A Critique of Top-down Independent Levels
Models of Speech Production: Evidence
from Non-plan-Internal Speech Errors*

TREVOR A. HARLEY

University of Dundee, Scotland

A number of speech errors are examined which are difficult to account for by
top-down serial processing models of speech production which have indepen-
dent levels of processing. In particular, most of these errors are characterized
by the presence of an interfering element which is external to the utterance
under current construction. This paper has two main aims: to classify these er-
rors, and to examine the constraints upon them. It is found that phonological
similarity between the target and intrusion is a major determinant of error
occurrence. The consequences for models of speech production are discussed
in a framework consisting of the architecture and control structure of those
models. A particular model to account for these data is proposed, consisting
of a spreading activation lexical network.

For some time, it has been recognized that the study of slips of the tongue can
provide a great deal of information about models of speech production, and
the value of this approach has been confirmed by recent work (e.g., From-
kin, 1973, 1980; Garrett, 1975, 1976, 1980a, 1980b, 1982). This research has
concentrated on what may be called the lower levels of the speech produc-
tion system—those processes involved with the construction of a syntactic

*Acknowledgements: This research is based upon the author’s doctoral thesis at the
Psychological Laboratory, University of Cambridge, England, and was funded by the Social
Sciences Research Council of Great Britain. I am grateful to Margaret Grubb and Aileen Sand-
dilands, who typed earlier versions of this manuscript. I would like to thank for their criticisms
of earlier drafts of this paper Brian Butterworth, Anne Cutler, Andrew Ellis, David Good,
Donald Norman, and two anonymous referees. Any errors, obtusities, etc. remaining are of
course wholly due to me. Correspondence and requests for reprints should be sent to the
author, whose current address is: Department of Applied Psychology, UWIST, Llyn-y-Grant,
Penylan, Cardiff, CF3 7UX, Wales, United Kingdom.
framework, stress allocation, lexical selection, morphophonemic translation, and articulatory planning—and has led to a model which takes as its input a single message (roughly defined as the meaning of the intended utterance) and translates this via various levels of processing, one at a time and in a top-down fashion, into an output consisting of a string of phonemes ready for articulation. This general model is summarized in Figure 1. and can be characterized as a top-down serial processing (TDSP) model which has independent levels of processing.

In this model, speech errors are accounted for by the malfunctioning of one of these processing levels in three possible ways: by the faulty accessing of items for a particular slot at that level (e.g., word substitutions), by the blending of alternative choices for a particular slot at that level (e.g., word blends), or by the misordering of units represented at that level (e.g., spoonerisms). The errors dealt with are explained by reference to processes which occur after the formulation of the intended utterance or message, and thus which are concerned with processing the planned utterance. In practice, there is little difficulty in deciding what constitutes the “intended” utterance; usually it corresponds to a high-level message or semantic specification, although there may also be other levels of intention (e.g., pragmatic).

Butterworth (1982) suggests a classification of errors based on that of Meringer and Mayer (1895) into three categories, depending on the source of the interfering element:

1. plan-internal errors, which are caused by interference between elements comprising the specification of the intended utterance;
2. alternative-plan errors, which are caused by interference from an alternative formulation of the intended thought;
3. competing-plan errors, which are caused by interference from an unintended thought (see also Baars, 1980).

However, because of the difficulty of operationally defining some of these terms, and because the error data to be discussed do not always support these distinctions, it is more convenient to distinguish between plan-internal errors, as defined above, and non-plan-internal errors, which are errors where the source of the intrusion is external to the intended utterance under current construction. Types of non-plan-internal errors, in addition to alternative and competing plan errors, will be discussed. Thus, blends and classical word substitutions (as discussed by Fromkin, 1973) would be characterized as plan-internal, because the determinants of these errors are immediately derived from the specification of the intended utterance; however, this is not the case for all word substitutions, as this paper will show. (By
Construction of a Syntactic Frame
with word slots and prosodic information.

Content Word Selection
Phonological form of content words is retrieved from the lexicon.

Affix and function word formation

Phonetic Segment Specification

Instructions to articulators

Figure 1. The Standard Top-Down Serial Processing Model of Speech Production. Based upon Clark and Clark (1977); Fromkin (1971); Garrett (1980). Levels of processing are independent, and lexical access is seen as occurring by semantic search of phonologically arranged material, so that semantic and phonological information cannot interact, after Fay and Cutler (1977).
“utterance” is meant the clause or sentence for output that is under current construction, and what is “intended” is found by asking the speaker.) The Top-Down Speech Production (TDSP) model is based upon plan-internal error data, and there has been very little discussion of the non-plan-internal types. It is the primary aim of this paper to analyze a corpus of candidates for the class of non-plan-internal errors, and to examine their implications for the TDSP model.

For non-plan-internal errors, it is particularly important to distinguish between the distal cause or determinant of the error which is necessary (but not necessarily sufficient) for that error to occur, and optional facilitators, factors which are not in themselves necessary but which both increase the probability of an error occurring and determine the precise locus of the intrusion, the most important of which is phonological facilitation, or phonological similarity between the intended and produced elements. This distinction is similar to, but not the same as, that drawn by Cutler (1980, 1982), who distinguishes between the distal cause of the slip and the proximal mechanism which is responsible for the actual form the error takes. This can be illustrated by reference to Freudian slips (Freud, 1901/1975). Freud, unlike most modern authors, attributed the cause of nearly all speech errors to categories (2) and (3) above, in particular identifying the source of the intrusion as a repressed unconscious thought. However, the Freudian data appear to show phonological facilitation, so that the actual locus of the intrusion is determined by the phonological form of both the target and the intrusion (thus Freud said that such errors were "over-determined"; Freud, 1975: p. 102). Although the determinant of a Freudian slip is the unconscious thought, it is the facilitating factor which increases the probability of the error being manifested and which determines the precise location of the intrusion.

A demonstration of the existence of phonological facilitation would necessitate some alteration to the TDSP model, because according to that model phonological information should only be available after lexical access, a process which occurs relatively late in the system: for example, in mapping from the functional to the positional level in Garrett’s (1975) model. Yet the determinant of the error for many of the errors to be discussed, and for Freudian slips, is above this level of processing.

