

Investigation of phonological encoding through speech error analyses: Achievements, limitations, and alternatives*

Antje S. Meyer

Max-Planck-Institute für Psycholinguistik, Postbus 310, NL 6500 AH Nijmegen, Netherlands; and Catholic University Nijmegen, Comeniuslaan 4, POB 9102, 6500 HC Nijmegen, Netherlands

Abstract

Meyer, A.S., 1992. Investigation of phonological encoding through speech error analyses: Achievements, limitations, and alternatives. *Cognition*, 42: 181–211.

Phonological encoding in language production can be defined as a set of processes generating utterance forms on the basis of semantic and syntactic information. Most evidence about these processes stems from analyses of sound errors. In section 1 of this paper, certain important results of these analyses are reviewed. Two prominent models of phonological encoding, which are mainly based on speech error evidence, are discussed in section 2. In section 3, limitations of speech error analyses are discussed, and it is argued that detailed and comprehensive models of phonological encoding cannot be derived solely on the basis of error analyses. As is argued in section 4, a new research strategy is required. Instead of using the properties of errors to draw inferences about the generation of correct word forms, future research should directly investigate the normal process of phonological encoding.

Introduction

The formulation of an utterance can be broken down into two components, namely the generation of the meaning and syntactic structure of the utterance on the one hand, and the creation of its form on the other hand (see, for example,

*I am very grateful to Andrew Crompton, Aditi Lahiri, Pim Levelt, and two anonymous reviewers for helpful comments on an earlier draft of this paper. In addition, I would especially like to thank Lyn Frazier and Harry van der Hulst for extensive discussions of the issues considered in this paper. Requests for reprints should be sent to A.S. Meyer, Max-Planck-Institute for Psycholinguistics.

Butterworth, 1980, 1989; Fay & Cutler, 1977; Fromkin, 1971; Garrett, 1976, 1982, 1988; Kempen & Huijbers, 1983; Levelt, 1989). The present paper is concerned with the second of these components, which will be called phonological encoding. Phonological encoding is defined as the set of processes creating the form of an utterance on the basis of syntactic and semantic information. It includes the retrieval of stored word forms from the mental lexicon, the determination of the rhythmical structure and intonation contour of the utterance, and the creation of a phonetic representation, which is taken to be the input to the articulatory component.

The available evidence about phonological encoding stems mainly from analyses of sound errors. Naturally, there is considerable variation in the results of analyses performed by different researchers on different corpora and in different languages. However, there are certain core properties of sound errors that have been discovered in many corpora. The discussion of the error evidence in section 1 will focus on these regularities and their interpretation (for more comprehensive reviews of the findings see, for instance, Levelt, 1989, or Shattuck-Hufnagel, 1983, 1987). In section 2, two prominent models of phonological encoding will be described, which are largely based on speech error evidence, namely Shattuck-Hufnagel's scan-copier (Shattuck-Hufnagel, 1979, 1983, 1987) and Dell's spreading activation model (Dell, 1986, 1988).

In section 3, it will be shown that in spite of the large number of error analyses that have been carried out, it has not been possible to characterize in detail which form representations speakers create or how they create them. It is, for instance, still unknown, whether speakers create a single phonological representation for each word or several representations, and what the planning units are out of which these representations are constructed. In addition, contrary to what has often been maintained, sound errors only provide very limited evidence about the process of phonological encoding.

There are a number of methodological reasons for the fact that error analyses have only been partially successful. It is, for instance, difficult to establish exactly how often different types of errors arise because they differ in how likely they are to be detected by listeners and because many errors can be classified in several ways. However, as will be shown, even if these problems are minimized, speech error analyses can still not provide more than a global characterization of phonological encoding.

It appears that in order to gain a better understanding of phonological encoding, a new research strategy should be adopted. So far, most researchers have taken a highly data-driven approach, working up from properties of errors to explanations of the generation of error-free speech. A more fruitful strategy for future research might be to start from a working model of correct phonological encoding, to derive new hypotheses about the generation of correct speech from the model, and to devise methods to test them empirically. How such a working model could be attained and tested is discussed in section 4.

1. Evidence from sound errors

Sound errors are utterances that deviate from the speaker's intention in the placement or identity of one or more phonological segments not corresponding to a complete morpheme of the target utterance (see (1) to (7)). Some sound errors, such as (1) and (2), are most naturally described as exchanges of segments or segment clusters. Examples (3) and (4) can be analysed as incomplete exchanges. In (3), which is a sound anticipation, [l] replaces [r] and is repeated in its target position. In (4), which is a perseveration, [g] appears in its target position and replaces the onset of the following word. In all of these errors, the sounds of words are ordered incorrectly, taking positions that were meant for other sounds. In addition to such ordering errors, there are so-called non-contextual errors, which cannot be explained by reference to the immediate utterance context. Some sound substitutions (such as (5)) and many additions (such as (6)) and deletions (such as (7)) belong to this category.¹

- (1) *heft lemisphere* (left hemisphere)
- (2) *fleaky squoor* (squeaky floor)
- (3) a *leading list* (a reading list)
- (4) *gave the goy* (gave the boy)
- (5) a *transformational rule* (a transformational rule)
- (6) *enjoyding it* (enjoying it)
- (7) *split bain* (split brain)

A large number of analyses of sound errors have been carried out. Two general questions have guided most of these analyses, namely, first, how the speakers' representations of word forms can be characterized and, second, how these representations are retrieved or constructed. In this section, I will review what sound errors reveal concerning these issues.

An important property of sound errors is that they are almost always phonetically well formed (e.g., Boomer & Laver, 1968; Wells, 1951). Errors rarely yield illegal sound sequences, and misplaced sounds are usually phonetically accommodated to their new environment, or the environment is accommodated to the intruding sound following the rules of the language in question (Berg, 1987; Fromkin, 1971, 1973; Garrett, 1976, 1980; Stemberger, 1983a, 1985a). For instance, when an English stop consonant moves from a word-initial to a word-internal position or vice versa, it loses or acquires aspiration, as appropriate for its new environment (Fromkin, 1973). When a syllable-final voiced consonant is replaced by a voiceless one or vice versa, the length of the preceding vowel is adjusted accordingly (Shattuck-Hufnagel, 1985a). To give a final example, when, as in (8), the first segment of a noun changes from vowel to consonant or vice versa, the preceding article is modified accordingly.

¹Unless indicated otherwise, the errors stem from the Appendix to Fromkin (1973).

(8) *a meeting arathon (an eating marathon)*

Linguistic theory distinguishes between a fairly abstract phonological representation of a word and a more detailed phonetic representation (e.g., Browman & Goldstein, 1986; Chomsky & Halle, 1968; Mohanan, 1986). The phonetic well-formedness of sound errors shows that they arise before the phonetic form of the utterance is created and before illegal sound sequences are edited out or changed (but see Stemberger, 1985a). Thus, speakers apparently construct both phonological and phonetic representations of utterances; and most sound errors arise during, and provide evidence about, the creation of the former representation.

It should be noted that not all sound errors are phonetically well formed. Ill-formed sequences regularly occur in tongue-twister experiments (Butterworth & Whittaker, 1980), and also, at least occasionally, in spontaneous speech (Buckingham & Yule, 1987; Fromkin, 1973; Hockett, 1967; Stemberger, 1983a). A possible reason why they appear to be so rare is that listeners often fail to notice violations of phonetic rules (see, for instance, Marslen-Wilson & Welsh, 1978, or Warren, 1970, for experimental results supporting this supposition; see also Cutler, 1981). The occurrence of phonetically ill-formed errors can either be explained by assuming that these errors arise at the phonological level, but that the processes that modify or delete ill-formed sequences have failed to apply, or that they arise during later planning processes.

The mere fact that sound errors regularly occur allows for an important conclusion about the process of phonological encoding. Apparently, word forms are not retrieved from the mental lexicon as fully specified units, but are constructed by selecting and combining certain sublexical units. If they *were* retrieved as single entities having no internal structure, errors involving parts of words could not arise.

The units out of which phonological representations are created can be determined by classifying the error units, that is, the sounds and sound sequences by which errors deviate from intended utterances. The most frequent error units are single segments, accounting for 60–90% of the errors. Another large class of error units, appearing in approximately 10–30% of the errors, are sequences of two adjacent segments, either two consonants, or a vowel and a consonant (see, for instance, Berg, 1985, 1988; Boomer & Laver, 1968; Fromkin, 1971; Nootboom, 1969; Shattuck-Hufnagel, 1983; Shattuck-Hufnagel & Klatt, 1979). Thus, word forms are apparently composed out of phonological segments and possibly certain segment sequences.²

²A number of studies have investigated whether some segments are more likely to be involved in errors than others, and whether certain classes of segments are “weak”, that is, replaced more often than replacing others, while other classes are “strong”, that is, replace other segments more often than being replaced. Though the evidence is not entirely clear, there do not seem to be any particularly error-prone, or strong or weak segments (e.g., Shattuck-Hufnagel & Klatt, 1979). In a speech error elicitation experiment, Levitt and Healy (1985) observed that high-frequency segments were more likely to replace low-frequency segments than the reverse.

