrasts with the report by Wheeldon and Levelt (1995) of a syllable-match effect when Dutch participants monitored their own inner speech for syllable targets, which suggests the possible involvement of perceptual factors in this study. The syllable frequency effects reported by Levelt and Wheeldon (1994) have also turned out to be irreplicable when confounds between syllable frequency and phoneme frequency were controlled for (Hendricks & McQueen, 1996; Levelt et al., in press).

Because evidence for the psychological reality of syllabic units is clearly stronger for perception than production, it was important that syllable effects in the present research be attributable to production rather than perception. For example, a reading task would not suffice because any

---

[3] However, Bachoud-Levy et al. (1998) also failed to find the usual latency differences between mono- versus bi-syllabic words controlled for length, and the causes of this non-effect are presently unclear. A collaborative research project is now underway, aimed to isolate the factors that are responsible for the presence-absence of the number of syllables effect in French, English and Spanish. Factors such as the use of a base-line versus overtraining method, proportion of monosyllabic fillers, stress placement, global slowing of latencies, and word frequency (which was suggested to us by Ludovic Ferrand) are already under evaluation. Data obtained up to now suggest stress placement as the most promissory intervening variable.

effect of number of syllables could be ascribed to the perception rather than production side of this task. In order to avoid such confounds, the present experiments adopted the picture naming task. Because perceiving a picture has nothing to do with perceiving phonology or orthography, a number-of-syllables effect could only reflect the production side of this task, i.e., retrieval of the name and its phonology, and not the perception side.

The Absence of Articulatory Buffering in NST

Many recent studies have found left-to-right preparation effects (Meyer, 1990, 1991; Meyer & Schriefers, 1991; Sevald & Dell, 1994; Sevald, Dell & Cole, 1995), generally taking the form of greater priming from primes that share the initial part of the phonological target word than from primes that share the final part. NST is able to explain these results under the assumption that sequential decisions can be executed in advance, during the preparation time, and can be facilitated by primes presented prior to production of the target word, i.e., inhibitory interactions between sequence nodes can be facilitated by preparatory priming. When the moment to start activating nodes for word production comes, sequential decisions are already "biased", greatly reducing the time for the appropriate sequence node to reach threshold.

However, the interpretation most widely proposed for these phenomena assumes a serial stage that precedes buffering of phonological material into an articulatory buffer. Articulation may not start until at least a whole phonological word is buffered. Priming occurs because less material remains to be buffered when the initial parts of the target word have been primed, a standard assumption in models of both reading aloud and word production (Coltheart, Curtis, Atkins & Haller, 1993; Dell, 1986; Levelt, 1989; Seidenberg & McClelland, 1989).

Up to now, only one study provides a clear differentiation between these views of processes that precede onset of articulation in reading aloud. Kawamoto, Kello, Jones, and Bame (1998) were able to independently measure vocal latency and duration of the first consonant for monosyllabic words whose vowel had a regular or irregular pronunciation. Although a small part of the effect of regularity on latency reflected vowel duration, the major part reflected the duration of the first consonant, indicating that complexities of later-coming parts of a word may be resolved after articulation has begun (thereby influencing word duration but not latency).

By examining effects of syllable structure complexity in initial, medial and final parts of a syllable, the present study will provide data directly related to the question of the criterion to start pronunciation and the existence of articulatory buffering in a purely production task.

GENERAL METHODOLOGICAL ISSUES

Despite ongoing controversies concerning the details of how pictorial stimuli are processed (see for example Glaser, 1992, for a review), most theorists agree that standard picture naming proceeds via the following ordered steps. Step 1 involves perceptual analysis and recognition of the picture, which implies access to the semantic code for the object. In Step 2, this semantic code is used to select an abstract verbal label. Step 3 involves ordered activation of the phonological units that guide articulation of the verbal label. We wanted the present experiments to tap into Step 3 while controlling for processes in Steps 1 and 2. We discuss these control procedures next, first for Step 2 processes and then for Step 1 processes.

Step 2 Control Procedures

Step 2 is complicated by problems known as name-underspecification: the fact that many different labels can apply to any given object (e.g., Levelt, 1989, p. 201). For example, a picture of a "dog" can be labelled at different levels of generality as a dog, an animal, or a Dalmatian. Even at the same level of generality, selection among alternative names is possible, e.g., a tap can be called a faucet, and different viewers often disagree on what name best applies to a given picture (see Snodgrass & Vanderwart, 1980).

To solve these name-underspecification problems, we first trained participants on the target name they should use for each picture in the present experiments. This training occurred in a "delayed naming task," where participants first saw a printed name and then the picture that they were instructed to associate with the target name. During picture presentation they also prepared to say the name as quickly as possible after appearance of a "go" signal. Our goal here was twofold: first and foremost, to ensure that participants knew the correct name to use for a particular picture and could pronounce the name correctly; and secondly, to check for between-condition differences in the time required for low level articulatory and acoustic processes, including possible voice key sensitivity differences across different sets of words.

However, we did not intend to use delayed naming latencies to draw strong conclusions regarding articulatory processes for two reasons. First, given our primary goal of training subjects on the selected picture-name pairings, we chose long and fixed delays, in order to minimise distraction. We were aware that such long and fixed delay intervals may induce participants to repeatedly access the verbal label while waiting for the "go" signal which is a different and complicating factor (see McRae, Jared, & Seidenberg, 1990; Savage, Bradley, & Forster, 1990).

Second, we chose the disappearance of the picture as the signal for participants to produce the prepared name. Pilot participants found it difficult to attend to and learn the relation between name and picture if they had to pronounce the name either as soon as possible after presentation of the picture or after a tone with the picture still visible. Because of perceptual after-effects, we recognise that picture disappearance was not the best "go" signal for assessing articulation time. Nonetheless, our delayed naming latencies provided a secondary check on the effectiveness of our main control for articulatory factors: balancing initial phonemes across conditions.

Step 1 Control Procedures

Perceptual analysis and recognition of depicted objects in Step 1 can be affected by the fact that objects and depicted objects can vary in visual complexity and familiarity or frequency of encounter. To control for these factors, we introduced a name-picture matching task, variants of which have been used to assess these aspects of picture recognition in many other studies (Jescheniak & Levelt, 1994; Schriefers, Meyer, & Levelt, 1990; Theios & Armhein, 1989). This control procedure required that our participants first learn the names for the set of experimental and filler pictures. They then saw one of the names, immediately followed by a picture, i.e., the same sequence of events as in the delayed naming task. However, unlike delayed naming, half the trials involved a name paired with the picture for an inappropriate (filler) object, and participants had to respond as quickly as possible via keys labelled yes and no on the computer keyboard to indicate whether the name matched the picture.

Unknown to the participants, experimental words always matched their pictures, and therefore always required yes responses, and correct yes responses in this task can be argued to involve Step 1 processes but not Step 2 or Step 3 processes under two accounts of the name-picture matching task. One account (after Theios & Armhein, 1989) is that participants store the verbal concept for the name as soon as it appears in the name-picture matching task, and after the picture appears, they call up the verbal concept linked to the picture in the training block (Glaser, 1992). They then compare the presented verbal concept with the picture-linked verbal concept, and respond yes for a match, and no for a mismatch. This comparison process always results in a match for experimental items, so that under this account, the matching process and the selection and execution of yes responses should not differ across experimental conditions that are defined by phonological characteristics of the names. Consequently, name-picture matching latencies index the time required to

perceive the picture and access its semantic code without involving phonological retrieval processes under this account.

Under a second account of name-picture matching, when participants see the name, they call up an image of its associated picture, and then compare this image with the subsequently presented picture. Because phonological retrieval processes play no role in this account, name-picture matching latencies should again only reflect factors related to perceptual analysis and object recognition, e.g., picture complexity or object frequency. In short, accounts 1 and 2 involve "conceptual matching" because they require neither ordered retrieval of a word's phonological units nor articulation of its first segment. These conceptual matching accounts therefore exclude Step 3 of the standard picture naming task. Also excluded under conceptual matching accounts is Step 2 (the process of selecting an abstract label or verbal concept) because the verbal concept has already been presented and resides in short term memory. This means that subtracting name-picture matching latencies from standard picture naming latencies should eliminate the time required to perceptually analyse and recognise the experimental pictures under conceptual matching accounts, so that these "corrected latencies" should provide an index of phonological retrieval time that is unrelated to picture complexity and frequency of encounter. Therefore, the rationale behind this control is that, under conceptual matching accounts, the standard picture naming latencies could be distorted by inequalities in picture/conceptual processing. These factors should also influence name-picture matching latencies, and their influence should be eliminated when subtracting the latter from the former. Our predictions are then mainly for corrected latencies and standard latencies are taken only as an approximation of these. Both types of latencies should show the same effects if all other factors are equal.

A third possibility is that phonological factors affect name-picture matching latencies. Under this account of name-picture matching, participants store the name in phonological form as soon as it appears, and after the picture appears, they retrieve the phonological form of the word that was linked to the picture in the training block. They then compare the presented and retrieved phonological forms, and respond yes for a match, and no for a mismatch. This phonological matching process therefore involves phonological retrieval of the experimental names (although not articulation of their initial segments), and should differ across experimental conditions that are hypothesised to affect phonological retrieval, i.e., onset complexity and number of syllables of the experimental names. If participants engage in this strategy, the phonological effects should be similar in standard naming and name-picture matching. Under this phonological matching account, any effects of onset

complexity and number of syllables in standard picture naming should disappear when picture-name matching latencies are subtracted from standard picture naming latencies, resembling delayed naming latencies in overall pattern.

