A Dual System for Producing Self-Repairs in Spontaneous Speech: Evidence from Experimentally Elicited Corrections

CAREL VAN WIJK AND GERARD KEMPEN

University of Nijmegen, The Netherlands

This paper presents a cognitive theory on the production and shaping of self-repairs during speaking. In an extensive experimental study, a new technique is tried out: artificial elicitation of self-repairs. The data clearly indicate that two mechanisms for computing the shape of self-repairs should be distinguished. One is based on the repair strategy called *reformulation*, the second one on *lemma substitution*. W. Levelt's (1983, *Cognition*, 14, 41–104) well-formedness rule, which connects self-repairs to coordinate structures, is shown to apply only to reformulations. In case of lemma substitution, a totally different set of rules is at work. The linguistic unit of central importance in reformulations is the major syntactic constituent; in lemma substitutions it is a prosodic unit, the phonological phrase. A parametrization of the model yielded a very satisfactory fit between observed and reconstructed scores. © 1987 Academic Press, Inc.

How do speakers correct inappropriate or erroneous utterances they produce spontaneously? In this paper we develop a new theoretical approach to this process based on an original technique: experimental elicitation of self-repairs. It allowed us to verify Levelt's (1983) well-formedness rule, which establishes a connection between self-repairs and syntactic coordinate structures. More important, however, is the discovery that, in addition to a syntax-based correction mechanism, there exists a second mechanism which hinges on a prosodic unit called phonological phrase (cf. Nespor & Vogel, 1982, 1983).

The paper is organized as follows. First, we give an overview of the current state of psycholinguistic theorizing on language production, in particular with respect to self-repairs (Sections 1 and 2). Then we present the new technique and the data it generated in an extensive experimental

study (Sections 3 and 4). Section 5 is devoted to a description of our model. Further empirical support gained from the experiment is reviewed in Section 6. The final section outlines some theoretical implications. Details concerning our taxonomy of self-repairs, and the parametrization of the model are presented in separate appendices.

1. A PSYCHOLINGUISTIC MODEL OF LANGUAGE PRODUCTION

In the 1970s, theorizing about the mental processes underlying language and speech production has progressed rapidly, most notably through the work of Merrill Garrett (1975, 1980). The model he put forward has gained sufficient influence to be called the standard model of language production (see Bock, 1987, for a discussion on some of its basic tenets). In this paper, too, we take it as a starting point. We first present the model briefly, using the terminology proposed by Kempen and Hoenkamp (1987).

Sentences are produced by the four modules or stages depicted in Fig. 1. They have access to the mental lexicon, and their outputs are continuously watched by a central monitor. The first, conceptual, module prepares the meaning contents to be conveyed to the listener. These conceptual representations are assumed to be nonlinguistic—i.e., language independent—to a large extent. The second, lexico-syntactic, module converts conceptual input into syntactic form by building functional linguistic structures (surface syntactic trees). Their terminal nodes are *lemmas,* i.e., abstract lexical items not yet containing any sound information (Kempen & Huijbers, 1983). The third, morpho-phonological, module retrieves and processes the phonological forms (*lexemes*) associated with the lemmas; it makes morpho-phonological adjustments and computes an intonation contour spanning the whole utterance (Van Wijk
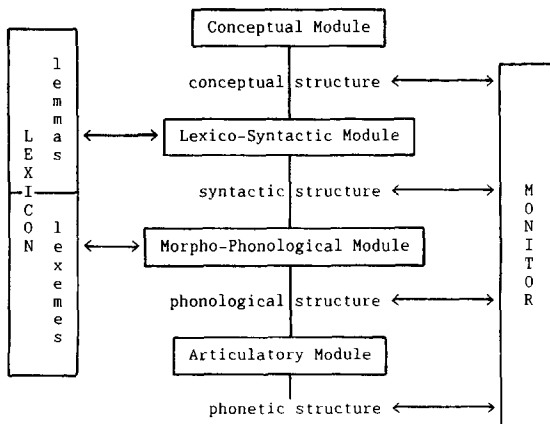

FIG. 1. A global model of language production.

& Kempen, 1985). Lemmas as well as lexemes are looked up in the mental lexicon. The fourth, articulatory, module transforms the phonological representation of an utterance into a phonetic one which controls the activity of the speech organs. The monitor is a supervisory agent observing the flow of information between modules, capable of intervening when certain special events take place. Needless to say, the monitor has no linguistic knowledge of its own and exerts control merely by sending and receiving messages.

The major source of evidence for the model has been the analysis of speech errors produced spontaneously during everyday conversations (see Cutler, 1981, for a discussion on methodological issues). It was not until recently that the natural companion of errors, their repairs, received closer attention in the literature.

2. LEVELT'S WELL-FORMEDNESS RULE FOR RETRACING REPAIRS

Levelt (1983) has proposed a classification of self-repairs in terms of their cause. The major types he distinguishes are *appropriateness repairs* (A-repairs) and *error repairs* (E-repairs). In case of an A-repair the speaker substitutes an utterance which is adequate in itself, by one which she judges more appropriate semantically or pragmatically. In (1), for instance, the repair leads to specifying a location more accurately.[1] An E-repair serves to correct a (morpho-)phonological, lexical, or syntactic error. In (2) an erroneous lexical item is replaced.

(1) We beginnen [rechts] op het, *wat   rechts*
    We start   right   on the somewhat right
    op het papier
    on the paper

(2) En   [boven] de grijze bol   een, OF *rechts van* de
    And over   the grey sphere a   or right of the
    grijze bol   een paarse bol
    grey sphere a   purple sphere
                              ((1) and (2) from Levelt, 1983)

Since we are interested not so much in the motives behind a self-repair as in the formal relationship of the repair text to the original utterance, we have developed an alternative typology based on a dichotomy between *retracing repairs* and *nonretracing repairs*. In retracing repairs,

[1] In the examples we use the following conventions: the target of the repair ("reparandum") is placed between square brackets, []; the moment of interruption is indicated by a comma; editing terms are rendered in capitals, the repair text (i.e., the text replacing the reparandum) in italics. Many of the examples presented in the various sections were taken from a corpus of spontaneous self-repairs collected by the first author. The remaining ones are quoted from the literature.

the speaker interrupts her ongoing speech more or less abruptly, back-tracks to an earlier point of the utterance, and repeats it in a fully or partly modified form (e.g., (1) and (2)). (This type of repair has dominated the literature, almost to the exclusion of nonretracing repairs where the reparandum is replaced by the repair text without any backtracking. Examples are given in Appendix A.) In the remainder of this paper we are chiefly concerned with retracing repairs.

After having interrupted the ongoing utterance, how far will the speaker backtrack? Levelt (1983) proposed a *syntactic* solution to this problem. He formulated a well-formedness condition based on the idea that structural commitments derived from the original, interrupted utterance determine where the restart may occur. More specifically, the original utterance and the repair text bear the same formal relationship to each other as the members of a coordination.

Consider the constructed example (3) where a lexical error is corrected. The repair is well-formed if it can be changed into a grammatical coordination as follows: (a) delete any editing expressions (e.g., NO); (b) complete the interrupted constituent (e.g., with *road*); and (c) insert the connective *and* at the point of interruption (in some cases a different connective is more appropriate, e.g., *or*). In (3a) the speaker has backtracked to the beginning of the prepositional phrase because the corresponding coordination is grammatical (see (3b)). Retracing to the position marked by an asterisk is precluded, since the corresponding coordination is ill-formed (see (3c)).

(3) a There you can park * the car at the [left-hand side] of the,
      NO at the *right-hand side* of the road
    b There you can park the car at the left-hand side of the
      road and at the right-hand side of the road
    c There you can park the car at the left-hand side of the
      road and the car at the right-hand side of the road

In his extensive corpus of spontaneously produced self-corrections, Levelt found only a very small number of violations of the well-formed-ness rule. However, the value of this observation is diminished by the following two considerations.

1. The number of potentially critical cases in a corpus of spon-
   taneous self-repairs is relatively small. In many sentences,
   the well-formedness rule leaves open several possible targets
   for retracing, irrespective of the position of the interrupt.

2. In the majority of spontaneous self-repairs, the speaker halts
   during or immediately after the pronunciation of a wrong
   word (in Levelt's corpus up to 75%). In most of these cases,
   the speaker backtracks no further than either the reparandum

itself or the beginning of the constituent the reparandum be-
longs to. Neither strategy will ever produce a violation of the
well-formedness rule.

In order to overcome both limitations, we have devised an experimental
procedure for eliciting repairs artificially. It allowed for controlling (a) the
sentence structures used by the speakers and (b) the position where the
ongoing utterance is interrupted. The crucial dependent variable was the
backtracking target chosen by the speakers in the various experimental
conditions.

3. METHOD

3.1 Introduction

The procedure was essentially a picture description task. Line drawings representing
everyday events were displayed on a CRT screen under the control of a computer. Subjects
described the depicted events using sentences of somewhat standardized form. During the
pronunciation of the descriptive sentence, some aspect of the picture might change, neces-
sitating an update of the description. The speakers had been instructed to react to the picto-
rial change as quickly as possible. The moment at which the pictorial change took place
could be varied relative to the time interval during which the descriptive sentence was
pronounced. By carefully selecting these moments, we could bring speakers into a situation
where the favorite strategy of immediate one-word retracings was out of reach. We could
also independently vary the changing pictorial attribute and the sentence form (e.g., active
versus passive). This enabled us to place the prospective reparandum early or late in the
description.

We occasioned violations of the well-formedness rule by arranging the visual scenes in
such a way that a natural description would often include noun phrases of the following
format:

[NP1 Art 1 N1 [PP1 Prep 1 [NP2 Art 2 N2]]]

i.e., containing a prepositional phrase as a postnominal modifier. Suppose that, due to some
pictorial change, N2 has to be replaced by, say, N3. The speaker who attempts to make this
correction *immediately after* having pronounced N2, may confine herself to retracing over a
distance of one word and say N3 (e.g., (4)). The well-formedness rule also allows further
retracings: to Art 2, Prep 1, or even Art 1. The reader can verify this by applying the rule,
for instance, to (5a): replace the editing term UH by the connective *and (en)*, and check the
grammaticality of the resulting coordination (see (5b)). However, after a *delayed interrupt*
such as (6a) and (7a), the well-formedness rule dictates retracing to the beginning of NP1,
that is, up to and including Art 1 ((7b) is grammatical, (6b) is not).

(4) the man with the [moustache], . .UH. . *glasses* greets the clown

(5) a the man with the [moustache], . .UH. . the *glasses* greets the clown
   b the man with the moustache and the glasses greets the clown

(6) a * the man with the [moustache] greets the, . .UH. . with the *glasses* greets the
      clown
   b * the man with the moustache greets the clown and with the glasses greets the
      clown

(7) a the man with the [moustache] greets the, . .UH. . the man with the *glasses* greets
      the clown
   b the man with the moustache greets the clown and the man with the glasses greets
      the clown

The ideas sketched in this section formed the basis of four pilot experiments intended to test the well-formedness rule. The results were very clear. Subjects frequently produced repairs whose form went against Levelt's rule. At the same time, the deviations displayed a very systematic pattern, suggesting that a different correction mechanism was at work, perhaps in addition to the mechanism which is responsible for the well-formedness rule. The findings obtained in the pilot studies motivated the specifics of the experimental setup described below. We need not summarize the outcomes of the pilots, since they showed the same trends as are apparent in the main experiment to be presented here.

