

Monitoring and self-repair in speech*

WILLEM J.M. LEVELT

Max-Planck-Institut für Psycholinguistik

Abstract

Making a self-repair in speech typically proceeds in three phases. The first phase involves the monitoring of one's own speech and the interruption of the flow of speech when trouble is detected. From an analysis of 959 spontaneous self-repairs it appears that interrupting follows detection promptly, with the exception that correct words tend to be completed. Another finding is that detection of trouble improves towards the end of constituents. The second phase is characterized by hesitation, pausing, but especially the use of so-called editing terms. Which editing term is used depends on the nature of the speech trouble in a rather regular fashion: Speech errors induce other editing terms than words that are merely inappropriate, and trouble which is detected quickly by the speaker is preferably signalled by the use of 'uh'. The third phase consists of making the repair proper. The linguistic well-formedness of a repair is not dependent on the speaker's respecting the integrity of constituents, but on the structural relation between original utterance and repair. A bi-conditional well-formedness rule links this relation to a corresponding relation between the conjuncts of a coordination. It is suggested that a similar relation holds also between question and answer. In all three cases the speaker respects certain structural commitments derived from an original utterance. It was finally shown that the editing term plus the first word of the repair proper almost always contain sufficient information for the listener to decide how the repair should be related to the original utterance. Speakers almost never produce misleading information in this respect.

It is argued that speakers have little or no access to their speech production process; self-monitoring is probably based on parsing one's own inner or overt speech.

*I am indebted to Ger Desserjer, Irma van Loon, and Marion Keiren who assisted substantially in coding and data analysis, and to Peter Rahner who developed the computer analysis programs. I acknowledge many helpful comments from Eve and Herb Clark, Gary Dell, Paul Eling, Merrill Garrett, Claus Heeschen, Wolfgang Klein, Ewald Lang, Sieb Nooteboom, and Graham Richardson. Reprint requests should be sent to Willem J.M. Levelt, Max-Planck-Institut für Psycholinguistik, Berg en Dalseweg 79, Nijmegen, The Netherlands.

An adequate theory of the organization of natural language will need to depict how a natural language handles its intrinsic troubles. Such a theory will, then, need an account of the organization of repair (Schegloff, *et al.*, 1977).

The purpose of the present paper is to analyze a large corpus of spontaneous self-repairs as a step towards developing such a theory. In the course of developing the theoretical framework it will become increasingly clear that what is to be developed is a theory of monitoring as well as a theory of repairing. We will find that not only the detection of trouble requires monitoring, but also that the production of the repair proper requires the speaker to have access to structural properties of the original utterance. Both aspects of self-repairs can be handled on the assumption that a speaker is continuously parsing his own inner or overt speech. After detecting an error or inappropriateness, the speaker will, in some way, 'transfer' structural properties of the original utterance to the correction. This creates systematic dependencies between original and new utterance. It will become apparent in the course of this paper that these dependencies are quite similar to those between two conjuncts in a coordinate construction, and those between question and answer. This suggests that these, on first view, quite distinct phenomena are due to a common source in language production. A unified treatment might involve the language comprehension or parsing system in all three of these cases. Monitoring one's own or an interlocutor's speech may thus provide the speaker with structural constraints to be implemented on the next utterance, be it a repair, a conjunct, or an answer.

By transferring and reusing structural properties of previous speech the speaker may at the same time gain in fluency, and establish discourse coherence to the advantage of the listener.

Since it is as yet a matter of great unclarity how this transfer of structural information is realized in the mind of the speaker, we will proceed cautiously. Most of the following sections will deal with detailed analyses of a large corpus of self-repairs in speech. These analyses will concentrate, firstly, on the monitoring process which occasionally induces the speaker to interrupt the flow of speech, secondly on the ways in which the speaker signals trouble to the listener by means of editing terms such as 'uh' or 'sorry', and thirdly on features of restarting for the repair proper, features which make it possible for the listener to decide very rapidly on how the new utterance is to be related to the interrupted one.

These descriptive and analytic sections (3-6) will be flanked by shorter theoretical ones: Section 2 outlines a framework for the analysis of production and parsing processes involved in self-repairs, the closing Section 7 relates this framework to some of the main empirical findings of this study.

Section 1 will be devoted to introducing the corpus of repairs, and some of their characteristic properties.

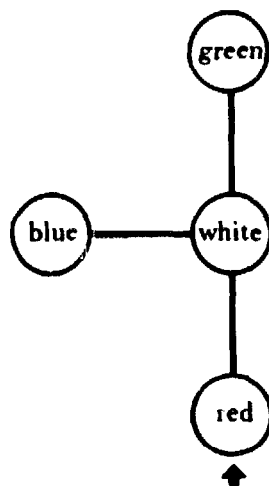
1. The corpus of repairs

The main data basis for the present analysis is a corpus of 959 repairs which were spontaneously made by adult Dutch-speaking subjects who had been asked to describe certain visual patterns. The setting was an experiment designed to study speaker's linearization strategies, i.e., the ways in which speakers order complex information for expression. Results of this experiment and similar ones have been reported elsewhere (Levelt 1981, 1982*a, b, c*).

All patterns to be described consisted of colored dots, connected by arcs having one of two orthogonal directions. The subjects had been asked to describe patterns in such a way that a listener who had seen a range of examples of such patterns would be able to draw the pattern from the tape-recorded description.

A typical example of such a pattern plus an actually occurring repair in a subject's description of it is given in Figure 1. There were 53 subjects in this experiment and there were 53 different patterns, each described once by each

Figure 1. *Example pattern and repaired utterance.*



"Van witte stip rechts eh linksaf naar blauwe stip"
 from white dot right uh left to blue dot

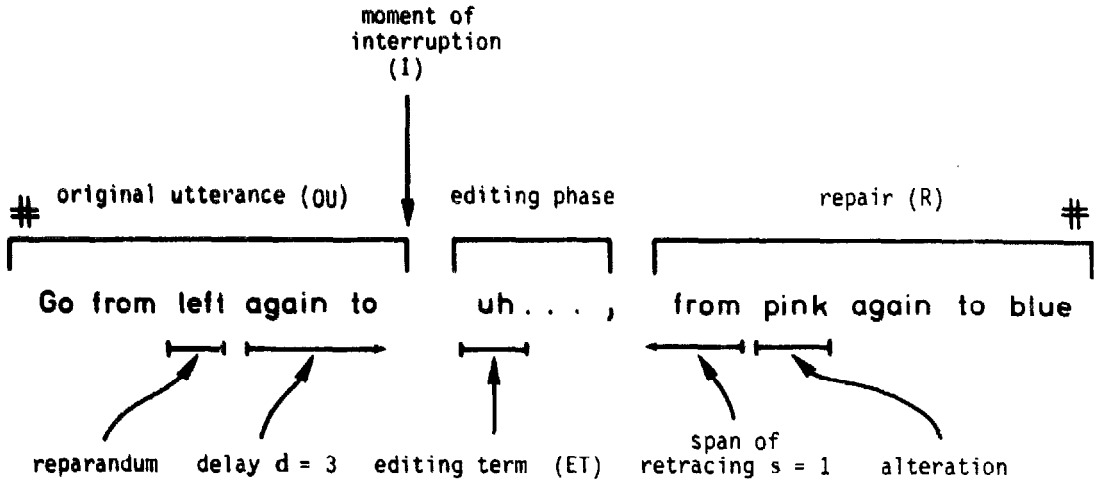
subject. The 959 repairs thus appeared in a total of 2809 pattern descriptions, i.e., at a rate of about one per three descriptions. The average number of repairs per subject was 18.1, ranging from 1 to 49, with a standard deviation of 10.3. All repairs are available on audiotape and were (non-phonetically) transcribed for further analysis. For some parts of the analysis it was necessary to work from the tapes themselves instead of from the transcripts.

The most common type of self-repair in the data is exemplified in (1):

(1) *Go from left again to, uh ..., from pink again to blue*

Such a repair typically consists of three parts. The first part is what we will call the *original utterance* (OU); in the example it is *Go from left again to*. The OU contains the trouble spot or *reparandum*, the item to be repaired. (In (1) this is *left*.) The reparandum can be anything, ranging from a single speech sound to a whole stretch of text. We define the OU to range from the last sentence boundary before the reparandum (i.e., # *Go ...* in (1)) to the *moment of interruption* (I), the point at which the flow of speech is interrupted for 'editing' (Hockett 1967) of some sort. Speech may become interrupted right after the reparandum, or even within the reparandum, but one can also observe delayed interruption. In (1) interruption occurs three syllables (*again to*) after the reparandum (*left*). We will call this the *delay* (d) of interruption, in (1) it has the value 3 (syllables). The second part we will call the editing phase, a shorter or longer period of hesitation (*uh ...*, in (1)) which may or may not contain an *editing term* (ET), (*uh, rather, well*, etc.). The third part is the *repair* (R) proper. We will use this term in spite of the fact that it invites a narrow interpretation: the correct version of what was wrong before. As will be observed shortly, there are many repairs where there is nothing wrong to start with; also many repairs are not correct themselves, sometimes leading to a staggering of additional repairs. Such staggering is not unusual: there are 159 cases in our data where 2 or more repairs cluster in one utterance. (This agrees well with the amount of staggering found by Dietrich (1982) in learners of German.) Still, each component of these multiple repairs will be analyzed independently.

A repair can start directly at the reparandum, or the speaker can retrace to an earlier point. In (1) the *span* of retracing (*s*) is one syllable: the speaker restarts at the preposition (*from*) before the reparandum; the value of *s* is 1 in this case. The repair usually contains an alteration with respect to OU. In (1) the alteration is *pink*, plus whatever prosodic changes were made. If no morphemes are changed, added, or deleted, one has to do with what will be called a *covert repair* (C). The most minimal form of covert repair that we will take into account in the analysis is the case where after the interruption and editing phase, the utterance is continued where it broke off (i.e., zero

Figure 2. *The structure of repair. Some technical notions.*

alteration). So even an unmistakable hesitation involving an editing term like 'uh' is included in the category. Quite common are covert repairs where the same word is repeated without change (*go to a red, red node*).

We take a repair to end at the first sentence boundary after the alteration (i.e. ... *blue #* in (1)). For covert repairs this will be taken to be the first # after the repeated part, and for mere hesitations the first # after the hesitation. Figure 2 summarizes most of the notions introduced so far. It should be obvious that cases can be constructed (and in fact observed!) where some of these notions do not apply, or need further specification. There are, for instance, repairs where the editing expression is distributed over the repair (*to blue, uh, yellow rather*), where the repair is more like an expansion, see example (43), and many other deviant cases. The presently introduced notions, however, suffice for a very substantial part of the analyses; we prefer to discuss additional qualifications only where they are really needed.

2. Perception, production and central control in self-repair

Self-correction in speech results from a complicated interplay of perceptual and productive processes. In order to make a repair, the speaker must, firstly, notice some trouble and interrupt his or her flow of speech, and, secondly, create a new utterance, which takes care of the trouble and its potential consequences for the listener. A theory of the speaker should give an account of both these aspects of the repairing process, but there is, as yet, an enorm-

ous lack of constraints on such accounts. Let us consider the issues in turn.

How does the speaker come to detect a source of trouble in what he is saying? There are two widely different ways in which this could be conceived. The first one is that the speaker has direct access to particular components of the production process. By applying certain criteria to the outputs of these components, alarm signals may result when these criteria are not met to a sufficient degree. The speaker may then decide to stop and reconsider his production. Let us call this 'the production theory of monitoring'. Laver's (1980) theory of monitoring is, in part, a production theory. The second way is to assume that the speaker has no access to the components of production, but only to the final result of the process. The speaker would parse this output (actually, his 'inner speech', see below) as in normal language understanding. He would then be able to detect any structural deviances which he might as well have detected in somebody else's speech, and he can moreover compare the derived message with his original intention. This will be named the 'perceptual theory of monitoring'.

Though there is not yet sufficient evidence to make an informed choice between these two alternatives, we are inclined to take the perceptual point of view. Arguments in support of this choice will be put forward as the paper develops, and the issue will be taken up again in the General Discussion. At this point we will restrict ourselves to mentioning two reasons for preferring a perceptual theory of monitoring. The first is that a perceptual theory avoids unnecessary 'doubling' of devices. It is known that normal language users are perfectly well able to monitor the speech of others for phonetic, syntactic, semantic and pragmatic properties and for distortions (cf., Cohen, 1980; Cole, 1973; Cole and Jakinick 1980; Foss, 1969; Marslen-Wilson and Tyler, 1980; Marslen-Wilson and Welsh, 1978, among others). It is more economical from the theoretical point of view, to assume that the same capabilities are used in monitoring one's own inner or overt speech.

The second reason is the one stressed by Bock, in her extensive review of sentence formulation research (1982), namely that the speaker presumably has no access to intermediate processing results, but only to his communicative intention on the one hand, and to the final products of his formulation process, on the other. In other respects, language production is an 'underground process' (Seuren 1978). Though it should be noted that empirical evidence whether for or against this point of view is still incomplete, it conforms with what Nisbett and Wilson (1977) observed more generally for cognitive processes, namely that only the end products of cognitive operations are accessible for attention, not the processes themselves. Several authors have stressed that the sheer speed (Herriot, 1970; Levelt, 1978) and parallel nature of formulation processes (Kempen and Huybers, 1983) testifies to

their being both automatic and independent of central control. Such characteristics would, of course, be incompatible with the premisses of a production theory of monitoring.

The suggestion that monitoring and editing one's own speech involves the mechanisms of normal speech comprehension is not particularly new. Similar suggestions can be found in Garrett, 1980; Hockett, 1967; Hoenkamp, 1980 (See Reference Note 4), and Laver, 1973, and much earlier in Wernicke's paper (1874) where one can read that "Apart from the lack of understanding the patient thus also has aphasic symptoms in speaking, determined by the failure of this correction unconsciously exercised by the sound image".

To complete the theoretical picture, it is necessary to give a short, and admittedly incomplete, listing of the processing components involved in formulating and repairing, as well as of the degrees of central access the language user has to the outputs of these components.