Apparently, the representations of word forms capture not only which segments, but also which phonological features the words include. Evidence for the representation of phonological features comes from two sources. First, there are errors like (9) that are best described as movements of individual features. Such feature errors are not observed very frequently; probably less than 5% of all sound errors are feature errors (e.g., Berg, 1985; Fromkin, 1973; Shattuck-Hufnagel, 1983; Shattuck-Hufnagel & Klatt, 1979), but they occur often enough to require an explanation. Second, interacting segments in sound errors show a strong tendency to share more features than expected on the basis of a chance estimate (Fromkin, 1971; García-Albea et al., 1989; Garrett, 1975; Nooteboom, 1969). In most sound errors, the displaced and displacing segments differ by only one feature. This phonemic similarity effect is observed in interactions of consonants as well as in interactions of vowels. The features shared most frequently by interacting consonants are manner and voice (see Berg, 1985; Kupin, 1982; MacKay, 1970; Nooteboom, 1969; Shattuck-Hufnagel & Klatt, 1979). In vowel errors the feature tense is more likely to be shared than backness (Shattuck-Hufnagel, 1986). The fact that vowels almost always interact with vowels and consonants with consonants (e.g., Fromkin, 1971; Garrett, 1975; MacKay, 1970) can also be seen as an instance of the general tendency of errors to involve similar segments, though there are other accounts for this finding (see below).³

(9) *glear plue sky* (clear blue sky)

Recently, it has been argued that sound errors involve smaller planning units than segments or features. Using electromyography, Mowrey and MacKay (1990) traced the motor activity during the production of tongue-twisters. As expected, they found some variation in the motor patterns associated with different realizations of one and the same correct utterance. However, there was much more variability in the patterns associated with errors. For instance, analyses of deletions of [l] in "Bob flew by Bligh Bay" showed that in most cases there were still traces of the motor activity associated with [l] at the moment when it should have been pronounced. Similarly, in most [l]-additions, stronger or weaker traces of the motor activity typically associated with [l] were registered, but not the full activity pattern. Thus, at the motor level sound errors do not seem to be all-or-none, but graded events.

One interpretation of these findings is that the error units are phonological segments, as has traditionally been assumed, and that the competition of

³The phonemic similarity effect is one of a number of similarity biases in sound errors. Another similarity bias is the repeated phoneme effect. Two segments that are followed or preceded by identical segments are far more likely to interact with each other than two segments that are followed or preceded by different segments (see, for example, MacKay, 1970). This effect has been found for vowel and consonant errors. It is strongest when identical segments are direct neighbours of the interacting segments, but it is also obtained from more remote identical segments (Dell, 1984; Shattuck-Hufnagel, 1986; Stemberger, 1990).

phonological forms affects the planning at subsequent levels using smaller planning units. An alternative hypothesis, favoured by Mowrey and MacKay, is that sound errors do not arise at the phonological level, but at a lower planning level, and that they are not transpositions of complete segments or features, but of units controlling small sets of motor units. On this account, the main reason why sound errors usually appear to be segmental errors is that the articulatory irregularities that can be discovered by means of electromyography are not noticed by listeners. On the basis of the available data it is not possible to decide between these hypotheses.

In addition to the segmental and subsegmental structure, the word forms speakers create also capture the syllabic structure of the words. Complete syllables rarely function as error units, probably in less than 5% of all sound errors (e.g., Shattuck-Hufnagel, 1983), but there is evidence from a number of other sources to support the representation of syllabic structure. First, when error units comprise two or more segments, these segments practically always belong to the same syllable. There are no errors in which the error unit includes the last segment of one syllable and the first segment of the next syllable. Moreover, the segments of complex error units usually belong to the same syllable *constituent*. A syllable can be divided into an onset, which comprises the pre-vocalic segments, and a rhyme, which includes the remaining segments. The rhyme can further be divided into a vocalic nucleus and a post-vocalic coda. By far the most common complex error units are onset clusters (see (10)), but complex nuclei (see (11) and (12)) are also regularly found. Second, the segments constituting complex syllable constituents of the target words typically stay together in errors. There are errors, such as (13)–(15), in which only one segment of an onset cluster or a complex nucleus is displaced (see Stemberger, 1983b; Stemberger & Treiman, 1986), but such cases are less frequent than replacements of complete complex constituents (Shattuck-Hufnagel, 1983, 1986; but see Kubozono, 1989, for counter-evidence from Japanese errors). Occasionally, errors are observed in which the error unit includes nucleus and coda (i.e., the rhyme), or onset and nucleus of a syllable (Nooteboom, 1969; Shattuck-Hufnagel, 1983). By contrast, there are no errors in which two segments belonging to different complex syllable constituents form an error unit. For example, the second consonant of an onset cluster and the following vowel never form an error unit (for other evidence on the psychological reality of syllable constituents see Fowler, 1987; Treiman, 1983, 1984, 1986).

(10) *sloat thritter* (throat slitter)

(11) *serp* is *souved* (soup is served)

(12) *ho blairer* (hair blower) (from Shattuck-Hufnagel, 1986)

(13) *sprive* for perfection (strive for perfection)

(14) *carn* – corn cobs (from Stemberger, 1983b)

(15) they [muy] – they may be moving (from Stemberger, 1983b)

Additional evidence for the coherence of syllable constituents in errors stems from analyses of word blends and haplogies. Word blends are combinations of words, which are typically related in meaning (see (16)–(18)). MacKay (1972) found that in blends of polysyllabic words the breakpoint tends to fall between rather than within syllables (see (16)). If breaks occur within syllables, they tend to fall between syllable constituents, most often between the onset and the rhyme, as in (17). By contrast, errors like (18), in which the segments of a single onset cluster are separated from each other, are rare.

(16) *recoflect* (recognize/reflect)

(17) *Irvine is quite clear* (close/near)

(18) *what shromkin said* (she/Fromkin)

Haplogies are errors in which parts of intended utterances are missing (see (19)–(22)). As Crompton (1982) has pointed out, the missing part usually corresponds to one or more complete syllables (as in (19)) or syllable constituents (as in (20) and (21)). However, a complex syllable constituent does not function as a unit in *all* deletion errors. In some errors, such as (22), only one segment of a complex constituent is missing (see Stemberger & Treiman, 1986).

(19) *tremenly* (tremendously)

(20) *shrig souffle* (shrimp and egg souffle)

(21) *it is too dailed* (detailed)

(22) *below the gottis* (glottis)

Another piece of evidence for the representation of syllabic structure is a positional constraint on complete and incomplete sound exchanges (see examples (23)–(25)), described by Boomer and Laver (1968) in the following way:

Segmental slips obey a structural law with regard to syllable-place; that is, initial segments in the origin syllable replace initial segments in the target syllable, nuclear replace nuclear, and final replace final. (p. 7)

(23) *mell wade* (well made)

(24) *bud begs* (bed bugs)

(25) *god to seen* (gone to seed)

This syllable-position constraint is observed in the majority of sound errors in the English, German, Spanish, and Dutch corpora that have been analysed (Fromkin, 1971, 1973; Garcia-Albea et al., 1989; MacKay, 1970; Motley, 1973; Nootboom, 1969; Shattuck-Hufnagel, 1983, 1987; Stemberger, 1982; but see Abd-El-Jawad & Abu-Salim, 1987, on Arabic and Kubozono, 1989, on Japanese errors). The description of certain errors as sound exchanges implies that the segments' positions are specified independently of the segments themselves (see also Fry, 1969; Garrett, 1975). If there were no independently defined positions, one would

expect that each of the two displaced segments could be inserted anywhere in the utterance, independently of the position of the other displaced segment. Instead, each segment is confined to the position vacated by the other segment involved in the exchange. An explanation for the observation that the interacting segments usually stem from corresponding syllable constituents is that the positions to which segments are assigned correspond to syllable constituents, and that each segment is marked as eligible for association to one type of syllable constituent (e.g., Baars & Motley, 1976). Accordingly, most models of phonological encoding assume that during the creation of word forms a set of phonological segments is retrieved, syllable frames are created, whose positions correspond to syllable constituents, and the segments are associated to the syllable constituents.

Recently, however, Shattuck-Hufnagel (1985b, 1987, this volume) has pointed out that most evidence for the positional constraint stems from errors involving word onsets. Shattuck-Hufnagel (1987) found that 66% of all consonant errors in her corpus occurred in word-onset position, whereas only 33% of all consonants in normal adult speech appear word-initially. The preference for word-onset positions was more pronounced for ordering errors (i.e., sound anticipations, perseverations, and exchanges) than for non-contextual errors; 82% of the misorderings and 43% of the non-contextual errors occurred in word-onset position. In most ordering errors, the interacting segments or clusters both stem from a word-onset position (Shattuck-Hufnagel, this volume; see also Fromkin, 1977; Garrett, 1975, 1980; but see García-Albea et al., 1989, for counter-evidence from Spanish errors). Thus, the segments of a word onset have a higher error rate than segments of other word positions, and they prefer to interact with each other rather than with word-internal or word-final segments. In analyses of hearing errors, it has been found that the edges of words, especially the beginnings, are perceived more accurately than the middle parts (Browman, 1978; Garnes & Bond, 1980; see Cohen, 1980, for the results of a shadowing study). Thus, errors in word onsets are probably more likely to be detected than errors in other word positions, but given the magnitude of the word-onset effect it is unlikely to be due exclusively to a listener bias.⁴

Shattuck-Hufnagel's analysis raises the question of whether it is necessary to assume both a word-based and a syllable-based constraint, or whether the movements of segments can be captured by reference to word position alone. Evidence bearing on this issue could come from errors involving word-internal and word-final segments. However, as such errors are scarce, there is little information about error rates and preferred interactions. It has been claimed that syllable onsets are more error prone than codas (e.g., MacKay, 1970), but it is

⁴What makes the word-onset effect particularly intriguing is that in sound errors word onsets are particularly vulnerable, whereas in malapropisms and TOT states they are more likely to be correct than segments in other word positions (e.g., Browman, 1978; Brown & McNeill, 1966; Fay & Cutler, 1977; Rubin, 1975).

uncertain whether this still holds once word onsets are excluded from analysis. An important observation is that vowels usually interact with other vowels and consonants with other consonants (see, for instance, Fromkin, 1971; Garrett, 1975; MacKay, 1970). This tendency *can* be explained by reference to syllable frames, in which the positions for onsets, nuclei, and codas are marked. However, it also suffices to assume the existence of frames that do not represent syllables, but include positions reserved for vowels and positions reserved for consonants (see Stemberger, 1990). This confines vowels and consonants to vocalic and consonantal positions, respectively, but it does not rule out interactions between onset and coda segments. A third possibility is to attribute the tendency of vowels and consonants to interact with segments of the same class to the general preference of segments to interact with phonologically similar rather than with dissimilar segments. No counts seem to be available of how often word-internal and word-final consonants move from their target positions to corresponding positions in new syllables, or how often they assume different syllable positions. If there is no syllable-position constraint, but only a word-onset constraint, frames could be postulated that include one distinguished position, the word onset, followed by an undifferentiated set of positions for the remaining segments of the word, or possibly by a sequence of consonantal and vocalic positions. Given frames divided into onset and "remainder", it would not be necessary to assume the existence of syllable frames.