Thus, if we find evidence of phonological facilitation, then it will be necessary to explain how this apparent interaction can occur in a TDSP model. Selective reporting and interpretation difficulties in Freud’s corpus make it unsuitable for analysis (Ellis, 1980); therefore, it is also the purpose of this paper to evaluate the evidence for phonological facilitation in a new corpus of errors, with particular attention to the non-plan-internal errors. The availability of phonological information and the interaction of levels in speech production is also discussed by Dell and Reich (1981). This paper
makes three improvements upon their analyses: First, we consider here a much wider range of error types; second, the error corpus has been collected according to stricter criteria; third, a better method for estimating chance rates is used. These points will be discussed in greater detail in the appropriate sections.

The general organization of this paper is as follows: First, we discuss our error corpus and the criteria used for collecting it; second, we provide a classification of errors which may be non-plan-internal, and a discussion of their locus of origin—i.e., their determinants; then, we examine the evidence for phonological facilitation in these errors, and review the consequences for the TDSP model. Alternatives for explaining the data are considered, and a model for accounting for the data is proposed.

THE ERROR CORPUS

The errors to be discussed here are examples drawn from a corpus of (currently) 1129 errors collected by the author at the Universities of Cambridge and Dundee. Each example was witnessed, and recorded as soon as possible after its occurrence with as much of the preceding context and conversation as possible. In addition, the speakers were questioned as to what they thought the cause of the error was, what they had intended to say, what they had been thinking, whether or not they had noticed the slip at the time of speaking, and what was in the environment that might have intruded into their utterance. These points constitute what is referred to as the context of the error. It is also noted whether the speakers corrected themselves. The following abbreviations are used in the text:

$S$ = speaker who made the error  
$X, Y, Z$ = other participants in the conversation  
$U$ = actual utterance  
$T$ = target utterance  
$T_1, T_2$ = targets, if more than one—e.g. blends  
$C$ = relevant context  
$P_1, P_2$ = refers to speakers’ turns  
/ = alternatives in a word blend  
// = noticeable pause

In discussing the following errors, the major problem is classification. It is here that the total context is particularly useful, as it often permits disambiguation between alternative interpretations. In fact, it is the availability of this information which in the first place enables certain errors to be recognized as non-plan-internal, whereas otherwise they would be classified as
plan-internal. For example, in one case a speaker said "Is that box heavy or slow?" when what was intended was "Is that box heavy or light?" Without the context this would be classified as a random semantically related word substitution (both the target and intruding words being concerned with measurement), but this interpretation changes when it is known that the speaker was moving a box labelled 'SLO-cooker'. More context is available than with corpora such as that of Fromkin (1973), thus enabling a better classification of the errors and preventing possible misinterpretations. The lack of context may explain why errors of the type discussed here have received so little discussion elsewhere in the literature.

In general, this methodology is similar to that of Meringer and Mayer (1895), who pioneered speech error research and who recorded a large amount of data for each error. It is still unknown whether some of the variables underlying this data (e.g., speaker's age and sex) have any effect on error occurrence. However, we will concentrate upon contextual information which is relevant to the hypothesized cause of the error, and therefore we will not include context where it is irrelevant. The emphasis will be upon examining the systematicities underlying speech errors to analyze the cognitive mechanisms responsible for speech production. With this aim, speech errors have a number of advantages: they are naturalistic, are not obtrusive prior to the error being made, occur at all levels of processing, they are not cognitively penetrable (Pylyshyn, 1980)—indeed many, especially low-level errors, do not appear to be influenced by social factors (although, see the following and Motley, 1980), do not involve generalizing from abnormal populations, and do show extensive systematicities.

**NON-PLAN-INTERNAL ERRORS**

We will discuss three main classes of non-plan-internal errors, together with a number of subclasses. The main types are the cognitive intrusions (comprising environmental contaminants, topic based errors, and the high-level intrusions), the blending errors, and the content-addressible errors.

1. Cognitive Intrusions

The determinant of these errors is the intrusion of the cognitive representations and procedures responsible for planning speech and other high-level cognitive systems. They can be distinguished from each other by the subjects' reports, but as they are constrained by similar sorts of factors their separation is not always clear, and, consequently, this classification should not be maintained too rigidly.
Environmental contamination. These errors are characterized by the intrusion of a lexical representation of something in the speaker’s environment into the intended utterance, with the result usually being manifested as a word substitution. We give four examples:

1. T: Get out of the car.
   U: Get out of the clark.
   C: As S was getting out of the car he glanced up at a shop front which had "Clark’s" printed on it. He reported that he was not aware of this or thinking about this as he was speaking.

2. T: I read about them this week in NME.
   U: I read about them this week in ATV.
   C: As S was speaking, he glanced at the television where a board with the words “ATV Land, Birmingham” was moving from left to right across the screen. “NME” is an abbreviation for the name of a magazine, “ATV” the abbreviated name of a television company.

3. T: Why not a plain white dress?
   U: Why not a bee?
   C: S was talking about dress-making. The other person was switching the television off, but a she did so the television announcer said, “The next program is ‘Bee in my Bonnett.’”

4. T: They’re doing...
   U: She’s doing...
   C: The speaker was talking while looking at a picture of a woman in the newspaper; however, he said that he was thinking about what he was saying and not about the photograph.

In these errors, we know from the speakers’ reports that the intruding element is irrelevant to the past, current, or projected conversation, and was not something the speaker was explicitly thinking about, or indeed was often something he or she was not aware of. There are examples in the corpus which show that the interfering stimulus may initially be processed by either the visual or auditory channels, and may already have either phonological or orthographic status, or neither.

Garrett (1980a) discussed two similar examples and suggests an explanation in terms of roving attention. Such an account leaves the precise mechanisms involved unspecified, and does not explain how the intruding stimulus attains a phonological representation in the cases where it does not initially possess one. Neither does it specify at what stage in the TDSP model the intrusion occurs.

It is easier first to account for the case where the intrusion is already in a suitable phonological or orthographic form. If the basic mechanism is along the lines of Garrett’s roving attention, we have to explain why so little...
of the environment actually does intrude, as this type of error is relatively rare. One solution is in terms of the facilitators—the intrusion will occur at the first suitable opportunity, and most of the time there is no conjunction between the determinants (e.g., a mechanism like roving attention) and enough suitable facilitators (such that the intrusion can occur in an utterance where there will be a minimum of disturbance). We have to assume that the ability of the determinant to interfere decays over time unless there are suitable facilitators available. This is especially clear with the example involving abbreviations. We will examine the evidence for phonological facilitation in these errors in a later section.