3.2 Materials

The stimuli were line drawings depicting two persons: an actor (*the man*) and an object (*the clown*). The actor performed one of two actions, *push away* (*wegduwen*) or *wave at* (*toezwaaien*) the clown. Both Dutch verbs contain a separable particle (*weg, toe*). The actions were represented by two different positions of the actor's arm (pushing: horizontal, toward the object; waving: vertical, upward). The two persons had a neutral appearance, demanding no further elaboration of the NPs describing them. In its most simple form, a description could be: *de man duwt de clown weg* (*the man pushes the clown away*). The persons could receive special attributes rendered as either an adjective or a prepositional phrase. Depending upon hairstyle, the man had to be called *bald* (*kaal*) or *neat* (*net*). The clown was *happy* (*blij*) or *sad* (*droef*) depending on the shape of his mouth. The man's face could be supplied with a *moustache* (*snor*) or *glasses* (*bril*); the clown could carry a *sac* (*zak*) or a *bag* (*tas*). In (8) we give the format of the description together with the Dutch words (mostly monosyllabic) filling its slots. Optional parts, corresponding to pictorial elements which may be absent from a drawing, are within parentheses.

(8)

the (adj) *man*	(PP)	Verb	*the* (adj) *clown*	(PP)	Particle
nette	met de bril	duwt	droeve	met de zak	weg
kale	met de snor	zwaait	blije	met de tas	toe

A trial started with the display of only one person, either the man or the clown. After registration of the onset of the subject's response via microphone and voice key, the second person was displayed alongside of the first one. This two-step buildup of each picture served to elicit sentences in passive and active voice. Presenting the actor first led to an active sentence (cf. (8)). When the object was displayed first, the speakers constructed a sentence in passive voice (cf. (9)). Notice that, due to Dutch word order rules, the sentence-final PP is always followed by a verbal element: a particle or a past-participle.[2] This implies that for repairs of the second PP the same restrictions hold as for those of the first PP. That is, backtracking to the beginning of the PP suffices after an immediate interrupt, but once the pronunciation of the verbal element has started, retracing to the beginning of the NP is obligatory.

(9)

the (adj) *clown*	(PP)	*is by the* (adj) *man*	(PP)	Past-Participle
droeve	met de zak	nette	met de bril	weggeduwd
blije	met de tas	kale	met de snor	toegezwaaid

---

[2] A past participle may be regarded as consisting of two words: a particle and a past participle. This is indicated by the possibility of inserting a word between them. For instance, *De clown zal door de man weg worden geduwd* (*The clown will by the man away be pushed*).

The stepwise buildup of pictures introduced an undesirable asymmetry between active and passive trials. In active trials, the initial picture fragment revealed actor as well as action: remember that the action was recognizable from the actor's arm position. The initial picture fragment displayed in passive trials, however, showed the clown without any action clues. This active-passive asymmetry was easy to remove: we presented the actor (the man) without an arm; during the second step the arm and the object (the clown) were added simultaneously. In fact, we used all three display possibilities in the experimental design:

		Step 1	*Step 2*
S (subject)	:	actor	+ action + object
SV (subject verb)	:	actor + action	+ object
PAS (passive)	:	object	+ actor + action

Repairs were induced by changing certain pictorial aspects during the subjects' pronouncing the descriptive sentence. Three types of changes were introduced:

1. Substituting an attribute, which induces a *SUB-repair:* The entailed reparandum was either an adjective (10), a prepositional object noun (11), or a verb (12).
    (10) the man pushes away the [happy] clo-, the *sad* clown
    (11) the clown with the [bag], with the *sac* is waved at by the man
    (12) the bald man [pushes away] the, *waves at* the clown

2. Deleting an attribute, which induces a *DEL-repair:* The most natural linguistic response to a deletion turned out to be a lexical substitution. In the pilot studies, many subjects found it difficult to leave out the adjective or PP whose pictorial counterpart had been deleted. They preferred special lexical items denoting the absence or disappearance of an attribute. For example, in (13) the adjective *neat* is replaced by *ordinary (gewoon)*; in (14) the preposition *with (met)* is substituted by *without (zonder)*. Thus, the adjective or PP was not actually removed from the description. It follows that DEL-repairs were much like SUB-repairs, although within the PP the preposition rather than the noun was substituted.
    (13) the [neat] man pu-, the *ordinary* man pushes away the clown
    (14) the man waves at the clown [with the bag], *without* bag

3. Adding an attribute, which induces an *ADD-repair:* The linguistic material to be inserted into the description was either a prenominal adjective (15) or a postnominal PP ((16) and (17)). In the former case, a retracing repair will ensue, which is governed by the well-formedness rule. In the latter case, however, the speaker sometimes managed to locate the interrupt immediately after the NP's head noun *(man, clown)*. Then she was free to continue by simply adding the PP, without any backtracking. This led to repairs of the nonretracing type, as is further attested by a pause and a deviating intonation contour (16).
    (15) the [clown], the *happy clown* is pushed away by the man
    (16) the clown, . . . *with the bag* is waved at by the man
    (17) the [man] waves, the man *with the glasses* waves at the clown

The pictorial and sentential material to be used in the experimental trials was composed as follows. Notice first that all sentences contain two NPs and that there are four NP shapes: (1) Art + NP, (2) Art + Adj + NP, (3) Art + N + PP, and (4) Art + Adj + N + PP. This implies that 16 structurally different sentences types are possible (disregarding, for the moment, active versus passive voice; cf. (8) and (9)). In order to keep the number of trials within manageable proportions, we applied several reductions. First of all, we did not use any sentences containing two PPs; this reduced the total number of sentence types from 16 to 12. For each of these we listed all logically possible ways of SUBstituting, ADDing and DELeting a pictorial element. (Note that verbs could only be substituted, not deleted or added. We did not distinguish V1 → V2 from V2 → V1). The second reduction consisted of

TABLE 1
Number of Structurally Different Descriptions for the Various Repair Types and
To-Be-Repaired Linguistic Elements

	SUB	DEL	ADD	
Verb	9	—	—	9
Prepositional phrase	8	8	8	24
Adjective	6	6	6	18
Total	23	14	14	51

removing all ADDs which would yield sentences with two PPs or two adjectives. The third reduction only applied to substitutions and deletions of adjectives. There (and only there) we excluded original sentences with double adjectives. Table 1 gives an overview of the 51 remaining repair types. It indicates, for example, that there are 8 structurally different sentence types to which a PP can be added. To the total of 51 we added 9 "catch items" where no pictorial change occurred and no repair was needed.

The resulting set of 60 items had to be used in three display conditions (S, SV, and PAS; see above). This made up a total of 180 items to be presented to every subject. The various pictorial attributes were distributed over these items in such a way that all attributes occurred equally often, both before and after the picture change. Moreover, the two directions a SUBstitution could go (e.g., V1 → V2 and V2 → V1) were selected with equal frequency.

For each item we needed repairs after a delayed interrupt and after an immediate interrupt. We therefore presented the items under two temporal conditions: once with a "late" pictorial change, once with an "early" change. Presumably, a late pictorial change would yield more delayed interrupts than an early change. We ran a separate pilot experiment enabling us to estimate the average interval needed by subjects to reach certain points in their descriptive utterance. For example, it took our subjects an average of 775 ms to proceed from the first to the second article in *de man met de snor* (*the man with the moustache*). On this basis, we decided that an "early" change from a moustache to glasses should take place after 625 ms (775 minus 150), a "late" change after 925 ms (775 plus 150). In this way we hoped to obtain reasonable numbers of immediate and delayed interrupts.

The above method yielded specifications for 360 trials, each consisting of two pictures (except for the "catch items," where one picture was sufficient), one display condition (S, SV, or PAS), and one interrupt interval ("early" or "late"). They were placed in random order and distributed over four blocks of 90 trials. These blocks were presented to the subjects in four different balanced orders.

3.3 Procedure

The experiment was run under the control of a PDP11/45 computer. Onset of vocal responses (descriptive utterances) was registered via a microphone and a voice key. Line drawings were stored on a computer disk as a list of vectors, and reproduced electronically on a Vector General display. During each trial, the vectors defining the drawings were stored in four separate buffers: one for the actor (the man), one for the action (his arm), one for the object (the clown), and one for the changing part (either actor, action, or object). The change of a pictorial attribute was accomplished by replacing one of the three former buffers by the fourth one. The display could be refreshed very quickly without disturbing side effects such as blinking. All responses were tape-recorded.

Subjects were told they had to perform a dual task: observing and speaking. The purported goal of the experiment was to study the effect of language production upon the efficiency of visual perception. Emphasis was laid on detecting all changes in the display, and reporting this as quickly as possible by interrupting the ongoing description and incor-

porating the pictorial change in an updated description. Responses were tape-recorded under the pretext that, later on, speed and accuracy of detection were to be measured. We had to give this deceptive information in order to divert the subjects' attention from the linguistic shape of their descriptions. Subjects were not informed that the responses would be analyzed linguistically, nor were suggestions made on how to incorporate a pictorial change into the interrupted description. After the session, the real purpose of the experiment was disclosed. Participants spoke at a normal rate without unduly long pauses: the two-step presentation of each stimulus did not pose any specific problems. The suggested active and passive sentence frames (see (8) and (9)) were followed virtually unanimously.

Each session consisted of four parts. First, the experimenter (always the first author) introduced the "man" and the "clown" with the various combinations of attributes, and suggested the words that could be used to refer to the persons and their attributes. Second, the subject responded to a series of 24 (or 48) drawings without changes (only SV and PAS), thus becoming familiar with the experimental equipment, the sentence formats, and the words referring to the attributes. Third, the subject went through a series of 20 trials *with* changes (S, SV and PAS) in order to get an impression of the actual task. Fourth, the 360 trials of the experiment were run without any further practice. There was a break of about 10 min between the second and third block of 90 trials. Each trial consisted of the following program steps:

1. Display the starting signal "***"; wait for subject to press button.

2. Display first part of drawing; wait for voice-key to trigger.

3. Start clock; complete the drawing; if clock matches the interrupt interval, replace one of the three buffers by the fourth; wait until clock matches the maximum interval. (The maximum interval was 1 s longer than the average time needed to pronounce the entire descriptive utterance. These averages had been obtained during a pilot study.)

4. Clear screen; go to 1 for next trial.

3.4 Description of the Corpus

In Section 3.1 we introduced the syntactic construction selected to put the well-formedness rule to a critical test: the NP containing a postnominal PP. Of the 360 trials, 144 aimed at repairing a PP within an NP (6 × 24, cf. the second row of Table 1). We present only analyses of responses delivered in these trials.