A. Message construction

This component generates, orders (Levelt, 1981), and delivers elementary intentions or messages to be formulated. These can be propositions, truth claims, summonses, etc.

It draws on situational givens, task requirements, detected trouble, motivations and a large base of long-term knowledge. The generation of these messages is highly accessible to central control; it is this subject material to which a speaker directs his primary attention while speaking. In other words: the messages usually 'pass through' working memory, and they will stay available for some time for comparison with the actual speech output (see below).

B. Formulating

The formulating component generates phonetic strings, i.e. necessary and sufficient instructions for the motor executive programs, normally articulation, to be carried out.

It draws on the messages constructed (see A), and, as we will argue, on information about previous self- or other-produced utterances which results from perceptual parsing or monitoring.

Phenomenologically, only the eventual 'inner speech' is accessible to attention. The precise nature of this 'inner speech' is unclear. For the present purposes we will equate it with the just mentioned phonetic strings.

The formulating component consists of at least four subcomponents:

B₁. Lexicalization

Accessing lexical items is the task of this subcomponent. It is, more specifically, engaged in two activities, which are not necessarily coupled:

(i) Retrieving phonologically unspecified lexical items, or lemmas (the latter term is from Kempen and Huybers, 1983) which are specified semantically, as well as in terms of syntactic category and functional syntactic properties (grammatical roles, roughly as in Bresnan, 1978). These lemma retrieval activities take as input concepts and their relations in the message, as well as lexical information from previous messages still available in working memory.

(ii) Retrieving word forms. These are the phonological shapes of the triggered lemmas. They are retrieved by taking corresponding lemmas as input, but word form information still available from previous discourse may also be effective (Levelt and Kelter, 1982).

Though there exists abundant evidence that these two lexical retrieval activities are not necessarily coupled or even simultaneous (compare for example Bock, 1982; Garrett, 1980; Kempen and Huybers, 1983; Levelt and Maassen, 1981), there is also solid recent evidence that the activities are not independent (Dell and Reich, 1981). As a consequence, form similarity of lexical items may affect the retrieval of their lemmas, and semantic similarity of lemmas may affect word form retrieval.

B₂. Functional frame builder

This subcomponent generates what Garrett (1975; 1980) calls the 'functional level representation', i.e., clause-type units in which the lemmas are arranged in correspondence with their functional syntactic restrictions, and where sentence accent is marked. Necessary input for this component to work on is, firstly, the conceptual relations in the message, including thematicity, modality, perspective, and secondly lemmas retrieved (cf. B₁ (i)).

B₃. Morphological frame builder

This subcomponent produces morphological phrases, having access to the developing functional level representations and their word forms. Garrett (*op. cit.*) calls the result 'positional level representation': it is a phonologically (both segmentally and suprasegmentally) specified string, containing all closed class elements such as pronouns and inflections. There is nothing to exclude the possibility that B₃ operates 'in tandem' with B₂, i.e., in almost parallel fashion, delivering its output phrase by phrase (cf., Bock 1982).

B₄. Phonetic coder

The phonetic coder creates a 'phonetic string' on the basis of B₃'s output. It is a running string of instructions to the articulatory apparatus. Though the

nature of these instructions is quite enigmatic, we will equate it with the subjective notion of 'inner speech'. Inner speech is probably chunked in phonological phrases of some sort, and it is readily accessible to central attention.

C. Articulating

This component translates phonetic strings into overt speech. The major point of importance for a theory of monitoring is that overt speech can be, and normally is, monitored by the speaker for at least some features, such as pitch and vowel quality (as appears, among others, from the delayed auditory feedback literature).

The processes of message construction, formulating, and articulating are essential components of all speech production. A theory of repairing involves, over and above these two further components, a parser and a monitor:

D. Parsing

The parser is a cover term for the sum total of procedures available to a language user for understanding spoken language. Though the normal input here is heard (overt) speech, and the normal output some representation of the intended message of the speaker, the parser is a far more powerful and flexible instrument. In a perceptual theory of monitoring the parser should not only be able to draw on overt, auditorily available speech, but it should be able to parse inner speech as well. It can then compare the derived message with whatever is still available of the original (input) message. The parser can, moreover, derive information other than the intended message, such as linguistic aspects of the speech string: whether particular phonemes of words are spoken, whether particular referents are mentioned, the parser can detect syntactic and prosodic features, voice qualities, etc. In short, a large variety of aspects of parsed information is accessible to attention.

It should, for completeness' sake, be added that the parsing mechanisms will, in part, have access to the same sources of information as the formulating mechanism. The lexicon is an obvious common source, but there may be more, dependent on how much truth there is in an analysis-by-synthesis theory of parsing.

E. Monitoring

The monitor, finally, performs two functions. The first one is a matching function: it compares parsed aspects of inner and outer speech with (i) the

intentions, and the message sent to the formulator, and (ii) criteria or standards of production. The former comparison (i) entails checking whether what was said corresponds to what was intended (cf., Laver, 1973). In the analyses to follow we will have to make an important distinction here: there may be a mismatch between message and speech (for instance when the color concept in the message is BLUE, but the word produced is *red*), or the message itself may need some qualification in the light of the original intention and the context in which the utterance was produced (so, for example when there is potential ambiguity: *From there you go to the left* may then become replaced by *From the green node you go to the left*).

The other comparison (ii) done by the monitor is with standards of production; it has to do with the detection of speech errors, syntactic flaws, etc, but also with maintaining standards of rate, loudness, and other prosodic aspects of speech (cf., Labov, 1970 and Laver, 1973). The following Section 3 will give a far more detailed account of the potential targets or foci of monitoring.

Let us now turn to the second function the monitor performs: it is to create instructions for adjustment. If some mismatch is detected which surpasses certain criteria, the monitor makes the speaker aware of this, or in other words: an alarm signal is sent to working memory. The speaker can then take action on the information received. This can range from quite fluent adjustments, such as in the loudness and rate of speech delivery, to a complete halt-and-restart action. This restarting is not neutral with respect to the interrupted utterance, it usually reinstalls some of the parsed properties of the original utterance. This can be to the benefit of the listener, who has to solve a 'continuation problem', i.e., how to relate the repair to the original utterance. Only these latter stop- and restart activities are the subject of the present study.

In conclusion, then, we prefer to assume that repairing speech involves a perceptual loop: the self-produced inner or overt speech is perceived, parsed and checked with respect to intentional and contextual appropriateness, agreement of intended and delivered message, and linguistic correctness. When trouble is detected, central corrective action is taken. This action is based on the character of the trouble, the still available parsing results (such as wording and constituent structure of the original utterance), and the estimated consequences for the listener. This controlled action can be based on any information the speaker can have central access to. We assume that this is, in principle, possible for the in- and outputs of all five components A-E above, but less so, or not at all for the information exchanged between the subcomponents of B, the formulator. The central control property of self-repair makes it subject to the usual limitations of working memory.

3. Some foci of monitoring and types of repair

It was remarked above that the parser, and hence the monitor has potential access to a large range of aspects of the speech produced. With Laver (1973), the distinction was further made between monitoring for intention and meaning of what is spoken, and monitoring for linguistic deviancy.

The present section reviews the major foci of monitoring that have to be distinguished for an insightful analysis of our corpus of self-repairs. The speaker may attend to any of the following aspects of what he or she is saying; they are formulated as questions:

3.1. *Do I want to say this now?*

The speaker may, while speaking, change his mind and realize that he better expresses another message than the one he is currently formulating. A common cause of this is the speaker's linearization problem (Levelt, 1981; 1982a): if the speaker intends to express some complex state of affairs it is necessary to decide on what to say first, what to say next, etc., that is on the ordering of messages. While speaking, the speaker may realize that another arrangement of messages would be easier or more effective. Example (2) from our corpus is an instance where such a state of affairs presumably triggered the speaker to repair.

- (2) We gaan rechtdoor offe ... We komen binnen via rood, gaan dan
We go straight on or ... We come in via red, go then
 rechtdoor naar groen
straight on to green.

The speaker realizes that another idea than the current one has to be expressed first and interrupts his speech to start anew. The current message is replaced by a *different* one. We will call such repairs *D-repairs*. They are quite infrequent in our corpus, we counted 10 of them (1%).

3.2. *Do I want to say it this way?*

Even if there is no doubt in the speaker's mind about the information to be expressed at a particular moment in discourse, the message may still vary dependent on the contextual information which is taken into account. The message can be more or less appropriate given what was previously said (or better: remembered to have been said), given the social and perceptual features of the interlocution situation, etc. The speaker may, while speaking,

become aware that the way he expresses the intended information (idea, concept, proposition) needs qualification in view of the context of expression. This awareness can either be based on the speaker's parsing his 'inner speech', and comparing the derived intention to the original one, or the speaker could, alternatively, do the monitoring at the message-level, i.e., testing the message under construction *before* any verbal formulating has taken place. As we will see, there is strong evidence that the former is often the case; this does not exclude, however, that the latter may occur as well. In all cases, we will speak of *appropriateness-repairs*, or *A-repairs*.

There are three important aspects of appropriateness that a speaker may monitor for. They are potential ambiguity given the context, the use of appropriate level terminology, and coherence with previously used terms or expressions. Let us consider these in turn.

Potential ambiguity of reference is presumably the occasion for the repair in (3):

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

Here the speaker noticed that *the middle* could also mean the middle of the pattern instead of the middle of the paper, and that there is no way for the hearer to know. The idea to be expressed therefore needs more explicit phrasing. (Notice that we take the reparandum here to be *the middle*, and that the alteration consists of an addition (*of the paper*)). Repairs intended for *ambiguity reduction* abound in our data. They often have to do with demonstratives (*From that one, the blue one, you go left*), or referentially ambiguous deictic expressions. A-repairs having to do with ambiguity of reference will be coded *AA-repairs*. There are 46 of them in the corpus (5%).

The case of monitoring for appropriate level terminology is clearly exemplified by (4):

- (4) ... met een blauw vlakje, een blauw rondje aan de bovenkant
with a blue spot, a blue disc at the upper end

Clearly, a blue disc is a blue spot, but the former term is somewhat more precise: the speaker is trying to find the *appropriate level* for expressing the core of the concept to the hearer. This type of repair usually goes from a less to a more precise term. A-repairs which shift the level of terms will be indicated as *AL-repairs*. The corpus contains 129 of them (13%).

The third aspect of appropriateness the speaker may monitor for is coherence with previous text, especially previously used terminology. An example in case is (5):

- (5) Ga je een naar boven, is uh ... kom je bij geel
Go you one up, is uh ... come you to yellow

Though it is not known what would have followed *is*, it should be an NP expressing the concept of a yellow node. This would have been correct and unambiguous. But after the 'you go up' in the previous clause, it is certainly more coherent to use another verb of motion at this place. Such repairs are quite frequent, there are 47 cases in the corpus (5%). They are coded as *AC-repairs*. For completeness's sake it should be added that for some repairs it is impossible to determine unambiguously whether the speaker makes a level-adaptation for a term, as in (4), or establishes coherence, as in (5). Such doubt would arise, for instance, if the speaker of (4) had been using the term 'disc' in all previous discourse (which, by the way, was not the case). Such repairs we coded as *ALC-repairs*. There are 68 such cases (7%) in our data.

It is important to notice that monitoring for ambiguity, and for coherent and appropriate level terminology is not monitoring for *error*. In the examples (3)–(5) above the OU was correct given the concepts to be expressed, the repairs were only made to express the same ideas more appropriately. Adding together the different types of appropriateness monitored for, there are 290 A-repairs in the corpus, i.e., 30% of all repairs.

3.3. Am I making an error?

When a speaker has no doubts about the idea expressed, or the appropriateness of the formulation, trouble may still arise. A speaker may discover that what he is saying contains an error of some sort, often to his own surprise. An example is (6):

- (6) Rechtdoor rood, of sorry, rechtdoor zwart
Straight on red, or sorry, straight on black

Here the speaker made a lexical error, the OU contained an erroneous color term, and the repair replaces it by the correct one. It is, of course, not completely decidable that the speaker meant 'black' to start with, a perceptual error may have been involved. We will have more to say about this in Section 4. Here it suffices to notice that lexical errors can involve almost any

lexical item, color words, direction terms, prepositions, articles, etc. Some further examples from the corpus are given in (7) to (9):

- (7) Linksaf naar, ... rechtsaf herstel naar blauw
Left to , ... right correction to blue
- (8) Ga dan naar de verkeer ..., naar de andere kant
Go then to the wrong ..., to the other side
- (9) Sla linksaf bij knooppunt, naar knooppunt blauw
Turn left at node , to node blue

For many of these and similar errors, a perceptual cause is unlikely. It is rather more probable that, in the theoretical terminology of the previous section, the formulator (B) was given the right input message, but that the wrong lexical item(s) got activated and phonetically realized as output.

Lexical error repairs are very frequent: there are 369 of them in the corpus (38%). We call them *EL-repairs*. But there are still other errors than lexical ones a speaker may become aware of. We have distinguished two further types of error-repairs: syntactic and phonetic ones.

In a syntactic repair (*ES-repair*), the speaker starts a syntactic construction which leads into a deadlock and which is subsequently repaired. An example is:

- (10) En zwart ... van zwart naar rechts naar rood
And black ... from black to right to red

Here a prepositional phrase is needed to describe the source of the next move, but the speaker started with an NP. Sometimes, syntax becomes fully scrambled for some reason, and the speaker starts all over again. These cases are not very frequent (N = 22, i.e., 2%).

Phonetic repairs (*EF-repairs*) are far less frequent than the literature on speech errors may suggest. There are no more than 8 of these in our data, i.e., only 1% of all repairs fall in this category. An example is given in (11):

- (11) Een eenheid, eenheid vanuit de gele stip
A unut , unit from the yellow dot

The sum total of error repairs, or *E-repairs* in the corpus amounts to 399, a share of 42%. It should be kept in mind that many speech errors, roughly half of them, are never repaired. We will return to this below.