In considering this hypothesis, however, other evidence should be taken into account. One relevant observation has already been discussed above, namely the tendency of complex syllable constituents to function as units in errors. Unfortunately, this evidence, like the evidence concerning the positional constraint, stems largely from errors involving word onsets. Another observation supporting the assumption of syllable frames is that sound errors are systematically affected by the stress pattern of the words in which they appear. Two trends can be distinguished. First, segments of stressed syllables are more likely to be involved in errors than segments of unstressed syllables.⁵ Second, a segment prefers to move from its target syllable to a syllable with the same stress value rather than to a syllable with a different stress value; that is, a segment from a stressed syllable tends to move to a new stressed syllable and a segment from an unstressed syllable to a new unstressed syllable (e.g., Boomer & Laver, 1968; Fromkin, 1971, 1973; Garrett, 1980; Nootboom, 1969; Shattuck-Hufnagel, 1983).⁶

The interpretation of these findings is complicated by the fact that the syllable

⁵This finding might in part be due to the fact that errors in stressed syllables are more likely to be detected than errors in unstressed syllables (see Browman, 1978; Games & Bond, 1980).

⁶In an analysis of German within-word errors, MacKay (1971) found that segments from stressed syllables were more likely to replace segments from unstressed syllables than the reverse, in particular when the unstressed preceded the stressed syllable (see also Berg, 1990). However, this finding was not replicated for English errors of the same type (Shattuck-Hufnagel, 1983).

carrying the main stress is often also the first syllable of the word. In monosyllabic words this is necessarily the case, but it also holds for many polysyllabic words. For instance, in Shattuck-Hufnagel's (1987) corpus, 60% of the polysyllabic words involving an error were stressed word-initially. Shattuck-Hufnagel (1985b, 1986) examined whether segment movement could be better predicted by reference to the word position or the stress value of the target syllable. She concluded that for movement of onset consonants word position was the more influential factor, but that stress also had a significant effect. By contrast, for vowel movement, stress was the more influential factor. Vowels of stressed syllables were more vulnerable than vowels of unstressed syllables, and both types of vowels preferentially moved to syllables with the same stress value as their target syllable (see also Berg, 1990; Shattuck-Hufnagel, this volume).

These findings show that the stress pattern of the utterance is represented at the moment when a sound error arises. In current linguistic theory, stress is usually taken to be carried by syllables, not by segments (see, for instance, Liberman & Prince, 1977). This view is also supported by speech error evidence. When vowels move from stressed to unstressed syllables or vice versa, the stress pattern of the utterance is usually maintained (see, for instance, Berg, 1990; Fromkin, 1971; Shattuck-Hufnagel, 1986; Stemberger, 1983a), indicating that stress value is not a property of a vowel itself, but of the position it takes. If stress values are linked to syllables, it is difficult to explain the effects of stress on sound errors without assuming that the syllabic structure of the utterance is represented. Hence, indirectly, the effects of stress support the assumption that the frames to which segments are associated in phonological encoding capture the syllabic structure of the words.

To summarize, though the results of the large number of speech error analyses performed by different investigators on different corpora by no means match in all details, there are a number of findings that have been replicated in many studies and that have led to important conclusions about phonological encoding. First, the fact that sound errors arise at all shows that word forms are assembled out of smaller units, rather than being retrieved from the mental lexicon as single entities. Second, the phonetic well-formedness of the errors indicates that they arise during the creation of a phonological representation, rather than during phonetic encoding or articulation. Third, given that the majority of error units can best be described as segments or segment sequences, the units out of which phonological representations are created are probably segments and maybe certain segment clusters. Fourth, the phonemic similarity effect and the occurrence of feature errors show that the subsegmental structure of words is also represented. Fifth, the fact that segments often exchange positions in errors indicates that the positions must be specified independently of the segments that fill them. Finally, the tendency of complex syllable constituents to function as coherent units in errors, the syllable-position constraint, and the effects of stress

on movement errors suggest that the syllabic structure of the utterance is represented at the moment when sound errors arise. One should, however, keep in mind that much of this evidence stems from analyses of errors involving word onsets, and that much less is known about the movement patterns and coherence of segments outside the word onset. In the next section, it will be shown how these generalizations are fleshed out in two models of phonological encoding, namely in the models proposed by Shattuck-Hufnagel (1979, 1983, 1986, 1987) and Dell (1986, 1988; for related models see Berg, 1988; Harley, 1984; MacKay, 1982, 1987; Stemberger, 1985b).

2. Models of phonological encoding

2.1. *Shattuck-Hufnagel's scan-copier*

Shattuck-Hufnagel (1979, 1983, 1986) assumes that during phonological encoding an ordered set of sublexical units is associated to the ordered positions of independently created syllable frames. Because speech errors do not provide conclusive evidence about the representation of complex syllable constituents, two versions of the model are proposed. In one, word forms are created out of segments, segment sequences, and zero segments. Syllables have three slots each, corresponding to the syllable constituents onset, nucleus, and coda. Each slot accepts a single segment, a segment sequence, or a zero segment as insert. Zero segments take the onset or coda positions of syllables that begin or end in vowels. In the other version of the model, there are no units corresponding to segment sequences, but only units corresponding to single segments and zero segments. Complex syllable constituents are represented as sequences of two or three segments. The syllable frame includes a separate slot for each segment.

The model presupposes the generation of the syllable frames for a stretch of speech, probably a phrase, and the retrieval of the corresponding ordered set of sublexical units. It describes the association of the segments to the positions of the syllable frames. A scan-copier is proposed, which selects the correct insert for each slot from the set of retrieved units and copies the units into the slots. This is done sequentially, proceeding slot-by-slot and unit-by-unit from the beginning of the utterance to its end. As soon as a given unit has been inserted into a slot, it is marked by a check-off monitor as "used". A second monitor inspects the developing representation and deletes or edits sequences that are likely to be the result of errors, such as sequences in which a particular segment is repeated several times.

Sound errors arise when slots are filled by wrong units, and the monitor fails to notice this. In a sound exchange, a unit is inserted into a slot too early, and the segment that should have taken that slot is inserted into the slot that was meant to

be filled by the anticipated segment. Both segments are correctly checked off as "used" as soon as they have been associated to a position. Anticipations and perseverations are more complex errors in that they involve not only wrong placements of segments, but also failures of the check-off routine. In an anticipation, a unit is erroneously inserted into a slot preceding its target slot, is not checked off, and is inserted again into the appropriate slot. Similarly, a perseveration occurs when a unit is inserted into its target slot, is not checked off, and is later inserted again into another slot. The model correctly predicts that most sound errors are segmental errors, as segments and possibly certain segment sequences are the sublexical units out of which word forms are composed. The model cannot explain the occurrence of feature errors.

In order to explain the syllable-position constraint on sound errors, it is assumed that the positions of the syllabic frames are labelled as onset, nucleus, and coda positions and that the sublexical units are labelled correspondingly as onset, nucleus, and coda units. A segment that may appear in two syllable positions (such as English stop consonants) is represented twice with different labels. In filling a given syllable position, the scan-copier only considers units that are marked as appropriate for that type of position. Thus, in correct utterances and in errors, each segment can only be linked to one type of position. In order to explain why a segment preferentially moves to a syllable with the same stress value as its target syllable, Shattuck-Hufnagel (1983) suggests that slots and inserts might not only be marked as to their positions in the syllable, but also as belonging to stressed or unstressed syllables. It should be noted that the labelling of segments and positions is only necessary to account for the constraints on movements of segments in errors, but not to explain the ordering of segments in correct utterances. As segments and positions are already ordered before they are linked to each other, and as the association is a left-to-right one-to-one mapping process, no additional information is necessary to link the sublexical units to correct positions.

Important observations that Shattuck-Hufnagel's (1979, 1983) original model cannot explain are the heightened error risk of word onsets and their tendency to interact with each other rather than with segments outside word onsets. However, a recent extension of the model includes an account of these findings (Shattuck-Hufnagel, 1987). In order to create a well-formed utterance, the speaker must integrate information about word forms and syntactic information, as they jointly determine the stress pattern and rhythm of the utterance. In Shattuck-Hufnagel's extended model, the insertion of sublexical units into positions of frames is part of the process by which lexical and syntactic information are combined. First, a frame is created with two positions for each word, one for the word onset and one for the "rest of the word". As the metrical structure of an utterance only depends on "rests of words", but not on word onsets, "rests of words" are associated to their positions before onsets. The most common type of error arising during this

process is that an onset unit is linked to the onset position of a wrong word. In the next processing step, a frame is created with segment-sized positions, and the segments of the words become available and are associated to these positions. During this process, all segments of a word are about equally likely to be involved in errors.