Sixteen subjects, undergraduate psychology students of the University of Nijmegen, participated individually in sessions lasting between 90 and 120 min. In all, they underwent 2304 (16 × 144) trials where they attempted to repair a prepositional phrase. In 2060 cases (89.4%) the attempt was successful. The pictorial change went unnoticed in only 44 cases (1.9%). The remaining failures were due to a variety of factors such as spontaneous errors preceding the pictorial change, a voice-key not working appropriately, etc. In addition, the subjects produced 52 PP repairs spontaneously in other trials. These were added to the corpus, now containing 713 SUB-repairs, 664 DEL-repairs, and 735 ADD-repairs. The number of DEL-repairs is lower mainly because subjects sometimes failed to produce the correct repair text: instead of replacing the preposition *with* by *without*, they removed the entire PP from the NP. The number of repairs after a delayed interruption was 388 for SUB-repairs, 384 for DEL-repairs, and 462 for ADD-repairs. Immediate repairs were somewhat less frequent: 325 for SUB-, 280 for DEL-, and 273 for ADD-repairs. In display conditions S, SV, and PAS we obtained 692, 714, and 706 repairs, respectively. The sentence-initial PP elicited 1065 repairs, the sentence-final one 1047.

All experimental sessions were recorded on tape. Complete transcriptions of the responses were typed into a computer file and coded for various characteristics. The central dependent variables were *point of interrupt* and *retracing target*. The former was encoded

as either "immediate" or "delayed." In an immediate interrupt, the speaker halted at or before the end of the PP. When the original utterance was interrupted at a later point (e.g., during/after the tensed verb, participle, or particle), the response was classified as delayed. Within the domain of possible retracing targets we distinguished two classes: retracing toward the beginning of the NP of which the PP made part, or retracing no further back than the beginning of the PP. We dubbed them "complete-NP" and "PP-only," respectively.

On the basis of the well-formedness rule and the reasoning in Section 3.1, one expects complete-NP restarts after all delayed interrupts. A PP-only retracing would imply a violation of the rule. After an immediate interrupt, both PP-only and complete-NP retracings are permitted. But the latter are inefficient in the sense of overshooting the nearest target designated by the well-formedness rule. Thus we grouped the 2112 PP repairs into four groups. Two were "suboptimal": (1) PP-only retracings after a delayed interrupt, and (2) complete-NP retracings after an immediate interrupt. Both remaining groups were "optimal": (3) PP-only retracings after an immediate interrupt, and (4) complete-NP retracings after a delayed interrupt. Suboptimal repairs are exemplified by the ill-formed cases in (18)–(20) and by the overshoots in (21)–(23).

(18) * the man with the [glasses] pushes, with the *moustache* pushes away the clown

(19) * the clown [with] the bag is pu-, *without bag* is pushed away by the man

(20) * the [man] waves, *with the glasses* waves at the happy clown

(21) the man with the [mou-], the man with the *glasses* pushes away the clown

(22) the man waves at the clown [with] the bag, the clown *without* bag

(23) the [clown], the clown *with the bag* is pushed away by the bald man

For each subject, we expressed the number of complete-NP and PP-only restarts as a percentage of the total number of repairs in the various categories we distinguished. We used a nonparametric statistical test—Wilcoxon's signed-ranks test for matched pairs—since the percentage scores were expected to be distributed extremely asymmetrically (approaching either 0 or 100), and since this test is sensitive to between-subject inconsistencies in the direction of an effect.

For most analyses, we subdivided the corpus into a number of smaller sets. As a result, some subjects failed to have a score on one of the two variables in a paired comparison. Missing scores were typically due to individual differences in speaking rate and, in case of DEL-repairs, deviations from the prescribed format. The exact number of paired comparisons on which a test is based, is indicated in the bottom row of each table. In all tests, a one-tailed level of significance was specified. Whenever we report an average score, it was computed on the basis of the paired observations only (including ties); the observations which happened to have no counterpart in a paired comparison were left out.

As already hinted at in Section 3.1, we found massive deviations from well-formedness and substantial proportions of overshoots in the four pilot studies. However, in the various sections of the experimental design the proportions of "suboptimal" responses were very different. The shape of repairs appeared to depend critically on the immediate context of the reparandum for one part, and on the location of the interrupt for another. Consequently, we redirected our study from a straightforward test of Levelt's well-formedness rule to a detailed examination of factors contributing to the choice of a "suboptimal" restarting point. The data presentation is organized according to the location of interrupts. In Section 4 we look at context effects after *immediate* interrupts. The data presented in this section give rise to the model described in Section 5. Section 6, which is devoted to context effects after *delayed* interrupts, provides additional support for the model.

4. MAIN RESULTS: CONTEXT EFFECTS AFTER IMMEDIATE INTERRUPTS

Before going into a detailed examination of our corpus of elicited re-pairs, we assessed the influence of incremental presentation of the pic-tures. A comparison between the S and SV display modes on all the vari-ables to be discussed in the Sections 4 and 6 revealed no significant dif-ferences. In fact, the numerical values of the scores were often virtually identical. We concluded that the visibility of the action did not influence form of repair in the SV condition. Differences between PAS and SV cannot be attributed to this factor. In the sequel, we combine responses in the S and SV modes into one ACT (= active) mode.[3]

The analysis of contextual determinants in this section concentrates on PP-repairs after *immediate* interrupts. (In Section 6 it is shown that es-sentially the same effects occur after delayed interrupts.) We discuss two context effects: the effect of the grammatical category of the word *fol-lowing* the reparandum (Section 4.1) and the effect of an adjective *pre-ceding* the reparandum in the same NP (Section 4.2).[4]

During the analyses it proved necessary to set apart a special subgroup of repairs: those SUB- and DEL-repairs where the interrupt was located *before* (but no *at*) the end of the to-be-modified PP. Actually, the inter-rupts were typically located *before* the prepositional object noun (see (24) and (25)). This subgroup, which comprised 13% of all sentence-final SUB- and DEL-repairs after immediate interrupts (n = 306; no such cases were observed in sentence-initial position), is characterized by the total absence of complete-NP restarts. This feature distinguishes them from repairs with an interrupt *at* the end of the PP. The latter group did attract a fair proportion of complete-NP restarts dependent upon the im-mediate context of the reparandum. Sections 4.1 and 4.2 are devoted to these context effects.

(24) the man waves at the clown with the, with the *sac*

(25) the clown is pushed away by the man [with] the, *without* glasses

---

[3] Display mode *did* have an effect upon the speed of detecting pictorial changes and, consequently, upon the place of the interrupt. For sentence-initial PPs, SV attracted more immediate interrupts than either S or PAS; in case of sentence-final PPs, S tended to elicit more immediate interrupts than SV or PAS. However, this effect is irrelevant to the central hypotheses of the present study.

[4] We also checked for the influence of other contextual characteristics, but none of these proved effective. The overshoot percentage of sentence-initial PPs appeared insensitive to the presence of an adjectival modifier within the final NP, and the percentage of overshoots of sentence-final PPs was not influenced by an adjectival modifier within the initial NP. This implies that the percentage of overshoots in SUB- and DEL-repairs was a function of the *immediate context* only.

4.1 Effect of Following Word

For sentence-initial PPs, the word directly following was a main verb in active sentences (26), and an auxiliary verb in passives (27). For sentence-final PPs, it was a verb particle in actives (26), and a past participle in passives (27).

(26) the (adj) man (PP) *main verb* the (adj) clown (PP) *particle*

(27) the (adj) clown (PP) *is* by the (adj) man (PP) *past-participle*

The results are presented in Table 2. In sentence-initial PPs, actives scored higher on suboptimal repairs than passives for SUB- and DEL-repairs (SUB: $W- = 0, p < .005$, 3 ties; DEL: $W- = 8, p < .05$, 3 ties). In sentence-final PPs, the difference was in the *opposite* direction: passives scored higher than actives for SUB- and DEL-repairs (SUB: $W- = 0, p < .005$, 2 ties; DEL: $W- = 17, p < .025$, 2 ties). There were no significant differences between the scores of ADD-repairs. Apparently, in SUB- and DEL-repairs the occurrence of a complete-NP restart was somehow dependent upon the grammatical category of the following word: overshoots were more likely to occur before a content word (finite or participle form of main verb) than before a function word (auxiliary, particle). In ADD-repairs, this effect was absent.

The Table also shows that the proportion of complete-NP restarts is very much higher for SUB- and DEL- than for ADD-repairs. Overall, the three percentages were 39, 44, and 14, respectively.

4.2 Effect of Preceding Word

The NP to which a to-be-repaired PP belongs may contain an adjective ($+$ adj), or not ($-$ adj; cf. (26) and (27)). What is the effect of a preceding adjective upon the likelihood of a complete-NP restart?

The results are presented in Table 3. For both sentence-initial and sentence-final PPs, SUB-repairs had a higher overshoot score when there was *no* adjective (sentence-initial PP: $W- = 8, p < .01$, 4 ties; sentence-final PP: $W- = 7, p < .05$, 7 ties). DEL-repairs showed a trend in the

TABLE 2

Percentage of Complete-NP Restarts in Repairs of Sentence-Initial and Sentence-Final PPs for Active (ACT) and Passive (PAS) Sentences (Immediate Interrupts)

	Sentence-initial PP			Sentence-final PP		
	SUB	DEL	ADD	SUB	DEL	ADD
ACT	67	68	26	8	20	6
PAS	35	44	21	62	55	13
No. of subjects	14	12	13	15	15	14

TABLE 3

Percentage of Complete-NP Restarts in Repairs of the Sentence-Initial and Sentence-Final PPs for NPs with (+ Adj) or without (− Adj) an Adjective (Immediate Interrupts)

	Sentence-initial PP			Sentence-final PP		
	SUB	DEL	ADD	SUB	DEL	ADD
+ Adj	45	50	27	12	21	9
− Adj	63	66	20	23	37	2
No. of subjects	16	13	13	16	12	14

same direction (sentence-initial PP: $W- = 10, p = .07$, 4 ties; sentence-final PP: $W- = 11, p = .09$, 3 ties). ADD-repairs yielded somewhat higher scores for NPs with an adjective (no significant differences, however). In sum, SUB- and DEL-repairs produced more overshoots when the constituent did *not* contain an adjective. And again this effect did not occur in ADD-repairs.

4.3 A Summary

The most striking outcome of our experiment is the systematic difference between ADD-repairs on the one hand and SUB- and DEL-repairs on the other. The behavioral pattern of these correction types can be summarized as follows.

1. ADD-repairs produce a considerably lower proportion of complete-NP restarts (overshoots) than SUB- and DEL-repairs.

2. ADD-repairs are not sensitive to properties of the left-hand context (+ adj vs − adj), whereas SUB- and DEL-repairs are.

3. ADD-repairs are not sensitive to properties of the right-hand context (content vs function word), whereas SUB- and DEL-repairs are.

The context effects on SUB- and DEL-repairs are of the following nature:

a. *Left-hand context:* if the NP contains a prenominal adjectival modifier (+ adj), the postnominal PP is less often repaired by means of a complete-NP restart.

b. *Right-hand context:* if the to-be-corrected PP is followed by a content word, more complete-NP restarts are made than in case of a function word following.

(Intuitively, both adjectives and other content words exert a kind of "repelling force" upon the backtracking target, pushing it out of their way.)

In the next Section we first lay the foundations for a model encompassing both types of self-corrections and their empirical properties, and then show how a parametrized version of the model can accurately reconstruct the data. In Section 6, the data presentation is resumed for repairs made after a delayed interrupt. From these, additional support for a two-strategy model is derived.