The cases where the intruding stimulus is not already in a phonological form are more difficult to explain, as some mechanism must allow the intruding item to achieve phonological status. In all of these cases, the subjects reported that they were not consciously thinking verbally about the intrusion, so some process other than verbal thought must be responsible. We suggest that the cognitive systems responsible for processing environmental information have access to the lexicon and can prime entries there. This spreading activation network will serve as the basis of our explanation for a number of error types.

It is difficult to locate the locus of the determinant of these errors. They are clearly non-plan-internal, and concerned with processes external to those of the message construction level. In terms of Garrett (1980a), these errors have an early source yet a late consequence. Environmental contamination appears to involve competition between elements, and bears some resemblance to attentional dyslexia (Shallice & Warrington, 1977), and the Stroop effect (Stroop, 1935), in that competing representations somehow combine in the speech stream.

(1b) Topic based errors. The following errors are characterized by the hypothesized intrusion of material from earlier in the conversation (but not in the utterance being planned) into the intended utterance.

(5)  
T:  At least they'll be good for books.
  U:  At least they'll be good for *boxes*.
  C:  S was helping the person he was speaking to load books into boxes, and they were worried about whether or not the boxes would be strong enough to hold them.

(6)  
T:  I haven't a clue what went wrong there.
  U:  I haven't a *cue* what went wrong there.
  C:  S was being taught how to play snooker, and was being told to pay particular attention to how she held the cue. After a particularly spectacular miscue, she said *U*. 
(7) T: I just fancy some tea and muesli.
   U: I just fancy some bacon and muesli.
   C: The person S was talking to had just said, “I don’t fancy anything fried for breakfast.” He said that as he spoke he was not aware of thinking about fried food.

(8) T: Just after Kepler moved, you know...
   U: Just after Prague moved, you know...
   C: S had just been talking about the Thirty Years War and the defenestration of Prague.

(9) T: I’m going to eat a yoghurt.
   U: I’m going to read a yoghurt.
   C: S had just been discussing literature but the word “read” had not been used.

These errors are the result of a perseveration of a lexical item (a word or a related concept) which had either been used earlier in the conversation or is an obvious associate to an earlier concept, with the range of the perseveration being from within the same clause (such errors being conventional whole word perseverations), to some turns or minutes. These errors are mentioned by Meringer and Mayer (1895).

The classification of these errors can sometimes be difficult, because in errors such as (5) and (6) what had been talked about in the past is still present; therefore, there is some danger of confusing these errors with environmental contaminants. The speakers’ introspections provided by the total context enables some means of providing a fairly unambiguous classification, but it is possible that similar cognitive mechanisms are responsible for both types of errors. These errors show a similar range of facilitators to the environmental contaminants.

These errors are clearly non-plan-internal, and the similarities between this type and the environmental contaminants, the fact that such intrusions are always whole words, and the fact that they show phonological facilitation, suggest that they arise at the lexical selection stage. Schenkein (1980) demonstrates that speakers must hold representations of fairly concrete features of prior discourse throughout conversations which can either be incorporated into or used to influence the form of new productions. For example, Schenkein reports conversations in which turn sequences recur in such a way that earlier speech act forms, lexical, syntactic, and prosodic features can be reflected in new utterances. The system must have a record of the past conversation to construct new conversation. If such information about preceding conversation is stored, it is perhaps not surprising that it can accidentally interfere under the appropriate conditions.
(1c) **High-level intrusion errors.** This class represents disparate error types sensitive to similar constraints, but here not even the speakers' reports are reliable at distinguishing them. To some extent it is a residual class of cognitive intrusions (see Norman, 1981).

(10)  
T: I want to cut out the elephant on the back of that.  
U: I want to *cook* the elephant on the back of that.  
C: P was in a kitchen cooking with some other people. He wanted to make conversation but was unsure whether to talk about cooking or about a picture of an elephant on the back of a box in the kitchen.

(11)  
T: Do you want to play tennis tomorrow?  
U: Do you want to play *physical* tennis tomorrow?  
C: S reported that at the same time as speaking he was thinking to himself: "I ought to do some physical exercise."

(12)  
T: I've read all my library books.  
U: I've *eaten* all my library books.  
C: S was hungry and thinking about preparing some food.

(13)  
T: Have you got your car in?  
U: Have you got your *key* in?  
C: S reported that he had also considered adding to the utterance the reason why he wanted to know whether or not X had his car in: because he wanted a lift to fetch a key.  
*NOTE*: Here the target was ungrammatical.

(14)  
T: You eat the kippers?  
U: You eat the *catters*?  
C: X was repeating a story someone else had told her:  
"Y: You don't know how much you hurt me.  
Z: You eat the kippers in the fridge last Thursday."

The speaker whom X was talking to reported that he then thought: "Kippers in the fridge for the cat?" and was wondering whether he was going to ask this in a following conversational turn.

These errors all involve the intrusion of some element that is both outside the primary intended utterance and concerned with message planning; thus, they can be said to be high-level errors. Ilene Butterworth's (1982) classification is inadequate, because not all high-level errors are due to interference between competing messages: high-level errors are not all caused by only one type of determinant. For some errors, the determinant is competition between intended utterances, whereas others are due to interference from inner speech, relevant to the utterance but not intended for output—either because it is involved with planning that utterance or considering what to say later in the conversation.

High-level intrusions are difficult to account for by a TDSP model for two reasons. First, they originate at the message level, and thus contradict...
Garrett's (1982) claim that speech production in general and speech errors in particular can be studied without reference to message level processes. Second, if they show evidence of phonological facilitation, then high-level processes appear to be constrained by low-level factors—in a TDSP model message processing should be complete before phonological information is accessible. This will be discussed later.

2. Content-addressible Errors

There are a number of whole word substitutions where the target and substitution show a semantic relationship which is of a rather more vague type that usually seen in (e.g., Fay & Cutler, 1977). These are errors where the substitution is due to an association rather than a strict semantic relation. Because the relation between the target and intrusion is so tenuous, we will consider these as non-plan-internal, but the lack of a clear definition of the error types and the continuum along which they fall suggests that there is a problem here. As we shall account for all content-addressible word substitutions (including the more common types) in terms of a spreading activation network, then the dichotomy between plan-internal and non-plan-internal types starts to break down. Consider the following examples.

(15) T: They even sleep on the wing.
    U: They even fly on the wing.
    C: S was talking about swifts.

(16) T: Head for the west coast and die of skin cancer.
    U: Head for the west coast and die of sun cancer.
    C: S was talking about moving to a place where the incidence of skin cancer due to sunshine was very high: he reported that he was not thinking of this as he spoke.