The separation of the word onset from the rest of the word provides for an explanation of the word-onset effects in errors, but it is not particularly plausible linguistically. For the determination of the metrical structure of an utterance, not only word onsets, but all syllable onsets are irrelevant (see, for instance, Selkirk, 1984). If metrically relevant information is to be processed first, syllable rhymes should precede onsets. Yet, in Shattuck-Hufnagel's model the complete "rest of the word", including rhymes and word-internal syllable onsets, is associated to its position before the word onset.

2.2. *Dell's spreading activation model*

In Dell's (1986) model, as in Shattuck-Hufnagel's, word forms are generated by inserting sublexical units into the slots of independently created frames. But whereas Shattuck-Hufnagel's model presupposes the availability of the sublexical units to the scan-copier, Dell's model describes their retrieval within a spreading activation framework.

The linguistic units participating in phonological encoding are morphemes, syllables, rhymes, segment clusters, segments, and features. The nodes representing these units are connected to form a hierarchical structure, in which each unit is linked to its constituents. As in Shattuck-Hufnagel's model, segments and clusters are marked as onset, nucleus, or coda units. Again, there are zero-segments, which take the onset or coda position in syllables beginning or ending in a vowel.

The nodes are processing units, which can be activated to a greater or lesser extent. The links between the units are bidirectional. When a particular unit is activated, it spreads some of its activation to all units to which it is connected, and these units in turn feed part of their activation back to the unit that activated them in the first place. Activation decays over time so that unbounded spreading of activation from one node to all other nodes of the network is avoided.

When the form of a monosyllabic morpheme is created, activation spreads from the morpheme node to its syllable node, and from there to the corresponding segment and cluster nodes, which become gradually more and more highly activated. At the same time, a syllable frame with ordered onset, nucleus, and coda slots is created. After a certain time interval, these slots are filled by whatever onset, nucleus, and coda units are the most highly activated. Provided that no error occurs, these are the units that are being activated by the morpheme

whose form is to be created. The three slots of the frame are filled in parallel. Upon their insertion into the frame, the selected units are tagged as being part of the phonological representation, and their activation level is reduced to zero so that they will not be immediately selected again. However, as the tagged units are still receiving some activation from activated superordinate and subordinate nodes, their activation quickly rebounds from zero and then gradually decays.

The syllables of polysyllabic morphemes are encoded in succession. For each syllable, the syllable frame is created and filled by suitable segments. From the activated morpheme node, activation spreads in parallel to all of its syllable nodes and to the corresponding segments and clusters. In addition, the first syllable node initially receives an extra boost of activation. Because of this additional activation, its segments and clusters become more strongly activated than all other segments and clusters and are therefore selected when the syllable frame is to be filled for the first time. Then the activation levels of the first syllable and its segments and clusters are set to zero, and the second syllable starts to receive extra activation. When the syllable frame is to be filled for the second time, the segments of the second syllable emerge as the most highly activated units and fill the slots, and so on, until all syllables of the morpheme have been encoded.

Dell (1988) has recently outlined a modification of his model, which instead of one syllable frame with the slots onset, nucleus, and coda, assumes several frames (called wordshapes, as only monosyllabic words are considered). These frames are sequences of C- and V-slots and are of variable length. A word node activates not only a set of segments, as in the earlier version of the model, but also a wordshape. From the wordshape, activation flows to so-called phoneme category nodes. The CVC-wordshape, for example, connects to the phoneme category nodes for pre-vocalic consonants, vowels, and post-vocalic consonants. The phoneme category nodes in their turn activate all segments of their respective categories. Thus, the segments of a word receive activation via two routes, directly from the word node and via the wordshape and phoneme category nodes. As soon as a segment reaches a selection threshold, it is selected as part of the phonological representation. The assumption of several types of frames is linguistically more plausible than the assumption of only one type because differences in the syllabic structure of various words can be represented and because it is not necessary any more to postulate zero-segments as fillers for onset and coda positions of syllables that begin or end in a vowel. In the new version of the model, such syllables simply do not have onset or coda positions.

Sound misorderings arise when segments or clusters are the most highly activated units of their categories at the wrong moment and are therefore associated to incorrect positions. For example, an anticipation like (26) occurs if the onset of the second word is more highly activated than that of the first word at the moment when the syllable frame is to be filled for the first time. Similarly, a perseveration like (27) occurs if the onset of the verb is still more highly activated

than that of the noun when the noun is to be phonologically encoded. Such irregularities in the activation of segments can arise because each unit in the mental lexicon connects to many others from which it receives variable amounts of activation. Usually, the input from the morpheme level is strong enough to override these random influences so that correct segments are selected in the right order, but sometimes wrong segments win out, and errors arise.

(26) *hinch* hit (pinch hit)

(27) gave the *goy* (gave the boy)

Many properties of sound errors are explained in similar ways in Dell's as in Shattuck-Hufnagel's model. Errors are usually phonetically well formed because they arise during the creation of a phonological representation and thus before the phonetic form of the utterance is specified. The main error units are segments and certain segment sequences because these are the units that are selected and combined to form phonological representations. Feature errors cannot be explained, but there is an account for the phonemic similarity effect on segmental errors: phonologically similar segments have links to identical feature nodes, through which they activate each other so that their activation levels become similar and they become more likely to be confused with each other than dissimilar segments that do not activate each other via shared subordinate nodes. The syllable-position constraint on movements of segments is explained, as in Shattuck-Hufnagel's model, by the assumption that a segment is labelled according to the position it may take. No account is offered for the effects of word onset on sound errors.

3. Limitations of speech error analyses

Typically, speech errors are not analysed because they are particularly interesting or important as such, but because they are assumed to reveal how correct speech is created. Section 1 discussed a number of properties of sound errors and the inferences about normal phonological encoding they invite. In section 2, it was shown how these inferences are fleshed out in two models of phonological encoding. Despite numerous investigations of sound errors and the existence of speech-error-based models of phonological encoding, it will be argued here that the creation of utterance forms is not well understood. Moreover, it seems unlikely that it will be understood unless new research methods are employed.

Understanding phonological encoding would imply, among other things, knowledge of the structure and content of the representations of utterance form that speakers create. Using the results of sound error analyses, certain aspects of these representations can be broadly characterized, but many important questions remain unanswered. For example, given that complex syllable constituents func-

tion as units in most, but not all errors, it cannot be decided whether they are represented as single entities, or as segment sequences, or perhaps in both ways. We also know little about the representation of phonological features. The phonemic similarity effect on segmental errors indicates that words are represented in terms of their features as well as in terms of their segments. But the radically different rates of segment and feature errors suggest that segments and features are represented in different ways. To give a final example, little is known about the frames to which the segments are associated; the available evidence does not reveal whether there are only syllable frames, or only word frames, or perhaps both types of frames.

Why has it not been possible to specify word forms in more detail? In section 1, two important methodological problems of speech error analyses were mentioned. One problem is that the diagnosis of sound errors depends on listener judgments. This is a problem because some errors probably have higher detection rates than others (e.g., Cutler, 1981; Ferber, 1991). For instance, it is likely that errors in word onsets and in stressed syllables are noticed more easily than errors in other word positions. A second problem is the notorious ambiguity of errors (see also Cutler, 1988). For example, many displacements of segments can either be characterized as movements from a given syllable position to a corresponding position in another syllable, or as movements from a given word position to a new word position of the same type. Listener strategies and the ambiguity of errors conspire to conceal the true distribution of different error types in speech production.

A third problem is that certain classes of errors that one would need to analyse in order to obtain a more complete picture of phonological encoding are hardly ever observed. Speakers must create representations of the intonation contour and stress pattern of their utterances, and they must construct phonetic representations. But errors of stress, intonation, or phonetic encoding are seldom observed (e.g., Cutler, 1980; Fromkin, 1977). One possibility is that speakers rarely commit these types of errors; an alternative is that such errors do, in fact, regularly occur, but that listeners usually do not notice them. As mentioned, there is experimental evidence suggesting that listeners might often fail to notice violations of phonetic rules (Cutler, 1981; Marslen-Wilson & Welsh, 1978; Warren, 1970). Cutler (1980) has pointed out that if a speaker places stress on the wrong word, the focus of the sentence will be affected; but unless this has dramatic pragmatic consequences, the listener will not notice an error, but will understand the utterance to mean something slightly different from what was intended by the speaker. For similar reasons, errors of intonation might appear to be rare. If a wrong intonation contour is applied to a sentence, the listener might misunderstand the speaker (e.g., as being ironic instead of sincere) rather than noticing a speech error.

There are ways of minimizing these problems. By using taped corpora instead

of transcripts, listener strategies can be minimized, and reasonable samples of unambiguous errors can be obtained by collecting large error corpora. An efficient way of acquiring speech error corpora is to induce errors in experimental settings, for instance by using the SLIPS procedure introduced by Baars et al. (1975),⁷ or by asking subjects to produce tongue-twisters, such as “She sells sea shells on the sea shore” (see, for instance, Butterworth & Whittaker, 1980; Kupin, 1982; Levitt & Healy, 1985; Shattuck-Hufnagel, 1987, this volume). Problems of such techniques are that some of the normal planning processes might be omitted or altered and that the articulation might be more difficult than in spontaneous speech. Therefore, the results of error induction experiments must always be validated by comparison to the results of analyses of errors in spontaneous speech.