5. INTERPRETATION: A DUAL SYSTEM FOR REPAIRING SPEECH UTTERANCES

We take as starting point the four-stage model of language generation introduced in Section 1. In our repair elicitation procedure, we assume the conceptual module passes a warning message to the monitor when a pictorial change has been detected. In response, the monitor sends an interrupt signal to all other modules, causing them to break ongoing activities and to take correction measures.

Basic to our model is the existence of two different repair strategies: *reformulation* and *lemma substitution*. Reformulation means computing a new syntactic structure corresponding to a modified meaning content— new in the sense that all or part of the structural elements of the original tree have been replaced by other elements. In case of lemma substitution, the lexico-syntactic module decides that the shape of the syntactic tree need not be revised in response to the modified meaning content: replacing one lemma by another one suffices. Presumably, lemma substitution is more efficient than the more powerful, but also more complicated, reformulation strategy. It is useful not only in describing abruptly changing scenes, as in our experiment, but also when correcting erroneous lexicalizations—a frequently occurring type of speech errors (e.g. (28)). Without going into the problem of how the lexico-syntactic module chooses between revision strategies, we assume that *ADD-repairs result from reformulation* and *SUB- and DEL-repairs from lemma substitution*.

> (28) What I've done here is torn [together] three, . . . UH torn
> *apart* three issues that . . .
>
> (from Garrett, 1975)

Upon receiving the revised syntactic structure, the morpho-phonological module locates the restarting point, i.e., the lexical item (lemma) whose phonetic realization will mark the beginning of the repair text. We assume that after a reformulation, this point is the first lemma of the revised part of the syntactic tree. Kempen and Hoenkamp (1987) have worked out the details of a reformulation mechanism which, as a matter of fact, also generates coordinate structures: the repair text after an editing term (e.g., *eh, no, or*) is generated in exactly the same way as the second member of a coordination after the connective (e.g., *and, or*). It follows

that Levelt's well-formedness rule is automatically accounted for. In this paper we cannot describe the workings of the reformulation/coordination mechanism. However, one feature of its output is important to note, namely that all revised parts of the syntactic tree occupy positions to the right of those parts which did not need any alterations in response to the modified meaning content (cf. the obvious fact that the second member of a coordination always follows the first member). This implies that, after a reformulation by the lexico-syntactic module, the morpho-phonological module need not do any backtracking, but simply continues processing at the first lemma of the revised syntactic tree.

After a lemma substitution, the morpho-phonological module does have to retrace to an earlier point in the syntactic structure, namely, at least as far as the substituted lemma. As a matter of fact, most speakers prefer to retrace even a little more (see (28)). The actual restarting target appears to depend on a combination of a syntactic and a prosodic factor. The former relates to boundaries between *major syntactic constituents:* there is a tendency to retrace to the beginning of the major constituent the reparandum belongs to. The prosodic factor is of a similar nature: it designates boundaries between *phonological phrases* as suitable points for restarting. As (29) shows, the two types of boundaries often coincide (major constituents are placed between round brackets; phonological phrase boundaries are marked by "/"). However, long major constituents are split up into several phonological phrases (e.g., the last constituent of (29)). And when a major constituent consists of no more than a single "nonprominent" word, it may be combined with an adjacent constituent into one phonological phrase (e.g., the first constituent of (29)).

> (29) (he)(told)/(his children)/(yesterday)/(a funny story/about
>   dragons)

Prominent words are content words of the following categories: nouns, verbs, many adverbs, and adjectives used predicatively. All other words are nonprominent (unless under special conditions, e.g., when carrying sentence accent). An informal (and incomplete) procedure for partitioning a sentence into phonological phrases looks as follows.

> (30) Traverse the sentence from left to right. When a prominent
>   word is encountered, put a boundary symbol after the *pre-*
>   *ceding* prominent word (if any). However, when the word
>   following that boundary symbol is nonprominent, then
>   move the symbol over that word if:
>   a. it belongs to the same major constituent as its prede-
>   cessor and completes the constituent, OR
>   b. makes up a major constituent of its own.

>NB: In Case (30a), the symbol is moved obligatorily; in
>   Case (30b), the move is optional.

This procedure delimits phonological phrases each containing exactly
one prominent word. We will call this word the *head* of the phonological
phrase. For extensive theoretical and empirical discussions of the notion
of phonological phrases, see Selkirk (1980), Nespor and Vogel (1982,
1983), Gee and Grosjean (1983) and Van Wijk (1987). In (31) we give the
segmentation of our experimental sentences into major syntactic constit-
uents (between round brackets) and into phonological phrases (boundary
symbol "/").

(31) a Active sentences:
>   (*the* adj *man* / *with the* noun)/(main verb)/(*the* adj *clown* /
>   *with the* noun)(particle)#
>   b Passive sentences:
>   (*the* adj *clown* / *with the* noun)/(aux verb)(*by the* adj *man* /
>   *with the* noun)/(past participle)#
>   NB: 1. The phonological phrase boundary preceding
>   the nonprominent auxiliary verb is optionally
>   placed after it (cf. (30b))
>   2. No phonological boundary marker is put at the
>   end of the sentence. This is due to the retro-
>   spective placement of markers (see (30)).
>   3. The final position is occupied by the symbol #.
>   This "end of message" marker is appended to a
>   sentence during lexico-syntactic processing (cf.
>   Van Wijk & Kempen, 1985). It is *not* a phono-
>   logical phrase boundary marker.

The rules determining restarting targets for the morpho-phonological
module can now be stated easily. In Table 4 we distinguish three positions
the interrupt can take relative to the head and end of the phonological
phrase the reparandum belongs to. For each of them, the Table specifies
two targets where morpho-phonological processing can be resumed. The
target mentioned in the left-hand column is the default alternative; but
under special circumstances relating to characteristics of the repar-
andum's context, the second alternative is preferred. When applying the
rules, one should go through the table from top to bottom and apply the
first matching rule.

The predictions from the model outlined so far can be summarized in
two decision trees governing the selection of retracing targets. The tree in
Fig. 2 applies to ADD-repairs only. Figure 3, which is largely a restate-
ment of Table 4, deals with SUB- and DEL-repairs.

TABLE 4

Restarting Targets for the Morpho-Phonological Module at Three Different Interrupt
Positions; Apply the Rules in the Order Given

Rule	Interrupt position	Target	
		Default	Marked
R1 a	Before head of phonological phrase that reparandum belongs to	Nearest boundary (phonological phrase or	S1 Next word
b	Before end of phonological phrase that reparandum belongs to	major constituent) preceding substituted word	S2 Substituted word S3 See R2-default
R2	At or beyond end of phonological phrase that reparandum belongs to	Nearest major constituent boundary preceding substituted word	S4 See R1-default

Note. The choice between alternatives S2 and S3 seems to depend on the type of error: S2 is selected typically after *phonological* errors, S3 after *lexical* errors (cf. Section 7.2).

When comparing the predicted and observed restarting points, one has to bear in mind that the language production modules which carry out the reformulation and lemma substitution strategies are critically dependent on information concerning the exact position where the ongoing utterance was interrupted. Since the language modules cannot themselves access the articulatory output, they have to rely on the monitor reporting back such information (cf. Fig. 1). We assume that the monitor sometimes errs by a few syllables, either in backward or in forward direction. Such "backward shifts" and "forward shifts" of the subjective interruption point relative to the objective interrupt cause the speaker to deviate from the predictions derivable from the decision trees. The modules concerned cannot but respond to subjective interrupt positions.

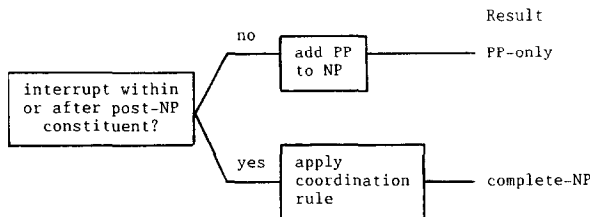In case of ADD-repairs, the monitor's inaccuracy has the following

FIG. 2. Decision tree for applying the reformulation strategy (ADD-repairs).
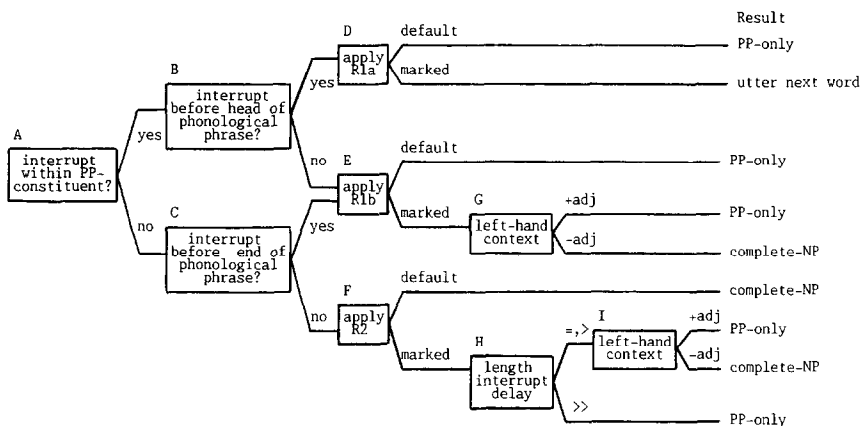
FIG. 3. Decision tree for applying the lemma substitution strategy (SUB- and DEL-re-pairs). The symbols " =, >, >>" designate length of the interrupt delay: " = " means immediately following the PP, ">" and ">>" correspond to " = verb" and ">verb" as explained in Section 6.1.

consequences. A forward shift of the subjective interruption point after an immediate interrupt constitutes a delayed interrupt and causes the lexico-syntactic module to decide in favor of a complete-NP instead of a PP-only restart (overshoot). After a backward shift, on the other hand, a delayed interrupt may look like an immediate one, and an ill-formed re-sponse is the overt result. The same inaccuracy of the monitor causes discrepancies between predicted and observed retracing targets in SUB- and DEL-repairs as well.

When linking the decision tree in Fig. 3 to the data,[5] one has to keep in mind various assumptions made earlier in this section (see (31)) as re-gards boundary marking in the reparandum's right-hand context.

> 1. Sentence-final particles are not surrounded by boundary markers "/"; so rule R1b would have to apply invariably for all repairs of the final PP in active sentences (cf. (32)). However,

[5] Rule R1a only applies when the interrupt occurs before the prepositional object noun. As we reported in the introduction to Section 4, we had to set apart these cases because they never led to a complete-NP restart. Now we can see how this observation follows from rule R1a. Selection of the default restart target causes a PP-only restart to show up as a so-called covert repair (*the man with the, with the glasses*). Selection of the marked contin-uation implies uttering the next word, typically accompanied by a hesitation (lengthening of vowels, pausing; e.g., *the man with theeeee . . . glasses*). As a matter of fact, we did not include the latter repair type in our analyses. Since rule R1a is relevant only to this minor subcategory of extremely early interrupts, we restrict our account of experimental results to the cases where rules R1b and R2 apply.

after delayed interrupts the "end of message" marker # is occasionally mistaken for a phonological phrase boundary, and rule R2 is selected instead.

2. Main verbs (finite or participle) are always preceded by a marker, leading to application of R2 both after immediate and delayed interrupts (cf. (33a–b)).