(17) T: You could have a tent up.
    U: You could have a flag up.
    C: S was talking about a lodger living in a tent in the garden. S reported that he could think of no reason for the error apart from the association he thought of afterwards: both tents and flags are made of similar materials, often canvas.

(18) T: I tried making some parchment last night.
    U: I tried making some patchwork last night.
    C: S said that she had two crafts, one of which involved making patchwork quilts, the other simulated pieces of parchment.

(19) T: You'll need some portage help.
    U: You'll need some cartilage help.
    C: S was speaking to a friend who was soon due to have a cartilage removed from his knee, which meant that he was not meant to put any
strain on the knee. As he was just about to carry some heavy objects about, S offered to help. From S's reports, we know that there had been no previous mention of the word "cartilage," that he was not thinking about it as he spoke, and he was not planning to say anything about it. Additionally, it is possible that a synonym for "portage," such as "carting," may have competed and therefore facilitated the error; however, the speaker did not report this.

As these errors appear to be related to orthodox whole word substitution (Fay & Cutler, 1977), showing similar types of constraints, and as there is no reason to support that they occur at a different level of processing other than lexical access, then the errors probably reflect something of the organization of the lexicon. The errors suggest that the lexicon contains or has access to information in addition to strict semantic relations and phonological information: There is evidence of access to a much wider range of semantic type information, and of episodic type information (18, 19), and this justifies their inclusion in a discussion on non-plan-internal errors. In fact, these errors are of the type that we would predict to occur if the lexicon were not like a dictionary but if it contained information more appropriate to an entire semantic network, which would also account for the whole-word cognitive intrusion errors. We shall consider this issue in detail at a later point, and we shall also defer our discussion of why items are erroneously substituted.

3. Alternative Plan Errors and Blends

Speakers can make errors because there are two or more alternative ways of conveying the same message, and these alternatives become merged into one utterance (see Brown, 1980; Butterworth, 1982; Fay, 1980, 1982; Meringer & Mayer, 1895). In general, this class of errors represent the range of blends from the word level, where the lexicon outputs two alternative items for one lexical slot, through MacKay's (1972) synonymic intrusions at the syntactic level, where alternative syntactic formulations of the same message blend, to Butterworth's class of alternative plan errors, where both syntax and lexical choice differ for the same intended message. Thus word blends and syntactic blends will be plan-internal, whereas alternative plan errors involve different specifications and are therefore non-plan-internal. We give examples of all classes.

(3a) Word blends.

(20) fire/flames—flire
(21) shut/locked—shlocked
(22) hypothesis/syndrome—hydrome
(23) heard/knew—herknew (/hanju/)
(3b) Syntactic blends.

(24) T1: It depends where they place their limits.
T2: It does depend where they place their limits.
U: It depends where it does place their limits.

(25) T1: The disease will run its course.
T2: The disease must run its course.
U: The disease will rust its course.

Note: There has been some adjustment here, as the blend is between two sentences which have not run exactly in parallel.

(26): T1: What are they?
T2: What are those?
U: What are they those?

(27) T1: You feel ever so good when you do work.
T2: One feels ever so good when one does work.
U: You feel ever so good when one does work.

(3c) Alternative plan.

(28) T1: I got up at 8:52.
T2: I felt fine at 8:52.
T3: I woke up feeling fine at 8:52.
U: I felt up fine at 8:52.

(29) T1: The sky is blue.
T2: The sun is shining.
U: The sky is shining.

(30) T1: They quail every time I phone.
T2: They shake in their boots every time I phone.
U: They quail every time I get in my boots.

(31) T1: They’re really losing money/profits.
T2: They’ve really got problems.
C: S reported that the primary target in T1 was “money,” but he also thought of “profits”; this caused a hesitation.
U: They’re really losing // problems.

Word blends are already well documented in the literature (see, e.g., Fromkin, 1971; Garrett, 1980a; MacKay, 1972). Shattuck-Hufnagel (1979) accounts for them by positing the confusion of two lexical outputs in a buffer. Thus, we should note that there is at least some parallelism implicit in the TDSP model as formulated by both Fromkin (1971) and Garrett (1976), because both accounts see items output in parallel from the lexicon. We still have to account for the failure of the lexicon to output a single item, and the most plausible explanation is that the input to the lexicon is underspecified, so that, for example, if the input consists of a list of semantic features, this list is not sufficient to discriminate between two items which both satisfy this input, and therefore both are output. In some cases (e.g., 22), there is
no obvious strict semantic relationship between the two alternatives, so that the underspecification or sign-flipping of semantic features (e.g., + FEMALE for + GIRL or − MALE for + MALE) cannot be an adequate explanation. In these cases however, the two items appear to be interchangeable in that context (see also Fromkin, 1971).

Thus, we can propose four sources for a blending error:

1. Two items both of which satisfy the input criteria (i.e. the message level specification) are simultaneously output by the lexicon;
2. A type of alternative plan error where in a particular conversational context there are two alternative ways of saying the same thing so that there are two semantic specifications generated for one lexical slot;
3. Where the semantic specification selects only one semantically appropriate item, but a phonologically related item is also output; these might be expected to occur but if they do they are very rare. The only possible example in the corpus is (32):

(32) anachronism + acronym − anacronym

An alternative interpretation is precluded by the subject's report that as he was thinking of the target, "acronym" intruded.
4. Where blending occurs due to the attempted correction of an error (e.g., word substitution). The only example in the corpus is (33):

(33) You keep getting poor averages—harvages—harvests.
See also Stemberger (1982).

Type (3) contradicts the Fay and Cutler (1977) model, because, on that account, if the selection procedure reaches the target, the processing would stop; if, on the other hand, it reaches the phonologically related word first, then why is this not rejected? Semantic search is concerned only with semantic information. The model of Butterworth (1982) suffers from similar problems; those of Dell and Reich (1981) and Stemberger (1983) do not.

Meringer and Mayer (1895) and Butterworth (1982) also describe alternative plan errors—errors where two ways of conveying the same message but using different syntactic constructions and lexical items blend. Butterworth (1982) considers the extent to which Fromkin's (1971) or Garrett's (1976) model can accommodate the type of parallelism necessary to explain these errors within the framework of a TDSP model, and concludes that, as it cannot, such a model is unsatisfactory. Thus, to account for alternative plan errors we would have to drop the serial processing requirement, and permit processing of parallel representations from the syntactic level (and before) to the phonological level. This is because, as Butterworth notes, the locus of blending for syntactic blends is apparently determined by phono-
logical and not syntactic criteria. (In word blends, however, there are some syntactic constraints: the word class of the two blending words tends to be the same.)