EXPERIMENT 1: ONSET COMPLEXITY AND NUMBER OF SYLLABLES

Experiment 1 tested Predictions 1 (number of syllables effect), 2 (onset complexity effect), and 3 (additive onset and syllable effects) for standard picture naming latencies, and for corrected latencies, i.e., standard picture naming latencies minus name-picture matching latencies. Mono- and bi-syllabic words starting with a simple or complex onset were selected, and presented in three different tasks: a delayed naming task, a standard naming task and a name-picture matching task. As described in the preceding section, the goal of the delayed naming task was to give subjects training on the chosen name-picture pairings, and the goal of the name-picture matching task was to correct for extraneous perceptual influences (of Step 1) on the latencies obtained in the standard picture naming task.

Method

Participants. Thirty participants were recruited from the University Cooperative Housing Association in Westwood, CA or from the summer school courses at UCLA. All were native English speakers or had used English continuously over the past 15 years. Each participant was paid $7 for participating.

Materials and design. With help from sources such as Corbell (1986) and Snodgrass and Vanderwart (1980), we selected 112 depictable words for Experiment 1. All referenced common objects and had moderately high frequency of usage in Francis, Kucera, and Mackie (1982). The 56 items shown in Appendix A served as experimental words and crossed onset complexity and number of syllables in a 2 (onset complexity) × 2 (number of syllables) factorial design. Words were matched on initial phoneme, initial-syllable stress and mean word frequency across the four conditions (see Appendix A). We also attempted to match the complexity of the vowel nucleus and coda in the first syllable of words across the four conditions. When this was not possible, we made sure that the less complex vowel nucleus or coda occurred in the bisyllabic or complex onset conditions, a conservative procedure that could only work against our hypotheses. To prevent awareness of our experimental variables, the remaining 56 "filler" words resembled experimental words in frequency but differed in stress pattern, number of syllables and syllable structure.

After the 112 words were selected, an artist (J.S.) drew a picture for each word by hand, and scanned these line drawings into independent computer files which were subsequently modified via an editing program for clarity. For example, the stimulus PLUM was coloured purple to preclude the response APPLE. Colour was also used to indicate cases where a part rather than the entire picture was to be named, e.g., the TRIGGER of a GUN.

Procedure.   Participants sat in a quiet room facing a computer screen, and the experimenter sat behind them and to their right in order to minimise distraction. The words were presented via PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993) running on a Macintosh Quadra 605. Vocal responses of the participants stopped the timer via a microphone-voice key system and the computer automatically recorded latencies to the nearest millisecond.

A typical experimental session lasted 40 minutes and consisted of 366 trials in three blocked conditions: 122 trials of delayed picture naming, 122 trials of standard picture naming, and 122 trials of name-picture matching. General instructions described the three tasks in broad overview for the participants. Participants were instructed to respond as quickly and as accurately as possible. More specific instructions preceded each block, followed by 10 practice trials. Then came the 112 experimental and filler pictures in random order for each block based on the following randomisation program: On trial 1, the program selected between experimental (E) or filler (F) at random. If E was selected, the program first chose between the four experimental conditions at random, and then selected one of the stimulus words in that condition at random. Because trial 1 was an experimental word, trial 2 was a filler selected at random from among the 56 fillers. On trial 3, the program again chose at random between E and F, and sampled at the word level without replacement, until all of the filler and experimental words had been chosen. These randomisation procedures ensured that every 8 trials included stimuli from all four experimental conditions intermixed with four fillers.

On delayed picture-naming trials, events proceeded in the following order: a fixation point for 500 ms, a name in capital letters for 1000 ms, and the picture appropriate to that name for 1000 ms during which participants were encouraged to encode the name-picture combination for two reasons: they had to use these same names in a subsequent naming task involving the same pictures; and they had to produce the appropriate name as quickly as possible after the picture disappeared from the screen. An error was recorded via the computer keyboard for anything other than a fluent production of the correct word, e.g., wrong names, dysfluencies, and extraneous noises.

On standard picture-naming trials, a 500 ms fixation point preceded a 1000 ms picture. Participants were instructed to name the picture as soon as possible after presentation, with errors scored as above. If participants produced a name other than the target name used in the delayed naming block, the experimenter reminded them of the desired name prior to the next trial.

On name-picture matching trials, a 500 ms fixation point preceded a 1000 ms name which preceded a 1000 ms picture, and participants were instructed to judge whether the name matched the picture, responding *yes* or *no* as quickly as possible via labelled keys on the computer keyboard. Index fingers of each hand rested on these keys such that the *yes* key always corresponded to the participant's dominant hand. Because names preceding the 56 filler pictures were rearranged (i.e., the name of a filler picture preceded a different and incongruent filler picture), the name failed to match the picture on half the trials in this task, and required *no* responses. The experimental words always required *yes* responses. The constraints of our randomisation procedure ensured that no more than two trials in a row required the same response in the name-picture-matching task. The computer automatically scored the accuracy of each response, together with the latency of correct responses. Participants pressed the space bar to proceed to the next trial.

Results

Latencies smaller than 100 ms or greater than 2000 ms were considered outliers and discarded. Table 1 presents mean errors and latencies for correct responses in the four conditions in each task in Experiment 1, and corrected latencies, i.e., standard picture naming latency minus name-

TABLE 1

Mean Latencies (in ms) and Number and Percentage of Errors as a Function of Onset Complexity and Number of Syllables for the Delayed Picture Naming, Standard Picture Naming, and Name-Picture Matching Tasks in Experiment 1

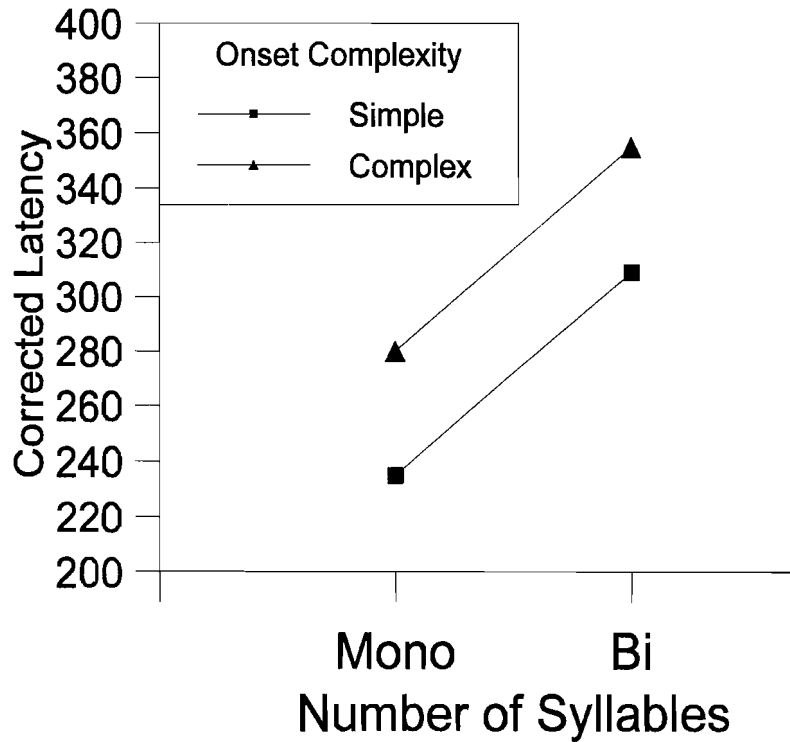| | | Number of syllables | | | |
| | | Latency | | Errors | |
Task	Onset complexity	Mono	Bi	Mono	Bi
Delayed	Simple	396	399	5 (1.19%)	12 (2.85%)
naming	Complex	407	399	2 (0.47%)	7 (1.66%)
Standard	Simple	768	859	46 (10.95%)	59 (18.8%)
naming	Complex	868	899	86 (20.47%)	79 (18.8%)
Name-picture	Simple	532	550	13 (3.09%)	7 (1.66%)
matching	Complex	587	543	11 (2.61%)	6 (1.42%)

Figure 3.   Mean corrected naming latency (by participants) as a function of onset complexity and number of syllables in Experiment 1.

picture matching latency, are presented in Figure 3. We analysed these measures independently for each task using 2 (onset complexity) × 2 (number of syllables) by-subject ANOVAs (F_1) and by-item ANOVAs (F_2). All P values smaller than 0.15 are reported.

Block 1: Delayed Naming.   Outliers comprised 2.5% of the data in the delayed naming task. The analyses on mean latencies indicated no main effect of onset complexity [$F_1(1,29) = 1.31$, MSe $= 626.92$, $P > 0.15$, $F_2 < 1$], or number of syllables, both Fs < 1. There was no reliable onset × number of syllables interaction [$F_1(1,29) = 3.04$, MSe $= 278.25$, $P = 0.09$, $F_2 < 1$].

Errors comprised 1.54% of experimental trials in the delayed naming task. Analyses of error means indicated a main effect of number of syllables [$F_1(1,29) = 7.25$, MSe $= 0.001$, $P = 0.01$; $F_2(1,52) = 4.43$, MSe $= 0.001$, $P = 0.04$], but no main effect of onset complexity [$F_1(1,29) = 1.82$, MSe $= 0.001$, $P > 0.15$, $F_2(1,52) = 1.97$, MSe $= 0.001$, $P > 0.15$], and no interaction, both Fs < 1.