3. Auxiliaries leave open two possibilities: they are either preceded or followed by a marker. (In the former case they have the same effect as a main verb, but not in the latter; cf. (34a–b).)

After having determined a restart rule, one finally has to choose between the default and marked target. This selection is dependent upon characteristics of the left-hand context (the presence or absence of an adjective) and of the right-hand context (length of interrupt delay). These features of the ongoing utterance are assumed to govern the decisions made at nodes labeled G, H, and I in Fig. 3.

(32) . . . main verb / *the clown* / *with the* noun particle #

(33) a *the man/ with the* noun / main verb / *the clown* particle #
     b . . . *is by the man* / *with the* noun / past-participle #

(34) a *the clown* / *with the* noun / *is by the man* / past-participle #
     a *the clown* / *with the* noun *is* / *by the man* / past-participle #

In order to derive numerical predictions from the model, we estimated values for the following four parameters:

a. the chance for the monitor to identify correctly the actual position of the interrupt with respect to the end of the PP-constituent that the reparandum belongs to;

b. the probability that a phonological phrase boundary marker precedes the auxiliary in passive sentences (cf. (34));

c. the probability that the "end of message" symbol # is interpreted as a phonological phrase boundary (cf. (32));

d. the probability of selecting the marked option in case both the default and the marked ones are open.

In Appendix B we describe the parametrization of our model in detail. In order to test the model's goodness of fit we applied a statistical technique developed by Mosteller (cf. Torgerson, 1958). Results were very satisfactory. Overall, tests showed no significant differences between predicted and observed scores. There was only one exception, which concerned a specific category to be discussed and explained in the following section.

6. ADDITIONAL RESULTS: REPAIRS AFTER
DELAYED INTERRUPTS

In this section, the data for PP-repairs after a *delayed* interrupt are discussed. We combine SUB- and DEL-repairs into one category. Results for sentence-initial and sentence-final NPs are reported separately. In the former group, we look at the relation between context effects and length of interrupt delay (Section 6.1). In the latter group, we had to set apart nonretracing from retracing repairs (cf. Section 2). Results pertaining to both repair varieties are presented in Section 6.2. Some noteworthy observations on so-called double repairs are discussed in Section 6.3. Section 6.4 finally summarizes the major findings.

6.1 Sentence-Initial PPs: Relation between Context Effects and Delay of Interrupt

The repairs of sentence-initial NPs after a delayed interrupt were classified as either " = verb" or ">verb." All active sentences interrupted within or at the end of the main verb counted as = verb, and so did all passives interrupted before the first article of the sentence-final NP (see (35)). In both sentence types, about half the trials were classified as = verb. Interrupts at more remote positions were coded as >verb. Only very few interrupts were delayed beyond the sentence-final NP.

(35)

	= verb	>verb
ACT: *the* (adj) *man* (PP)	verb	*the* (adj) *clown* (PP) particle
PAS: *the* (adj) *clown* (PP)	*is by*	*the* (adj) *man* (PP) past-participle

The results presented in Table 5 reflect the right-hand context effect. Scores for SUB- and DEL-repairs were hardly affected by length of interrupt delay. Within active sentences they even decreased a little. For = verb interrupts, the difference between sentence types approached significance (81 vs 72%, $W- = 17, p < .08$, 5 ties). Reformulations scored

TABLE 5
Percentage of Complete-NP Restarts in Repairs of Sentence-Initial PPs in Active and Passive Sentences for = Verb and >Verb Interrupts

	SUB + DEL		ADD	
	= Verb	>Verb	= Verb	>Verb
ACT	81	79	75	90
PAS	72	76	68	94
No. of subjects	16	15	16	11

TABLE 6

Percentage of Complete-NP Restarts in Repairs of Sentence-Initial PPs as a Function of
Place of Interrupt (= Verb or >Verb) and Length of NP (+ Adj or − Adj)

	SUB + DEL		ADD	
	= Verb	>Verb	= Verb	>Verb
+ Adj	68	76	77	88
− Adj	83	75	76	93
No. of subjects	16	16	16	16

significantly higher on complete-NP restarts in case of >verb than = verb interrupts. This was true for both sentence types (ACT: $W- = 8, p < .025$, 6 ties; PAS: $W- = 0, p < .01$, 4 ties).

The scores needed to assess the effect of an adjective making part of the NP (the left-hand context) are presented in Table 6. Within the combined group of SUB- and DEL-repairs, the presence of an adjective significantly lowered the complete-NP score for = verb interrupts ($W- = 9, p < .02$, 5 ties). This effect completely disappeared when the interrupts were delayed further (>verb). ADD-repairs showed no effect at all.

6.2 Sentence-Final PPs: Complete-NP Restarts after Delayed Interrupts

In sentence-final position, PP-repairs took two different forms. In a fair number of repairs the particle or past participle was absent in the repair text (see Examples (36) and (37)). We considered these to be instances of a *nonretracing* repair type. (The repair text serves a function comparable to PP-extraposition or right-dislocation; cf. Appendix A.) In the remaining repairs the sentence-final verbal element was repeated; these represent clear cases of retracing repairs. The results for both repair varieties, retracing and nonretracing, are examined separately.

Most subjects had a clear preference for one of the two forms. Half of them produced retracing self-repairs almost exclusively; the remaining subjects hardly ever retraced (49% of the repairs were retracings; $n = 558$). This reduced the sample considerably. For most comparisons the number of subjects dropped to about 9.[6]

(36) de clown wordt door de [man] weggeduwd   UH *met*
the clown is   by   the man away-pushed uh   with
de bril
the glasses

---

[6] This preference for a specific strategy was a local phenomenon, i.e., restricted to repairs of the sentence-final PP made after completion of the descriptive utterance. The two groups of subjects did not behave differently in any other respects.

(37) de man zwaait de clown [met] de tas toe NEE de
     the man waves the clown with the sac at   no the
     clown *zonder* tas
     clown without bag

The results for retracing repairs are presented in Table 7. The differ-
ence between sentence types was significant for SUB- and DEL-repairs
($W-$ = 0, $p < .001$, no ties) but not for ADD-repairs. The sample com-
bining passive and active sentences showed a significantly higher score
for $-$adj NPs within SUB- and DEL-repairs ($W-$ = 10, $p < .05$, 1 tie).
Again, ADD-repairs failed to produce any effect.

Nonretracing repairs came in two varieties: PP-only (36) or complete-
NP (37). They yielded a pattern of results notably different from the one
obtained for retracing repairs. First, for the combined sample of SUB-
and DEL-repairs, the proportion of complete-NP repairs was not higher
in $-$adj than in $+$adj NPs (53% ($+$adj) vs 56% ($-$adj)). Second, *all* re-
pair types produced a significantly smaller proportion of complete-NP
restarts in actives than in passives (see Table 8; SUB + DEL: $W-$ = 1,
$p < .01$, 1 tie; ADD: $W-$ = 0, $p < .025$, 2 ties).

6.3 Determinants of Double Self-Repairs

In a fair number of trials, subjects produced a *double self-repair,* i.e., a
repair of their own (failing) self-repair. In this experiment, a double repair
consists of a short-distance retracing followed by retracing to a more re-
mote target. In (38)—(40), an initially ill-formed repair is thus made well-
formed; in (41)–(43) an "optimal" PP-only retracing is replaced by a
"suboptimal" (but still grammatical) overshoot. Double repairs can be
taken as reflections of Subjects' monitoring of the well-formedness of
their repairs.

(38) the man with the [moustache] pushes the, with the, the man
     with the *glasses* pushes away the clown

(39) the happy clown [with] the bag is, without bag is wa-, the
     happy clown *without* bag is waved at by the bald man

TABLE 7
Percentage of Complete-NP Restarts in Retracing Repairs of Sentence-Final PPs as a
Function of (a) Sentence Type and (b) Length of NP

(a)	SUB + DEL	ADD	(b)	SUB + DEL	ADD
ACT	36	74	$+$adj	38	82
PAS	77	89	$-$adj	53	84
No. of subjects	10	8		11	10

TABLE 8
Percentage of Complete-NP Restarts in Nonretracing Repairs of the Sentence-Final PPs as a Function of Sentence Type

	SUB + DEL	ADD
ACT	44	56
PAS	69	83
No. of subjects	9	8

(40) the [man] pushes, with the gla-, the man with the *glasses* pushes away the sad clown

(41) the sad clown is waved at by the man with the [moustache], with the, the man with the *glasses*

(42) the clown [with] the sac, without, the clown *without* sac is waved at by the man

(43) the [man] . . . with the glasses, the man with the *glasses* pushes away the happy clown

Table 9(a) lists the percentages of PP-only restarts which were transformed into complete-NP restarts. ADD-repairs were in best agreement with the well-formedness rule: they obtained a relatively high score (22%) after a delayed interrupt, and a low score (0%) subsequent upon an immediate interrupt. Thus it looks as if well-formedness and avoidance of overshoots were observed more carefully in ADD-repairs than in SUB- and DEL-repairs.

That ADD-repairs were more sensitive to place of interrupt than SUB- and DEL-repairs is made clear by single as well as double repairs. For the

TABLE 9
Percentage of Double Repairs (a) after Delayed and Immediate Interrupts and (b) for Sentence-Initial PPs as a Function of Place of Interrupt

	SUB + DEL	ADD
(a) Delayed	11	22
	($n = 240$)	($n = 110$)
Immediate	5	0
	($n = 335$)	($n = 233$)
(b) Immediate	5	0
	($n = 138$)	($n = 95$)
= Verb	7	13
	($n = 46$)	($n = 30$)
> Verb	14	46
	($n = 36$)	($n = 11$)

former see Table 5; for the latter Table 9(b). After a >verb interrupt only 1 out of 7 SUB- and DEL-repairs was doubled (14%); for ADD-repairs chances rose to 1 in 2 (46%).

We finally mention a remarkable effect discernible in Table 10, which represents double repairs of sentence-final PPs. ADD-repairs showed a comparable percentage of doubles in the retracing and nonretracing subgroups (19 and 23%), whereas in SUB- and DEL-repairs there was a large difference between the doubling scores in retracing and nonretracing cases (7 vs 19%).

6.4 A Summary

The results reported in this section have further validated the decision to distinguish between two different repair strategies. After delayed interrupts, ADD-repairs were marked off from SUB- and DEL-repairs in the same manner as after immediate interrupts: they produce considerably fewer PP-only restarts (suboptimal repairs), and are insensitive to properties of the left-hand (+adj vs −adj) and right-hand (content vs function word) context. But in addition three new distinguishing features emerged:

> 1. ADD-repairs are very sensitive to the length of the interrupt delay (=verb vs >verb) whereas SUB- and DEL-repairs are not.
>
> 2. After delayed interrupts, ADD-repairs give rise to double repairs (thus restoring well-formedness) more often than SUB- and DEL-repairs.
>
> 3. After immediate interrupts, ADD-repairs do not develop into double repairs. Thus overshoots are avoided. SUB- and DEL-repairs do show double repairs.