Though this completes our listing of the speaker's foci of monitoring, some additional cases should be mentioned. There can be no doubt that a speaker

monitors for prosodic features of his speech. We have not observed any cases where this monitoring led to a halt-and-restart procedure. This type of monitoring will therefore not be dealt with in this article. There are, however, two 'default' categories of repairs which must be taken into account. We already mentioned the category of covert repairs (*C-repairs*): they are characterized by either just an interruption plus editing term (N = 167, 17%), or the repeat of one or more lexical items (N = 69, 7%). Examples are (12) and (13):

- (12) Dan rechtsaf, uh grijs
 Then right , uh grey
- (13) En aan de rechterkant een oranje stip, oranje stip
 And at the right side an orange dot , orange dot.

Covert repairs are problematic data in that it is almost always impossible to determine what the speaker is monitoring for. For (12) and (13) it may or may not be the case that the color term was attended to. Since nothing gets changed in the end there is no basis for deciding. It is even impossible to decide whether a covert repair results from a 'false alarm' of the monitor, though this is surely a theoretical possibility. Though *C-repairs* abound in the corpus (N = 236, 25%), we will make only very limited use of them just because the target of the repair is unclear. What many of the covert repairs, those of type (12), do tell us, however, is that monitoring can take place before the utterance is overtly expressed. This is an argument, though not a sufficient one, for the assumption that some level of 'inner speech' is accessible to attention. Other evidence for this comes from Dell (1980, see Reference Note 1), who showed that subjects could monitor inner speech for speech errors.

Finally, there is a small set of 'repairs' which are so completely confused that they defy any systematic categorization other than 'rest category'. These *R-repairs* count up to 24 (2.5%).

This brings us to a remark on scoring. All repairs were scored by at least two independent judges (trained student assistants, see acknowledgements). The largest subset of repairs judged by the same pair of judges contained 514 cases. The *initial* scoring on the above categories corresponded between the judges in 73% of the cases. Non-corresponding cases were always discussed between the two judges to see whether agreement could be reached. In the negative case a third judge (usually the present author) was consulted in order to decide on a final code. The full corpus of repairs and their final codes are available on request. As will appear shortly, the final codes involve many more aspects of the repairs than the ones discussed in the present section.

The next three sections will successively treat the three phrases of repair: the OU and its interruption, the use of editing terms, and the factual construction of repair.

4. Interrupting the utterance and the occasions for repairs

This section will start with the expression of what will be called the *Main Interruption Rule*. We will then proceed to consider evidence for and against it. In the course of these analyses evidence for some of the above levels and types of monitoring will be presented.

Main Interruption Rule.

Stop the flow of speech immediately upon detecting the occasion of repair.

This rule has been explicitly suggested and discussed by Nootboom (1980) in his analysis of the repairs in the Meringer (1908) corpus. The rule says that if trouble of any of the sorts discussed above is detected, processing is simultaneously interrupted in all components of the production apparatus. The term 'immediately' is not intended to exclude a constant latency from detecting to interruption, it only means that this latency is quite short (in the order of 200 ms or less), and about equal for message construction, formulation, and articulation.

One corollary of the rule is that linguistic structure is ignored in the process of interruption: any moment in the flow of speech is a potential place for interruption. On first view there seems to be good evidence for this. Speakers frequently interrupt right after the reparandum, even if it doesn't complete a phrase, like in Example (5). They even interrupt the reparandum itself, as in (14):

- (14) We kunnen rechtdoor naar het ge..., naar het oranje kruispunt
We can straight on to the ye..., to the orange node

Here the color name (geel – *yellow*) is interrupted before the final consonant.

One might argue that the speaker always detects the trouble before or during overt production of the reparandum. He then either stops immediately, as in (5) or (14) and in covert repairs, or decides to complete the linguistic unit(s) he is working on—thus producing delayed, but linguistically motivated moments of interruption. The data, however, are full of counterexamples. Also in delayed interruptions we find a multitude of cases in which the point of interruption is not a phrase boundary (15), and not even a word boundary (16):

- (15) En boven de grijze bol een, of rechts van de grijze bol een
And over the grey sphere a , or right of the grey sphere a
 paarse bol.
purple sphere.
- (16) Rechtdoor naar kruisp ... oh nee sorry, rechtsaf naar kruispunt rood
Straight on to no- ... oh no sorry, right to node red

In (15) the reparandum is 'boven' (*over*), but the interruption is delayed till after 'een' (*a*), which is not a phrase boundary. In (16) the reparandum is 'rechtdoor' (*straight on*) and interruption is delayed to *within* the word 'kruispunt' (*node*). The latter example, and similar ones in our data, contradict the rule proposed by Nootboom (1980) in his analysis of the Meringer data: "although a speaker sometimes stops before the word, or sometimes even before the syllable against which the error is made is completed, he never stops in the middle of another word". In our terms: within-word interruptions should only occur within the reparandum itself. There can be no doubt that Nootboom's rule holds for the Meringer data as published. It is likely, however, that Meringer, who wrote down repairs from memory, didn't notice delayed within-word interruptions, or didn't bother to be so precise about them.

Though examples such as (15) and (16) show that delayed interruptions may violate phrase and word boundaries, they do not suffice to counter a more statistical argument. It may still be the case that a speaker *prefers* to complete syllables, words and phrases before interruptions. There may be a more than random incidence of phrase, word, and syllable boundaries at moments of interruption. In order to test this, three analyses were done on the data. The first one concerns the distribution of phrase or constituent boundaries, the second one word boundaries, and the last one within-word phonological boundaries.

Constituent boundaries.

There are 235 repairs in the data (25%) in which interruption is delayed by one or more syllables after the reparandum (excluding, of course, covert repairs where this is mostly undecidable). Each of these repairs was coded as to whether the interruption occurred at a surface constituent boundary or not. Potentially problematic here were delayed repairs in which the interruption occurred *within* a word. These 30 cases were left out of the analysis. For the remaining 205 repairs it was almost always clear what should count as a completed constituent. There was a heavy preponderance of PP's and NP's in the data, and there can be little doubt about how to code structures of the following general form: (Prep)-(Det)-(Adj)-N (as in 'naar het gele punt', *to*

the yellow point; 'vanuit dat punt', *from that point*, etc.). In all these cases we coded only the point after N as completing a constituent. There were two cases of possible doubt. The first one concerns phrases like 'van geel' (*from yellow*), where we took the color name to be nominalized, yielding the structure Prep-N. The second one concerned cases of post-positioning of color names ('naar een punt groen'—*to a point green*). Though this is 'ungrammatical' in Dutch, it is quite frequent in some subjects' descriptions. These cases we interpreted as being of the form (PP)-(Det)-N-Adj, and we took the constituent to be completed only after the Adj, not after N. One uncertainty here was whether a subject interrupting after N had intended to produce an Adj at all. But both the repair made, and the general practice of the speaker hardly ever left any doubt about this.

How to evaluate statistically whether the incidence of constituent boundaries in the thus coded data is over chance level? The following procedure was applied. For each of the 205 repairs under concern we took at random another utterance from the same pattern description in which the subject made the repair. This 'comparison utterance' (i.e., produced by the same speaker in describing the same pattern, though another part of it) was now analyzed as follows: we counted the number of words in the repair's OU, i.e., from the beginning to the moment of interruption. There were 6 cases where it was impossible to do this, because the beginning of the OU could not be determined unambiguously; these were left out of consideration. For each of the remaining 199 delayed repairs we then counted an equal number of words from the beginning of their comparison utterance. This point was then coded in the comparison utterance as completing a surface constituent or not, applying exactly the same code that was used for the repairs. We felt that this would give us a fair estimate of the incidence of constituent boundaries for subjects and patterns described. Finally, a McNemar test was applied to the pairs of codes obtained in this way. It turns out that the samples do differ statistically ($p < 0.05$, one-tailed): the incidence of constituent boundaries is somewhat greater at interruption points in repairs than at corresponding places in the comparison utterances. The corresponding values are as follows: for delayed interruptions the points of interruption coincide with a constituent boundary in 66% of the cases, as compared to 58% for the corresponding places in the comparison utterances. There are two alternative explanations for this result: (i) The Main Interruption Rule is wrong: there is a tendency in speakers to finish the current phrase after detecting trouble. (ii) The Main Interruption Rule is correct, but detection of trouble tends to occur towards or at the end of phrase. How to distinguish between these two alternatives?

If alternative (i) holds one would predict a difference between immediate and delayed interruptions. Immediate interruptions, i.e., where a speaker

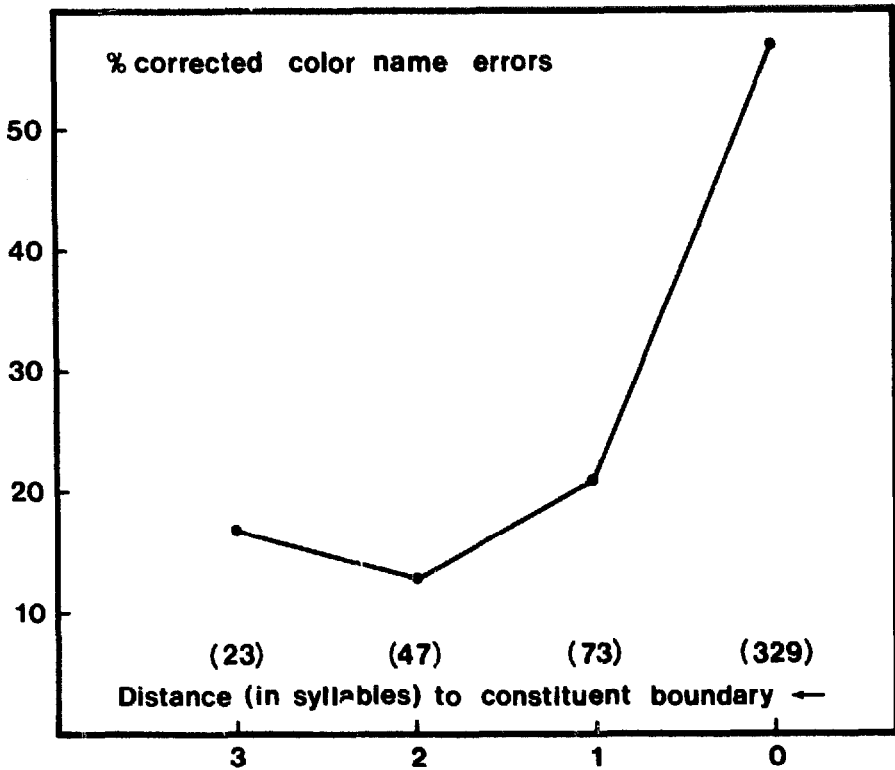
does not 'talk on', should show fewer phrase boundary constraints than delayed interruptions, where speakers *may* have talked on after detecting trouble in order to complete a phrase. In other words, one would predict a lower incidence of phrase boundaries at points of immediate interruption than at points of delayed interruption. If alternative (ii) holds, there should be no difference of the just mentioned type.

In order to test this we undertook to also code immediate repairs for phrase boundaries at the point of interruption. Repairs where within-word interruptions had been made were again dismissed, and exactly the same constituent boundary criteria were used as in the analysis of delayed interruptions. For 400 cases of immediate interruption (excluding within-word interruptions), we found 74% phrase boundaries at the point of interruption. This clearly, is not less than the 66% found for delayed repairs (neither is it significantly more: $p < 0.10$ by chi-square test). Immediate interruptions respect constituent boundaries at least as much as delayed interruptions (the 74% differs significantly from the 58% in the comparison utterances: $p < 0.001$, by chi-square). Alternative (i) can thus be disposed with: it appears that it is the *detection* of trouble that tends to interact with phrase structure. The speaker seems to have enhanced attention for 'trouble' towards the end of phrases.

Since this is an important claim about the monitoring process, we tried to obtain independent evidence for it. If it is correct to say that the speaker's attention for trouble increases towards the end of constituents, one would predict that error detection chance relates to position of the error within the constituent: constituent-final errors should have higher detectability than constituent-non-final errors.

In order to test this, we took the homogeneous set of 218 color repairs in our data (i.e., overt repairs where one color name became replaced by another one). These are, therefore, cases where the speaker detected the error. In order to estimate detection chance for different positions in the constituent, it was necessary to calculate the number of non-detected color errors. This was done by going through the complete set of 280⁹ pattern descriptions and by checking every color name for correctness *vis-à-vis* the pattern described. This analysis yielded 254 cases of non-corrected color naming errors. In other words, the average detection chance for color name errors can be estimated at $218/(218 + 254)$, i.e., 46% for our speakers. The issue here is how detection chance varies with position within the constituent. For each of the color errors (repaired and non-repaired ones) we determined the number of syllables between color name and constituent boundary (end of constituent), and for each of these error positions we computed the ratio of corrected errors to corrected plus non-corrected errors. Figure 3 shows the distribution

Figure 3. *Detection proportion for color name errors, for different positions of error within constituent. Numbers in parentheses depict the total number of color name errors for that position.*



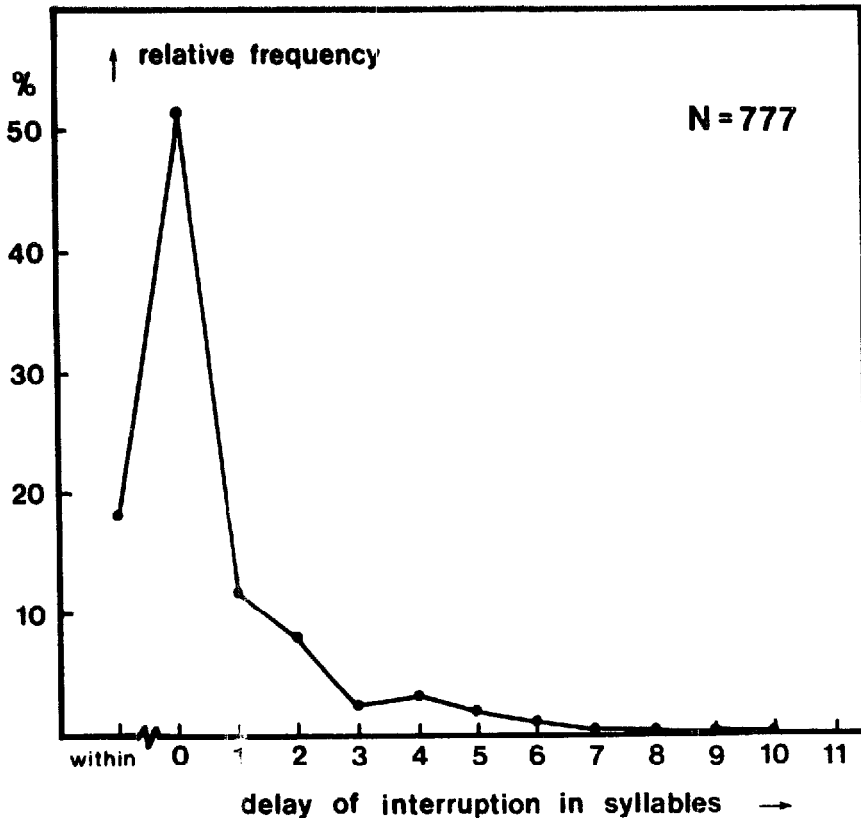
of these detection chance estimates. The results clearly confirm the claim that trouble detection increases towards the end of the constituent; the range is from about 15% in non-final position to 57% for phrase-final color terms.