Block 2: Standard Naming.   Outliers comprised 1.13% of the data in the standard naming task. Mean latencies showed a main effect of onset complexity [$F_1(1,29) = 39.19$, MSe $= 3713.37$, $P < 0.01$; $F_2(1,52) = 4.47$, MSe $= 12155.38$, $P = 0.03$], and of number of syllables [$F_1(1,29) = 33.60$, MSe $= 3353.18$, $P < 0.01$; $F_2(1,52) = 3.32$, MSe $= 12155.38$, $P = 0.07$]. There was also an onset complexity × number of syllables interaction by subjects [$F_1(1,29) = 16.35$, MSe $= 1688.64$, $P < 0.01$], but not by items [$F_2(1,52) = 1.42$, MSe $= 12155.38$, $P > 0.15$].

Errors comprised 16.07% of experimental trials in the standard naming task. Error means showed a main effect of onset complexity by subjects [$F_1(1,29) = 25.58$, MSe $= 0.006$, $P < 0.01$], and a marginal effect by items [$F_2(1,52) = 3.3$, MSe $= 0.022$, $P = 0.07$]. No other effects reached $P < 0.15$.

Block 3: Name-picture Matching.   Outliers comprised 0.23% of the data in the name-picture matching task. Analyses of mean manual latencies indicated no main effect of number of syllables [$F_1(1,29) = 1.27$, MSe $= 3955$, $P > 0.15$, $F_2(1,52) = 1.04$, MSe $= 2613.6$, $P > 0.15$], but a main effect of onset complexity [$F_1(1,29) = 5.57$, MSe $= 3264.85$, $P = 0.02$; $F_2(1,52) = 3.68$, MSe $= 2613.6$, $P = 0.06$]. There was also an onset complexity × number of syllables interaction [$F_1(1,29) = 15.89$, MSe $= 1803.18$, $P < 0.01$; $F_2(1,52) = 4.97$, MSe $= 2613.6$, $P = 0.03$], such that manual latencies were longer for monosyllables with complex onsets than for any other stimulus condition, which did not differ among themselves (see Table 2).

Errors comprised 2.2% of experimental trials in the name-picture matching task, and an ANOVA on these data indicated no reliable effects, despite a trend towards reduced accuracy in both monosyllabic conditions [$F_1(1,29) = 2.72$, MSe $= 0.002$, $P = 0.11$; $F_2(1,52) = 3.1$, MSe $= 0.001$, $P = 0.08$].

TABLE 2

Mean Latencies (in ms) and Number and Percentage of Errors as a Function of Onset Complexity and Number of Syllables in Experiment 2 using Identical Stimuli to Experiment 1 in the Delayed and Standard Naming Tasks

| Task | Onset complexity | Number of syllables | | | |
| | | Latency | | Errors | |
		Mono	Bi	Mono	Bi
Delayed	Simple	374	387	2 (0.47%)	2 (0.47%)
naming	Complex	398	397	1 (0.23%)	8 (1.90%)
Standard	Simple	780	856	45 (10.71%)	64 (15.23%)
naming	Complex	882	926	73 (17.38%)	75 (17.85%)

Corrected Latencies.   Mean corrected latencies are shown in Figure 3. There were main effects of number of syllables by subjects [$F_1(1,29)$ = 20.61, MSe = 8030.19, $P < 0.001$], and by items [$F_2(1,52)$ = 6.37, MSe = 10075.37, $P = 0.01$], and of onset complexity by subjects [$F_1(1,29)$ = 7.7, MSe = 7891.93, $P = 0.01$], but not by items [$F_2(1,52)$ = 1.809, MSe = 10075.37, $P > 0.15$]. There was no onset complexity × number of syllables interaction (see Figure 3; both Fs < 1). Appendix A shows the value of the corrected latency obtained for each experimental item.

Discussion

Standard naming latencies in Experiment 1 supported Predictions 1 and 2 of NST: standard picture naming latencies showed effects of both onset complexity and number of syllables of the word to be pronounced. Response errors did not qualify interpretation of the latency data. However, the greater effect of onset complexity for monosyllables failed to support Prediction 3 (additive effects of onset complexity and number of syllables). Corrected latencies showed a clear 74 ms effect of number of syllables, and a 45 ms effect of onset complexity by subjects, consistent with Predictions 1 and 2, and there was no hint of an onset complexity × number of syllables interaction, consistent with Prediction 3. In short, corrected latencies supported NST Predictions 1–3 under a conceptual matching account of how participants carry out the name-picture matching task. Corrected latencies did not support the alternative phonological matching account of the name-picture matching task: corrected latencies were not equivalent for words with one versus two syllables, or for syllables with complex versus singleton onsets, nor did corrected latencies resemble delayed naming latencies in overall pattern.

Despite these fits with NST predictions under a conceptual matching account, several problems in Experiment 1 require discussion. The main effect of onset complexity in the name-picture matching data suggest that variables other than phonological structure, e.g., picture complexity, may have affected the monosyllabic, complex onset condition. The high error rates in the standard naming task for both mono- and bi-syllabic words with complex onsets (see Table 1) raises a similar issue, and suggests the need for replication with (at least partially) different stimuli.

Another issue concerns the limited statistical power in the by-item analyses in Experiment 1, which was especially troublesome in the case of the large (45 ms) but unreliable onset complexity effect with corrected latency measures. Several factors contributed to this power problem. One concerned the strong constraints on item selection, which limited the number of items per condition. The between-item statistical analyses required when intrinsic characteristics of the items define the experimental

conditions, as in the present research, also contributed to the power problem. For by-item analyses, between-item designs are less sensitive than within-item designs, increasing the difficulty of discriminating experimental effects from noise (see the General Discussion section). A third contributor was data loss due to outlier and error discard in the latency analyses.

One surprising aspect of Experiment 1 data for which we have no ready explanation is the large number of errors for bi-syllabic words in the delayed naming block. Perhaps this number-of-syllables effect reflects the fixed, 2000 ms preparation interval in this task. If some subjects required more than 2000 ms to memorise the name-picture pair for some items and retrieved the phonology only after the reaction signal appeared, this would introduce effects of phonological complexity on delayed naming data. This possibility again suggested the need for replication of this experiment, which was one of the goals of Experiment 2.

To summarise, Experiment 1 showed that both onset complexity and number of syllables influence standard picture naming latency and, when name-picture matching latency is used to correct for factors unrelated to language production, both variables have additive effects.

EXPERIMENT 2: ONSET, VOWEL NUCLEUS AND CODA COMPLEXITY

Experiment 2 employed the same general procedures as Experiment 1, but had three new goals. One was to replicate the findings of Experiment 1 with the same and different stimuli in order to test the robustness of onset complexity and number of syllables effects. A second goal was to test NST Predictions 4 and 5 that complexity of the vowel and coda will not affect vocal latency and will not interact with the onset complexity effect. However, we could find very few depictable words with CVVCC and CCVVCC structures, where VV represents a complex vowel (a diphthong). This made it impossible to cross two levels of complexity for onsets, vowels, and codas in a full factorial design. Instead, we created three independent, partial factorial designs consisting of three overlapping sets of words in Experiment 2. Item Sets 1 and 2 contained words with CVC, CVV(C), CVCC, CCVC, CCVV(C) and CCVCC structures, where constituents in parentheses are optional. Words with CVC, CVV(C), CCVC, and CCVV(C) structures factorially crossed onset and vowel nucleus complexity while matching other factors, and words with CVC, CVCC, CCVC, and CCVCC structures crossed onset and coda complexity with other factors matched.

Moreover, we were also able to exactly replicate procedures for the delayed and standard picture naming tasks of Experiment 1, by using Item

Set 3, which included the full set of stimuli from Experiment 1, acting either as experimental items or as fillers for the other designs in Experiment 2. In short, Item Set 3 tested NST Predictions 1–3 for delayed and standard naming latencies. However, Item Set 3 stimuli that were acting as fillers for the other designs in Experiment 2 were assigned *no* responses in the name-picture matching task, whereas those acting as experimental items were assigned *yes* responses, making infeasible an exact replication of name-picture matching procedures of Experiment 1.

Method

Participants. The 30 participants in Experiment 2 were a different sample from the same population as in Experiment 1.

Procedure. Procedures were identical to Experiment 1 except in the following respects. A typical experimental session lasted 45 minutes and consisted of 396 trials, with 132 experimental trials plus 6 practice trials in each of the three blocks: delayed naming, standard naming, and name-picture matching.

Materials and analytic design. Materials for Experiment 2 appear in Appendix B and had the same general characteristics as in Experiment 1. However, experimental words in Experiment 2 fell into three overlapping item sets that we later treated separately in our statistical analyses. Item Set 1 consisted of 55 monosyllabic words matched for initial phoneme and frequency across 5 different conditions (11 words per condition) corresponding to the structures CV(C), CVV(C), CVCC, CCVC, and CCVV(C).

Item Set 2 was a partially overlapping set of 36 CV(C), CVV(C), CVCC, CCVC, CCVV(C), and CCVCC words, with 6 words per structure matched for frequency and initial phonemes. The reason for creating Item Set 2 was that we could find only 6 CCVCC words that were depictable. In order to test for effects of coda complexity, we therefore had to combine these 6 CCVCC words with 6 CV(C), CVV(C), CVCC, CCVC and CCVV(C) words matched for frequency and initial phonemes from Item Set 1. Item Set 2 formed the basis for two partially independent 2 × 2 designs, one that crossed onset and vowel nucleus complexity (conditions CV(C), CVV(C), CCVC, and CCVV(C)), and the other that crossed onset and coda complexity (conditions CV(C), CVCC, CCVC, and CCVCC).