On the basis of further analyses of large reliably registered error corpora it should be possible to obtain estimates of the relative frequencies of different classes of errors and to answer some of the open questions concerning the nature of the representations of utterance forms. For instance, it can probably be determined how likely the segments of a complex syllable constituent are to stay together in errors or to be separated from each other and how often segments outside word onsets move from their target positions to corresponding versus different positions in new syllables. Such distributional data are necessary to decide whether the frames to which the segments are associated encode the syllabic structure of the utterance. Whether substantial corpora of errors involving stress and intonation and of phonetic errors can be collected, and what can be learned from such corpora, remains to be seen.

In order to understand phonological encoding we must not only find out *which* form representations are constructed, but also *how* they are constructed. It must be determined which functionally different planning processes are to be distinguished and how these processes are coordinated with each other in time. Sound errors allow for a rough estimate of the relative order of certain planning processes. For instance, the effects of syllable position and stress on displacements of segments show that the syllabic structure of an utterance is generated before, or at about the same time as the segments are retrieved and ordered. Similarly, the phonetic well-formedness of most errors indicates that the phonetic

⁷In SLIPS experiments, subjects are presented with series of word pairs. On most trials, no overt reaction is required, but occasionally a word pair is accompanied by an auditory signal prompting the subject to say the word pair aloud. Such target word pairs are preceded by a series of biasing pairs designed to induce particular errors. For instance, if the target pair “deal back” is preceded by “big dumb”, “bust dog”, and “bet dart”, subjects are likely to make the onset exchange error “beal dack” (example from Dell, 1988). This paradigm has been widely used to test output biases, such as the tendency of sound errors to result in existing words or syntactically well-formed sequences of the language (Baars & Motley, 1976; Baars et al., 1975; Motley & Baars, 1976; Motley et al., 1981, 1983), but also to investigate the repeated phoneme effect (Dell, 1984) and properties of onset cluster errors (Stemberger & Treiman, 1986).

form of an utterance is specified after its segments have been assigned to positions in word or syllable frames.

However, to build a satisfactory model, far more detailed information about the time course of various processes is required. The results of additional error analyses might firmly establish the existence of both word and syllable frames, but they will not reveal whether the two types of frames are created at the same time or, as Shattuck-Hufnagel (1987) has proposed, at different times. Similarly, even if we know exactly how often segment and feature errors arise, we still cannot decide whether the segmental and subsegmental structures of words are created in parallel or in succession. To give a final example, speech errors show that segments are associated to positions in planning frames, but they reveal nothing about the time course of this process. Shattuck-Hufnagel (1983) assumes that the association is a strictly serial process, proceeding position by position from the beginning to the end of the utterance. By contrast, in Dell's model (1986), the segments within a given syllable can be associated to their positions in any order, but the positions of successive syllables must be filled in sequence. These assumptions have important implications for other features of the models. In Shattuck-Hufnagel's model, the order of a word's segments must be stored in the mental lexicon; whereas in Dell's view, the order of the segments within a syllable is not stored, but established when the segments are associated to the ordered slots of syllable frames. Speech errors do not tell us anything about the time course of the association of segments to positions, nor do they reveal whether the order of segments within syllables is stored in the mental lexicon or generated during phonological encoding.

A model of phonological encoding should not only describe the temporal coordination of various planning processes, but should also specify their functions. Unfortunately, speech errors do not provide any functional information. Sound errors reveal that word forms must be created out of smaller units, but not why this is necessary. Given that there is only one correct phonological representation for each word, one might expect word forms to be stored and retrieved as units without internal structure. The function of the mapping of segments to positions in frames is also unknown. Both Dell (1986) and Shattuck-Hufnagel (1979, 1983) assume that it serves to establish the appropriate surface order of segments. However, in Shattuck-Hufnagel's model the segments are already ordered before the association, so it is unclear why they should be ordered again. Other functions have been ascribed to the mapping process, for instance, that it might be part of the transfer of information from a lexical to a phrasal processor (Shattuck-Hufnagel, 1987), or that it might be necessary to create surface phonological forms out of stored lexical representations (Levelt, 1989). Sound errors do not convey which of these hypotheses, if any, is correct.

To explain certain properties of errors, other processes have been postulated whose functions are also quite obscure. For instance, connectionist models assume that activation spreads not only from higher to lower-level units (i.e., from words

to syllables, segments, and features), but also back from lower- to higher-level units (e.g., Dell, 1985). This explains a number of characteristics of sound errors, such as the phonemic similarity effect on segmental errors, but it is not clear what the function of the upward spread of activation could be. One can speculate about it. For instance, there might only be one lexicon for language production and comprehension, and upward connections might be crucial for the comprehension process. But speech errors do not provide any evidence in support of this, or any other, functional hypothesis.

To sum up, analyses of sound errors are usually carried out on the assumption that errors provide evidence about the representations of utterance forms and the way they are generated. However, so far only a very broad characterization of certain parts of the form representations has been achieved. Maybe some properties of the representations can be further specified on the basis of additional evidence. However, in the light of what has been achieved so far, it seems unlikely that sufficient evidence will ever be obtained from speech errors to characterize all levels of representation, including, for instance, the representations of intonation and stress. Furthermore, errors provide little evidence about the planning processes involved in phonological encoding, revealing little about the temporal coordination of different processes and nothing about their functions. Thus, it appears that the evidence that can be gained from speech errors is far more limited than has often been assumed.

4. Implications for future research

In order to understand phonological encoding a new research strategy seems to be required. The main strategy pursued so far has been to start from existing speech error corpora and to construct theories that are tailored to account for the characteristics of errors, but that are also supposed to explain phonological encoding in error-free speech. As we have seen, a number of important properties of sound errors have indeed been accounted for; but we still do not understand the normal process of phonological encoding very well, mainly because sound errors fail to provide the necessary evidence. It might be more fruitful to use as a starting point a working model of correct phonological encoding, to derive new hypotheses from that model, and to find ways to test them empirically. The evolving theory should be tailored to explain normal phonological encoding, but it should, ideally, also explain the properties of errors. Thus, instead of a highly data-driven approach that relies on one type of evidence and primarily aims at the explanation of a fairly infrequent type of behaviour, namely errors, a more theory-driven approach is advocated that can draw on many different types of data and aims directly at the explanation of the true object of arguing, the generation of correct speech.

How could a working model of phonological encoding be derived? There is little psycholinguistic evidence on which such a model could be based, but there is a large body of relevant linguistic research. Detailed theories of lexical representations have been proposed (e.g., Clements, 1985; Clements & Keyser, 1983; Goldsmith, 1976; Halle & Mohanan, 1985; McCarthy, 1979; Prince, 1983; Selkirk, 1984; van der Hulst, 1984, 1989; van der Hulst & Smith, 1986). In addition, there are theories on the derivation of the stress pattern and rhythm of utterances (e.g., Gussenhoven, 1984; Kaisse, 1985; Nespor & Vogel, 1986; Pullum & Zwicky, 1986; Selkirk, 1984) and descriptions of the intonation contours of various languages and their conditions of application (e.g., Bolinger, 1985; Cruttenden, 1986; Hart & Collier, 1975; Liberman & Pierrehumbert, 1984). Finally, there is ample phonetic evidence on the acoustic realization of intonation contours and stress patterns (e.g., Collier & Gelfer, 1983; Cooper & Paccia-Cooper, 1980; Cooper & Sorensen, 1981; Fujimura, 1981; 't Hart et al., 1990; Nakatani et al., 1981; Ohala, 1978). Obviously, phonological and phonetic theories are not processing theories, but they can contribute to such theories by offering hypotheses about the representations speakers might create.

Given that linguistic theories do not entail hypotheses about the generation of representations, the processing assumptions of a working model of phonological encoding must be independently motivated. Levelt (1989) has recently proposed a model that is much wider in scope than the speech-error-based models discussed in this paper. It not only specifies the retrieval of stored phonological forms, but also the creation of phonological words and phrases, the generation of stress patterns and intonation contours, and the selection of articulatory commands. The model's structural assumptions are largely linguistically motivated, whereas its processing assumptions are based primarily on psychological considerations. A central assumption of the theory is, for instance, that speakers generally attempt to minimize their memory load. Therefore, representations are created in a piecemeal fashion, and partial representations are handed over to the following processor as soon as possible. According to the theory, speakers do not construct and store the phonological representation of an entire complex sentence before starting to determine its phonetic form and to select appropriate motor commands. Instead, phonetic encoding and articulation begin as soon as possible, maybe as soon as the first phonological word of the utterance has been created. A second consequence of the strategy of minimizing memory load is that, in creating representations, speakers rely as much as possible on local information rather than looking far ahead. For instance, when assigning stress to a particular word, speakers probably consider the stress pattern of preceding words, the lexical stress pattern of the word under consideration and maybe of the following word, but most of the time they do not look further ahead than that.

How could such a working model of phonological encoding be tested? Given the complexity of the process under consideration, it is unlikely that one method

can be found that can be employed to address all these questions. Most likely, a variety of methods must be used to investigate different aspects of phonological encoding. A number of suitable techniques are described below.