This independent evidence makes it safe to conclude that the sensitivity of the monitoring process fluctuates with constituent structure: attention for self-produced speech is enhanced towards the end of constituents.

At this point it suffices to notice that, as far as constituent structure is concerned, the Main Interruption Rule can be maintained: the speaker stops immediately upon detecting trouble. But detection, in its turn, depends in part on the position of trouble in the constituent being produced.

Word boundaries

Among the immediate interruptions there were 142 within-word ones; this amounts to 26% of the sum total (N = 542). All other 74% are between-word interruptions. To evaluate this number one needs an estimate of average

Figure 4. *Distribution of interruption moments in syllables after reparandum.*

word length: If average word length is l phonemes, then a rectangular distribution of interruption points would predict an average occurrence of $100/l\%$ between-word interruptions. For $100/l\% = 74\%$ this would amount to $l = 1.35$ phonemes, a number which differs by an order of magnitude from reality (which should be 4 or 5 phonemes for this type of discourse). A first conclusion should therefore be that word boundaries are taken into account in a speaker's immediate interruption. Just as in the case of constituent boundaries we should now raise the question whether this indicates a deviation from the Main Interruption Rule in that speakers prefer to complete a word even after having discovered trouble, or that word ends are more sensitive moments for detecting trouble.

An argument for the word-completion hypothesis (i.e., against the Main Interruption Rule) could be made if the tendency to complete a word would turn out to vary irrespective of detection chance. Such might be the case if we compare trouble words and non-trouble words. In immediate interrup-

tions the speaker interrupts within or right after the trouble word; in delayed interruptions the flow of speech stops within or after a 'neutral' word. Since delayed interruptions are less frequent than immediate ones (see Figure 4 for the distribution of moments of interruption in our data), we must assume that detection chance slopes down accordingly. But there is no good reason to suppose that the *within-word* detection distributions are different for trouble words and neutral words. If words' ends are places of increased attention for trouble, that should hold for trouble and neutral words alike. One might argue that trouble words only *become* trouble words some time after they start, i.e., detection chance might be somewhat lower at the beginning of trouble words than at the beginning of neutral words, but that can only strengthen the argument to be made. The null-hypothesis is, therefore, that the proportion of within-word interruptions is the same for neutral words and trouble words. If, however, some significant difference is found, there is reason to assume that speakers prefer to complete the word in at least one of these cases, i.e., independent of the detection chance.

To test whether such is the case, we computed the proportion of within-word interruptions for immediate and delayed interruptions, respectively. Among 542 immediate interruptions there were 142 within-word cases, i.e., 26%; of the 235 delayed interruptions 30 were within-word, i.e., 13%. This difference is highly significant (Two tests were run. A χ^2 -square test gave $\chi^2 = 16.392$, $p < 0.001$. We also computed the two fractions of within word interruptions for each subject individually. A McNemar-test on these paired fractions yielded $\chi^2 = 20.891$). The conclusion should thus be that there are cases where speakers prefer to complete a word after detection of trouble, namely where the word is a neutral one. The consequence here is that the Main Interruption Rule needs a qualification in the direction of Nootboom's original conjecture: speakers sometimes tend to complete words after detection of trouble.

How can this exception to the Main Interruption Rule—speakers' tendency to complete words—be explained? One explanation would be that speakers have less control over the 'later', more peripheral stages of speech production: a word, once triggered, tends to run its own course in a more or less autonomous way. An argument against this explanation is that no such thing is apparent in the even more peripheral stages of phonological and/or articulatory planning: the speaker can easily interrupt a word at phonologically odd places, as will be shown shortly. It can, moreover, not account for the just observed fact that interruption within a word occurs twice as often in trouble words than in neutral ones: both should show the same peripheral 'inertia'.

An alternative explanation is a pragmatic one. One might conjecture that by interrupting a word, the speaker signals to the hearer that that word is

Table 1. *Within-word and after-word interruptions for appropriateness and error repairs*

| Type of repair | Interruption | | | | Average delay <i>d</i> in syllables | Total |
|----------------|--------------|------------|-------------|------------|-------------------------------------|------------|
| | Immediate | | Delayed | | | |
| | Within-word | After-word | Within-word | After-word | | |
| A-repair | 20 (7%) | 155 (53%) | 10 (3%) | 105 (36%) | 2.85 | 290 (100%) |
| E-repair | 91 (23%) | 193 (48%) | 19 (5%) | 96 (24%) | 3.79 | 399 (100%) |

wrong. So far, however, the distinction 'trouble' / 'neutral' was confounded with immediate *versus* delayed interruption. What has to be shown still is that also for immediate interruptions it is the case that words that are wrong are more often interrupted than other words. It should be remembered that not every reparandum is erroneous. In an appropriateness repair (A-repair, see Section 3.2) the reparandum is correct but needs some qualification which may or may not lead to its replacement. An error-repair (E-repair, cf., Section 3.3), however, involves an erroneous word (such as *blue* instead of *red*) which has to be undone as soon as possible. The communicative status of an inappropriate word or phrase thus differs markedly from one that is plainly erroneous. One would therefore predict that speakers will be more likely to interrupt an erroneous word than a not fully appropriate one. To test this, we checked for all immediate and delayed A- and E-repairs whether a within-word or an after-word interruption was made. By necessity, an immediate within-word interruption is an interruption of the reparandum itself. For the delayed interruptions also the delay *d* was determined, i.e., the number of syllables between reparandum and moment of interruption. Table 1 summarizes the findings. It is evident from this table that appropriateness repairs, as opposed to error repairs, are seldomly made by interrupting the reparandum within a word (7% and 23%, respectively). This is in support of the pragmatic hypothesis: it is all right to interrupt a word which needs total replacement because it is erroneous, but it is not good practice to interrupt a correct word which only needs further specification. If interruption is delayed, there are no erroneous words left to be interrupted even in the case of E-repairs. In that case within-word interruptions for E-repairs also drop to the low level of 5%. So, the more general rule seems to be that correct words should not be interrupted, and this holds equally well for correct trouble words (i.e., in A-repairs) as for neutral words (i.e., in delayed interruptions). Interrupting a word signals that that word is wrong.

Taken together, these results indicate one factor which qualifies the Main Interruption Rule: *only erroneous words may be interrupted upon detection of the occasion for repair.*

Still, one might argue that this cannot be the whole story: most erroneous words do *not* get interrupted. No less than 77% of them are completed. But at this point we prefer to maintain the rule, and assume that in most cases trouble is not detected until the end of the reparandum. In the 23% cases where it *is* detected earlier, immediate interruption follows if the word being spoken is erroneous itself. If it is not erroneous, but merely inappropriate or even correct (in case of delayed detection), the word tends to be completed. This state of affairs would account for most of the relevant data (for 93% to be precise). In other words: Nooteboom's all-or-none rule for words is generally correct, except for words that are erroneous themselves.

Within-word phonological boundaries

A final unit we should consider is the 'phonological word'. If a speaker makes a within-word interruption one wonders whether this can be at any place, or whether the speaker respects certain phonological boundaries. This is sometimes hard to test objectively. We listened again carefully to all within-word interruptions (N = 172), asking ourselves the question: is whatever the speaker pronounced up to interruption a phonologically possible Dutch word? The question was answered in the affirmative for 105 of the 172 cases. Example (17) is one of them:

- (17) Boven het groe ... nee ik zit fout . Links van het groene rondje
Over the gree ... no I am wrong. Left of the green disc

Here the speaker interrupts within *groene* (*green*), producing *groe*. Though the latter is not a word in Dutch it is a *possible* word. Judgments of this sort are often less than certain. Still, we came up with 67 cases of *impossible* words. They often involved interruption during the initial consonant or consonant cluster of a word, as in (18):

- (18) ... zit een [v] ... een horizontale lijn
... is a [v] ... a horizontal line

Here, the [v] is almost surely the initial consonant of 'verticale'. We counted 36 cases of initial consonant or initial consonant cluster interruptions. The other most frequent case was interrupting a lax vowel, as in (19):

- (19) ... naar die roze [bə] ... naar dat roze bolletje
... to that pink [bə] ... to that pink sphere

There were 22 such cases in the data. Of the remaining 9 cases no less than 5 involved interruptions of the last consonant cluster in 'oranje'. Taken together, a substantial number of word-interruptions (39%) violate phonological boundaries of Dutch. This result is hard to evaluate statistically, but surely demonstrates that phonological boundaries are not sacrosanct in self-initiated interruptions.

Returning now to the Main Interruption Rule, the only clear qualification that had to be made was that a speaker tends to complete non-erroneous words, i.e., neutral or merely inappropriate ones, after detection of trouble. The observed tendency to respect surface phrase boundaries should be seen as a property of the monitoring process itself; it is not due to delaying interruption after detection of trouble.

This section will be completed by considering one special category of lexical errors in our data in more detail.

Color word repairs

There is one especially frequent type of error in the data, the use of an erroneous color term. Among the overt repairs there are no less than 218 cases of color name repairs, i.e., repairs where one color name becomes replaced by another one. Among the covert repairs there were 69 that involved hesitation before the color word and/or repetition of the color word. An analysis of these repairs may shed some light on a long-standing issue in speech-production research: is 'lexical trouble' semantically or phonologically caused? Butterworth (1980), for instance, argues from analyses of speech accompanying gestures that prelexical hesitations in the fluent phase of speech are caused by problems of retrieval of phonological form, rather than meaning: they are typically accompanied by iconic gestures which are meaning-related. Garrett (1981, see Reference Note 2), in a very detailed analysis of normal and aphasic word finding problems suggests that prelexical hesitation is quite generally form-related: "the hesitation arises not out of a search for a lexical item which satisfies conceptual constraints, but rather out of processes which retrieve items from the form-based inventory". Still, Garrett does recognize that meaning-based prelexical hesitations may arise, especially where (existing) phrasal constraints allow for different lexical items. The latter is precisely the rule for color words: the maximal constraint is that there should be a color term in a particular slot, but there is no further syntactic restriction on which color term it should be.

What evidence is there for form-based versus meaning-based trouble in the color word repairs? One source of evidence to consider is the character of the substitutions made by the speakers. If trouble is meaning-based, i.e., has to do with lemma-selection (cf., Section 2, sub-section B₁ Lexicalization),

one would expect (i) a significant rate of exchange errors (see below), and (ii) substantial meaning relations between the erroneous color name and the target color name. If, however, trouble is form-related, one would expect to find (iii) a phonological relation between the erroneous color name and the target color name. Finally, a phonological explanation could be supported indirectly by comparing color name errors and prelexical (i.e., pre-color-name) hesitations. If the latter are largely form-based, as the literature suggests, then one should find (iv) some degree of similarity between the distribution of color name errors and pre-color-name hesitations. Let us consider these four sources of evidence in turn.

(i) Exchange errors. According to Garrett (1975) a strong argument in favor of defining a level of functional representation, under direct control of the message, is the existence of so-called exchange errors. Exchange errors arise at lemma selection. A word (or even sequence of words) intended for one phrase ends up in another phrase: the message level conceptual units thus become expressed in the wrong surface phrases (e.g., *Why was that horn blowing its train?*, where two NP's have exchanged materials). Are such cases to be found in our data? It would, in fact, suffice to find cases like 20:

- (20) Rechts van paars ligt eh van wit ligt paars
Right of purple is eh of white is purple

There are two NP's in the intended (and true) sentence (*Right of white is purple*), but the N of the second NP shifts to the position of the first NP. Still there is no full exchange: the error is corrected before this could happen. The question thus is, are there instances of speakers anticipating the next color name? We found 28 cases of overt color repairs in which the erroneous color word was the one to be mentioned next. Example (20) is one of those, and so is (21):

- (21) Ingang naar geel eh naar grijs. Doorgaan naar geel
Entrance to yellow eh to grey. Go on to yellow

Are these really anticipations, or are they a statistical artifact? There are 218 cases of color name replacement. Of these 31 cannot be taken into account for the present analysis for various reasons (e.g., within-word interruption makes the original color name ambiguous, such as 'b...' for either *blue* or *brown*, or the repair is an appropriateness one as in 'blanco' → 'wit'). For the remaining 187 cases one can state the following. Since there are 11 color names used in these spatial descriptions, there are in each case 10 alternative erroneous color names. Under the null-hypothesis there is a chance of 0.1 that the erroneous name happens to be the color to be mentioned next. The expected number of cases is thus 18.7. This is significantly ($\chi = 5.139$ $p <$

0.05) different from the 28 cases we found. So we are inclined to believe that some of the color word errors are due to anticipating a subsequent color. For these cases the production process failed to derive from the message the correct functional level representation. It should be noted that these anticipations do not necessarily respect sentence boundaries. Example (20) does, but (21) doesn't. Seventeen of the 28 cases were like (21).

Though this finding supports the view that color naming errors arise at the level of lemma selection, the number of cases explained by such anticipations is fairly small. It is, therefore, important to consider the next source of evidence.