However, there are two main objections to revising the TDSP model in this way (Butterworth, 1982). First, there could be a combinatorial explosion which would result in an unacceptable proliferation of representations throughout the various levels of processing. Second, as blending errors could occur at each stage, we should then find a range of otherwise unpredicted errors. There may be a number of constraints on sentences and clauses, such as Garrett's (1980) upon grammatical types and clause boundaries, and Butterworth argues that, to preserve sensitivity to these constraints, the representation of the competing utterances in parallel must be very different from that of two sentences intended to be spoken successively. Both of these criticisms have rejoinders. For the first objection, we can point to the fact that, as the probability of an error occurring is very low, we are unlikely to encounter many situations where there are more than two simultaneously competing strings, and in any case, the number permissible will only be limited by the architectural capacity of the system. For example, there may be a capacity limitation such that more than two or three competing representations cannot be maintained in parallel. Thus, no combinatorial explosion is necessary.

The second objection, that of the nature of the competing representations, can be answered by reiterating that the architecture of the system (briefly, the stable global aspects of the system) is unknown. Thus, there are no known relevant constraints upon the nature of the representation of maintained parallel alternatives. In any case, we find a range of errors which suggest that interaction between competing representations is indeed to some extent unconstrained. Consider example (34) from Butterworth (1982).

(34) 

T1: I really hate to get up in the morning.
T2: I really like to stay in bed in the morning.
U: I really like to—hate to get up in the morning.

If T2 is in some manner inhibited so that it can only be manifested by a range of Fromkin-type mechanisms (word and segment movements), as Butterworth argues, we should expect that the following range of error outcomes could be derived:

(35) 

U: ...really stay to get up...
    ...have to stay up...
    ...state to get up...
    ...to get bup...

Nevertheless, some such errors are indeed found: for example, (28) is from the author's own corpus. Thus, Butterworth's claim needs revision. In gen-
eral, cases where alternative syntactic constructions blend without regard to syntactic constraints or sentence position are difficult to account for both by Fromkin and Garrett, and by Butterworth, who predicts that:

...if such errors do occur, then the model will have to be radically modified...to preserve this constraint [segment movements rarely cross clause boundaries—Garrett, 1975] it must be assumed that the representation of (T1 and T2) is quite different from the arrangement of two sentences intended to be spoken successively.

(Butterworth, 1982; p. 84)

But this should not be unexpected, as the two alternatives are intended to be simultaneous, not successive. There is no reason to assume any radical modification apart from the two representations at each level being stored separately but with access to one another.

The argument about whether or not the speech production system can support parallel processing will eventually be resolved by our knowledge of the architecture of this system. There are three possible ways of accounting for apparent parallelism (see Harley, 1984):

(a) True parallel processing, where more than one process is running simultaneously. This is of two subtypes: parallel processing by the same system (e.g., two syntactic processors or a massively parallel lexicon), or the parallel operation of, say, the syntactic and semantic systems;

(b) A time sharing central processor which switches rapidly between processing the alternatives, so that at any one time the system is only actually processing one input, yet seen over a longer period of time more than one alternative appears to have been processed;

(c) The serial processing of two or more alternatives within each level of processing, holding the result of the first application of the process in a buffer until the completion of the second, only after which can processing at the next level commence; we shall refer to this as “concurrent processing.”

Anderson (1976) has suggested that unless we make certain assumptions we cannot distinguish on behavioral grounds alone between serial and parallel processing of types (b) and (c) (but see Pylyshyn, 1980). This is not the case if we know the architecture of the system. The architecture can be seen to impose limitations on the processing due to the hardware; for example, if the speech production system has only one central processor, we can rule out true parallel processing (a) as an alternative. It describes the relatively stable, global aspects of the system, referring to a level of description incorporating the software or hardware, or both, as necessary. Currently,
there are no data upon this point. We cannot rule out the possibility that true parallel processing does in fact occur in speech production.

Let us now consider the possible locus of blending for alternative plan errors. Butterworth (1982) demonstrates that blending of this type is conditioned by phonological similarity. However, processes other than simple merging appear to be involved. Consider again (30):

(30)  
T1: They quail every time I phone.  
T2: They shake in their boots every time I phone.  
U: They quail every time I get in my boots.

(Note that if the subject sometimes does not or cannot report all the information available, an alternative T1 which gives a better account of this error may be "They quail every time I get on the phone.") This sentence respects syntactic constraints: for example, "their" is changed to "my," "The... every time I phone" is shared, and editing and accommodation occurs so that "get" is inserted appropriately into the final sentence. This suggests that there is either late editing, or interaction between levels, or that the blending occurs at a high level. If Butterworth (1982) is correct and such errors are conditioned phonologically, then we have two alternatives: either blending actually occurs at the phonological level, in which case the two alternatives must be processed in parallel from the message level, or the solution is one of a different control structure, so that high level processes are sensitive to low level descriptions. Butterworth (1982) discusses the evidence for each of these alternatives, and concludes that a hierarchical control structure is possible and economical.

NON-PLAN-INTERNAL ERRORS: DISCUSSION

We have examined a number of errors where the error determinant is external to the elements of the specification of the intended utterance. At the very least, the TDSP model will need extension to explain these error types. We can no longer explain all speech errors without reference to message level processes, and the high-level intrusion errors suggest that we cannot discuss the speech production system isolated from other cognitive systems. Nevertheless, as the data currently stand, we need only to postulate a few more processes: without a formal demonstration that these error types are constrained by apparently plan-internal mechanisms, there is no special problem for the model. We therefore intend to show that these error types can be phonologically facilitated. Furthermore we hope to demonstrate that phonological facilitation is also found in plan-internal errors.
PHONOLOGICAL FACILITATION

To demonstrate that phonological facilitation occurs, that is to demonstrate that errors which are determined (i.e., caused) at a level of processing above the phonological level are conditioned by phonological similarity between the target and intrusion, we need to show that the probability of a speech error occurring is greater if the target word and intruding or substituting word are phonologically more similar (according to suitable criteria) than we would expect by chance. Therefore, we have first to establish suitable criteria, and then to evaluate what the chance rate would be.