A special position was occupied by nonretracing repairs of sentence-final PPs. (These occurred after completion of the original descriptive sentence and the repair text did not include the particle or past participle.) Within this subgroup, the distinction between ADD- vs SUB- and

TABLE 10

Percentage of Double Repairs Made after Delayed Interrupts for Sentence-Final PPs as a Function of Repair Type

SUB + DEL		ADD	
Retracing	Nonretracing	Retracing	Nonretracing
7	19	19	23
(n = 101)	(n = 57)	(n = 21)	(n = 48)

DEL-repairs broke down: they all produced a substantial proportion of double repairs and were equally insensitive to the left-hand context effect. This suggests that these nonretracing repairs are best considered as reformulations. PP-only ADD-repairs exemplified by (36) are extrapositions comparable to (44). The SUB- and DEL-type corrections such as (37) are cases of right-dislocation which can be accommodated under the well-formedness rule in terms of *gapping* (cf. 45)). Since the nonretracing repairs transform into a gapping construction when the connective *and* is substituted for the editing term, and since gapping constructions are cases of coordination, it follows that the well-formedness rule is applicable (see Pijls & Kempen, 1986). What we cannot explain yet is the fact that actives attracted far more nonretracing PP-only repairs than passives (see Table 8). Although this observation resembles the right-hand context effect displayed by lemma substitutions, it needs a syntactic rather than a prosodic explanation.

(44) Ik heb   de [blouse] aan *met de bloemetjes*
       I have the shirt   on with the flowers

(45) a. Deze politiek heeft de relatie   met de [Russen]
        This policy   has   the relation with the Russian
        verstoord NEE met de *Chinezen* BEDOEL IK
        troubled no   with the Chinese mean   I

     b. Deze politiek heeft de relatie   met de Russen
        This policy   has   the relation with the Russian
        verstoord en   met de Chinezen
        troubled and with the Chinese

7. DISCUSSION

In the previous section we worked out a distinction between two mechanisms for computing the linguistic shape of self-repairs in spontaneous speech: one based on the repair strategy called *reformulation*, the second one on *lemma substitution*. Levelt's (1983) well-formedness rule, which connects self-repairs to coordinate structures, was shown to apply only to reformulations. In case of lemma substitutions a totally different set of rules, summarized in Table 4, appears to be at work. Reformulations are represented in the experimental data by ADD-repairs, lemma substitutions by SUB- and DEL-repairs. The linguistic unit of central importance in reformulations is the major syntactic constituent; in lemma substitutions this role is played by phonological phrases. The point where speech production is resumed after the interrupt is determined by the lexico-syntactic module in case of reformulations, by the morpho-phonological module after lemma substitutions.

In this final section we elaborate on the dual character of self-repairs and investigate its relevance to other phenomena reported in the speech error and repair literature.

7.1 Lemma Substitutions Preserve Structure

In a large proportion of appropriateness repairs, a so-called prespecification is added (cf. (46)). In our terminology they would be called ADD-repairs. Error repairs leave the original sentence structure intact; nothing is changed except for the substitution of an erroneous element. We would name them SUB-repairs. That the repair strategies underlying the two types are indeed different is nicely illustrated by Levelt's corpus of self-repairs. He mentions two cases where a postnominal prepositional phrase was corrected. Of the two, the ADD-repair was well-formed (47), whereas the SUB-repair was not (48).

(46) We beginnen [rechts] op het, *wat   rechts*
     We start   right   on the somewhat right
     op het papier
     on the paper

(47) We beginnen in het [midden] met, . . . in het *midden van*
     We start   in the middle   with   in the middle of
     het papier met een blauw rondje
     the paper with a   blue   disc

(48) *Rechts van [paars] ligt, UH van *wit*   ligt paars
      Right of   purple is   uh of white is   purple
                                     ((46)–(48) from Levelt, 1983)

This observation supports our claim that lemma substitution is the basic mechanism which underlies lexical error repairs—probably the most frequent type of retracing repair in spontaneous speech. Further doubts about syntactic interpretations of lexical error repairs, i.e., treating them as reformulations, are raised by the following two observations.

  1. Lexical error repairs do not show ellipsis. We have been unable to find any examples of lexical error repairs of the retracing variety which contain sentence-internal deletions (cf. the invented repair (49)). Invariably, all pronounced words inbetween reparandum and interrupt are repeated in the repair text, without any signs of elliptical formations. Sentence-final repairs with an apparent deletion such as (37) are no counterexamples. They are nonretracing repairs of the right-dislocation type, and in the previous section we have shown that these do result from reformulation.

  2. Lexical error repairs do not show pronominalization. This is another indication that the repair text simply copies the words uttered be-

tween reparandum and interrupt. In (50), for example, *read the book* is said, instead of *read it*. Of course, pronominalization may occur in clear cases of reformulation such as (51) where a prespecification is added, and (52) which belongs to the nonretracing category.

> (49) Ik ga   dan bij de [rode] cirkel li-, NEE bij de *blauwe*
>   I   turn then at the red   circle le- no   at the blue-one
>   linksaf
>   left

> (50) I would like to see it now that I've [written] the book, UH
>   *read* the book   (from Garrett, 1975)

> (51) I tell you [all my secrets] *nearly all of them*

> (52) I saw [their son] yesterday YOU KNOW *the one with the*
>   *blond hair*

7.2 A Third Repair Strategy?

In our repair elicitation paradigm, the cause of the repair—a perceived pictorial change—is external to the language production modules (cf. Fig. 1). In everyday life, too, many self-repairs are occasioned by changes in the perceptual world or by "changes of mind." In case of speech errors, however, self-corrections—if any—follow upon some temporary malfunctioning of the language production system. In (50), for instance, the lemma for *write* instead of *read* was activated. This error must be attributed to either the conceptual module, the lexico-syntactic module, or the mental lexicon; anyhow, it does not disclose a change of mind in the speaker. Another type of lexical error is caused by exchanges of lexemes rather than lemmas (cf. (53) and (54)). Presumably, such "malapropisms" can be traced down to either the mental lexicon or the morpho-phonological module. Finally, there are phonological errors most probably originating in the articulatory module (see (55)–(57)).

> (53) Vorig jaar was deze atleet   de snelste van zijn
>   Last year was this athlete the fastest of   his
>   [contingent], van zijn *continent*
>   contingent   of   his continent

> (54) They haven't been [married], . .UH *measured* with the
>   precision you're using   (from Garrett, 1980)

> (55) En   dan een kwartier   voor   tijd deze aanval
>   And than a   quarter-of-an-hour before end this   attack
>   van [Meinheim], *Mannheim*
>   of   Meinheim   Mannheim

> (56) We wilden een statie[portet], *portret* hebben
>   We wanted a   state-portrait portrait have

(57) Can you tell me when the next [mus] to, *bus* to Monticello
     leaves?

(from Shattuck-Hufnagel, 1979)

Nooteboom (1980) observed that phonological errors are much more
often handled by direct replacement than lexical errors (compare (53)
with (55)). And even after a delayed interrupt there are very few re-
tracings beyond the substituted word (cf. (57)). In Table 4 we proposed to
subsume phonological error repairs under rule R1 as a marked case.
However, it may prove more adequate to treat phonological error repairs
as a separate correction system characterized by the near absence of de-
layed interrupts and of backtracking beyond the reparandum. Thus we
could do justice to the fact that phonological errors differ from lexical
errors both in their origin and in the way they are corrected. Following
this line of argument, there would be a third, articulatory correction
strategy alongside of the syntactic and prosodic ones.

7.3 The Phonological Phrase and the Morpho-Phonological Module

The model we proposed in Section 5 critically depends on the assump-
tion that the phonological phrase functions as the main processing unit
within the morpho-phonological module. In this context it is important to
remember an essential feature of the procedure for partitioning sentences
into phonological phrases (see (30)). While going through a sentence from
left to right, it deposits boundary markers *in a retrospective manner:*
when hitting upon a prominent word, it places a marker after the *pre-
ceding* prominent word. This implies that at least *two* phonological
phrases are being processed at the same time (except at the beginning of
an utterance, of course). Direct psychological evidence in favor of this
claim is not available. But indirect support may be inferred from a type of
speech errors called "combined-form exchanges" by Garrett (1975). He
attributes them to the "positional level of processing" which, in our ter-
minology, makes part of the morpho-phonological module. Examples of
combined-form exchanges from Garrett's corpus are given in (58).

(58) Fancy getting your *model* re*nosed*
     She's already *trunk*ed two *pack*s
     It just *sound*ed to *start*
     I'm not in the *read* for *mood*ing
     He didn't get a*well* so *long*

Garrett observed that the exchanged elements always stem from
nearby positions in the sentence. At most, two words can intervene, none
of them ever being a content word. Bierwisch (1981) hypothesized that
such errors are constrained in terms of metrical planning units: the utter-

ance domain over which elements are exchanged contains at most one relative accent peak. Pursuing this line of thought, we suggest that the morpho-phonological module can process no more than two phonological phrases simultaneously. This would limit the exchange domain to adjacent phonological phrases.

It is tempting to speculate on the function served by phonological phrases. One possibility which easily springs to mind is that they chop up a sentence into chunks of roughly comparable size. Presumably, the resulting segments are more suitable as input to the articulatory module than major syntactic constituents which diverge so widely in length. Another possibility relates to the computation of intonation contours—a job the morpho-phonological module is heavily involved in (Van Wijk & Kempen, 1985). Whatever the true reason of its existence, in virtue of its association with the morpho-phonological module, the phonological phrase may become a useful source of constraints on interpretations of language production phenomena. A recent case in point is provided by the "performance structures" which Gee and Grosjean (1983) inferred from sentence-internal pausing data. Since these structures are, to a large extent, describable in terms of phonological phrases, we place their origin in the morpho-phonological module (Van Wijk, 1987). However, in view of the great potential for psycholinguistic theorizing held by the notion of phonological phrases, we feel that its incomplete definition is a serious disadvantage. Hopefully, this obstacle can be removed in concerted action by linguists and psychologists.

APPENDIX A
Nonretracing Self-Repairs: Some Basic Types

Nonretracing repairs come in many varieties. It is often hard to tell them apart from normal cases of incremental sentence production (e.g., (1)). That indeed a repair is at issue, is discernible on the basis of prosodic characteristics such as pitch contour and pausing pattern (intonational errors; cf. Cutler, 1980; Van Wijk & Kempen, 1985). A clearer example is given by (2). The end of the utterance is marked intonationally, and an editing expression precedes the adjoined constituent. The interspersed elements need not be restricted to editing terms. In (3), for example, the conversational partner butts in with a short question.

  (1) I want you to be happy . . . tonight

  (2) How did things go after the accident? . . . I MEAN between you and John

  (3) I'm going to leave school soon (interviewer: are you?) WELL when I'm sixteen anyway

                                                              (from Rogers, 1978)

Usually it is the syntactic structure which bears out that a repair has been made. We briefly discuss four categories which make up the large majority of nonretracing repairs in our collection. (Most of them were appropriateness repairs.)

The first group is called *tensed verb reduplication*—a construction which is considered unacceptable by prescriptive grammars but occurs frequently in spontaneous speech

(Jansen, 1981, 1985). This "portmanteau construction" (in Kroch & Hindle's, 1982, termi-
nology) is characterized by a constituent belonging to two successive clauses or sentences
(e.g., (4) and (5)). The first clause need not be completed syntactically or semantically
before the switch from the first to the second clause is made (e.g., (6)). When embarking
upon the second clause, speakers do not always repeat the lexical material of the first one
but sometimes seize the opportunity to make an A-repair. For example, in (4) and (6) a
pronoun is replaced by a more definite referent; in (5) a more specific main verb is intro-
duced.