(ii) Meaning relations between error and target. If we list target \rightarrow error pairs in the order of frequency of occurrence (and cutting off at $n = 4$) we obtain the following: PINK \rightarrow *orange* ($n = 19$), PINK \rightarrow *purple* (8), ORANGE \rightarrow *pink* (7), ORANGE \rightarrow *red* (7), YELLOW \rightarrow *green* (7), PURPLE \rightarrow *blue* (7), ORANGE \rightarrow *yellow* (6), RED \rightarrow *blue* (6), BLUE \rightarrow *green* (6), GREEN \rightarrow *red* (6), RED \rightarrow *brown* (5). With the exception of RED \rightarrow *blue* and GREEN \rightarrow *red*, all of the above pairs are instances of color similarity. Depending on the leniency of one's definition of color similarity, between one-third and more than one-half of the errors are of this sort. This clearly shows that trouble in color word repairs is to a substantial degree meaning-based. Still, it could be form-based as well, which is the next point to consider.

(iii) Phonological relations between error and target. Adopting the definition that two color names are phonologically related if they have the same initial consonant, the following names in the sample are related: 'groen', 'geel' and 'grijs'; 'blauw' and 'bruin'; 'rood' and 'rose', i.e., three clusters of names. For each of these colors as target we counted the number of errors within the cluster. There were 25 such form-related errors in total. This number was compared with that derived from the null-hypothesis, namely, that all errors (whether within the cluster or not) are equally likely to occur; their probability, then, is 0.1, since there are 10 possible erroneous names for a target color. The difference turned out to be non-significant ($\chi^2 = 3.152$, 1df, $p < 0.10$). Hence, there is little evidence for form-based color naming errors in the sample. A slight tendency for such errors to occur can be explained by reference to the work of Dell and Reich (1981) who showed that lemma and word form retrieval are not fully independent. The system 'leaks', in that lemma selection can be affected by phonological similarity. It seems therefore that trouble in lemma selection is the major source of color name errors.

(iv) Errors and hesitations. Earlier we cited Garrett's conjecture that pre-lexical hesitations are form-based. If, as seems to be the case, color naming

Table 2. Overt and covert repairs for different target colors

| Target color | groen <i>green</i> | blauw <i>blue</i> | rood <i>red</i> | geel <i>yellow</i> | rose <i>pink</i> | grijs <i>grey</i> | paars <i>purple</i> | zwart <i>black</i> | bruin <i>brown</i> | oranje <i>orange</i> | wit <i>white</i> | Total |
|----------------------------------------|-----------------------|----------------------|--------------------|-----------------------|---------------------|----------------------|------------------------|-----------------------|-----------------------|-------------------------|---------------------|------------|
| Percent nodes of color in the patterns | 14% | 13% | 12% | 12% | 8% | 8% | 8% | 8% | 7% | 6% | 3% | 100% |
| Overt repairs | 21 (11%) | 16 (9%) | 19 (10%) | 12 (7%) | 38 (20%) | 15 (8%) | 16 (9%) | 9 (5%) | 11 (6%) | 23 (12%) | 7 (4%) | 187 (100%) |
| Covert repairs | 1 (1%) | 8 (12%) | 7 (10%) | 10 (14%) | 7 (10%) | 10 (14%) | 8 (12%) | 2 (3%) | 5 (7%) | 9 (13%) | 2 (3%) | 69 (100%) |

errors are largely meaning-based, one would expect covert hesitations and overt repairs to exhibit different patterns of distribution. In order to test this supposition, Table 2 was composed. The 11 target colors did not appear with equal frequency in the visual patterns. The top row of Table 2 depicts their proportion of occurrence in the patterns to be described by the subjects. The second row gives the number of naming errors for each of the target colors. Comparison of the two rows by χ^2 yields a highly significant difference ($\chi^2 = 60.466$, 10 df, $p < 0.001$), which leaves no doubt that some colors create more naming problems than others. The main contributor to χ^2 is *pink* (35.27) and the next one is *orange* (12.43). Though this finding is most likely to be accounted for in terms of strong perceptual similarities which obtain between these colors and others in the set, there is another explanation which should not be ruled out before it has been given due considerations: 'Rose' (*pink*) and 'oranje' (*orange*) are among the least frequent color names in Dutch. Is it the case that naming error rate is related to color name frequency? In order to test this conjecture, the *relative* error rate for each color name was computed by dividing the row-two value in Table 2 by the corresponding row-one value. This relative error rate was then correlated with word-frequency in Dutch (using values taken from Uit den Boogaart, 1975, Table B, sub T schr.). The Spearman rank correlation was found to have a non-significant value of $r = -0.38$.

Now consider the distribution of the covert repairs, i.e., the 69 pre-lexical hesitations before color words, the distribution for which is presented in the third row of Table 2. There is, again, a significant difference with the row-one distribution ($\chi^2 = 21.888$, 10 df, $p < 0.02$), with *orange* (2.20) and *gray* (1.82) as the main contributors. The difference between row two and row three, i.e., between overt and covert repairs, falls short of significance ($\chi^2 = 16.465$, df = 10, $p < 0.10$). But there is still reason to suppose that the distribution of pre-lexical hesitations has a different basis from that of overt naming errors, in view of the fact that the rank-correlation between the relative hesitation rates and word-frequencies turns out to be $r = -0.74$, which is significant at the $p < 0.01$ level.

These results lead to the following general conclusions with respect to color naming trouble in the present corpus of repairs. Errors in color naming are largely due to speakers' failure to select the correct lemma. This conclusion follows from the existence of exchange errors (anticipations), and from the perceptual relatedness between error and target color. Word-form retrieval is not an important factor in the generation of errors. There is no noticeable phonological similarity between target and error name, and no noticeable relation to word frequency. Pre-lexical hesitations, however, do

show a strong word frequency effect. This finding supports the notion that they occur largely on account of trouble in word-form retrieval, as has been proposed elsewhere in the literature.

5. The use of editing terms

In this section we consider what happens right after the speaker interrupts the flow of speech, but before the repair proper (overt or covert) is initiated. The most characteristic phenomena at this point are the pauses and the use of 'editing expressions' (Hockett 1967) or editing terms (ET's). Not much is known about the use and functions of these terms. James (1972, 1973) gives analyses of some interjections which typically occur in covert repairs, especially 'uh', 'oh' and 'ah', and shows how these differ semantically. However, only her interpretation of 'uh' is relevant in the present context, since the other English editing terms do not correspond in simple ways to Dutch ones. According to James, 'uh' expresses that something was temporarily forgotten, but is now in the process of being retrieved, as in the covert repair (22):

(22) *I saw .. uh.. twelve people at the party*

This interpretation of the use of 'uh' is not incompatible with the suggestion initially made by Maclay and Osgood (1959) regarding the general function of 'fillers', namely that they serve to prevent interruption by the interlocutor, or to keep the floor.

Du Bois (1974) analyzes several interjections which occur in overt repairs, such as *that is*, *rather*, and *I r..ean*. He suggests that *that is* occurs to specify a referent, especially a pronoun:

(23) *He hit Mary .. that is .. Bill did*

This is the case which was described above as repairing for ambiguity reduction (AA-repairs). *Rather*, according to Du Bois, is 'nuance editing', getting closer to the intended meaning:

(24) *I am trying to lease, or rather, sublease my apartment*

This is exactly the case discussed earlier of looking for a more appropriate level term (AL-repairs). *I mean*, Du Bois suggests, indicates that an all out mistake is being corrected. This is, in our terminology, an E-repair:

(25) *I beg to present to you my half-warmed fish, I mean, my half-formed wish ...*

Editing expressions apparently differ in the semantic and/or pragmatic function they perform. In the following, we will try to relate our speakers' use of editing terms to the *raison d'être* of their repairs.

The editing terms used by our subjects are, in decreasing order of frequency: 'uh', 'of' (literally *or*), 'dus' (literally *thus* or *therefore*), 'nee' (*no*), 'sorry' (N.B. this is Dutch), and a large variety of infrequent expressions ('nou ja', 'wat zeg ik', 'ik bedoel') which will be called 'other'. Also classified as 'other' are cases where 2 or more of the above terms are combined, like in 'uh nee', or where these terms combine with less frequent ones, like in 'uh nou ja'. There are 74 such cases in the data. In almost half of the repairs (42%) no ET is used.

Table 3 gives the distribution of these ET's over the different occasions for repairs: Appropriateness, Error, Different, and Covert repairs, as well as the Rest-category of unclassifiable repairs. Also given are the subcategories distinguished earlier for A- and E-repairs. The rightmost column of the table shows the distribution of all 959 repairs over the different types of repair. It shows that 290 repairs in the total set are appropriateness repairs, 399 error repairs, etc.

The last but one column from the right shows the number of cases in which no editing term is used. As noticed, this is so for 42% of all repairs. For A-repairs this percentage is much higher: 72%. A-repairs apparently elicit relatively few ET's, whereas E-repairs and C-repairs are comparatively high on the use of editing terms. E-repairs elicit more than twice as many editing terms than A-repairs (62% versus 28%). This testifies again to the special status of A-repairs. Appropriateness repairs are not made for *correction*, but for *further specification*. This is also expressed in the type of ET. The most frequent term used for A-repairs is 'dus'. This connective normally presupposes the correctness of the previous propositions and introduces some consequence or state of affairs which is compatible with it. 'Dus' is, on the other hand, absolutely never used for error repairs. There are three highly frequent ET's for error repairs: 'nee', 'sorry' and 'of'. 'Nee' (*no*), normally implies denial of what was previously said. It is used only 3 times among the appropriateness repairs; all other 17 uses of 'nee' in our data are for repairing lexical errors. 'Sorry', as used in Dutch, involves a slight form of excuse. Again, we observed only 2 such cases among the appropriateness repairs. Of the 12 remaining cases 9 were used after an error. The speaker seems to apologize for having been out of control, for having said *orange* for *purple*, or *horizontal* for *vertical*, etc; the word delivered unexpectedly didn't express the intended concept. For 'of' (*or*), finally, we find a similar imbalance. Of the 47 uses of 'of' in our data only 9 appear in A-repairs, but 32 in repairs of lexical error. It is as if this ET is used to indicate disjunction. Though the normal use of

Table 3. *Distribution of editing terms over types of repair*

| Type of repair | Editing terms | | | | | | | Total number |
|-------------------------|---------------|---------|----------|---------|---------|-----------|-----------|--------------|
| | 'Uit' | 'Of' | 'Dus' | 'Nee' | 'Sorry' | Other | None | |
| Appropriateness (total) | 24 (8%) | 9 (3%) | 27 (9%) | 3 (1%) | 2 (1%) | 16 (6%) | 209 (72%) | 290 (100%) |
| AA-repairs | 6 (13%) | 1 (2%) | 3 (7%) | 0 (0%) | 1 (2%) | 2 (4%) | 33 (72%) | 46 (100%) |
| AL-repairs | 8 (6%) | 2 (2%) | 18 (14%) | 1 (1%) | 0 (0%) | 9 (7%) | 91 (71%) | 129 (100%) |
| AC-repairs | 4 (9%) | 2 (4%) | 4 (9%) | 1 (2%) | 1 (2%) | 3 (6%) | 32 (68%) | 47 (100%) |
| ALC-repairs | 6 (9%) | 4 (6%) | 2 (3%) | 1 (1%) | 0 (0%) | 2 (3%) | 53 (78%) | 68 (100%) |
| Error (total) | 93 (23%) | 32 (8%) | 0 (0%) | 17 (4%) | 9 (2%) | 95 (24%) | 153 (38%) | 339 (100%) |
| EL-repairs | 87 (24%) | 32 (9%) | 0 (0%) | 17 (5%) | 8 (2%) | 95 (26%) | 130 (35%) | 369 (100%) |
| ES-repairs | 3 (14%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 19 (86%) | 22 (100%) |
| EF-repairs | 3 (38%) | 0 (0%) | 0 (0%) | 0 (0%) | 1 (13%) | 0 (0%) | 4 (50%) | 8 (100%) |
| D-repairs | 0 (0%) | 2 (20%) | 0 (0%) | 0 (0%) | 0 (0%) | 3 (30%) | 5 (50%) | 10 (100%) |
| Covert repairs | 170 (72%) | 2 (1%) | 0 (0%) | 0 (0%) | 3 (1%) | 27 (11%) | 34 (14%) | 236 (100%) |
| Rest category | 0 (0%) | 2 (8%) | 3 (13%) | 0 (0%) | 0 (0%) | 17 (71%) | 2 (8%) | 24 (100%) |
| Total number | 287 (30%) | 47 (5%) | 30 (3%) | 20 (2%) | 14 (1%) | 158 (16%) | 403 (42%) | 959 (100%) |

'of' is inclusive (like it is for *or*), there is a special morphological variant of 'of' in Dutch: 'offe' whose uses are far more limited and exclusive. The sentence 'Jan offe Piet komt' simply cannot mean that both John and Peter come, it can only be read as a correction. This variant of 'of' is frequent in our data, though it is often hard to distinguish the two forms by ear: it all turns around a final schwa.

Let us now turn to the most frequent ET, 'uh'. It is used in 30% of the repairs, and is almost as frequent as all other ET's together. Table 3 shows that it is particularly used in covert repairs (in 72% of them). No more than 8% of the A-repairs, but 23% of the E-repairs contain 'uh'. James' interpretation of *uh* in English, we saw, was that the speaker marks something that is temporarily forgotten and is now in the process of being retrieved. If the same would hold for Dutch, it would correctly take account of the very frequent use of 'uh' in the covert repairs, but it is less clear how it would distinguish between the uses of 'uh' in appropriateness and error repairs. In both cases something may have been temporarily forgotten, resulting in error for the E-cases, and in need for further specification in the A-cases.

There is probably a far more marked characteristic in the use of 'uh': it is a symptom of the *actuality* or *recency of trouble*. In covert repairs the trouble is 'still on' at the moment of interruption. This is almost never the case for A-repairs. As was shown in the preceding section, appropriateness repairs almost never interrupt *within* a word, i.e., for immediate interruptions the reparandum is less recent at the moment of interruption than, on the average, in C- and E-repairs; the latter also often interrupt within the reparandum (cf., Table 1).