The first is the easier. In the following we will take three measures of phonological similarity, well-known in the literature for their importance in lexical phonology (e.g., determining output in the tip-of-the-tongue state: Brown and McNeill, 1966; Dell & Reich, 1981; Fay & Cutler, 1977):

(a) first consonant in the stressed syllable (in most cases the initial phoneme);
(b) equal number of syllables with identical stress pattern;
(c) cases which share both (a) and (b).

The second point, that of estimating the probability of two words being similar according to these criteria, is notoriously difficult in speech production. We have adopted the following method. It is argued that this is one of the best available methods because it is based on the same sort of data base as the error corpus, and provides a very conservative estimate; therefore, it is likely that the probabilities generated have a wide range of applicability. A number of transcripts were made from tape-recordings of natural conversations between speakers in Cambridge, England, where most of these errors were recorded, with in fact the speakers being responsible for many of the errors in the corpus. These transcripts were used to generate a list of all the content words in these transcripts. Types were taken rather tokens (each word was recorded on the list only once, but this actually makes a negligible difference). The words were classified on their occurrence as noun, verb, or adjective in the conversation.

To measure (a), the initial stress syllable consonant was obtained for each word in the list. As the three lexical classes did not show any significant difference ($\chi^2 = 40.9, 38$df, $p > .3$, ignoring those cells containing expected frequencies of less than 5), the three classes were combined to give a total pool of 1035 words. The frequencies were then converted to probabilities for each consonant occurring, giving the greatest probability of 0.122. As the most conservative possible estimate, this figure will be used throughout. As an additional check, words were randomly paired from the list, and this gave a mean probability of .06 for 2908 pairings. Hence we take $p$(initial stress syllable phoneme) = .122.
To measure (b), random pairings of noun-noun, verb-verb, and adjective-adjective were taken from the master list of all words, and the number of matches for identical stress and syllable pattern recorded. Again, there was no significant difference between word class ($\chi^2 = 4.92, 2\text{df}, p > .05$), the results were combined. This gives a chance probability for (b) of almost exactly one third, $p(\text{identical stress and syllable pattern}) = .33$.

Finally, to measure (c), the probability of two words sharing the same initial consonant in the first stress syllable and having the same stress and syllable pattern, random samples were taken from the word lists. This gave a chance estimate of $p = .025$. Full details of these analyses are available from the author.

If we use these chance estimates to analyze the above error types, we find the probability of the phonological facilitation occurring by chance to be as follows (see Table I).

<table>
<thead>
<tr>
<th>Error type</th>
<th>(a) 1st cons base ($p = .122$)</th>
<th>(b) stress ($p = .33$)</th>
<th>(c) both ($p = .025$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environmental</td>
<td>N.s ($N=2$)</td>
<td>&lt; .02 ($N=15$)</td>
<td>N.s ($N=1$)</td>
</tr>
<tr>
<td>Contamination</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topic Based</td>
<td>&lt; .02 ($N=12$)</td>
<td>&lt; .02 ($N=26$)</td>
<td>&lt; .001 ($N=9$)</td>
</tr>
<tr>
<td>High level Intrusion</td>
<td>.08 ($N=7$)</td>
<td>&lt; .01 ($N=21$)</td>
<td>&lt; .01 ($N=5$)</td>
</tr>
<tr>
<td>Content Addressible</td>
<td>&lt; .001 ($N=9$)</td>
<td>&lt; .01 ($N=18$)</td>
<td>&lt; .001 ($N=8$)</td>
</tr>
</tbody>
</table>

Classification is by error type with $N$, the number of errors in each error type, showing the probabilities used for calculating values for (a) two words sharing the same first consonant; (b) having equivalent stress pattern; (c) sharing (a) and (b). P-values were calculated using $\chi^2$ except where expected value $<5$, in which case the exact binomial has been used.

Thus it can be seen that there is considerable evidence for phonological facilitation in these non-plan-internal errors. This finding cannot be predicted by the TDSP model, because, according to this account, phonological information should only be available at the lexical access stage—well after the locus of the determinant of the error.

Combined Phonological and Semantic Errors

Phonological facilitation may also be found in plan-internal errors. There is another class of errors where we should not expect phonological facilitation to occur, and these are the plan-internal semantically based word substitu-
tions (Fay & Cutler, 1977). However, there appear to be word substitutions where the intruding word is both semantically and phonologically related to the target word (see also Aitchinson & Straf, 1982; Butterworth, 1982; Cutler, 1982; Dell & Reich, 1981; Shallice & McGill, 1978). The following are examples from the current corpus.

(36) electronic — electric
(37) directory — dictionary
(38) poke — peek
(39) bought — brought
(40) petrol — paraffin
(41) chromatography — crystallography

We will now establish whether or not there are statistically significantly more word substitutions that have combined semantic and phonological determinants than we would expect by chance. First, we need to clarify what we mean by two words being semantically similar. We will use two definitions of semantic similarity, strict and loose. Words will be classified as semantically related in the strict sense if they meet one of three conditions (see also Dell & Reich, 1981): (1) if they are antonyms (e.g., dry-wet), (2) if they are in coordinate relationship with each other (e.g., East-West), or, (3) if they are in direct super-ordinate relationship (e.g., poodle-dog). The loose sense is rather more free: in addition to these criteria, the words will be counted as semantically related if they are either synonyms or closely associated (e.g., determination-dedication). Second, using the above criteria (a) to (c), we need to be able to evaluate the phonological similarity of two words given that they are semantically related. To obtain the chance rate for these, the word lists discussed above were scanned for pairs of words that were semantically related according to the strict and loose definitions of semantic relatedness. In this way we obtain chance probabilities as follows: for the loose definition, for 150 pairs, (a) \( p = .087 \), (b) \( p = .407 \), (c) \( p = .047 \); for the strict definition, for 82 pairs, (a) \( p = .122 \), (b) \( p = .549 \), (c) \( p = .073 \).

A number of instances of combined errors were deleted from the analysis in case they should artificially inflate the significance levels. These included errors where the substituting word was a morphological derivative of the target, and where the only likely candidate for a semantic substitution is phonologically related. The evidence for phonological facilitation in these errors is shown in Table II.

Thus this evidence supports Dell and Reich (1981): there are a greater number of whole word substitutions in which the target and substitutions are related than could be expected by chance alone.