Item Set 3 was another partially overlapping set of words that crossed onset complexity with number of syllables in order to replicate the delayed and standard naming tasks in Experiment 1 as closely as possible. Item Set 3 contained the 56 experimental pictures from Experiment 1, with one

exception: because the picture for TRUNK was frequently misnamed *tree* rather than *trunk* in Experiment 1, Experiment 2 depicted *trunk* as an ELEPHANT TRUNK to reduce the total error rate in this condition. Some Item Set 3 words also acted as experimental items in Item Sets 1 and 2 and the remainder served the role of fillers for those item sets. Overall, there were 66 fillers, an even number required to achieve 50% *yes* and 50% *no* responses in the name-picture matching task. Since Experiment 2 contained 61 experimental words, we therefore added five fillers to the CCVCC condition that we defined as experimental words to enable our randomisation program to rotate between 66 fillers and 66 experimental words, but that were then excluded from the analyses. Randomisation constraints were set to choose 2 experimental words and 2 fillers every 4 trials, giving a maximum of 4 consecutive *yes* or *no* responses on name-picture matching trials.

Results

We first present results for the replication of Experiment 1 in Item Set 3, followed by results for the two partial factorial analyses in Item Set 2. We did not use Item Set 1 for analysing vowel nucleus effects because four words in Item Set 1 (*bull, pier, port,* and *plow*) had very high error rates, and long latencies that were approximately two standard deviations or more above the mean for their conditions. No items in Item Set 2 behaved this way, and Item Set 2 simplified the comparison of effect sizes and the assessment of onset × vowel nucleus complexity and onset × coda complexity interactions.

Replication of Experiment 1

Table 2 shows mean vocal latencies and errors for the replications of the delayed and standard naming tasks of Experiment 1. They were analysed by 2 (onset complexity) × 2 (number of syllables) ANOVAs.

Block 1: Delayed Naming. Outliers comprised 2.55% of the data in the delayed naming task and were discarded. Mean latencies showed an unexpected main effect of onset complexity by subjects [$F_1(1,29) = 8.97$, MSe = 964.45, $P < 0.01$], but not by items [$F_2(1,52) = 2.54$, MSe = 1237.04, $P = 0.11$]. No other effects or interactions were reliable at $P < 0.15$.

Errors comprised 0.77% of experimental trials in the delayed naming task, and error means showed no main effects, but an onset complexity × number of syllables interaction by subjects [$F_1(1,29) = 4.16$, MSe = 0.001, $P = 0.05$], and marginally by items [$F_2(1,52) = 3.3$, MSe = 0.001, $P = 0.07$].

Block 2: Standard Naming. Outliers comprised 1.72% of the standard naming data. Analyses of mean latencies indicated a main effect of onset complexity by subjects [$F_1(1,29) = 56.52$, MSe $= 3937.23$, $P < 0.01$], and marginally by items [$F_2(1,52) = 3.8$, MSe $= 20050.68$, $P = 0.057$]. There was also a main effect of number of syllables by subjects [$F_1(1,29) = 31.97$, MSe $= 3347.68$, $P < 0.01$], but not by items [$F_2(1,52) = 1.71$, MSe $= 20050.68$, $P > 0.15$]. There was no onset complexity × number of syllables interaction [$F_1(1,29) = 2.23$, MSe $= 3363.23$, $P = 0.14$; $F_2 < 1$].

Errors comprised 15.29% of standard naming trials, and ANOVAs on these data indicated a main effect of onset complexity by subjects [$F_1(1,29) = 6.06$, MSe $= 0.011$, $P = 0.02$], but no other reliable effects.

Analyses of Onset and Vowel Nucleus Complexity (Item Set 2 only)

Mean latencies and errors are shown in Table 3, and were analysed through 2 (onset complexity) × 2 (vowel nucleus complexity) ANOVAs.

Block 1: Delayed Naming. Outliers comprised 3.06% of the delayed naming data. Mean latencies showed no main effect of onset or vowel nucleus complexity, and no interaction [$F_1(1,29) = 2.73$, MSe $= 1427.14$, $P = 0.1$] (all other Fs < 1).

Errors comprised 0.69% of delayed naming trials, and were unaffected by the experimental factors.

Block 2: Standard Naming. Outliers comprised 0.97% of the standard naming data. Mean latencies showed a main effect of onset complexity by

TABLE 3

Mean Latencies (in ms) and Number and Percentages of Errors as a Function of Onset Complexity and Vowel Nucleus Complexity in the Three Tasks of Experiment 2: Delayed and Standard Picture Naming and Name-Picture Matching. Corrected Naming Latencies are the Latencies for Standard Naming Minus Picture-Name Matching

Task	Onset complexity	Vowel Nucleus Complexity			
		Latency		Errors	
		Simple	Complex	Simple	Complex
Delayed	Simple	382	396	1 (0.55%)	0 (0%)
naming	Complex	396	386	3 (1.66%)	1 (0.55%)
Standard	Simple	756	781	9 (5.00%)	24 (13.33%)
naming	Complex	846	873	23 (12.77%)	43 (23.88%)
Name-picture	Simple	532	570	4 (2.22%)	5 (2.77%)
matching	Complex	549	586	3 (1.66%)	6 (3.33%)

subjects $[F_1(1,29) = 21.57, \text{MSe} = 11524.37, P < 0.01]$, and marginally by items $[F_2(1,20) = 4.01, \text{MSe} = 14326.39, P = 0.059]$. The main effect of vowel nucleus complexity was marginally reliable by subjects $[F_1(1,29) = 3.43, \text{MSe} = 5983.66, P = 0.07]$, and there were no other reliable effects or interactions (all $Fs < 1$).

Errors comprised 13.75% of the standard naming trials. Error means were affected by onset complexity by subjects $[F_1(1,29) = 11.86, \text{MSe} = 0.02, P < 0.01]$, and marginally by items $[F_2(1,20) = 3.11, \text{MSe} = 0.016, P = 0.09]$. There was also a main effect of vowel nucleus complexity by subjects $[F_1(1,29) = 18.45, \text{MSe} = 0.015, P < 0.01]$, and marginally by items $[F_2(1,20) = 3.5, \text{MSe} = 0.016, P = 0.07]$. There was no onset × vowel nucleus complexity interaction.

Block 3: Name-picture Matching. Outliers comprised 0.13% of the name-picture matching data. Analyses of the latency data indicated a main effect of vowel nucleus complexity $[F_1(1,29) = 15.29, \text{MSe} = 2735.12, P < 0.01; F_2(1,20) = 4.27, \text{MSe} = 1883.06, P = 0.052]$, but no main effects of onset complexity, and no interactions. No effects were reliable in the error data (2.5%).

Corrected Latencies. Figure 4 shows the mean corrected latencies, which showed a main effect of onset complexity by subjects $[F_1(1,29) = 9.91, \text{MSe} = 17047.25, P < 0.01]$, and marginally by items $[F_2(1,20) = 3.4, \text{MSe} = 11698.98, P = 0.08]$. No other effects or interactions reached the 0.15 probability level.

Analysis of Onset and Coda Complexity Effects (Item Set 2 only)

Results are shown in Table 4, and were analysed through 2 (onset complexity) × 2 (coda complexity) ANOVAs.

Block 1: Delayed Naming. Outliers comprised 2.63% of the delayed naming data. There were no effects on mean latencies [onset complexity: $F_1(1,29) = 2.25, \text{MSe} = 2119.39, P = 0.14$, with all other P values > 0.15].

Errors comprised 0.55% of experimental trials in the delayed naming task, and showed a weak non-significant effect of coda complexity by items $[F_2(1,20) = 3.07, \text{MSe} = 0.001, P = 0.09]$ with all other $Fs < 1$.

Block 2: Standard Naming. Outliers comprised 1.66% of the standard naming data. Mean latencies showed a main effect of onset complexity by subjects $[F_1(1,29) = 30.03, \text{MSe} = 8044.34, P < 0.01]$, and marginally by items $[F_2(1,20) = 3.71, \text{MSe} = 17621.3, P = 0.06]$. There was also a main
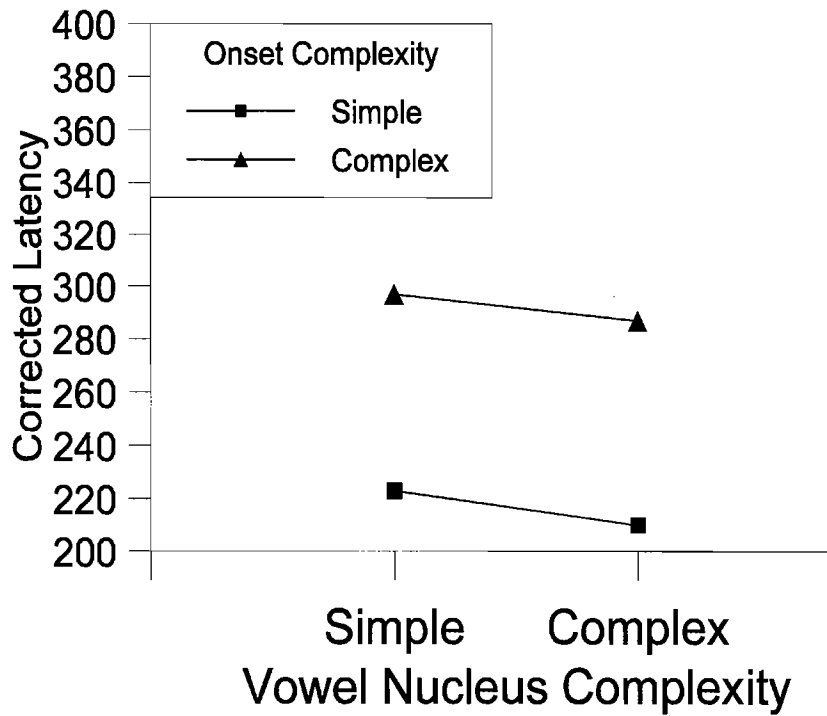
Figure 4. Mean corrected naming latency (by participants) as a function of onset and vowel nucleus complexity in Experiment 2.