In order to test this recency-hypothesis for the use of 'uh' it is necessary to deconfound recency and repair occasion (E, A, C, etc.). The largest homogeneous class of repairs are the EL-repairs (N = 369), the repairs of lexical error. These we categorized in terms of the span between trouble spot and interruption. This was done by first distinguishing the cases of immediate and delayed interruption, and within the immediate cases those that interrupt within the trouble word and those that do so after. For the delayed interruptions we furthermore computed the average span of delay, *d*, the number of syllables between lexical error and moment of interruption. These computations were done for lexical repairs with 'uh', with other editing terms, and with no editing terms. The results are given in Table 4. They give clear support to the recency hypothesis. If one computes the relative frequencies of using 'uh' in within-error, after-error and delayed interruptions, it turns out that these slope down from 33% *via* 24% to 15% of the repairs. This decrease is significant ($\chi^2 = 8.191$, 2 df, $p < 0.02$). The effect is even more marked if one compares the share of 'uh' in repairs that contain an editing

Table 4. *The use of 'uh' versus other editing terms for different delays of interruptions in repairs of lexical error*

| Editing term used | Interruption | | | Average delay <i>d</i> | Total |
|-------------------|--------------|------------|---------------------|------------------------|------------|
| | Immediate | | Within + after-word | | |
| | Within-word | After-word | | | |
| 'Uh' | 28 (32%) | 42 (48%) | 17 (20%) | 1.71 | 87 (100%) |
| Other | 12 (8%) | 77 (51%) | 63 (41%) | 4.33 | 152 (100%) |
| None | 45 (35%) | 55 (42%) | 30 (23%) | 3.77 | 130 (100%) |
| Total | 85 (23%) | 174 (47%) | 110 (30%) | 3.77 | 369 (100%) |

term. For within-error interruptions no less than 70% of the ET's used are 'uh'. For immediate after-error interruption this halves to 35% ($\chi^2 = 13.258$, 1 df, $p < 0.001$), and for delayed interruptions the share of 'uh' is only 21% (the step from 35 to 21% yields $\chi^2 = 3.875$, 1 df, $p < 0.05$). In short, both the absolute and relative use of 'uh' decreases with delay of interruption. Finally, also the average delay *d* of interruption is shorter when 'uh' is used than in the case of other editing terms (1.71 and 4.33 syllables respectively).

In the previous section it was argued that delay of interruption is almost entirely due to delay of *detection*. The use of 'uh', therefore, signals that at the moment of trouble detection the source of trouble is still actual or quite recent in the flow of speech.

'Uh' has a special status among editing terms. It is the most frequent one, it appears preferably when the trouble is still 'on' at the moment of interruption. 'Uh', moreover, is probably also the only interjection, if not the only lexical item, which is universal across languages. There are, surely, phonetic variations in the sound of 'uh' between languages, but these may be largely due to the neutral position of the oral cavity for different languages. The conclusion thus seems to be that 'uh' is not a conventional lexical item like the other interjections, but a neutral sound produced when speech is interrupted at or close to trouble. This, of course, does not exclude the possibility that 'uh' acquires some form of derived lexical status related to this basic phenomenon; its meaning will then be close to what James suggested: "I have temporarily forgotten X". We may even have to allow for the possibility that this lexical status leads to phonological change, and generalizes to most uses of 'uh'. This may have happened in Swedish where 'uh' is not realized as schwa, but as [ɛ]. (This was brought to my attention by Jens Allwood.)

This should finish the analysis of the use of editing terms in our data.

But it does not at all exhaust the analysis of their significance in the making of repairs. Especially where appropriateness repairs are concerned editing terms range from clear correction signals such as 'of' (*or*), to terms that lead in parentheticals (*a-propos*), or expansions (such as *namely* or *to be sure*). Especially the latter may have been planned by the speaker to start with, and are not repairs in the sense that we are discussing here. The borderlines between these cases are fuzzy, however, and we will have to return to some of these issues in the next section where it is discussed how the repair proper is made by the speaker.

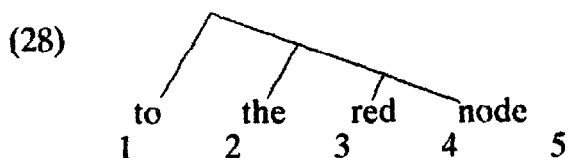
6. Making the repair

After having interrupted the flow of speech, the speaker will at some moment restart in order to make the factual repair. In this section we will analyze where speakers restart and why, and how they shape the repair. It has repeatedly been observed (Du Bois, 1974; Nootboom, 1980) that speakers almost always restart at phrase or constituent boundaries, with the exception that they may interrupt within or right after the reparandum and repair it immediately. Examples for these two cases from our data are (26) and (27):

- (26) Nog een keer naar rechts, naar links ...
Still one time to right , to left
- (27) Vanuit het groene ga je naar links, rechts
From the green go you to left, right

The speaker of (26) retraces to the prepositional phrase boundary, but this is not so in the otherwise very similar repair of (27). There the directional adverb is replaced immediately.

Though these and similar examples from our data almost always confirm the just mentioned observations in the literature, there are two reasons for analyzing these observations more deeply. The first one is that if one would restart at a random place of the original utterance (OU), the chances are very high that that will be a phrase boundary. The apparent contrast between speakers' interrupting at non-constituent boundaries, and restarting at constituent boundaries (and this contrast is repeatedly stressed in the literature) may for a large part be due to the syntactic structure of right-branching languages. In a right-branching language endings but not beginnings of embedded phrases tend to coincide. The phrase in (28) exemplifies this:



A new phrase starts at each of the transitions 1, 2, 3, and dependent on one's theory at 4. But there is only one position where all these phrases end, namely at 5. If one would interrupt this string at a random place, the best chances are that the interruption will violate a constituent boundary; if one would restart at a random position it is sure beforehand that it will be a constituent boundary. A more careful statistical analysis, as done for interrupting in Section 4, is thus necessary to show that the apparently obvious is in fact true.

The second concern is the distinction made between cases such as (26) and (27). It should not only be asked whether immediate repairs as in (27) do (statistically) more often violate the phrase boundary rule than cases like (26), but even if they are in fact different in that respect, one should still strive for the formulation of a more general rule from which both cases can be derived.

In the following we will first analyze the constituent boundary issue for both types of cases. Next we will proceed to formulating a more general well-formedness rule governing the speaker's retracing. Finally, several factors will be analyzed which may affect the speaker's retracing or restarting within the boundaries set by the well-formedness rule.

6.1. *Constituent boundaries*

As a starting point for the analysis we used the same repairs involving delayed (but not within-word) interruptions as had been analyzed in Section 4. When the speaker talks on after the reparandum, he cannot start the repair with the last word spoken. The argument that only the last word spoken can be repaired with neglect of the constituent boundary constraint thus does not apply to this set of delayed interruptions; it should therefore give the clearest test of the existence of such a constituent boundary constraint.

Scoring for phrase-initial positions in a sentence is somewhat harder than scoring for phrase-endings, just because the latter but not the former are usually multiply determined. We scored as phrase or constituent boundaries the beginning of a sentence, the point before the tensed element, and the point before any preposition. These are clear cases. Within prepositional phrases we scored as 'minus constituent' the boundary between preposition and adverbial in phrases such as 'naar/links' (*to/left*), but as 'plus constituent' the

transition from preposition to NP, as in 'naar/het groene punt' (*to/the green point*) and in 'naar/groen' (*to/green*) where the color name was used as a substantive. More problematic were the within-NP transitions. The transition from the article to what followed was scored 'plus', as in 'het/groene punt' (*the/green point*), but adjective-noun transitions were scored as 'minus'. The main reason was the frequent occurrence of adjective-noun reversals: some subjects quite often used constructions like 'een punt groen' (*a point green*) over and above the more standard 'een groen punt' (*a green point*). In our intuition the transition from *point* to *green* in the less standard form (*a point/green*) does not mark the beginning of a new constituent. Intuitions for the transition in (*a green/point*) are less pronounced, but given the equivalence of these constructions, both being NP's consisting of adjective and noun, we decided to code the latter case 'minus constituent' as well. It should be noticed that these are rather lenient criteria for constituency, but scoring for 'major constituents' only would have made the criteria even more arbitrary.

For each of the 205 delayed repairs we selected at random an unrepaired comparison utterance from the same pattern description of the same subject (and this comparison utterance was different from the ones analyzed in Section 4). For each repair we determined the number of words in the factual repair R (see Figure 2). If this number was n , we then took the last n words of the comparison utterance. Call this the 'comparison string'. Finally, the above scoring was applied to both R and the comparison string, i.e., we determined for both whether or not they initiated a constituent. A sign test was applied to the 205 pairs. We found a significant ($p < 0.0005$, one-tailed, McNemar-test), but not very large difference between the two samples. All repairs but two started at a phrase boundary, but no less than 89% of the comparison strings did as well. The test was based on no more than 14 cases in which comparison strings violated constituent boundaries where the repair didn't. The result confirms the constituent boundary constraints for repairs but reduces it to the level of only marginally surprising constraints.

Let us turn now to repairs involving immediate interruptions, such as (26) and (27). Consider first the cases where the speaker does not instantly replace the reparandum but rather restarts at an earlier word or with a wholly new construction. There are 280 such cases. Of these, all but one start at a constituent boundary. There are 262 repairs where immediate interruption is followed by immediate repair, as in (27). Of these no less than 96% also show the constituent boundary constraint, so they don't seem to form a special category. There were, in fact, just 10 exceptions in these data. Eight of them had the Noun-Colorname construction, with immediate repair of the color name. One example is given in (29):

- (29) ... naar hoek paars uh rose
to corner purple uh pink

The remaining two cases repeat a noun, and an adjective, respectively, without repeating the preceding article.

It is a mildly interesting observation that similar cases do not occur among delayed repairs, nor in immediate repairs where the speaker restarts before the reparandum. We will return to this issue in the context of the following well-formedness discussion. At this point some conclusions are warranted: (i) Speakers restart at constituent boundaries in making repairs, (ii) They also do so in making immediate repairs, (iii) It is almost impossible to violate the constituent boundary constraint given the right-branchingness of the language. The constraint is thus on the margin of vacuity.

6.2. A well-formedness rule for repairs

There is an additional problem with the constituent boundary constraint: it is not biconditional. One can construct repairs which are intuitively ill-formed but which observe the constraint; compare the invented cases (30) – (32):

- (30) *to the right is a green node, uh a blue*
 (31) *Did the man leave, uh the man enter?*
 (32) *With his sister he talked frequently, uh his mother he talked frequently*

We will now propose a new biconditional well-formedness rule for repairs. The rule applies to all repairs, except those involving syntactically or phonologically ill-formed constructions (of which there are only 22 in our corpus). The latter exceptions, however, are trivial as will become clear shortly. The rule is stated most easily if the original utterance (OU) is symbolized α , and the factual repair (R) by γ . It ignores the editing expressions used.

Well-formedness rule

A repair $\langle \alpha \gamma \rangle$ is well-formed if and only if there is a string β such that the string $\langle \alpha \beta \text{ and } * \gamma \rangle$ is well-formed, where β is a completion of the constituent directly dominating the last element of α . (**and* to be deleted if γ 's first element is itself a sentence connective).

It may be helpful to give a few applications of the rule before discussing its merits. Repair (33) is intuitively well-formed:

(33) $\overbrace{\text{to the right is a green,}}^{\alpha}$ $\overbrace{\text{a blue node}}^{\gamma}$

The rule pairs it with (34):

(34) $\overbrace{\text{to the right is a green}}^{\alpha}$ $\overbrace{\text{node}}^{\beta}$ $\overbrace{\text{and a blue node}}^{\gamma}$

Here β (*node*) completes the smallest current NP-constituent, i.e., the smallest constituent including *green*. The coordination (34) is clearly well-formed, and so is the corresponding repair. Consider another well-formed case, (35):

(35) $\overbrace{\text{Did you go right,}}^{\alpha}$ $\overbrace{\text{go left?}}^{\gamma}$

The rule correctly predicts well-formedness for the coordination in (36):

(36) $\overbrace{\text{Did you go right}}^{\alpha}$ ϕ $\overbrace{\text{and go left?}}^{\gamma}$

Here β is ϕ since the last word of α completes the current constituent (VP). A small change will turn (35) to ill-formed (37):

(37) $\overbrace{\text{Did you go right,}}^{\alpha}$ $\overbrace{\text{you go left?}}^{\gamma}$

This corresponds to the ill-formedness of (38), as predicted by the rule.

(38) $\overbrace{\text{Did you go right}}^{\alpha}$ ϕ $\overbrace{\text{and you go left?}}^{\gamma}$

This case is similar to (31) above. The reader can readily observe that (30) and (32) are also correctly predicted by the rule. In both cases $\beta = \phi$, so that only *and* should be inserted between α and γ . This results in ill-formed coordinations.

It was discussed above that eight out of ten constituent boundary violations in immediate repairs were of the form (29). How does the well-formedness rule treat such cases? Compare (39):

- (39) $\overbrace{\text{ga naar hoek paars,}}^{\alpha}$ $\overbrace{\text{rose}}^{\gamma}$
go to corner purple, pink

The rule pairs it with the coordination (40):

- (40) $\overbrace{\text{ga naar hoek paars}}^{\alpha}$ ϕ $\overbrace{\text{en rose}}^{\gamma}$
go to corner purple and pink

which is intuitively as well-formed as (39). The rule treats 12 of the 13 noted exceptions to the constituent boundary constraint in the corpus correctly. It also predicts that restarts at post-posed adjectives will not occur when interruption is delayed. Example (41) would be such a case:

- (41) $\overbrace{\text{ga bij hoek paars linksaf naar,}}^{\alpha}$ $\overbrace{\text{paars rechtsaf naar rood}}^{\gamma}$
go at corner purple left to, purple right to red

Contrary to (40) this repair sounds definitely ill-formed, and this is predicted from the rule, which pairs it with (42):

- (42) $\overbrace{\text{ga bij hoek paars linksaf naar}}^{\alpha}$ $\overbrace{\text{geel}}^{\beta}$ $\overbrace{\text{en paars rechtsaf}}^{\gamma}$
go at corner purple left to yellow and purple right
 $\overbrace{\text{naar rood}}^{\gamma}$
to red

The latter is ill-formed for any choice of β . Indeed, cases such as (41) are not found in our data, whereas repairs like (39) do occur. It should further be noticed that there is nothing in the γ of (41) itself which makes it an impossible repair proper, the ill-formedness only arises in its conjunction with the specific α . If the α had been 'ga van groen', (*go from green*), $\langle \alpha \gamma \rangle$ would have been a well-formed repair. Different from the constituent boundary rule, the well-formedness rule also hinges on the original utterance.