These errors can be seen as either phonologically-facilitated semantic errors, or semantically-facilitated phonological errors. Therefore, as it is not clear here what is the determinant, we can view both semantic and phonological factors as making equal causal contributions to the error outcome.
TABLE II
Evidence for combined Phonological and Semantic Determinants in Whole Word Substitutions

<table>
<thead>
<tr>
<th>Relation</th>
<th>N</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1st cons.</td>
<td>=stress</td>
<td>=both</td>
</tr>
<tr>
<td>Loose</td>
<td></td>
<td>(.087)</td>
<td>(.407)</td>
<td>(.047)</td>
</tr>
<tr>
<td>Errors</td>
<td>127</td>
<td>&lt; .001</td>
<td>&lt; .01</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Strict</td>
<td></td>
<td>(.122)</td>
<td>(.549)</td>
<td>(.073)</td>
</tr>
<tr>
<td>Errors</td>
<td>62</td>
<td>&lt; .001</td>
<td>N.s</td>
<td>&lt; .04</td>
</tr>
</tbody>
</table>

Probabilities are given for evaluation by both the loose and strict criteria of semantic relatedness (see text). Thus the table shows the evidence for phonological facilitation in semantically related substitutions according to criteria (a), (b) and (c). (See text and Figure 1). The base probabilities (for estimating chance) are also given.

Fay and Cutler (1977, p. 516) conclude that "the evidence from malapropisms argues strongly that there is indeed just one mental dictionary" for both speech production and comprehension, and for phonological and semantic information, and that this lexicon is arranged according to phonological principles (such as syllable structure, stress pattern, and initial phoneme), which is accessed in production by means of an address system of semantic features. In this model, semantic errors are due to either an underspecification of the semantic features, or an incorrect branching at some point in the accessing network, while phonological substitutions occur when the semantic addressing system points to a particular word, but then for some reason the system mistakenly selects a close neighbour in the lexicon. Because the lexicon is phonologically arranged, this word will be a phonologically related item. Because the relationship between sound and meaning in a language is virtually completely arbitrary, the semantic and phonological information should on this account be quite independent, and therefore combined errors cannot be accounted for. In addition, Fay and Cutler (1977, p. 516) admit that their proposal of how semantic addressing occurs in production is "quite unspecific, and the intent is only to provide an example of how the addressing might be accomplished to contrast it with access in comprehension."

Cutler (1982) has suggested that the semantic similarity between two phonological near-neighbours increases the probability of a malapropism occurring, because the semantic access of the intended word might also prime its neighbour (in some unspecified way). The way in which the lexicon is accessed semantically is unknown. In accessing a particular unknown word, it is postulated that the semantic address runs through a list of features which would end up by pointing to a space in the lexicon, and it is at this stage that a combined error would occur: prior to this point, according
to the principle of independent processing, the selection has been via semantic processes only, without reference to phonological information.

However, according to this model the semantic accessing must by now have primed all those other items which are semantically related to the target word. But, at this point, processes which would normally result in an ordinary malapropism occur: for some unknown reason, a near neighbour is erroneously output, and as the lexicon is phonologically arrayed, the output will be phonologically similar to the target. But, to explain the combined errors, the item which is mistakenly selected by this last stage must be that which is semantically primed. Therefore, the final accessing stage must occur via some property which the target and primed neighbour share. It is argued that this account amounts to the spreading activation model discussed below, but has none of the advantages of that model for accounting for the other error types and the actual processes involved.

Butterworth (1982) has proposed a model that differs from that of Fay and Cutler by dividing the accessing during production into two separate stages, so that accessing occurs via two independent lexicons, the first being accessed by a semantic specification and being arranged semantically, the output being a pointer to an item in the second lexicon, which is arranged phonologically. (Similar to the Fay and Cutler model, the same lexicons are used for both speech production and comprehension.) In accessing the semantic lexicon, semantic errors are made; whereas phonological errors arise in the accessing of the phonological lexicon. But, to explain the greater-than-chance number of combined errors, a further innovation is necessary, and in the Butterworth (1982) model this takes the form of an editor, operating in the fashion of that of Baars, Motley, and MacKay (1975). The editor checks the output phonology of the phonological lexicon, and rejects those errors which fail to satisfy the appropriate criteria; however, to explain why we ever get any errors we have to specify that often the editor fails to work: for example, we do get a not inconsiderable number of pure semantic substitution errors. Butterworth suggests that the primary mechanism of the editor is that the phonological stage is run through twice, and the editor detects any mismatch between the two outputs. However, the editor may be quite inefficient, because it fails to detect a large number of errors, and the phonological criteria it is using for acceptance cannot be a perfect match, because again then we would have to predict that either we find phonological errors or semantic errors, but again not both. Therefore, it must be the case that the editor checks for phonological similarity between the two processing runs, and not necessarily a perfect match. The manner in which this matching could proceed is unspecified. For further criticism of the filter model, see Dell and Reich (1980), and Stemberg (1983). The argument is based on simplicity, that editors are unnecessary, rather than on any logical or empirical grounds. It is difficult to envisage what data could either falsify or verify the editor model (Motley, Camden, & Baars, 1982).
Thus, it can be seen that no simple revision to the standard model can account for the existence of errors with combined determinants. It would be parsimonious if we could account for these errors in a similar way as for phonological facilitation of nonplan-internal errors.

**DISCUSSION**

The existence of non-plan-internal errors and of phonological facilitation demonstrates the unsatisfactory nature of TDSP models of speech production as currently envisaged. The error corpus discussed suggests that, in addition to the plan internal mechanisms usually discussed in the standard accounts, any model of speech production must also be able to account for the following data:

1. Irrelevant material from the environment can interfere with the construction of the current utterance. The material is often (but need not be) in orthographic form. (Examples 1–4)
2. Words or concepts used or especially prominent in the preceding conversation can intrude into the utterance. (Examples 5–9)
3. If there is more than one way of satisfying a conversational goal, then two different messages satisfying that goal can blend into a single utterance. (Example 10)
4. Internally verbalized thought can intrude into an utterance. (Examples 11–14)
5. If there are alternative ways of successfully conveying the same message or achieving the same pragmatic intent, these alternatives can blend. (Examples 20–31)
6. A much wider range of content-addressible errors than is usually recognized can occur, supplementing the strict semantic errors with a range based upon loose semantic relations, association, and personal encyclopaedic knowledge. These are affected by similar constraints, and cannot be distinguished as types on the current data, suggesting that they are either unitary data types or are processed by similar mechanisms. (Examples 15–19)
7. In addition to the above determinants of error types, the phonology of the target affects the probability of an error occurring and its locus. This has been called phonological facilitation.
8. There are more whole-word substitutions conditioned by both semantic and phonological factors than would be expected on the basis of chance alone.