(4) [That]'s the only thing he does is *fight* (from Kroch & Hindle, 1982)

(5) Ik [ken] alleen oom Bertus *kan* ik *me herinneren*
    I know only uncle Bertus can I remember

(6) [Hij] is vorig jaar is *D.* ermee gestopt
    He   has last year has D. with-it stopped

Three further syntactic constructions jumped at by speakers when attempting to correct
an utterance are *extraposition, adverb-over-verb,* and *right-dislocation* (see (7), (8)–(9), and
(10)–(17), respectively).
A clear example of repairing by extraposition concerns the prepositional phrase. In (7)
the PP follows a sentence tag with "question intonation," suggesting it was added as an
afterthought. The Adv-over-V construction in (8) and (9) verges on ungrammaticality; its
repairlike origin is attested by the pause preceding the adverb (cf. Jansen, 1980).

(7) Ik weet niet of   ik dat een [beperking] moet noemen HE? *van het*
    I know not whether I that a   restriction must call   he of the
    celibaat
    celibacy

(8) Ik heb [zes zeven jaar] alleen gezeten . . . *misschien*
    I have six seven years alone been   maybe

(9) Maar vroeger werd dat [gewoon] aanvaard . . . *schijnbaar*
    But formerly was that simply accepted   apparently
                                             ((8)–(9) from Jansen, 1980)

There are no reasons to restrict the term right-dislocation to corrections/expansions of
deictic pronouns (see (10)–(12), (14), (16), (17)). We include cases such as (13) and (15)
where nonpronominal reparanda are replaced. Against right-dislocation examples such as
(10) one might object that they were planned and do not represent corrections. This objec-
tion does not hold for (11), however, where a change is made from singular to plural, and is
implausible with respect to (12) and (13) where editing terms intervene. The reparandum
need not be sentence-initial as in (10)–(13), but may occur in other positions as well (see
(14)–(15)). Right-dislocations need not be appended at the end of the original utterance;
they may also be inserted mid-sentence, preferably at the boundary between finite clauses
(see (16)–(17)).

(10) [Dat] vond ik plezierig *het opbouwen van contacten*
     That found I pleasant the building-up of   contacts

(11) [Dat] moet in de menie *de hoekijzers*
     That must in the minium the wall-ties

(12) [Toen] hebben we eerst thee gedronken DUS *na   het eten*
     Then   have   we first tea drunk   that-is after the meal

(13) En   [naar links] een groen punt NEE OF *rechtdoor*
     And to   left   a   green node no   or   straight-on

                                                 (from Levelt, 1983)

(14) Volgens   mij ben ik [er]   niet voor in de wieg   gelegd voor *het casino*
     According-to me am I there not for   in the cradle put   for   the casino

(15) Zal   ik jou eens in [bad] stoppen? NEE in *bed* BEDOEL IK
     Shall I you   in bath put   no   in bed mean   I

(16) [Hij] heeft toegezegd *de minister zelf*   DUS dat . . .
     He   has   promised the minister himself that-is that

(17) Ook moet je   als sportarts   bedenken dat als je   [daar] nee tegen zegt
     Also must you as sports-doctor consider   that if you there no to   say
     tegen *dopinggebruik* dat . . .
     to   use-of-doping that

APPENDIX B

A Parametric Reconstruction of the Data

B.1 A Reconstruction of Lemma Substitution Scores

Figure i is a schematized version of the decision tree for applying lemma substitution rules (cf. Fig. 3). There are nine bifurcations (at nodes A to I) leading to nine different terminal nodes (denoted #1 through #9). Of the latter, only three are associated with a complete-NP restart, viz. #5, #6, and #8.

When traversing the schema (from node A to one of the terminal nodes), the direction to follow is usually chosen deterministically (e.g., an adjective is either present or not). In three cases, however, a probabilistic choice has to be made. These relate to the parameters summarized in Table i.
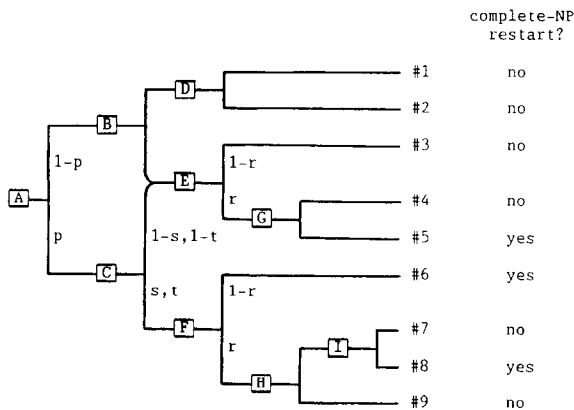

FIG. i. A schematized version of the decision tree for the application of lemma substitution rules presented in Fig. 3.

TABLE i

Parameters Needed for Applying Lemma Substitution Rules

---

p: The Monitor's accuracy in locating interrupt position

    *p*1: Probability that the interrupt is located at or beyond PP-boundary when it actually occurred at PP-boundary (immediate interrupt)

    *p*2: Probability that the interrupt is located at or beyond PP-boundary when it actually occurred within verb (= verb interrupt)

    *p*3: Probability that the interrupt is located at or beyond PP-boundary when it actually occurred beyond verb (>verb interrupt)

r: Probability that the marked restart target is chosen

s: Probability that in a passive the phonological phrase boundary precedes auxiliary verb

t: Probability that the "end of message" marker is read as a phonological phrase boundary

---

Note. p is subjected to the ordinal restriction $p1 < p2 < p3$.

The parameter p (at node A) denotes the probability that the interrupt is *perceived* to have occurred at or beyond the PP boundary. Of course, the accuracy of this perceptive judgment increases when the actual interrupt is delayed further. Therefore different, increasing values have been estimated for each of the three interrupt positions, "immediate," " = verb," and ">verb" ($p1$, $p2$, and $p3$, respectively).

Parameters s and t (node C) concern peculiarities of the phonological structure. The two different prosodic segmentations of passive sentences influence the choice of continuing with either E or F. The latter probability is denoted by s. The same procedure applies to the "end of message" marker in active sentences. If this marker is interpreted as a phonological phrase boundary, option F will be chosen. This probability equals t. Parameter r (nodes E and F) represents the probability of selecting the marked restart target.

The formulae for predicting scores from the decision tree and its parameters are presented in Table ii (for repairs of sentence-initial PPs) and in Table iii (for those of sentence-final PPs). First, we present formulae for all different sentence contexts which happened to occur in the corpus. These we call "simple cases." (Since we want to predict the percentage of complete-NP restarts, parameters are considered only when belonging to a path which leads to such a restart, i.e., #5, #6, and #8.) Then, equations are given for a number of "compound cases." This is necessary because the observed scores reported in the various Tables are always combinations of several different "simple cases." For example, the score for actives is based on the responses to actives with and without an adjective. In making these combinations we had to take into account the relative frequency with which the "simple cases" occurred in the experimental design: actives occurred twice as often as passives, and nominal phrases with an adjective were as frequent as those without one.

In order to compute the predicted scores we substituted the following values for the parameters:

$$p1 = .65 \qquad p2 = .85 \qquad p3 = .95 \qquad r = .20 \qquad s = .50 \qquad t = .33$$

These estimates were obtained informally. They do not necessarily represent an "optimal solution" (e.g., in a least squares sense).

The results of the parametrization are presented in Table iv. The estimated and observed scores correlated .98. As a single overall test of goodness of fit, we applied a technique developed by Mosteller (as reported in Torgersson, 1958, p. 186). This test uses the inverse-sine transformation in order to obtain a statistic which follows a χ^2 distribution. There

TABLE ii

Lemma Substitutions within the Sentence-Initial PP at Three Interrupt Positions
(Immediate, = Verb, or >Verb; Obtained Data Are Presented in Tables 2, 3, 5, and 6)

Interrupt position: *immediate* (directly after PP)

Simple cases

S1 Act +adj   : #3 #4 #6 #7: $(1 - r)p1$

S2 Pas +adj ϕ < aux: $S2 = S1$

S3 Act −adj   : #3 #5 #6 #8: $(1 - p1)r + (1 - r)p1 + rp1 = S1 + r$

S4 Pas −adj ϕ < aux: $S4 = S3$

S5 Pas +adj ϕ > aux: #3 #4 #3 #4: 0

S6 Pas −adj ϕ > aux: #3 #5 #3 #5: $r(1 - p1) + rp1 = r$

Compound cases

C1 Act : $.5*S1 + .5*S3 = (1 - r)p1 + .5r$

C2 Pas : $s[.5*S2 + .5*S4] + (1 - s)[.5*S5 + .5*S6] = s(1 - r)p1 + .5r$

C3 +adj: $.67*S1 + .33[s*S2 + (1 - s)S5] = .33(2 + s)(1 - r)p1$

C4 −adj: $.67*S3 + .33[s*S4 + (1 - s)S6] = C3 + r$

Interrupt position: = *verb* (within or directly after verb)

Simple cases

S7 Act +adj   : #3 #4 #6 #7: $(1 - r)p2$

S8 Pas +adj ϕ < aux: $S8 = S7$

S9 Act −adj   : #3 #5 #6 #8: $(1 - p2)r + (1 - r)p2 + rp2 = S7 + r$

S10 Pas −adj ϕ < aux: $S10 = S9$

S11 Pas +adj ϕ > aux: $S11 = S1$

S12 Pas −adj ϕ > aux: $S12 = S3$

Compound cases

C5 Act : $.5*S7 + .5*S9 = (1 - r)p2 + .5r$

C6 Pas : $s[.5*S8 + .5*S10] + (1 - s)[.5*S11 + .5*S12]$
$$= (1 - r)[(1 - s)p1 + sp2] + .5r$$

C7 +adj: $.67*S7 + .33[s*S8 + (1 - s)S11] = .33(1 - r)[(1 - s)p1 + (2 + s)p2]$

C8 −adj: $.67*S9 + .33[s*S10 + (1 - s)S12] = C7 + r$

Interrupt position: >*verb* (within postverb phrase)

Simple cases

S13 Act +adj   : #3 #4 #6 #9: $(1 - r)p3$

S14 Pas +adj ϕ < aux: $S14 = S13$

S15 Act −adj   : #3 #5 #6 #9: $r(1 - p3) + (1 - r)p3 = S13 + r(1 - p3)$

S16 Pas −adj ϕ < aux: $S16 = S15$

S17 Pas +adj ϕ > aux: $S17 = S7$

S18 Pas −adj ϕ > aux: $S18 = S9$

Compound cases

C9 Act : $.5*S13 + .5*S15 = .5[(2 - 3r)p3 + r]$

C10 Pas : $s[.5*S14 + .5*S16] + (1 - s)[.5*S17 + .5*S18]$
$$= (1 - s)(1 - r)p2 + .5[s(2 - 3r)p3 + r]$$

C11 +adj: $.67*S13 + .33[s*S14 + (1 - s)S17] = .33(1 - r)[(1 - s)p2 + (2 + s)p3]$

C12 −adj: $.67*S15 + .33[s*S16 + (1 - s)S18]$
$$= .33(1 - s)(1 - r)p2 + .33(2 + s)(1 - 2r)p3 + r$$

Note. In case of S5 and S6 the same nodes are reached twice via different routes (node B and node C).