TABLE 4
Mean Latencies (in ms) and Number and Percentage of Errors as a Function of
Onset Complexity and Coda Complexity in the Three Conditions of Experiment 2:
Delayed and Standard Picture Naming, and Name-Picture Matching.
Corrected Naming Latencies are the Latencies for Standard Minus Picture-Name
Matching

| Task | Onset complexity | Coda Complexity | | | |
| | | Latency | | Errors | |
		Simple	Complex	Simple	Complex
Delayed	Simple	382	376	1 (0.55%)	0 (0%)
naming	Complex	396	388	3 (1.66%)	0 (0.0%)
Standard	Simple	756	848	9 (5.00%)	33 (18.33%)
naming	Complex	846	938	23 (12.77%)	41 (22.77%)
Name-picture	Simple	532	582	4 (2.22%)	3 (1.66%)
matching	Complex	549	612	3 (1.66%)	2 (1.11%)

28

effect of coda complexity by subjects [$F_1(1,29) = 29.09$, MSe $= 8753.23$, $P < 0.01$], and marginally by items [$F_2(1,20) = 3.39$, MSe $= 17621.3$, $P = 0.08$]. There was no onset $\times$ coda complexity interaction (both Fs < 1).

Errors comprised 14.72% of experimental trials in the standard naming task, showing no interactions, but a main effect of onset complexity by subjects [$F_1(1,29) = 4.98$, MSe $= 0.022$, $P = 0.03$], but not by items [$F_2(1,20) = 1.76$, MSe $= 0.013$, $P > 0.15$]. There was also a main effect of coda complexity by subjects [$F_1(1,29) = 19.55$, MSe $= 0.021$, $P < 0.01$], and by items [$F_2(1,20) = 6.42$, MSe $= 0.013$, $P = 0.02$].

Block 3: Name-picture Matching. Outliers comprised 1.66% of the name-picture matching data. Analyses of the latency data indicated a main effect of coda complexity by subjects [$F_1(1,29) = 19.74$, MSe $= 4887.46$, $P < 0.01$], and by items [$F_2(1,20) = 7.85$, MSe $= 2464.88$, $P = 0.01$]. The main effect of onset complexity was marginally significant by subjects [$F_1(1,29) = 3.97$, MSe $=4039.11$, $P = 0.056$], but not by items [$F_2(1,20) = 1.49$, MSe $= 2464.88$, $P > 0.15$]. There was no onset $\times$ coda complexity interaction (both Fs < 1). Errors comprised 1.66% of the name-picture matching trials, and showed no main effects or interactions (all Fs < 1).

Corrected Latencies. Figure 5 shows mean corrected latencies. These data showed a main effect of onset complexity by subjects [$F_1(1,29) = 15.36$, MSe $= 8656.92$, $P < 0.01$], but not by items [$F_2(1,20) = 2.62$, MSe $= 14569.81$, $P = 0.12$]. Coda complexity exhibited a non-significant trend in corrected means [$F_1(1,29) = 2.2$, MSe $= 17093.04$, $P = 0.14$; $F_2 < 1$], with no other effects or interactions (all Fs < 1).

Discussion

Corrected latencies in Experiment 2, in general, supported all five NST Predictions. The analysis of Item Set 3 in standard naming resembled Experiment 1: main effects were comparable in size to Experiment 1 for onset complexity (86 ms versus 70 ms) and for number of syllables (60 ms versus 61 ms), and can be considered a successful replication of Experiment 1 despite the marginal significance of the by-items analyses in Experiment 2. Consistent with NST Prediction 3, the number of syllables $\times$ onset complexity interaction in the standard naming latencies of Experiment 1 disappeared in Experiment 2, in part because changing TREE TRUNK to ELEPHANT TRUNK reduced errors for monosyllabic words with complex onsets. In this regard, it is of interest that Dupoux (pers. comm., 1997) was able to replicate the present number of syllables and onset complexity effects in a subsequent standard naming task using the same stimuli as in Experiment 1, when participants were given
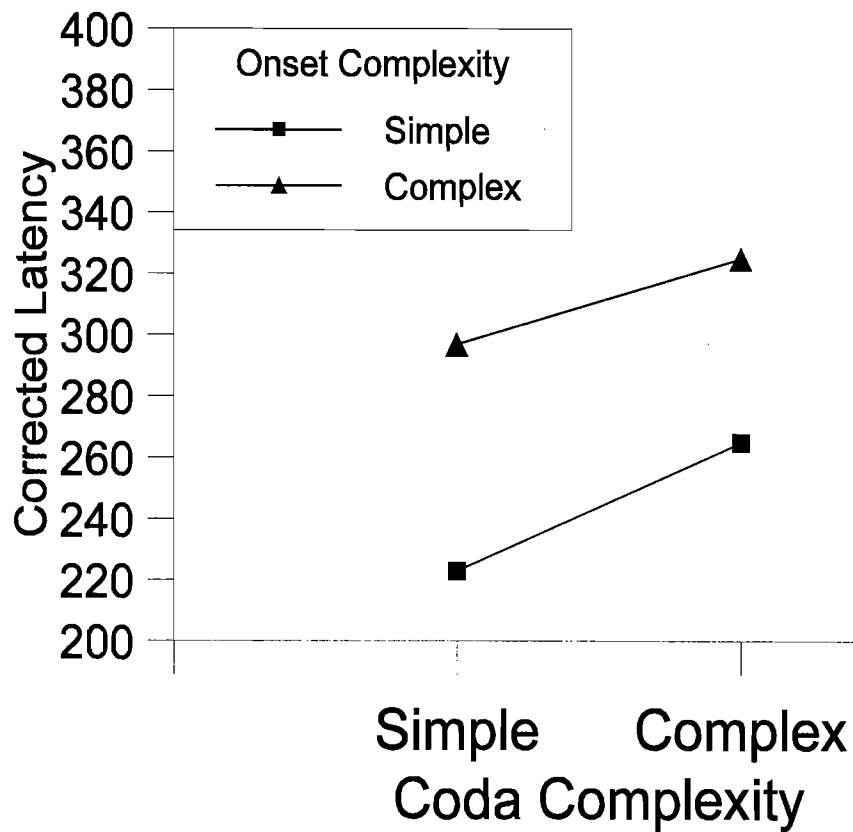
Figure 5.   Mean corrected naming latency (by participants) as a function of onset and coda complexity in Experiment 2.

additional training blocks in an initial delayed naming task. Moreover, there was no number of syllables × onset complexity interaction in the Dupoux data, suggesting that added training can override unwanted picture complexity effects. The extra training virtually eliminated errors in the standard naming task and both main effects but no interactions were observed, which conclusively rules out explanations of standard picture naming latencies in terms of number of errors.

Results for Item Set 2 also indicated an onset complexity effect that supported Prediction 2 of NST. Onset complexity had a clear effect on both standard naming latency (about 90 ms) and corrected latency (about 71 ms) with conditions CVC, CVV(C), CVCC, CCVC, CCVV(C) and CCVCC treated as two partial factorial designs that crossed onset complexity with vowel nucleus complexity, and onset complexity with

coda complexity. Consistent with NST Prediction 4, vowel nucleus complexity affected neither standard naming latency nor corrected latency, and although coda complexity reliably affected standard naming latencies, this effect was non-significant in the corrected measure. Moreover, onset complexity did not interact with either vowel nucleus or coda complexity, consistent with NST Prediction 5. Finally, the overall pattern of corrected latencies did not resemble delayed naming latencies in Experiment 2, undermining a phonological matching account of the name-picture matching task.

Despite these fits with NST predictions under a conceptual matching account, several aspects of Experiment 2 warrant further discussion. One is the nature of the coda clusters in Experiment 2. Both speech errors (MacKay, 1978; Treiman & Danis, 1988; Treiman, 1995; Stemberger, 1983; Shattuck-Hufnagel, 1986) and metalinguistic tasks (Treiman, 1984) indicate that liquids (/r/ and /l/) are sometimes grouped not with the subsequent consonant in a coda cluster, but with the preceding vowel, albeit less cohesively than is the case of diphthongs or complex vowels. This means that liquids in four of the six words with CVCC codas in Item Set 2 (see Appendix B) may group with the vowel rather than the coda, a possible basis for the small effect of coda complexity in Experiment 2. Contrary to this hypothesis, however, coda clusters in condition CCVCC contained no liquids, but the effect of coda complexity for CCVCC words (92 ms in standard naming, 28 ms in corrected latency) was equivalent or smaller than for CVCC words (92 ms in standard naming, 42 ms in corrected latency).

Other aspects of Experiment 2 that warrant discussion concern the Item Set 3 replication of Experiment 1, specifically, the effect of onset complexity on delayed naming (unlike Experiment 1) and the non-effects of number of syllables on delayed naming (unlike the effect of number of syllables on errors in Experiment 1). These differing outcomes in Experiments 1 and 2 suggest that delayed naming effects are unstable and without consequence for interpreting present results.