These demands can be collapsed so that a speech production model must, in addition to the usual constraints,
(a) Explicate high level processes such that more than one message can be constructed at any one time, and so that it is possible to plan the content of the intended message in more than one way. Hence, we need to include in our model of speech production some account of how messages are constructed, represented, and selected, and for this reason this will be referred to as the "high-level processes" requirement.

(b) The "phonological facilitation" requirement demands that the model explains phonological facilitation and accounts for the existence of combined errors.

Initially, requirement (a) places no restrictions on the lower levels of the speech production system, and therefore it might be thought that we could save the TDSP model by extending it so that we are also concerned with the formulation of the message. However, these errors, as well as various plan-internal errors, can also be phonologically facilitated. The intrusion or substitution often occurs when the phonological form of the intrusion is similar to or most closely resembles that of part of the phonology of the utterance. However, in a TDSP model, alternatives at the message level are eliminated before the phonological level and therefore should not be sensitive to descriptions at that level. We discuss a number of ways of explaining this (see also Harley, 1982).

First, we could modify the control structure of the speech production system: that is, we could alter the order in which processing proceeds, or which other levels a particular level has access to, the control of conditional branching, and so on. Butterworth (1980, 1982) prefers this solution, so that the top-down serial chain structure is replaced by a more complex one—in this case, a hierarchy where higher levels can access lower levels directly. Moreover, the order of processing may not always be fixed (see also Harley, 1982), so that, depending on the circumstances, particular phrases may have extra processing priority—thus, important concepts of the utterance can be translated into phonological code at the message stage, before the generation of the syntactic frame has occurred. Nevertheless, the model is still top-down, although a chain of control has been replaced by a hierarchical arrangement of the systems (see Butterworth, 1980). There is no evidence which motivates altering the basic top-down structure (but see Harley, 1984).

Second, the problem could be associated with the architecture of the speech production system—the capacity constraints and structural limitations which are placed on the model before the control structure can be considered. Alternatively, the architecture can be seen as the logical arrangement of the machine from the point of view of the software run on that machine, or the fixed capacities of the system (see Pylyshyn, 1980). Thus, we could permit parallel processing and the proliferation of outputs and
representations. Whereas modifying the control structure enables us to abandon the top-down requirements, altering the architecture changes the serial processing nature of the standard model. However, in modifying the hypothesized architecture, because of the proliferation of processors at each step, the control structure will necessarily become more complex, and will specify where the blending of outputs of each level can occur and whether interference can occur within a level. Therefore, this solution remains neutral as to whether or not interaction can occur between levels. Modifying the production architecture is very powerful, because it is difficult to constrain the proliferation which occurs when talking about parallel processing with a set of processors and an output buffer for each stage.

Third, we can postulate that the lexicon does not consist of a passive list-like arrangement of items, but is an active network where activation can spread between nodes, in the manner of Dell and Reich (1980, 1981); see also Harley (1982). Lexical entries are activated or primed if they, or a related concept, are accessed during production, and following this need less subsequent input to attain the threshold criterion. Following their initial activation, the amount of activation decays as it returns to its base level. For example, graphological input will prime all entries which are closely related to the target concept node either semantically, phonologically, orthographically, or via episodic memory, by spreading activation through the interlexical links. Thus environmental contamination is explained if “roving attention” (Garrett, 1980a) activates the appropriate nodes in the network, and topic based errors are accounted for by the base level of entries for some reason not returning to their base level sufficiently quickly. Facilitation is accounted for by activation spreading along the corresponding links. Obviously, the parameters and values (for example, the rate of decay of activation) remain to be investigated. It should now be clear that the above distinction between plan-internal and non-plan-internal errors is not germane when dealing with content-addressible errors, because it is hypothesized that, for these errors, the determinants are indistinguishable from the facilitators, so that whatever relevant input is available to the system may increase an item’s activation level.

To account for the data, the minimum change to the TDSP model that would be necessary is to incorporate the spreading activation lexicon. But, in addition, we have to account for the proliferation of parallel processing in and from the message level, so that the high level processes requirement is satisfied; this also necessitates changing the control structure or architecture in some way.

We now examine a model that can account for this data and which satisfies requirements (a) and (b) (see Figure 2).

To this model, we need to add a set of processes which are capable of formulating a relevant and fairly detailed specification of an utterance in a
Figure 2. A model of speech production. The message-semantic, syntactic, and lexical modules are acting asynchronously in parallel in cascade. As soon as any output is available, it is passed on to the next system. Apparent interaction (facilitation) is accounted for by the spreading activation lexicon. The output of this is a sequence of activated phonemes (see Dell & Reich, 1980). For further detail, see Harley (1984).
conversation. The content of the utterance will depend upon the goals the speaker is attempting to satisfy with that utterance (see, for example, Cohen & Perrault, 1979; Schank & Abelson, 1977). Thus, in casual conversation the speaker may wish to say something which is interesting and relevant (see Grice, 1975, 1978; and, for some more specific mechanisms, Schenkein, 1980), so that the intended message of the utterance being planned should satisfy these goals. The production system should then either eliminate or temporarily suspend all but one of these alternatives, and failure to do this could result in a blending error. Alternatively, competing elements could prime lexical entries if the high level has direct entry to the lexicon via semantic links, as indeed it must have. It is at this stage that competing plan errors arise. There may be more than one way of formulating the same message so as to satisfy the stated goals, and the system has to choose between these alternatives; failure to do this could result in alternative plan blends, as the architecture supports parallel processing. All propositions must be correctly specified, and can prime other entries in the lexicon via spreading activation. At this stage, lower-level processes take over, as in, for example, a refined version of Butterworth (1980).

CONCLUSIONS

It has not been the aim of this paper to provide a complete model of speech production; that would be an enormous undertaking. With our present level of knowledge, there are a number of alternative explanations, with insufficient evidence to decide between them. It has been our principal aim to demonstrate that a top-down serial processing model needs considerable refinement to account for natural language production. We have discussed a number of errors which have their locus of causation outside the utterance under current construction, and have shown that there is a distinction between the determinants of errors and factors which facilitate their occurrence. Future theories should pay more attention to these higher levels concerned with the construction and form of the intended message, and consider alternatives to top-down serial chain models. Furthermore, data collectors are going to have to take more notice of what information speakers have to offer about the errors that they make.

REFERENCES


TOP-DOWN MODELS OF SPEECH PRODUCTION