TABLE iii

Lemma Substitutions of the Sentence-Final PP at Two Interrupt Positions (Immediate or Delayed; Obtained Data Are Presented in Tables 2, 3, 7, and 8)

Interrupt position: *immediate* (directly after PP)

  Simple cases

    S19 Act +adj: $S19 = S5$
    S20 Pas +adj: $S20 = S1$
    S21 Act −adj: $S21 = S6$
    S22 Pas −adj: $S22 = S3$

  Compound cases

    C13 Act : $.5*S19 + .5*S21 = .5r$
    C14 Pas : $.5*S20 + .5*S22 = C1$
    C15 +adj: $.67*S19 + .33*S20 = .33(1 - r)p1$
    C16 −adj: $.67*S21 + .33*S22 = C15 + r$

Interrupt position: *delayed* (usually beyond particle/participle)

  Simple cases

    S23 Act +adj $\# \neq \phi$: $S23 = S5$
    S24 Pas +adj   : $S24 = S13$
    S25 Act −adj $\# \neq \phi$: $S25 = S6$
    S26 Pas −adj   : $S26 = S15$
    S27 Act +adj $\# = \phi$: $S27 = S7$
    S28 Act −adj $\# = \phi$: $S28 = S9$

  Compound cases

    C17 Act : $t[.5*S27 + .5*S28] + (1 - t)[.5*S23 + .5*S25] = t(1 - r)p2 + .5r$
    C18 Pas : $.5*S24 + .5*S26 = C9$
    C19 +adj: $.67[t*S27 + (1 - t)S23] + .33*S24 = .33(1 - r)[2tp2 + p3]$
    C20 −adj: $.67[t*S28 + (1 - t)S25] + .33*S26$
$$= .67t(1 - r)p2 + .33(1 - 2r)p3 + .67r$$

Note. The delayed interrupt is considered to be ">verb" in passives (S24 and S26, i.e., use p3), and " =verb" in actives (S27 and S28, i.e., use p2). See also Footnote 2.

proved to be no significant difference between the theoretical and empirical scores ($\chi^2(22) = 26.95, p > .20; N = 77$).[7]

B.2 A Reconstruction of Reformulation Scores

Figure ii presents a schematized version of the decision tree for application of reformulation rules (cf. Fig. 4). There is one choice point (node J) leading to two different terminal nodes (#10 and #11). Only the latter is associated with a complete-NP restart.

The single parameter p in Fig. ii reflects the monitor's accuracy of determining the location of the interrupt. Again, three values for p are estimated corresponding to the three interrupt positions (see Table v).

The predicted score simply amounts to the value of p at that specific interrupt position. The following values were assigned, again informally, to the levels of the parameter:

$$p4 = .25 \qquad p5 = .75 \qquad p6 = .90$$

---

[7] Degrees of freedom were equal to the number of reconstructed proportions (28) minus the number of parameters estimated (6: $p1$, $p2$, $p3$, r, s, and t). N denotes the average number of observations per proportion.

TABLE iv
Results of the Parametrization of Lemma Substitutions

Right-hand context				Left-hand context			
Table	Case	Est.	Obs.	Table	Case	Est.	Obs.
Sentence-initial							
(2) Act SUB	C1	62	67	(3) + Adj SUB	C3	43	45
DEL	C1	62	68	DEL	C3	43	50
Pas SUB	C2	36	35	− Adj SUB	C4	63	63
DEL	C2	36	44	DEL	C4	63	66
(5) = verb Act	C5	78	81	(6) = Verb + adj	C7	65	68
Pas	C6	70	72	− adj	C8	85	83
> verb Act	C9	77	79	> Verb + adj	C11	75	76
Pas	C10	77	76	− adj	C12	79	75
Sentence-final							
(2) Act SUB	C13	10	8	(3) + Adj SUB	C15	17	12
DEL	C13	10	20	DEL	C15	17	21
Pas SUB	C14	62	62	− Adj SUB	C16	37	23
DEL	C14	62	55	DEL	C16	37	37
(7a) Act	C17	33	36	(7b) + Adj	C19	40	38
Pas	C18	77	77	− Adj	C20	47	53

For sentence-initial NPs results are presented in Table vi. The correlation between estimated and observed scores was quite high: .99. Mosteller's test showed a very favorable outcome ($\chi^2(9) = 5.10$, $p > .80$; $N = 59$).

For sentence-final NPs results are given in Table vii. Note that for scores reflecting the effect of left-hand context after a delayed interrupt (see Table 7b), estimates were not equal to a value of p. Remember that the interrupt position differs for the two sentence types: ">" in active sentences, but ">>" in passives (cf. Footnote 2). Since in the experimental design active sentences occurred twice as often as passives, the predicted score equals $[.67 * p5 + .33 * p6]$ (C21).

Although the correlation remained rather high (.95), Mosteller's test showed a very significant discrepancy between the observed and predicted scores ($\chi^2(9) = 121.49$, $p < .001$; $N = 59$). Only for the cases where we could single out a unique set of reformulations, i.e., the retracing repairs after a delayed interrupt, predicted scores approached the observed ones (see the results in Tables 7a and b). In all other cases, the observed score remained far below the estimated value. As explained in Section 6.4, we attribute this finding to interference with a nonretracing repair strategy, namely, extraposition.
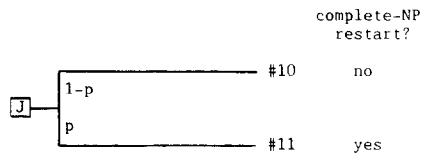

FIG. ii. A schematized version of the decision tree for the application of reformulation rules presented in Fig. 2.

TABLE v

Parameters Needed for Applying the Reformulation Rule

---

p: The Monitor's accuracy in locating interrupt position

    $p4$: Probability that the interrupt is located beyond PP-boundary when it actually occurred at PP-boundary (immediate interrupt)

    $p5$: Probability that the interrupt is perceived beyond PP-boundary when it actually occurred within verb (= verb interrupt)

    $p6$: Probability that the interrupt is perceived beyond PP-boundary when it actually occurred beyond verb (>verb interrupt)

Note. p is subjected to the ordinal restriction $p4 < p5 < p6$. In comparison with the estimates for lemma substitutions, a second restriction applies: $p1 > p4$, $p2 > p5$, and $p3 > p6$. This follows from the fact that cases where the interrupt is located *at* the PP-boundary do not count here.

TABLE vi

Results of the Parametrization for Reformulations of Sentence-Initial NPs

Right-hand context				Left-hand context			
Table	Case	Est.	Obs.	Table	Case	Est.	Obs.
(2) Act	p4	25	26	(3) +adj	p4	25	27
Pas	p4	25	21	−adj	p4	25	20
(5) = verb Act	p5	75	75	(6) = verb +adj	p5	75	77
Pas	p5	75	68	−adj	p5	75	76
>verb Act	p6	90	90	>verb +adj	p6	90	88
Pas	p6	90	94	−adj	p6	90	93

TABLE vii

Results of the Parametrization for Reformulations of Sentence-Final NPs

Right-hand context				Left-hand context			
Table	Case	Est.	Obs.	Table	Case	Est.	Obs.
(2) Act	p4	25	6	(3) +adj	p4	25	9
Pas	p4	25	13	−adj	p4	25	2
(7a) Act	p5	75	74	(7b) +adj	C21	80	82
Pas	p6	90	89	−adj	C21	80	84
(8) Act SUB + DEL	p5	75	44				
ADD	p5	75	56				
Pas SUB + DEL	p6	90	69				
ADD	p6	90	83				

REFERENCES

Bierwisch, M. (1981). Linguistics and language error. *Linguistics*, 19, 583–626.

Bock, K. (1987). Exploring levels of processing in sentence production. In G. Kempen (Ed.), *Natural language generation: New results in artificial intelligence, psychology and linguistics* (pp. 351–363). Dordrecht/Boston: Kluwer Academic Publishers.

Cutler, A. (1980). Errors of stress and intonation. In V. Fromkin (Ed.), *Errors in linguistic performance* (pp. 67–80). New York: Academic Press.

Cutler, A. (1981). The reliability of speech error data. *Linguistics*, 19, 561–582.

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

Garrett, M. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production. Vol. 1: Speech and Talk* (pp. 177–220). New York: Academic Press.

Gee, J., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, 15, 411–458.

Jansen, F. (1980). De sociolinguïstische benadering van spreektaalsyntaxis: het bijwoord achter de werkwoordelijke eindgroep. In G. Geerts & A. Hagen (Eds.), *Sociolinguïstische Studies 1* (pp. 131–153). Groningen: Wolters-Noordhoff.

Jansen, F. (1981). *Syntaktische konstrukties in gesproken taal*. Amsterdam: Huis aan de Drie Grachten.

Jansen, F. (1985). *A quantitative investigation of left dislocation and verbal repetition in spoken Dutch*. Manuscript, University of Amsterdam.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, 11, 201–258.

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, 14, 185–209.

Kroch, A., & Hindle, D. (1982). On the linguistic character of non-standard input. In *Proceedings of the 20th annual meeting of the Association for Computational Linguistics* (pp. 161–163).

Levelt, W. (1983). Monitoring and self-repair in speech. *Cognition*, 14, 41–104.

Nespor, M., & Vogel, I. (1982). Prosodic domains of external Sandhi rules. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations* (Pt. I, pp. 225–255). Dordrecht: Foris Publications.

Nespor, M., & Vogel, I. (1983). Prosodic structure above the word. In A. Cutler & R. Ladd (Eds.), *Prosody: Models and measurements* (pp. 123–139). Berlin: Springer-Verlag.

Nooteboom, S. (1980). Speaking and unspeaking: Detection and correction of phonological and lexical errors in spontaneous speech. In V. Fromkin (Ed.), *Errors in linguistic performance* (pp. 87–95). New York: Academic Press.

Pijls, F., & Kempen, G. (1986). Een psycholinguistisch model voor grammatische samentrekking. *De Nieuwe Taalgids*, 79, 217–234.

Rogers, S. (1978). Self-initiated corrections in the speech of infant-school children. *Journal of Child Language*, 5, 365–371.

Selkirk, E. (1980). Prosodic domains in phonology: Sanskrit revisited. In M. Aronoff & M.-L. Kean (Eds.), *Juncture* (pp. 107–129). Saratoga, CA: Anma Libri.

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

Torgerson, W. (1958). *Theory and methods of scaling*. New York: Wiley.
van Wijk, C. (1987). The PSY behind PHI: A psycholinguistic model for performance struc-
    tures. *Journal of Psycholinguistic Research,* **16,** 185–199.
van Wijk, C., & Kempen, G. (1985). From sentence structure to intonation contour—An
    algorithm for computing pitch contours on the basis of sentence accents and syntactic
    structure. In B. Müller (Ed.), *Sprachsynthese* (pp. 157–182). Hildesheim: Georg Olms
    Verlag.