GENERAL DISCUSSION

Results of the present experiments demonstrate effects of onset complexity and number of syllables on vocal latency in picture naming. However, we found no reliable effects of vowel nucleus or coda complexity on corrected latencies. Finally, onset complexity always exhibited additive effects (see Figures 3–5) with number of syllables, coda complexity, or vowel nucleus complexity.

These results are consistent with NST Predictions 1–5 and the mechanisms that gave rise to those predictions. That is, word and syllable

onset complexity had large effects on vocal latency in our data because both variables increase the number of time-consuming sequential decisions that are required in order to activate the left branch nodes in hierarchically organised phonological structures prior to the onset of overt pronunciation.

Our effects of onset complexity on vocal latency are also consistent with evidence from speech errors (Stemberger & Treiman, 1986) and metalinguistic tasks (Treiman, 1989) that initial consonant clusters operate as units at some level in a hierarchy of phonological retrieval processes. At the same time, however, present results go beyond the speech error and metalinguistic data by providing evidence for consonant clusters in an on-line speech production task. Our number-of-syllables effect also reinforces the surprisingly weak prior evidence for output syllables (see the Introduction), and suggests that the syllable is a psychologically real unit in language production (see also MacKay, 1987).

The null or small and non-significant effects of vowel nucleus and coda complexity on our corrected latency measures are also consistent with NST: Coda and vowel nucleus complexity do not affect the main phonological determinant of latency under NST (the number of sequential decisions required before activation of the left-most, bottom-most node in hierarchical phonological structures), but only affect the complexity of central and right branches of the tree. The lack of interaction between onset and vowel nucleus complexity, and between onset and coda complexity further suggests that onset clusters add to the complexity of the syllable tree in a manner independent of coda and vowel nucleus complexity, consistent with the top-down, left-to-right tree traversal activation process in NST.

This account of the differing effects of structural complexity for initial versus non-initial parts of a word supports the claim that no articulatory buffering of a whole phonological word occurs prior to the start of pronunciation (see also Kawamoto, Kello, Jones & Bame, 1998, for a review of priming and reading task results that comport with the NST framework). This account also provides an interesting explanation of some otherwise puzzling results in the production literature. Experiment 4 in Levelt and Wheeldon (1994) compared latencies in a symbol naming task for two classes of bisyllabic words that had identical initial syllables. However, one class had CV-CVC structure, and was shorter in length and structural complexity than the other class, which had CV-CCCVVC structure and was reliably longer in acoustic duration. The results showed that latencies for these two classes of words were virtually identical (681 ms versus 678 ms), which might seem surprising given that consonant clusters in initial syllables of monosyllabic and bisyllabic words strongly affect vocal latency. However, this null effect is consistent with NST

because a consonant cluster in the second syllable of a bisyllabic word does not affect its left branch complexity, and should not affect latency.

Experiment 2 in Santiago (1997) provides an interesting contrast with Levelt and Wheeldon (1994, exp. 4) that further supports the NST account. This experiment manipulated complexity of the first syllable of bisyllabic Spanish words with CVC-CV versus CCV-CV structure, while keeping structure of the second syllable constant (CV), in a reading task. In the results, latencies were longer for CCV-CV than CVC-CV words, consistent with the NST claim that first syllable complexity affects latency by increasing left branch complexity, whereas second syllable complexity neither increases left branch complexity nor affects latency. Of course, these results can only be viewed as suggestive regarding production processes because reading tasks also involve perceptual processes.

Further research is needed to address two types of power problem in the present research. One concerned the unreliability of some of our by-items measures. For example, number of syllables exhibited reliable by-subjects and by-items effects in corrected latency measures of Experiment 1, but onset complexity only exhibited a reliable by-subjects effect for corrected latency measures. However, three factors suggest that this unreliable by-items effect reflects the relatively small number of items per condition (e.g., 14 in Experiment 1, 6 in Experiment 2). One was that by-subjects analyses had greater sensitivity across-the-board than by-items analyses in Experiments 1 and 2: Never did item analyses give significant results when subject analyses did not. Second, when effects were highly reliable by subjects, item effects usually were at least marginally significant, i.e., with P values in the 0.05 to 0.10 range. Third, mean power computed using Cohen (1988) was much higher for by-subject than by-item analyses.[4] In Experiment 1, mean power to detect our onset complexity effect was 0.35 in our by-subject analysis, but only 0.23 in our by-item analysis. For our number of syllables effect in Experiment 1, mean power was 0.97 by subjects, and 0.52 by items. However, this by-items power issue does not seem serious enough to contradict our onset complexity effect, which was quite sizeable both in Experiment 1 (45 ms) and Experiment 2 (67 and 75 ms in the two partial factorials). We also replicated the onset complexity effect for different participants and partially different items

---

[4] To compare power for our by-subject and by-item analyses, we treated our main effects in the corrected latency data of Experiments 1 and 2 as paired t-tests and computed the mean power for the two tests. For instance, we broke down the onset complexity effect in Experiment 1 into two comparisons: monosyllables with simple onset versus monosyllables with complex onset, and bisyllables with simple onset versus bisyllables with complex onset. This procedure provided a correction for the smaller error variance in our by-subject analyses (which always involved a within subject design) than in our by-items analyses (which involved a between items design).

in Experiments 1 and 2, and this effect has been replicated in another lab (Dupoux, pers. comm., 1997), and in another language (Spanish) using a syllable reading task (Santiago, 1997; Santiago, Palma, & Gutiérrez, in press).

The second statistical power issue seems more serious and calls for caution in interpreting the non-effects predicted under NST and observed in Experiment 2 for vowel and coda complexity. For example, even though the effect of coda complexity was statistically unreliable both by-items and by-subjects, it was nonetheless sizeable (35 ms) for the corrected latency measure, suggesting that a more powerful design might yield a significant effect. Consistent with this hypothesis, mean power to detect the effect of coda complexity was only 0.28 by subjects and 0.12 by items (Cohen, 1988). Further research is therefore needed to determine whether present coda complexity results can be satisfactorily replicated, and to determine whether complexity of mid- and right-most branches delay initiation times to some extent, albeit less than word and onset complexity.

In what follows, we examine two alternative hypotheses for explaining present results: The Length Hypothesis (that the onset complexity effect is just a result of greater number of phonemes), and the Shape Frequency Hypothesis (that the relevant factor is the frequency of the abstract phonological frame of syllables with complex onsets).

Levelt (1992) and Roelofs (1997a) proposed that phonological units might be retrieved in parallel and associated in a sequential left-to-right manner to the word's structural frame. After segment-to-frame association, phonetic syllable programs are accessed and stored in an output buffer as a linear string. Only when the buffer contains one or more phonological words can articulation begin. Consequently, the more segments a word contains, the more time is needed before production onset. Under this Length Hypothesis (LH), how segments are structured in words of identical length should have no effect on latencies. For example, the LH predicts that consonant clusters at the beginning of a word should delay vocal latency to the same extent as consonant clusters at non-initial parts of the word, contrary to present data. Also problematic for the LH, is our Experiment 2 finding that longer words, i.e., complex vowel nucleus words, had shorter mean latencies than shorter words, i.e., simple vowel nucleus words. This finding is also difficult to explain under the hypothesis that diphthongs function as single segments (Stemberger, 1983, 1984; Levelt, 1989; Shattuck-Hufnagel, 1986) because diphthongs clearly function as two separate units under some circumstances (see MacKay, 1978). Additional analyses of the present results further contradicted predictions of the LH: In Experiment 2, CVVC and CVCC words were longer than CVC words, but had combined latencies that did not differ from CVC words, $F < 1$. Similarly, CCVVC and CCVCC words in

Experiment 2 were longer than CCVC words, but had combined latencies that did not differ from CCVC words, $F < 1$. The results of Santiago (1997; Santiago et al., in press) also contradict this LH prediction for Spanish participants reading CVC versus CCV syllables (e.g., *tar* versus *tra*) matched for length in phonemes, for phonemic and orthographic content, for syllable frequency in printed Spanish, and for spelling-to-sound regularity. As predicted under NST, reading onset times were longer for CCV than CVC syllables despite their equivalent length. Santiago (1997) also replicated this onset complexity result for bisyllabic Spanish words with CVC-CV and CCV-CV structures that did not differ in length, further contradicting the LH. Again, reading data may be argued to be irrelevant to this discussion, but they are suggestive.

Another clear prediction of the LH is that latencies will vary with word length in segments, regardless of how the segments are structured in syllables of the word. One key study that failed to find this predicted length effect is Levelt and Wheeldon's (1994) Experiment 4, which manipulated second syllable duration and complexity of bisyllabic CV-CVC versus CV-CCCVVC words while keeping constant the structure and content of the first (CV) syllable. Although the two classes of words differed greatly in acoustic duration and number of segments, latency remained constant in their data, contradicting the LH. Also contradicting the LH, Bachoud-Levy et al. (1998) found no effect of length *per se* on standard picture and symbol naming in French and English.