First, much can be learned about phonological encoding from analyses of spontaneous speech. As mentioned, there is an extensive linguistic literature on the phonological and phonetic properties of utterances. However, linguistic research has primarily investigated which types of utterances are permissible, universally or in particular languages. The empirical evidence often stems from a small group of trained speakers, who carefully read out a set of words or sentences. By contrast, surprisingly little is known about the characteristics of spontaneous speech. For instance, certain types of Dutch sentences can be pronounced with a particular intonation contour, the so-called "hat" pattern, which is characterized by the maintenance of a high pitch level between two pitch accents ('t Hart & Collier, 1975). Speakers *can* apply this intonation contour, but they need not do so in order to produce well-formed utterances. Similarly, English speakers can adjust the stress patterns of words to avoid clashes of stressed syllables (saying, for instance, *sixteen ábstract páintings* instead of *sixtéén abstráct páintings*), but it has been shown that they rarely do so (Cooper & Eady, 1986; see also Kelly & Bock, 1988).

Analyses of spontaneous speech reveal which of the options existing in a given language speakers typically take and allow for certain inferences about the underlying representations of utterance forms. For instance, it can be determined whether speakers regularly produce utterances whose intonation contours are best described as "hat" patterns, or utterances in which the lexical stress patterns of words are altered to avoid stress clashes. In addition, certain inferences can be drawn about the speaker's planning processes. In particular, it can be established whether certain postulated processes, like the creation of "hat" patterns or stress adjustment, take place at all and which information speakers take into consideration when performing them. For example, one might find that speakers create "hat" patterns, but only if the syllables carrying pitch accent are not separated from each other by more than a certain number of syllables or words. This would suggest that the processor takes some of the following context of a word into account when determining its pitch level but that its preview is limited to a certain stretch of speech.

An alternative to the study of spontaneous speech is to analyse utterances elicited in experiments in which subjects describe pictures, answer questions, paraphrase sentences, or recall sentences or texts from memory. An obvious advantage of the experimental approach, compared to analyses of spontaneous speech, is that the structure of the speakers' utterances can be systematically varied, as well as the linguistic and non-linguistic factors that are expected to affect utterance forms. A disadvantage is that some of the planning processes taking place in spontaneous speech might be altered or not take place at all.

Therefore, it is important to validate the results of utterance elicitation studies by comparing them to the results of analyses of spontaneous speech.

Another way to study phonological encoding is to test how long it takes speakers to generate certain types of utterances. Utterance initiation times provide evidence about the time course of the generation of utterances, since they reveal which planning activities are completed before utterance onset and which are executed afterwards. In one group of experiments, the difficulty of semantic and phonological encoding of different parts of utterances was varied, and it was tested which of these manipulations would affect utterance initiation times (Kempen & Huijbers, 1983; Levelt & Maassen, 1981; Lindsley, 1975, 1976). The results of these studies have led to the conclusion that speakers plan utterances further ahead on the semantic than on the phonological level; however, they do not tell us exactly how far ahead they plan on either level. In further research, it could be tested more systematically whether the difficulty of semantic and phonological encoding of words appearing in various sentence positions affects sentence initiation times. On the basis of such results it could be determined which sentence fragments are semantically and phonologically encoded before sentence onset.

Additional evidence concerning the issue of advance planning in sentence production comes from experiments by Ferreira (1991), who studied sentence initiation times and pauses when subjects reproduced sentences of varying length and syntactic complexity from memory (see also Balota et al., 1989; Sternberg et al., 1978).⁸ She found that the initiation times for sentences like *The big and hairy dog went to the pond that's next to the museum* depended on the number of phonological words in the subject of the sentence and, given a constant number of phonological words, on the syntactic complexity of the subject noun phrase. By contrast, the complexity of the object did not affect the initiation times but instead the probability and duration of pre-verbal pauses. Ferreira assumed that sentence initiation times and pause durations depended on the difficulty of translating the semantic-syntactic structures of the sentences into phonological representations. On this assumption, her results suggest that the size and nature of the sentence fragment that speakers phonologically encode before utterance onset do not correspond to a certain number of syllables or words, but are defined by the structure of the sentence. The available results do not reveal which aspect of sentence structure is relevant; that is, whether the encoded fragment is a

⁸Some evidence about when speakers plan different parts of a sentence can be gained from analyses of hesitations and pauses in spontaneous speech (e.g., Butterworth, 1980; Cooper & Paccia-Cooper, 1980; Ford & Holmes, 1978; Garrett, 1980; Gee & Grosjean, 1983; Goldman-Eisler, 1968; Kowall et al., 1985; van Wijk, 1987). However, as the distribution and durations of pauses and hesitations are determined by a number of factors, such as the speaker's communicative intentions, syntactic and semantic influences, and the difficulty of various planning processes, it is difficult to obtain evidence suitable to answer questions concerning phonological encoding.

syntactic unit (i.e., the subject) or a prosodic unit (i.e., a phonological or intonational phrase). Further experiments using the same paradigm should address this issue.

Utterance initiation times have also been analysed to investigate the time course of phonological encoding of individual words. Meyer (1991) ran a series of experiments, in which subjects first learned sets of word pairs, such as *dog–cat*, *prince–king*, *thief–cop*. On each test trial, the first member of one of the pairs (e.g., *dog*) was presented, and the subject named the second member of the pair (e.g., *cat*) as quickly as possible. The stimulus materials consisted of two types of sets. In so-called *heterogeneous* sets the response words were unrelated in form, whereas in *homogeneous* sets they had one or more segments in common. The shared segments either corresponded to the syllable onset (as in the above example) or to the rhyme (as in *boy–man*, *car–van*, *skin–tan*).

Naming times were shorter in the homogeneous than in the heterogeneous sets when the response words shared the onset, but not when they shared the rhyme. In the shared-onset condition, the subjects presumably created and retained a phonological representation of the redundant part of the response words, and on each trial only appended the non-redundant part of the word to it. This took less time than creating the complete phonological representations of the response words. In the shared-rhyme condition, such a strategy could apparently not be applied. However, in experiments using disyllabic response words, stronger facilitatory effects were obtained in homogeneous sets in which the response words shared the entire first syllable than in sets in which they only shared the word onset. In preparing for the response words, the subjects apparently could not deviate from the order in which words are normally encoded, which is in a left-to-right manner. In other words, the results suggest that the onset of a syllable must be phonologically encoded before its rhyme. In other experiments using the same paradigm it was shown that successive syllables of a word are likewise encoded sequentially, according to their order in the word (Meyer, 1990).

In order to obtain convergent evidence for the claim that word forms are created in a left-to-right fashion, Meyer and Schriefers (in press) tested this hypothesis again, using a picture–word interference paradigm. The subjects saw pictures of common objects, which they named as quickly as possible. Again, utterance initiation times were the main dependent variable. Together with the pictures, interfering words (IWs) were presented, which were either phonologically related or unrelated to the picture names. For each picture there were two related IWs. The two related IWs for pictures with monosyllabic names had the same onset and nucleus, or the same nucleus and coda as the picture names. The two related IWs for pictures with disyllabic names shared the first or second syllable with the picture names. In addition, the relative timing of picture and IW presentation was varied. The IWs were either presented so that the segments they

shared with the picture names began exactly at picture onset, or so that the shared segments began slightly before or after the picture onset.

The effect of the phonological relatedness of IWs and picture names depended on the timing of the IW presentation and on the word positions of the shared segments. When the IWs were presented so that the critical segments began before picture onset, utterance initiation times were shorter in the related than in the unrelated condition, provided that the shared segments appeared word-initially (i.e., corresponded to the onset and nucleus of monosyllabic words and to the first syllable of disyllabic words). By contrast, when the IWs were presented later, facilitatory effects were obtained in the related condition regardless of the word position of the shared segments. Presumably, when an IW was presented, its phonological representation in the mental lexicon was activated for a certain time period. When some of the segments of the IW were also included in the target, the selection of these target segments was facilitated, provided that this process took place soon enough after the presentation of the IW. A likely reason why the effect of shared word-initial segments appeared before the effect of shared word-final segments is that the encoding of the ends of words began later than the encoding of their beginnings. Thus, these results support the conclusion from Meyer's (1990, 1991) experiments that onset and rhyme of a syllable, as well as successive syllables of a word, are phonologically encoded in succession.

Because the effects of many different types of phonological relationships can be tested, the paradigms used by Meyer and by Meyer and Schriefers may be useful tools for further research on phonological encoding. The available results suggest that the onset of a syllable is encoded before its rhyme. In further experiments it should be tested whether there are smaller units within these syllable constituents that are encoded in succession, and in which order different parts of morphologically complex words are encoded. In addition, evidence could be gained about the types of representations speakers create. It could, for instance, be tested whether the naming responses in a picture-word interference experiment are facilitated by IWs that do not share any segments with the target names, but include the same number and types of syllables. This would indicate that subjects create representations of the syllabic structure of the words they are about to say. In a similar way it could be tested whether the representations capture which phonological features words are composed of.

Finally, phonological encoding can be investigated by means of so-called production priming experiments (see Dell & O'Seaghdha, this volume, for a review). In such experiments, subjects prepare for a certain word, phrase, or sentence. Depending on a cue presented after the preparation interval, they either have to produce the prepared utterance, or to change plans and do something else, like reading a new word or classifying a new stimulus as a word or non-word. It is tested how quickly subjects can switch from the planned to the new reaction. As the relationship between the two responses can be freely varied, this paradigm can be used for many different purposes.