The rule, moreover, gives a basis for distinguishing between real repairs on the one hand, and parentheticals and expansions involving real shifts on the other. Example (43) is a clear case of a well-formed expansion—it could

have been planned as such by the speaker, and delivered without interruption or hesitation:

- (43) $\overbrace{\text{He conquered Babylon,}}^{\alpha} \overbrace{\text{the great Alexander.}}^{\gamma}$

If one would erroneously treat this as a repair and apply the rule, the result would be ill-formed under the intended co-reference reading:

- (44) $\overbrace{\text{He conquered Babylon}}^{\alpha} \text{ and } \overbrace{\text{the great Alexander}}^{\gamma}$

A remark should be made, finally, about cases of syntactic or phonological error corrections. It is obvious that if α is ill-formed itself, any string $\alpha \beta$ and γ will be ill-formed, which could predict that no well-formed repair can be made in these cases. This is clearly false, and the rule should not be applied to these cases. For phonological corrections the rule still predicts correctly if one ignores the local ill-formedness in α .

The next issue should be whether there are repairs in the corpus which violate the rule. We checked all repairs in this respect, excluding only the rest-category (R-repairs), and the syntactic errors (ES-repairs), which left us with 913 cases. It should be kept in mind that well-formedness judgments are not glaringly trustworthy (cf., Levelt 1972). This is even more strongly the case for natural speech data with their heavy dependence on context. But repairs are particularly difficult to judge, because there is the additional complication that the speaker is in trouble, and often operating on the borderline of grammaticality anyhow.

With all these provisos made, the reader will be in the right state of mind to evaluate the following findings. We found 17 clearly problematic cases in the data. On inspection they could be categorized in four types; of these four types only the last two create problems for the well-formedness rule.

(i) Both the repair and the corresponding coordination are ill-formed. There are 4 such cases. An example is given in (45):

- (45) Links daarvan een, dat zwarte een rose rondje
Left thereof a , that black a pink disc

Here the demonstrative 'daarvan' (*thereof*) is further specified as 'dat zwarte' (*that black one*). The repair, however, sounds wrong, and so does the corresponding coordination (as in (46)):

- (46) Links daarvan een punt en dat zwarte een rose rondje
Left thereof a node and that black a pink disc

It should not be surprising that speakers occasionally make ill-formed repairs, the important point is that ill-formedness is correctly predicted from the rule, and that seems to be so for these cases.

(ii) Both the repair and the corresponding coordination are doubtful. Four cases were put in this category. An example is (47):

- (47) En naar links een groen punt, nee of rechtdoor
And to left a green node, no or straight on

In this example the speaker corrects 'naar links' (*left*) and replaces it by 'rechtdoor' (*straight on*). It is not a beautiful repair, and neither is the corresponding coordination, (48):

- (48) En naar links een groen punt en rechtdoor
And to left a green node and straight on

Since both repair and coordination are in the same ball park as far as well-formedness is concerned, these cases form no threat for the rule.

(iii) The repair is well-formed, but the corresponding coordination is doubtful. There are three such cases. An example is (49):

- (49) Dezelfde lijn, horizontale lijn aan de andere zijde
The same line, horizontal line on the other side

As in (45) a demonstrative (*the same*) gets further specified (*horizontal*). This sounds all right here, but the corresponding coordination is not excellent:

- (50) Dezelfde lijn en horizontale lijn aan de andere zijde
The same line and horizontal line on the other side

(iv) The repair is only doubtful, whereas the corresponding coordination is ill-formed. There are six cases of this sort in the data. A typical example is (51):

- (51) Daar boven de oranje een gele. Boven de, je begint met een
*There over the orange a yellow. Over the, one starts with an
 oranje
 orange*

Here the speaker starts a new sentence ('Boven de'—*Over the*), then interrupts, and repairs the *previous* sentence, i.e., the *yellow* there should have been *orange*. One can, maybe, do this in a repair: stop in the middle of a sentence, and start anew with a totally different sentence. This will usually not result in well-formedness of the corresponding coordination, *vide* (52):

- (52) Daar boven de oranje een gele. Boven de groene en je
There over the orange a yellow. Over the green and one
 begint met een oranje
starts with an orange

Another case in the present category is (53):

- (53) Rechts daarvan een, van dat rooie een geel en een groen rondje
Right thereof a , of that red a yellow and a green disc

This is, again, a further specification of a demonstrative. Though the repair is not beautiful itself, the corresponding coordination is worse:

- (54) Rechts daarvan een knooppunt en van dat rooie een geel en
Right thereof a node and of that red a yellow and
 een groen rondje
a green disc

The 9 cases contained in the last two categories may, eventually, bring us to one or two slight qualifications of the well-formedness rule. A first candidate would be certain ways of repairing demonstratives, such as (49) and maybe (53). A second qualification could be the general addition that one can always start a new sentence, whatever the place of interruption.

Meanwhile, however, one can safely conclude that the well-formedness rule is in very general accordance with the well-formedness intuitions for our data set. As was shown earlier, it further gives the right prediction for the small set of cases in the corpus which violated the constituent boundary constraint. In particular, it predicts that such cases (restarts at postponed adjectives) should not appear otherwise than in immediate repairs, which is also in accordance with the facts. Finally, the rule may help to distinguish between real repairs and expansions or parentheticals.

Two other aspects of the rule are still undiscussed. The first one is that the rule is only an indirect one: it predicts well-formedness of repairs from well-formedness of coordinations. In the final section of this paper we will argue that this is not a vice but a virtue of the rule: it links the explanatory principles of two quite divergent domains of phenomena in speech production, and it

will be argued that further expansions to still other domains are to be considered. The second aspect of the rule is its non-determinacy. It predicts well-formedness for various different repairs of a given OU, especially involving more or less backtracking, as in (55):

- (55) *The ball rolls to the left side, uh* { *corner*
 left corner
 the left corner
 to the left corner
 rolls to the left corner
 The ball rolls to the left corner

but it does not tell us how the speaker chooses from among such alternatives. In the following section various possible determinants of backtracking will be considered.

6.3. *Some determinants of restarting*

Among potential determinants for the way in which a speaker restarts to make the repair are the repair occasion (especially whether there was an error or an inappropriateness) and the delay between trouble spot and moment of interruption. We will further consider some potential restrictions on restarting that would be particularly advantageous to the listener: it should be maximally clear to the listener how the new information (the repair) should be related to the old information (the original utterance). It is an empirical issue whether, and to what degree the speaker behaves according to such restrictions.

The occasion for repair

It appears from Table 3 that the major repair classes are Appropriateness repairs (30%), Error repairs (42%) and Covert repairs (25%). Since there is no sure way to determine the source of trouble of the latter, covert repairs, we will limit the analysis to A- and E-repairs.

The way in which a repair is made is very different for A- and E-cases. This appears most clearly when one categorizes the ways of restarting as follows. A first type is where there is a single trouble word, and the speaker retraces to just that word and replaces it by a new item. Examples are (5), (7), (9), (11), (15), (16), (17), (49), and (53) above. Another example (from the corpus) is (56):

- (56) Weer naar links naar hetzelfde blanco kruispunt, wit kruispunt
Again to left to the same blank node, white node

Here the (correct) term 'blanco' is replaced by the more appropriate term 'wit' (*white*). Such cases amount to 42% of all Error and Appropriateness repairs. We will call these *instant replacements*. Notice that these do not require immediate interruption. Cases like (56) show delayed interruption, but instant replacement.

A second type is where the speaker restarts by retracing to, and repeating some word prior to the reparandum. Examples from the corpus are (3), (4), (8), (14), (20), (21), and (57):

- (57) Rechts naar een geel, naar een blauw
Right to a yellow, to a blue

Speakers do this in 35% of the E- and A-repairs. Let us call these *anticipatory retracings*: the speaker retraces to an element which anticipates the reparandum. The third type is the category of *fresh starts*. The speaker neither instantly replaces a trouble word, nor retraces to an earlier word in the OU, but restarts with fresh material. There are essentially two variants here. The first one is exemplified by (58): the speaker restarts by making a new construction, which does not copy part of the OU.

- (58) Rechttuit komen we eerst op een bruinachtig kruispunt, nee,
Straight on come we first to a brownny crossing point, no,
 da's een knooppunt
that's a nodal point

The second variant is where the speaker does use parts of the OU, but led in by a fresh beginning. An example is given in (59):

- (59) De weg begint met een, of die loopt door en die begint met een
The road begins with a, or it goes on and it begins with a
 groene kruising
green crossing

In this example the OU-part 'begint met een' (*begins with a*) is picked up by the speaker, but it is led in by a new element (*it goes on and it*). This variant will be called *pre-specification*.

It is, however, not always possible to make sharp distinctions between these two variants. In (58), for instance, the elements 'een' (*a*) and 'punt'

Table 5. *Ways of restarting for Appropriateness and Error repairs*

| | Instant repairs (retrace to and replace trouble word) | Anticipatory retracing (retrace to and repeat earlier word) | Fresh start (restart with fresh material) | Total |
|-----------------------------|----------------------------------------------------------------|----------------------------------------------------------------------|-------------------------------------------------|------------|
| Appropriateness- repairs | 88 (30%) | 73 (25%) | 129 (44%) | 290 (100%) |
| Error-repairs | 204 (51%) | 165 (41%) | 30 (8%) | 399 (100%) |

(*point*) from the OU reappear in the repair, but it is hard to decide whether the speaker is really copying these elements.

Table 5 gives the distribution of A- and E-repairs over the three types of restarting. The table shows marked differences between A- and E-repairs. Error repairs concentrate in the categories of *instant repairs* and *anticipatory retracings*, only 8% is realized as *fresh start*. Appropriateness repairs, however, are *fresh starts* in 44% of the cases. A different way of putting this is that E-repairs are more *conservative* than A-repairs: In 92% of the cases E-repairs leave the OU unaffected but for the erroneous element; nothing is changed or added that is not strictly necessary. A-repairs, however, can substantially affect the OU. The conservative way of repairing is followed in just over half the cases. A-repairs are further specifications of what has already been expressed. These specifications are often made by adding fresh materials. The most dominant way of doing this is to make what we called *pre-specifications*: the fresh specification starts the repair, and is then followed by a citation of the part of OU which needs the specification. Example (59) was such a pre-specification: 'oegint met' (*begins with*) needs further specification, the speaker restarts by making the specification, followed by *begins with*. Another example is (60):

- (60) We beginnen rechts op het, wat rechts op het papier
We start right on the, somewhat right on the paper

Here 'rechts' (*right*) is qualified by the pre-specification 'wat' (*somewhat*). We counted the number of pre-specifications among the A-repairs in the *fresh start* category and found 96 of them. This is no less than 33% of all A-repairs, and 74% of those in the category of fresh starts. (There are only 4 such pre-specifications for E-repairs.)

Moment of interruption

Is the manner of restarting dependent on the delay between trouble spot

Table 6. *Ways of restarting for different moments of interruption (A and E repairs)*

| | Instant repairs | Anticipatory retracings | Fresh starts | Total |
|--------------------------------|-----------------|-------------------------|--------------|------------|
| Immediate interruptions | | | | |
| within-word | 51 (46%) | 44 (40%) | 16 (14%) | 111 (100%) |
| after-word | 144 (41%) | 117 (34%) | 87 (25%) | 348 (100%) |
| total | 195 (43%) | 161 (35%) | 103 (22%) | 459 (100%) |
| Delayed interruptions | | | | |
| within-word | 14 (48%) | 8 (28%) | 7 (24%) | 29 (100%) |
| after-word | 83 (41%) | 69 (34%) | 49 (24%) | 201 (100%) |
| total | 97 (42%) | 77 (33%) | 56 (24%) | 230 (100%) |

and interruption? As was discussed, this delay strongly affected the use of 'uh' as an editing term. If the actuality or recency of trouble has an effect right after interruption, it may as well affect the way of restarting. Table 6 presents the three ways of restarting distinguished above for different types of interrupting the flow of speech. The types of repairs covered in the table are all 689 A- and E-repairs in the corpus.

The table shows a surprising absence of major effects. The sum-rows for immediate and delayed interruptions show almost identical distributions over the different ways of restarting. It is, in particular, generally not the case that delayed interruption would lead to more anticipatory retracing and less instant repair than immediate interruption. There is only a slight tendency in the table for within-word immediate interruptions to lead to more 'conservative' modes of restarting (i.e., instant and anticipation) than for afterword ones ($\chi^2 = 4.827$, $p < 0.05$). There is no comparable tendency for delayed within-word interruption.

Though Table 6 leaves no doubt about the absence of a *general* tendency to do less instant repairing in case of delayed interruption, there may still be types of repair for which such a relation holds. The largest homogeneous subclass of repairs in the corpus are the color name corrections where one color name is replaced by another one: there are 218 instances in the data. Table 7 shows the ways of restarting for these cases. It is immediately apparent from the table that speakers make far more instant repairs after immediate interruptions than after delayed interruptions (A test of instant *versus* non-instant repairs yields $\chi^2 = 12.069$, $p < 0.001$).