Present effects of onset complexity are consistent with NST, and are troublesome for production theories that do not propose special processing difficulty for onset clusters (Dell, 1986; Dell, Juliano & Govindjee, 1993; Eikemeyer & Schade, 1991; Hartley & Houghton, 1996; Levelt, 1992; Levelt & Wheeldon, 1994; Levelt et al., in press; Roelofs, 1997a, 1997b; Shattuck-Hufnagel, 1979). However, Dell's (1988) revised model suggests a somewhat different account of our onset and syllabic complexity effects that we call the Shape Frequency hypothesis. In Dell's theory, word shape nodes function like sequence nodes but are connected to one another in a chain-like manner. For example, the word *truck* has a word shape node representing its overall CCVC structure and this CCVC shape node is connected to a CC shape node representing the initial consonant cluster of the word, which in turn connects to a V node representing its vowel nucleus. This V node then connects to a C shape node representing its coda, which finally connects to an END node. Word shape nodes compete with one another for activation in a manner that allows word shape and syllable shape frequency to influence the rate and probability of activating their corresponding content nodes. Consequently, word shape frequency may contribute to our number of syllables effect because monosyllabic word shapes, e.g., CCV for *true*, are more frequent than bisyllabic word

shapes, e.g., CCV-CV for *truly*. Our onset complexity effect could likewise reflect syllable shape frequency because syllable structures starting with a consonant cluster, e.g., CCV for the syllable *true*, are less frequent than ones starting with a singleton, e.g., CVC for the syllable *tour*, both in English (Schiller, pers. comm., 1997), Dutch (Schiller et al., 1996) and Spanish (Justicia, Santiago, Palma, Huertas & Gutiérrez, 1996). However, our non-significant effects of coda and vowel nucleus complexity are difficult to explain in terms of shape frequency.

CONCLUSIONS

When people name depicted objects, structural complexity of the phonological representation of the name affects vocal latency if other factors are controlled. Both consonant clusters in syllable onset position and number of syllables in the word delay initiation of the naming response. In contrast, complex vowel nuclei do not affect production latency, while coda clusters show a small effect if any. These results fit NST and perhaps other theories as well, but pose problems for models that do not contain syllable and consonant cluster units, that do not assume hierarchic organisation within syllables, or assume that an entire phonological word must be buffered before onset of pronunciation. Present results therefore help constrain the space of theoretical alternatives in the study of phonological retrieval processes.

REFERENCES

Bachoud-Levy, A-C., Dupoux, E., Cohen, L., & Mehler, J. (1998). Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory and Language, 39*, 331–346.

Baumann, M. (1995). *The production of syllables in connected speech.* Unpublished doctoral dissertation, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

Berg, T. (1991). Phonological processing in a syllable-timed language with pre-final stress: Evidence from Spanish error data. *Language and Cognitive Processes, 6*, 265–301.

Boomer, D.S., & Laver, J.D.M. (1968). Slips of the tongue. *British Journal of Disorders of Communication, 3*, 2–12.

Brown, A.S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin, 109*, 204–223.

Burke, D., MacKay, D.G., Worthley, J.S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language, 30*, 237–246.

Carreiras, M., Alvarez, C. & de Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language, 32*, 766–780.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Cohen, J.D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments & Computers, 25,* 257–271.

Collard, R., & Povel, D.-J. (1982). Theory of serial pattern production: Tree traversals. *Psychological Review, 89,* 693–707.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993) Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 1000,* 589–608.

Corbell, J.-C. (1986). *The facts on file visual dictionary.* New York: Facts on File.

Costa, A., & Sebastián, N. (1996). La codificación fonológica en la producción del lenguaje: Adecuación del paradigma de facilitación implícita para el estudio de efectos silábicos. [Phonological encoding in language production: Use of the implicit priming paradigm to study syllabic effects]. *Cognitiva, 8,* 25–42.

Costa, A., & Sebastián, N. (1998). Abstract phonological structure in language production: Evidence from Spanish. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24,* 886–903.

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

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

Dell, G.S., Juliano, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science, 17,* 149–195.

Eikemeyer, H.-J. & Schade, U. (1991). Sequentialization in connectionist language-production models. *Cognitive Systems, 3,* 128–138.

Ferrand, L., & Segui, J. (1998). The syllable's role in speech production: Are syllables chunks, schemas, or both? *Psychonomic Bulletin & Review, 5,* 253–258.

Ferrand, L., Segui, J., & Grainger, J. (1996). Masked priming of word and picture naming: The role of syllabic units. *Journal of Memory and Language, 35,* 708–723.

Francis, N., Kucera, H., & Mackie, A.W. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Fromkin, V. (1971). The non-anomalous nature of anomalous utterances. *Language, 47,* 27–52.

Fudge, E.C. (1969). Syllables. *Journal of Linguistics, 5,* 193–320.

Fudge, E.C. (1987). Branching structure within the syllable. *Journal of Linguistics, 23,* 359–377.

García-Albea, J.E., del Viso, S., & Igoa, J.M. (1989). Movement errors and levels of processing in sentence production. *Journal of Psycholinguistic Research, 18,* 145–161.

Garrett, M.F. (1975). The analysis of sentence production. In G.H. Bower (Ed.), *The psychology of learning and memory,* Vol. 9 (pp. 133–177). New York: Academic Press.

Glaser, W.R. (1992). Picture naming. *Cognition, 42,* 61–105.

Gordon, P.C., & Meyer, D.E. (1987). Control of serial order in rapidly spoken syllable sequences. *Journal of Memory and Language, 26,* 300–321.

Greeno, J.G., & Simon, H.A. (1974). Processes for sequence production. *Psychological Review, 81,* 187–197.

Harris, J.W. (1983). *Syllable structure and stress in Spanish: A nonlinear analysis.* Cambridge, MA: MIT Press.

Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language, 35*, 1–31.

Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry, 20*, 253–306.

Hendricks, H., & McQueen, J. (1996). *Annual Report 1995.* Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

Hooper, J.B. (1972). The syllable in phonological theory. *Language, 48*, 525–540.

Houghton, G., & Hartley, T. (1995). Parallel models of serial behavior: Lashley revisited. *Psyche, 2*(25) [on-line journal].

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

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

Jordan, M.I., & Rosenbaum, D.A. (1989). Action. In M.I. Posner (Ed.), *Foundations of Cognitive Science* (pp. 727–767). Cambridge, MA: MIT Press.

Justicia, F., Santiago, J., Palma, A., Huertas, D., & Gutiérrez, N. (1996). La frecuencia silábica del español escrito por niños: Estudio estadístico. [Syllable frequency: A statistical study of written productions by Spanish children]. *Cognitiva, 8*, 131–168.

Kawamoto, A.H., Kello, C.T., Jones, R., & Bame, K. (1998). Initial phoneme versus whole-word criterion to initiate pronunciation: Evidence based on response latency and initial phoneme duration. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*, 862–885.

Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language, 37*, 295–311.

Klapp, S.T. (1995). Motor response programming during simple and choice reaction time: The role of practice. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1015–1027.

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

Lashley, K.S. (1917). General reviews and summaries: Sensory psychology of animals. *Psychological Bulletin, 14*, 276–283.

Lashley, K. (1951). The problem of serial order in behavior. In L.P. Jeffress (Ed.), *Cerebral Mechanisms in Behavior: The Hixon Symposium* (pp. 112–146). New York: J. Wiley & Sons.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

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

Levelt, W.J.M., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1–73.

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

MacKay, D.G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia, 8*, 323–350.

MacKay, D.G. (1972). The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology, 3*, 210–227.

MacKay, D.G. (1974). Aspects of the syntax of behavior: Syllable structure and speech rate. *Quarterly Journal of Experimental Psychology, 26*, 642–657.

MacKay, D.G. (1978). Speech errors inside the syllable. In A. Bell & J.B. Hooper (Eds), *Syllables and segments* (pp. 201–212). Amsterdam: North-Holland.

MacKay, D.G. (1981). The problem of rehearsal or mental practice. *Journal of Motor Behavior, 13*, 274–285.

MacKay, D.G. (1987). *The organization of perception and action.* New York: Springer-Verlag.

McRae, K., Jared, D., & Seidenberg, M. (1990). On the roles of frequency and lexical access in word naming. *Journal of Memory and Language, 29*, 43–65.

Mehler, J., Dommergues, J.Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior, 220*, 298–305.

Meijer, P.J.A. (1994). *Phonological encoding: The role of suprasegmental structures.* Unpublished doctoral dissertation, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

Meijer, P.J.A. (1996). Suprasegmental structures in phonological encoding: The CV structure. *Journal of Memory and Language, 35*, 840–853.

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

Meyer, A. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language, 30*, 69–89.

Meyer, A. (1997). Word form generation in language production. In W. Hulstijn, H.F.M. Peters, & P.H.H.M. Van Lieshout (Eds), *Speech production: Motor control, brain research and fluency disorders* (pp. 73–88). Amsterdam: Elsevier.

Meyer, A., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*, 1146–1160.

Pew, R.W., & Rosenbaum, D.A. (1988). Human movement control: Computation, representation, and implementation. In R.C. Atkinson, R.J. Herrnstein, G. Lindzey, & R.D. Luce (Eds), *Steven's handbook of experimental psychology: Vol. 2. Learning and Cognition* (pp. 473–509). New York: J. Wiley & Sons.

Pierrehumbert, J., & Nair, R. (1995). Word games and syllable structure. *Language and Speech, 38*, 77–114.

Povel, D.-J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 3–18.

Restle, F. (1970). Theory of serial pattern learning: Structural trees. *Psychological Review, 77*, 481–495.

Roelofs, A. (1997a). The WEAVER model of word form encoding in speech production. *Cognition, 64*, 249–284.

Roelofs, A. (1997b). Syllabification in speech production: Evaluation of WEAVER. *Language and Cognitive Processes, 12*, 657–693.

Roelofs, A., & Meyer, A. (1998). Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*, 922–939.

Romani, C. (1992). *The representation of prosodic and syllabic structure in speech production.* Unpublished doctoral dissertation, Johns Hopkins University.