Dell & O'Seaghdha (this volume) carried out production priming experiments to investigate the span of advance planning in sentence production. Their subjects prepared to say sentences, such as *The boxer removed the coat*. After a preparation interval, they either saw an asterisk and recited the sentence, or they saw a new stimulus (e.g., *coal*), which they named or, in a different experiment, classified as word or non-word (see also Levelt et al., 1991, for a similar methodology). Dell and O'Seaghdha found systematic reaction time differences between new words that were phonologically related to one of the words in the prepared sentence and new words that were unrelated to the words in the sentences. The direction of the difference (facilitation vs. inhibition in the related compared to the unrelated condition) depended on whether the new word was related to a word appearing at the beginning or at the end of the sentence. As Dell and O'Seaghdha argue, this suggests that at utterance onset, beginning and end of a sentence are in different states of preparation. Most likely, the phonological representation of the beginning of the sentence has been completed, whereas the representation of the end is only beginning to be activated.

Like the method used by Ferreira (1991), the response-priming paradigm can be used to test which parts of utterances are phonologically encoded before speech onset. In order to explore when non-initial parts of sentences are phonologically encoded, new words can be presented at different moments during the articulation of a sentence rather than before utterance onset. In addition, by varying the type of relationship between the new word and the critical word in a prepared utterance, it can be determined which properties of word forms (e.g., their segmental or subsegmental structure, their stress pattern, or syllable structure) have been retrieved. For instance, Meyer and Gordon (1985) asked subjects to prepare for syllable pairs, such as *up-ub*, which shared the vowel and differed in their codas. The new response required on some test trials was to say the syllables in reversed order. Meyer and Gordon found that it took subjects longer to initiate the new response when the final consonants of the syllables shared the feature voicing or place of articulation than when this was not the case. As they argue, this suggests that the representations of the syllables include specifications of their phonological features and that segments with shared features inhibit each other (see also Yaniv et al., 1990).

To summarize, the most common approach in the study of phonological encoding has been to start from sound errors and to construct a theory that primarily explains the properties of errors, but is also supposed to explain phonological encoding in error-free speech. This strategy has led to important insights about phonological encoding, for instance to the conclusion that word forms cannot be retrieved as units from the mental lexicon, but must be constructed out of certain sublexical units, that phonological segments are important planning units, and that frames are built to whose positions the segments are associated. Yet, it appears that no more can be expected from analyses of sound errors than a global characterization of the representations of utterance

forms speakers create. Moreover, sound errors reveal very little about how these representations are constructed.

For the further investigation of phonological encoding a new research strategy seems to be more promising, which is to directly investigate the normal process of phonological encoding rather than taking the detour via error evidence. It was illustrated in this section how one might proceed. The starting point could be a working model of phonological encoding. So far, little psycholinguistic research has been specifically directed at phonological encoding, but there is a large body of linguistic evidence suggesting hypotheses about the speaker's representations of utterance forms. Hypotheses about how these representations could be constructed can be deduced from general psychological principles.

A number of different ways of empirically testing such a model were discussed above. Important evidence can be obtained by analysing the phonological and phonetic properties of utterances produced spontaneously or in laboratory settings. Such analyses reveal what the products of the speaker's planning activities are and permit certain inferences about the corresponding representations and planning processes. Additional evidence can be gained from chronometric studies, for instance from analyses of utterance initiation times and from production priming studies.

As the existing methods will certainly not suffice to address all issues arising in the study of phonological encoding, new methods must be developed. I would expect that in research on phonological encoding, as in other areas of psycholinguistics, methodological and theoretical developments will go hand in hand; that is, that once researchers become interested in those aspects of phonological encoding about which speech errors fail to provide evidence, they will also find ways to empirically investigate them.

References

- Abd-El-Jawad, H., & Abu-Salim, I. (1987). Slips of the tongue in Arabic and their theoretical implications. *Language Sciences*, 9, 145-171.
- Baars, B.J., & Motley, M.T. (1976). Spoonerisms as sequencing conflicts: Evidence from artificially elicited errors. *American Journal of Psychology*, 89, 467-484.
- Baars, B.J., Motley, M.T., & MacKay, D.G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, 14, 382-391.
- Balota, D.A., Boland, J.E., & Shields, L.W. (1989). Priming in pronunciation: Beyond pattern recognition and onset latency. *Journal of Memory and Language*, 28, 14-36.
- Berg, T. (1985). Is voice suprasegmental? *Linguistics*, 23, 883-915.
- Berg, T. (1987). The case against accommodation: Evidence from German speech error data. *Journal of Memory and Language*, 26, 277-299.
- Berg, T. (1988). *Die Abbildung des Sprachproduktionsprozesses in einem Aktivationsflußmodell: Untersuchungen an deutschen und englischen Versprechern*. [The representation of the process

- language production in a spreading activation model: Studies of German and English speech errors.] Tübingen: Niemeyer.
- Berg, T. (1990). The differential sensitivity of consonants and vowels to stress. *Language Sciences*, 12, 65–84.
- Bolinger, D. (1985). *Intonation and its parts: Melody in spoken English*. London: Edward Arnold.
- Boomer, D.S., & Laver, J.D.M. (1968). Slips of the tongue. *British Journal of Disorders of Communication*, 3, 2–12.
- Browman, C.P. (1978). *Tip of the tongue and slip of the ear: Implications for language processing*. UCLA Working Papers on Phonetics, No. 42. University of California, Los Angeles.
- Browman, C.P., & Goldstein, L.M. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3, 219–252.
- Brown, R., & McNeill, D. (1966). The “tip of the tongue” phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 5, 325–337.
- Buckingham, H.W., & Yule, G. (1987). Phonemic false evaluation: Theoretical and clinical aspects. *Clinical Linguistics and Phonetics*, 2, 113–125.
- Butterworth, B. (1980). Some constraints on models of language production. In B. Butterworth (Ed.), *Language production: Vol 1. Speech and talk* (pp. 423–459). London: Academic Press.
- Butterworth, B. (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 108–135). Cambridge: MIT Press.
- Butterworth, B., & Whittaker, S. (1980). Peggy Babcock’s relatives. In G.E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 647–656). Amsterdam: North-Holland.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.
- Clements, G.N. (1985). The geometry of phonological features. *Phonology Yearbook*, 2, 225–252.
- Clements, G.N., & Keyser, S.J. (1983). *CV-Phonology: A generative theory of the syllable*. Cambridge, MA: MIT Press.
- Cohen, A. (1980). Correcting of speech errors in a shadowing task. In V.A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 157–163). New York: Academic Press.
- Collier, R., & Gelfer, C.E. (1983). Physiological explanation of FO declination. In M.P.R. Van den Broecke & A. Cohen (Eds.), *Proceedings of the Tenth International Congress of Phonetic Sciences* (pp. 354–360). Dordrecht: Foris.
- Cooper, W.E., & Eady, S.J. (1986). Metrical phonology in speech production. *Journal of Memory and Language*, 25, 369–384.
- Cooper, W.E., & Paccia-Cooper, J. (1980). *Syntax and speech*. Cambridge, MA: Harvard University Press.
- Cooper, W.E., & Sorensen, J.M. (1981). *Fundamental frequency in sentence production*. New York: Springer.
- Crompton, A. (1982). Syllables and segments in speech production. In A. Cutler (Ed.), *Slips of the tongue and language production* (pp. 663–716). Berlin: Mouton.
- Cruttenden, A. (1986). *Intonation*. Cambridge, UK: Cambridge University Press.
- Cutler, A. (1980). Errors of stress and intonation. In V.A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 67–80). New York: Academic Press.
- Cutler, A. (1981). The reliability of speech error data. *Linguistics*, 19, 561–582.
- Cutler, A. (1988). The perfect speech error. In L.M. Hyman & C.S. Li (Eds.), *Language, speech, and mind* (pp. 209–223). New York: Cumms and Helm.
- Dell, G.S. (1984). Representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 222–233.
- Dell, G.S. (1985). Positive feedback in hierarchical connectionist models: Applications to language production. *Cognitive Science*, 9, 3–23.
- Dell, G.S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321.
- Dell, G.A. (1988). The retrieval of phonological forms in production: Test of predictions from a connectionist model. *Journal of Memory and Language*, 27, 124–142.