Rosenbaum, D.A. (1985). Motor programming: A review and scheduling theory. In H. Heuer, U. Kleinbeck, & K.-M. Schmidt (Eds), *Motor behavior: Programming, control and acquisition* (pp. 1–33). Berlin: Springer-Verlag.

Rosenbaum, D.A. (1990). On choosing between movement sequences: Comments on Rose (1988). *Journal of Experimental Psychology: Human Perception and Performance, 16*, 439–444.

Rosenbaum, D.A. (1991). *Human motor control.* New York: Academic Press.

Rosenbaum, D.A., Gordon, A.M., Stillings, N.A., & Feinstein, M.H. (1987). Stimulus-response compatibility in the programming of speech. *Memory and Cognition, 15,* 217–224.

Rosenbaum, D.A., Inhoff, A.W., & Gordon, A.M. (1984). Choosing between movement sequences: A hierarchical editor model. *Journal of Experimental Psychology: General, 113,* 372–393.

Rosenbaum, D.A., Kenny, S., & Derr, M.A. (1983). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 86–102.

Rosenbaum, D.A., Weber, R.J., Hazelett, W.M., & Hindorff, V. (1986). The parameter remapping effect in human performance: Evidence from tongue twisters and finger fumblers. *Journal of Memory and Language, 25,* 710–725.

Santiago, J. (1997). *Estructura Silábica y Activación Secuencial en Producción del Lenguaje.* [Syllable Structure and Sequential Activation in Language Production]. Unpublished doctoral dissertation, Dept. Psicología Experimental y Fisiología del Comportamiento, Universidad de Granada, Spain.

Santiago, J., Palma, A., & Gutiérrez, N. (1999). Efectos de complejidad de la cabeza silábica en la lectura de sílabas aisladas [Effects of syllabic onset complexity on the reading of single syllables]. *Cognitiva, 11,* 45–66.

Savage, G.R., Bradley, D.C., & Forster, K.I. (1990). Word frequency and the pronunciation task: The contribution of articulatory fluency. *Language and Cognitive Processes, 5,* 203–236.

Schiller, N. (1997). *The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography.* Unpublished doctoral dissertation, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

Schiller, N. (1998). The effect of visually masked syllable primes on the naming latencies of words and pictures. *Journal of Memory and Language, 39,* 484–507.

Schiller, N., Meyer, A., Baayen, R.H., & Levelt, W.J.M. (1996). A comparison of lexeme and speech syllables in Dutch. *Journal of Quantitative Linguistics, 3,* 8–28.

Schriefers, H., Meyer, A.S., & Levelt, W.J.M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29,* 86–102.

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

Selkirk, E. (1982). The syllable. In H. van der Hulst & N. Smith (Eds), *The structure of phonological representations.* Part II (pp. 337–383). Dordrecht: Foris Publishers.

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

Sevald, C., Dell, G.S., & Cole, J.S. (1995). Syllable structure in speech production: Are syllables chunks or schemas? *Journal of Memory and Language, 34,* 807–820.

Shattuck-Hufnagel, S. (1979). Speech error evidence for a serial ordering mechanism in sentence production. In W.E. Cooper & E.C.T. Walker (Eds), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 295–342). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Shattuck-Hufnagel, S. (1986). The representation of phonological information during speech production planning: Evidence from vowel errors in spontaneous speech. *Phonology Yearbook, 3,* 117–149.

Shattuck-Hufnagel, S. (1987). The role of word-onset consonants in speech production planning: New evidence from speech error patterns. In E. Keller & M. Gopnik (Eds),

Motor and sensory processes in language (pp.17–51). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

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

Shattuck-Hufnagel, S. & Klatt, D. (1979). Minimal use of features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior, 18,* 41–55.

Simon, H.A. (1972). Complexity and the representation of patterned sequences of symbols. *Psychological Review, 79,* 369–382.

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

Stemberger, J.P. (1983). The nature of /r/ and /l/ in English: Evidence from speech errors. *Journal of Phonetics, 11,* 139–147.

Stemberger, J.P. (1984). Length as a suprasegmental: Evidence from speech errors. *Language, 60,* 895–913.

Stemberger, J.P. (1985). An interactive activation model of language production. In A.W. Ellis (Ed.), *Progress in the psychology of language: Vol. 3* (pp. 143–186). Hove: Lawrence Erlbaum Associates Ltd.

Stemberger, J.P. (1990). Wordshape errors in language production. *Cognition, 35,* 123–157.

Stemberger, J.P. & Treiman, R. (1986). The internal structure of word-initial consonant clusters. *Journal of Memory and Language, 25,* 163–180.

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

Theios, J., & Armhein, P.C. (1989). Theoretical analysis of the cognitive processing of lexical and pictorial stimuli: Reading, naming, and visual and conceptual comparisons. *Psychological Review, 96,* 5–24.

Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition, 15,* 49–74.

Treiman, R. (1984). On the status of final consonant clusters in English syllables. *Journal of Verbal Learning and Verbal Behavior, 23,* 343–356.

Treiman, R. (1986). The division between onsets and rimes in English syllables. *Journal of Memory and Language, 25,* 476–491.

Treiman, R. (1989). The internal structure of the syllable. In G. Carlson & M.K. Tanenhaus (Eds), *Linguistic structure in language processing* (pp. 27–52). Dordrecht: Kluwer Academic Publishers.

Treiman, R. (1995). Errors in short-term memory for speech: A developmental study. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21,* 1197–1208.

Treiman, R., & Danis, C. (1988). Short-term memory errors for spoken syllables are affected by the linguistic structure of the syllables. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14,* 145–152.

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

Yaniv, I., Meyer, D.E., Gordon, P.E., Huff, C.A., & Sevald, C.A. (1990). Vowel similarity, connectionist models, and syllable structure in motor programming of speech. *Journal of Memory and Language, 29,* 1–26.

Yngve, V.H. (1960). A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society, 104,* 444–466.

APPENDIX A

Experimental stimuli in Experiment 1 with frequency of use in parentheses (from Francis et al., 1982), and their corrected latencies (rounded to the nearest ms). Rows contain stimuli with identical initial segments

Monosyllables				Bisyllables			
Simple onset		Complex onset		Simple onset		Complex onset	
bulb	(10) 198	bread	(41) 209	bullet	(49) 387	breakfast	(55) 317
bull	(16) 399	brook	(3) 261	bubble	(25) 299	blackboard	(2) 446
bell	(36) 172	bride	(40) 361	basket	(19) 300	blanket	(39) 412
can	(12) 350	cliff	(11) 321	castle	(12) 228	cradle	(8) 274
cap	(22) 129	clock	(28) 160	candle	(23) 201	classroom	(23) 338
duck	(6) 281	drill	(21) 232	demon	(17) 510	dragon	(3) 298
fish	(33) 116	flag	(18) 185	feather	(19) 308	flashlight	(8) 246
fork	(20) 135	flock	(11) 367	fountain	(22) 280	flautist	(8) 451
ghost	(16) 217	grill	(11) 433	garlic	(4) 520	grenade	(9) 453
pill	(23) 298	plug	(23) 200	package	(25) 391	planet	(44) 355
pig	(14) 265	prop	(8) 326	peacock	(6) 436	printer	(4) 451
tie	(27) 226	trap	(27) 547	tiger	(9) 156	trigger	(11) 273
torch	(4) 367	trunk	(13) 186	tunnel	(12) 352	tractor	(31) 332
tusk	(3) 366	trash	(2) 301	turtle	(9) 165	treasure	(10) 327
Mean	(17.2)		(18.35)		(17.9)		(18.2)

APPENDIX B

Experimental stimuli in Experiment 2 with frequency of use (from Francis et al., 1982). Rows contain stimuli with identical initial segments. Item Set 1 is framed with a single line, Item Set 2 is framed with a double line. See text for description of syllable structure categories and item sets

CVC	Freq.	CVV(C)	Freq.	CVCC	Freq.	CCVC	Freq.	CCVV(C)	Freq.	CCVCC	Freq.
bat	18	bow	13	bulb	10	bread	41	blade	26		
bean	13	bear	24	bolt	9	brook	3	blouse	2		
bull	16	bowl	26	bomb	68	drill	21	bride	40		
fish	33	foam	59	fork	20	flag	18	fly	23	flask	5
pen	23	pie	19	pond	32	prop	8	plate	44	priest	33
pig	14	pear	8	pearl	9	plum	2	prow	1	prince	40
tea	29	tie	27	tank	30	truck	80	train	86	trunk	13
can	12	cage	11	cart	9	clip	8	crow	2	clamp	6
cap	22	cake	16	cork	10	cliff	11	crane	2	crank	1
cup	58	coat	52	colt	21	clock	28	crown	19		
pill	23	pier	18	port	24	plug	23	plow	12		
Mean set 1	23.27		24.81		22		21.09		23.36		
Mean set 2	21.3		23.3		18.3		21.16		26.33		16.33

APPENDIX C

Experimental stimuli in Experiment 2, item set 2, with their corrected latencies (rounded to the nearest ms)

CVC	RT	CVV(C)	RT	CVCC	RT	CCVC	RT	CCVV(C)	RT	CCVCC	RT
fish	107	foam	393	fork	58	flag	94	fly	307	flask	329
pen	196	pie	191	pond	423	prop	492	plate	273	priest	356
pig	230	pear	162	pearl	302	plum	341	prow	347	prince	427
tea	404	tie	125	tank	224	truck	207	train	160	trunk	166
can	260	cage	253	cart	382	clip	384	crow	325	clamp	433
cap	151	cake	266	cork	314	cliff	404	crane	380	crank	373