- Fay, D., & Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry*, 8, 505–520.
- Ferber, R. (1991). Slip of the tongue or slip of the ear? On the perception and transcription of naturalistic slips of the tongue. *Journal of Psycholinguistic Research*, 20, 105–122.
- Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language*, 30, 210–233.
- Ford, M., & Holmes, V.M. (1978). Planning units in sentence production. *Cognition*, 6, 35–53.
- Fowler, C.A. (1987). Consonant–vowel cohesiveness in speech production as revealed by initial and final consonant exchanges. *Speech Communication*, 6, 231–244.
- Fromkin, V.A. (1971). The non-anomalous nature of anomalous utterances. *Language*, 47, 27–52.
- Fromkin, V.A. (1973). Introduction. In V.A. Fromkin (Ed.), *Speech errors as linguistic evidence* (pp. 11–45). The Hague: Mouton.
- Fromkin, V.A. (1977). Putting the emPHAsis on the wrong sylLable. In L.M. Hyman (Ed.), *Studies in stress and accent* (pp. 15–26). Southern California Occasional Papers in Linguistics. Vol 4. Department of Linguistics, University of Southern California, Los Angeles.
- Fry, D. (1969). The linguistic evidence of speech errors. *BRNO Studies of English*, 8, 69–74. (Reprinted in V.A. Fromkin (Ed.) (1973), *Speech errors as linguistic evidence* (pp. 157–163). The Hague: Mouton.)
- Fuijamura, O. (1981). Temporal organization of articulatory movements as a multi-dimensional phrasal structure. *Phonetica*, 38, 66–83.
- 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.
- Garnes, S., & Bond, Z.S. (1980). A slip of the ear: A snip of the ear? A slip of the year? In V.A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 231–239). New York: Academic Press.
- Garrett, M.F. (1975). The analysis of sentence production. In G.H. Bower (Ed.), *The psychology of language and motivation* (Vol. 9, pp. 133–175). New York: Academic Press.
- Garrett, M.F. (1976). Syntactic processes in sentence production. In R.J. Wales & E. Walker (Eds.), *New approaches to language mechanisms* (pp. 231–256). Amsterdam: North-Holland.
- Garrett, M.F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production: Vol 1. Speech and talk* (pp. 177–210). New York: Academic Press.
- Garrett, M.F. (1982). Production of speech: Observations from normal and pathological language use. In A.W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19–76). London: Academic Press.
- Garrett, M.F. (1988). Processes in language production. In F.J. Newmeyer (Ed.), *Linguistics: The Cambridge Survey. Vol. III: Psychological and biological aspects of language* (pp. 69–96). Cambridge: Harvard University Press.
- Gee, J.P., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, 15, 411–458.
- Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. London, New York: Academic Press.
- Goldsmith, J. (1976). An overview of autosegmental phonology. *Linguistic Analysis*, 2, 23–68.
- Gussenhoven, C. (1984). *On the grammar and semantics of sentence accents*. Dordrecht: Foris.
- Haie, M., & Mohanan, K.P. (1985). Segmental phonology of modern English. *Linguistic Inquiry*, 16, 57–116.
- Harley, T.A. (1984). A critique of top-down independent levels of speech production: Evidence from non-plan-internal speech errors. *Cognitive Science*, 8, 191–219.
- Hart, J.'t, & Collier, R. (1975). Integrating different levels of intonation analysis. *Journal of Phonetics*, 3, 235–255.
- Hart, J.'t, Collier, R., & Cohen, A. (1990). *A perceptual study of intonation*. Cambridge, UK: Cambridge University Press.
- Hockett, C.F. (1967). Where the tongue slips there slip I. *To honor Roman Jakobson: Vol. 2*. The Hague: Mouton. (Reprinted in V.A. Fromkin (Ed.) (1973). *Speech errors as linguistic evidence* (pp. 93–119). The Hague: Mouton.)

- Kaisse, E.M. (1985). *Connected speech: The interaction of syntax and phonology*. New York: Academic Press.
- Kelly, M.H., & Bock, J.K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 389–403.
- Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, 14, 185–209.
- Kowall, S., Bassett, M.R., & O'Connell, D.C. (1985). The spontaneity of media interviews. *Journal of Psycholinguistic Research*, 14, 1–18.
- Kubozono, H. (1989). The mora and syllable structure in Japanese: Evidence from speech errors. *Language and Speech*, 32, 249–278.
- Kupin, J.J. (1982). *Tongue twisters as a source of information about speech production*. Bloomington: Indiana University Linguistics Club.
- Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W.J.M., & Maassen, B. (1981). Lexical search and order of mention in sentence production. In W. Klein & W. Levelt (Eds.), *Crossing the boundaries in linguistics: Studies presented to Manfred Bierwisch* (pp. 221–252). Dordrecht: Reidel.
- Levelt, W., Schriefers, H., Vorberg, D., Meyer, A.S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, 98, 122–142.
- Levitt, A.G., & Healy, A.F. (1985). The roles of phoneme frequency, similarity, and availability in the experimental elicitation of speech errors. *Journal of Memory and Language*, 24, 717–733.
- Lieberman, M., & Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In M. Aronoff & R.T. Oehrle (Eds.), *Language and sound structure: Studies in phonology presented to Morris Halle by his teacher and students* (pp. 156–233). Cambridge, MA: MIT Press.
- Lieberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8, 249–336.
- Lindsay, J.R. (1975). Producing simple utterances: How far to we plan? *Cognitive Psychology*, 7, 1–19.
- Lindsay, J.R. (1976). Producing simple utterances: Details of the planning process. *Journal of Psycholinguistic Research*, 5, 331–351.
- MacKay, D.G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia*, 8, 323–350.
- MacKay, D.G. (1971). Stress pre-entry in motor systems. *American Journal of Psychology*, 84, 35–51.
- MacKay, D.G. (1972). The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology*, 3, 210–227.
- MacKay, D.G. (1982). The problem of flexibility, fluency, and speed–accuracy trade-off in skilled behavior. *Psychological Review*, 89, 483–506.
- MacKay, D.G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer.
- Marslen-Wilson, W.D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10, 29–63.
- McCarthy, J. (1979). *Formal problems in Semitic phonology and morphology*. Doctoral dissertation, MIT, Cambridge, Massachusetts (distributed by Indiana University Linguistics Club, Bloomington).
- Meyer, A.S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language*, 29, 524–545.
- Meyer, A.S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language*, 30, 69–89.
- Meyer, A.S., & Schriefers, H. (in press). 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*.
- Meyer, D.E., & Gordon, P.C. (1985). Speech production: Motor programming of phonetic features. *Journal of Memory and Language*, 24, 3–26.
- Mohanan, K.P. (1986). *The theory of lexical phonology*. Dordrecht: Reidel.

- Motley, M.T. (1973). An analysis of spoonerisms as psycholinguistic phenomena. *Speech Monographs*, 40, 66–71.
- Motley, M.T., & Baars, B.J. (1976). Laboratory induction of verbal slips: A new method for psycholinguistic research. *Communication Quarterly*, 24, 28–34.
- Motley, M.T., Baars, B.J., & Camden, C.T. (1981). Syntactic criteria in prearticulatory editing: Evidence from laboratory-induced slips of the tongue. *Journal of Psycholinguistic Research*, 5, 503–522.
- Motley, M.T., Baars, B.J., & Camden, C.T. (1983). Experimental verbal slip studies: A review and an editing model of language production. *Communication Monographs*, 50, 79–101.
- Mowrey, R.A., & MacKay, I.R.A. (1990). Phonological primitives: Electromyographic speech error evidence. *Journal of the Acoustical Society of America*, 88, 1299–1312.
- Nakatani, L.H., O'Connor, J.D., & Aston, C.H. (1981). Prosodic aspects of American English speech rhythm. *Phonetica*, 38, 84–106.
- Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.
- Nooteboom, S.G. (1969). The tongue slips into patterns. In A.G. Sciarone, A.J. van Essen, & A.A. van Raad (Eds.), *Nomen: Leyden Studies in Linguistics and Phonetics* (pp. 114–132). The Hague: Mouton.
- Ohala, J.J. (1978). Production of tone. In V.A. Fromkin (Ed.), *Tone: A linguistic survey* (pp. 5–39). New York: Academic Press.
- Prince, A. (1983). Relating to the grid. *Linguistic Inquiry*, 14, 19–100.
- Pullum, G.K., & Zwicky, A.M. (1988). The syntax–phonology interface. In F.J. Newmeyer (Ed.), *Linguistics: The Cambridge Survey. Vol. 1: Linguistics: Foundations* (pp. 255–280). Cambridge, UK: Cambridge University Press.
- Rubin, D.C. (1975). Within word structure in the tip-of-the-tongue phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 14, 392–397.
- Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial-order 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: Erlbaum.
- Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In P.F. MacNeilage (Ed.), *The production of speech* (pp. 109–136). New York: Springer.
- Shattuck-Hufnagel, S. (1985a). Segmental speech errors occur earlier in utterance planning than certain phonetic processes. *Journal of the Acoustical Society of America*, 77 (Suppl. 1), S84–85.
- Shattuck-Hufnagel, S. (1985b). Context similarity constraints on segmental speech errors: An experimental investigation of the role of word position and lexical stress. In J.L. Lauter (Ed.), *Proceedings of the conference on the planning and production of speech in normal and hearing-impaired individuals: A seminar in honor of S. Richard Silverman*. ASHA Report, 15, 43–49.
- 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 of language* (pp. 17–51). Hillsdale, NJ: Erlbaum.
- Shattuck-Hufnagel, S., & Klatt, D.H. (1979). The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior*, 18, 41–55.
- Stemberger, J.P. (1982). The nature of segments in the lexicon: Evidence from speech errors. *Lingua*, 56, 43–65.
- Stemberger, J.P. (1983a). *Speech errors and theoretical phonology: A review*. Bloomington: Indiana University Linguistics Club.
- Stemberger, J.P. (1983b). The nature of /r/ and /l/ in English: Evidence from speech errors. *Journal of Phonetics*, 11, 139–147.

- Stemberger, J.P. (1985a). *Phonological rule ordering in a model of language production*. Bloomington: Indiana University Linguistics Club.
- Stemberger, J.P. (1985b). *The lexicon in a model of language production*. New York: Garland Publishing.
- 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. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117–152). New York: Academic Press.
- 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.
- van der Hulst, H. (1984). *Syllable structure and stress in Dutch*. Dordrecht: Foris.
- van der Hulst, H. (1989). Atoms of segmental structure: Components, gestures and dependency. *Phonology*, 6, 253–284.
- van der Hulst, H., & Smith, N. (1986). An overview of autosegmental and metrical phonology. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations. Part I* (pp. 1–45). Dordrecht: Foris.
- Van Wijk, C. (1987). The PSY behind PHI: A psycholinguistic model for performance structures. *Journal of Psycholinguistic Research*, 16, 185–199.
- Warren, R.M. (1970). Perceptual restoration of missing speech sounds. *Science*, 167, 392–393.
- Wells, R. (1951). Predicting slips of the tongue. *Yale Scientific Magazine*, 26, 9–30. (Reprinted in V.A. Fromkin (Ed., 1973), *Speech errors as linguistic evidence* (pp. 82–87). The Hague: Mouton.)
- Yaniv, I., Meyer, D.E., Gordon, P.C., 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.