It is not evident why this effect should appear for color name repairs. We checked whether there is an unusually large number of within-color word interruptions among the instant repairs. There are, however, 25 of them

Table 7. *Ways of restarting for color name repairs*

| | Instant repairs | Anticipatory retracings | Fresh starts | Total |
|-------------------------|-----------------|-------------------------|--------------|------------|
| Immediate interruptions | 81 (49%) | 78 (47%) | 6 (4%) | 165 (100%) |
| Delayed interruptions | 11 (21%) | 33 (62%) | 9 (17%) | 53 (100%) |
| Total | 92 (42%) | 111 (51%) | 15 (7%) | 218 (100%) |

among the 81 in the table, whereas there are 26 among the 78 anticipatory retracings. Also, there is no indication that immediate interruption after color names leads to extraordinarily large numbers of instant repairs. Table 6 shows 43% instant repairs for all A and E-repairs, Table 7 gives 49% for immediate color name repairs. The clearest tendency in Table 7 is for delayed interruptions *not* to yield instant repairs (only 21%).

Why would speakers evade instant repairing of color names after delayed interruptions? There is one obvious answer to this: the speaker tries to evade potential ambiguity. This can be exemplified by (61), which is a case of delayed instant repairing in the corpus:

- (61) Rijdend over een groen punt kom ik op een bruin punt, of
Running over a green point come I on a brown point, or
 rood, rood hè?
red, red uh?

It is quite normal in our data that an utterance contains two or more color names, as in (61). In the case of delayed interruption it is not self-evident any more which color name should be replaced (*green* or *brown* in the example). The speaker can disambiguate the situation by doing anticipatory retracing to, especially, the preposition, or by making a fresh start, which leaves no doubt about which color is meant. If this explanation is correct, it means that the effect observed is quite task-dependent, and is not expected to reappear for color name repairs in other (non-ambiguous) situations. Still this finding suggests to us that the speaker is aware of potentially ambiguous links between a repair and the original utterance. This will be further analyzed in the final part of this section.

So far the conclusion is warranted that there exists no direct relation between the manner of interrupting and the way of restarting. One should not conclude, however, that what followed the reparandum before interruption is ignored by the speaker. A repair can be elliptical in ways that are determined by what was said during the delay. In (61) the speaker says *red*, not

red point. This is a well-formed ellipsis (in Dutch), given the preceding *brown point*. So, the final shape of the repair may depend on the manner of interrupting, but the way the repair restarts doesn't.

Manners of restarting from the listener's point of view

After the speaker's interruption, the attentive listener faces the task of relating the repair to whatever was said prior to the interruption. An analysis of the listener's task may lead to further hypotheses about the constraints a speaker adheres to in the way he restarts his speech. The earlier stated well-formedness rule is, in our view, such a constraint: it helps the listener to proceed interpreting the utterance within the framework of the constituent structure at the moment of interruption.

Here some further constraints will be discussed. They should make it possible for the listener to decide on whether he should at all try to maintain, in part or in full, OU's interpretation, or whether he should rather interpret the repair as a fresh start, i.e., as a new utterance. Let us call this the 'continuation problem,' for the listener.

Given the on-line character of the listener's speech processing (see e.g., Marslen-Wilson and Tyler 1980), one would expect the hearer to make an early decision on the way in which the repair will have to be processed, given the OU. It would, in particular, be advantageous for the listener if the first word of the repair would already contain information on how to insert whatever is going to be said in the previous context. This is, in fact, feasible if the speaker would adhere to certain conventions on restarting. It will then become an empirical issue whether the speaker does follow these constraints.

There are two aspects of the first repair word (r_1) on which the hearer could capitalize in order to solve the continuation problem. Whether the hearer attends to these two aspects simultaneously (as would follow from Marslen-Wilson and Tyler's parallel-interactive theory) or sequentially, will be left undiscussed here, for lack of data on the listener's processing of repairs. The only point to be made is that these aspects may be powerful cues for the listener.

The first aspect is the *syntactic category* of r_1 . The listener might profitably apply the following rule: if the syntactic category of r_1 is equal to the syntactic category of the last word of OU before interruption (o_n), then o_n should be replaced by r_1 with maintenance of the constituency at o_n . The repair is interpretable as a continuation from the thus replaced o_n . By maintenance of constituency we mean maintenance of the existing syntactic commitments at the designated point (in this case o_n). The notion is essentially the same as Yngve's (1961, 1973)—see for a discussion Levelt (1974, vol. 3): if a speaker is producing a sentence's verb-phrase, and has uttered the article of the

object-NP, there are then three commitments: to complete S, to complete VP, and to complete NP. The precise definition of constituency type depends, of course, on one's linguistic theory. It is not necessary, however, for the present purposes to be more explicit at this point. The rule would cover a large number of cases: all repairs where the speaker interrupts immediately in or after the reparandum, and replaces it instantly. Examples in our corpus are (5), (7), (11), (27), (29), (49) and (62):

- (62) Van het groene rondje naar boven naar een roze ..., oranje rondje
From the green disc to up to a pink ..., orange disc

No less than 223 repairs in our corpus are of this sort (including 41 covert repairs involving repeats of the same word). In example (62) *pink* (o_n) and *orange* (r_1) have the same syntactic category (Adj), and the listener will replace *pink* by *orange* without changing the constituency type at *pink*: he or she will maintain the prepositional phrase under construction, replace the adjective in the interrupted NP and complete it with the subsequent noun (*disc*).

In order for this strategy to be effective, however, the speaker should adhere to the following constraint: *only restart at the same category if r_1 is to replace o_n* . Otherwise the hearer would be misled. We will shortly return to a test of this constraint.

The second aspect is the *lexical identity* of r_1 . The listener might follow this strategy: if r_1 is lexically identical to o_i (i.e., is the same word and of the same syntactic category), for any $1 \leq i \leq n$, then replace o_i by r_1 , and insert the repair from there on, maintaining the constituency at o_i but replacing whatever follows in OU. This strategy would lead to correct results for examples (1), (2), (3), (4), (6), (8), (13), (14), (20), (21), (26), (51), (52), (57), and (63):

- (63) Rechtsaf naar geel, eh naar wit
Right to yellow, uh to white

In (63), for example, r_1 (*to*) is the same word as o_2 (*to*) and the listener might instantly decide that whatever will follow is to replace the part of OU from *to*, maintaining the syntactic commitments at that point (i.e., 'this is going to be a prepositional phrase'). As our corpus contains 345 repairs of this sort, the strategy should thus be quite powerful. However, for this to work, the speaker should adhere to a constraint: *let r_1 only be identical to o_i for any i if the repair is to be inserted at o_i , replacing the rest of OU*. (If there is more than one word in OU identical to r_1 , we take the constraint to apply to the most recent one.) Let us now turn to a test of this constraint and the

Table 8. *Applicability and violation of word and category identity constraints in restarting*

| | Category identity constraint | | | | |
|---------------------------------|------------------------------|----------|----------------|-------|-----|
| | Applicable | | Not applicable | Total | |
| | Respected | Violated | | | |
| Word identity constraint | | | | | |
| Applicable | respected | 50 | 8 | 270 | 328 |
| | violated | 0 | 9 | 8 | 17 |
| Not applicable | | 149 | 7 | 458 | 614 |
| Total | | 199 | 24 | 736 | 959 |

previous one. For ease of reference the results of these tests are summarized in Table 8.

The first constraint (category identity) says that the syntactic category of r_1 should be different from that of o_n (the last word before interruption), except if r_1 is to replace it. We checked all our repairs on category identity between o_n and r_1 . There are 223 cases (i.e., 23%) of such category identity between o_n and r_1 . In these 223 cases are included 50 repairs where o_n and r_1 are not only category identical, but actually word identical (such as *to orange, uh... orange*). Of these 223 cases only 24 cases violate the convention, i.e., r_1 should *not* replace o_n . Examples are (64) and (65):

- (64) ... in het roze rondje naar, via het roze rondje naar beneden
 ... *in the pink disc to* , *via the pink disc to down*
- (65) Ga van links weer naar, of van roze weer naar blauw terug
Go from left again to , *or from pink again to blue back*

In both cases o_n and r_1 are prepositions. Example (64) is, indeed, likely to 'gardenpath' the listener: he might initially be inclined to replace *to* by *via*, creating '*in the pink disc via the ...*'. At this point, however, *pink disc* reappears leading to a rejection of this interpretation. It should be noticed that the prosody of *via the pink disc* exactly mimics that of *in the pink disc*, but is neither stressed on *in* nor on *via*. Example (65) is less likely to create problems. The hearer may not even start constructing '*go from left again from pink...*', not only because this is semantically unlikely, but because r_1 (*from*) is identical to a previous element of OU (o_2). The other strategy outlined above would thus provide the correct solution for this continuation

problem. Eight out of the 24 violations are of this sort, or in other words, there are only 16 (i.e., 7%) serious violations which can neither be treated by the one or the other listener strategy. Here we will leave undiscussed whether some priority relation holds between the two strategies. Of importance is the observation that speakers adhere quite strictly to the constraint on category identity.

The second constraint proposed (word identity) is that r_1 should only be identical to some word o_i of OU if the repair is to be inserted at o_i in OU. We tested all repairs for word identity of the sort specified in the constraint. There are no less than 345 such cases in the corpus (36%). Among these are the 50 cases mentioned above where r_1 equals o_n (like in *to orange, uh ... orange*). Of the 345 cases, 328 do entertain the replacement relation prescribed by the constraint. Of the 17 (5%) violations none respects the category identity constraint. One example of these few violations is given in (66):

- (66) Naar links naar het roze rondje, of naar rechts naar het roze rondje
To left to the pink disc, or to right to the pink disc

Here, there are two words in OU identical to r_1 , but r_1 does not replace the most recent one. Another case is given in (67):

- (67) En aan de onderkant op de lijn een rode stip, een verticale lijn
And at the bottom of the line a red dot, a vertical line

Here, *a vertical line* has to replace or specify *the line*, not *a red dot*. As noticed, cases like (66) and (67) are quite rare.

Taken together, there are 501 repairs in the corpus where the first word of the repair (r_1) is of the same syntactic category as the last word before interruption (o_n), or is identical to some word o_i in the original utterance, or both. In almost all of these cases the speaker intends the listener to take r_1 as a replacement for the corresponding element o in the OU, and to continue from there on, maintaining the constituency at o . There are no more than 24 repairs of this type which do not adhere to either the one or the other constraint, i.e., less than 5%. For the majority of all repairs (52%), the listener can thus safely apply the just mentioned replacement strategies.

It should be noticed that the two self-imposed constraints of the speaker do not follow from the well-formedness rule: the rule easily allows for repairs which violate the constraints, as is the case in most of the just mentioned examples of violation. The constraints have little to do with well-formedness, they should rather be interpreted as conversational conventions.

It is clear that the two strategies taken together are not sufficient for the

hearer to solve the continuation problem in all cases; there is still 48% of the repairs to be accounted for. There are three major types of continuation in this category. The first one is very similar to the two cases above: r_1 is of the same syntactic category as some $o_i \neq o_n$ of the OU, and is intended to replace it. An example is given in (68):

- (68) Sla linksaf bij knooppunt, naar knooppunt blauw
Turn left at node , to node blue

Here the preposition 'naar' (*to*) is to replace the preposition 'by' (*at*), i.e., there is a category identity and replacement relation, but it does not concern o_n (*node*) but an earlier occurring reparandum. These are, therefore, always cases of delayed interruptions and instant repair. There are 78 (8%) cases in the data where the conditions for the above two strategies are not fulfilled but where there is category identity between r_1 and an earlier element o_i of OU. Of these 78 cases 74 (95%) require replacement of this o_i by r_1 . Another way of looking at this is to consider it as an extension of the category-constraint; the more general formulation would then be: *if r_1 is category identical to some o_i of OU, then replace o_i by r_1 .* (In the case of more than one such category identity, this should apply to the most recent one.) This general condition is fulfilled for 301 repairs of the corpus (31%), and the replacement relation is violated in 28 cases. Of these, 8 are taken care of by the word identity constraint, which leaves 20 irreparable cases (7%).

The second major type of continuation still to be dealt with are hesitations where only an editing expression is used, but where there are no words repeated. There are 167 (17%) of these in the data. An example is (69):

- (69) Gaan dan rechtdoor naar, uh geel
Go then straight to , uh yellow

In these cases there is, usually, no category or word identity relation between r_1 and elements of OU, and the hearer's first strategy might then be to continue processing as if no interruption had taken place. This strategy is fool proof for this category, but will create problems for the third and last major type of case: fresh starts. There are a total of 165 (17%) of these in the corpus. An example is (70):

- (70) Rechtdoor naar, of de ingang is bruin
Straight to , or the entrance is brown

As in the former case of hesitations there is usually no word or category identity relation between r_1 and an earlier element. So, the hearer's problem will be to distinguish fresh starts from hesitations. So, for instance, the hearer should not interpret (70) as 'Rechtdoor naar de ingang...' (*Straight to the entrance...*), since that will garden path him into an erroneous reading. There are three potential cues the listener can use at the beginning of the repair proper to distinguish a fresh start from a hesitation. Two of them have to do with the use of editing terms. Table 9 gives the distribution of editing terms for the hesitations which do not involve repeats of words, and for the fresh starts in the corpus. The first cue is presence or absence of an editing expression. By definition all hesitations of the type under concern have an editing expression. But it is an empirical fact that 111 of the 165 fresh starts in the corpus, i.e., 67% of the fresh starts have none. So, if there is no ET, and r_1 has no category or word identity relation to an earlier item in OU then the hearer can safely conclude that the speaker is making a fresh start, and this takes care of two thirds of the fresh starts.

The second cue applies in the one third of cases where there is an ET. Table 9 shows that whereas 92% of the editing terms used in hesitations are 'uh', only 19% of the ET's in fresh starts are. On the other hand, 'of', 'dus', 'nee', and 'sorry' are exclusively used in fresh starts (cf., example (70)), never in hesitation. If the hearer would rigorously interpret cases of 'uh' as hesitations and all other ET cases as fresh starts, there would be no more than 7% error for the data in Table 9.

The third and last potential cue to be mentioned is prosody. Fresh starts and hesitations containing editing terms may well differ in the way they prosodically relate r_1 to o_n . One would expect sentential prosody to be maintained and continued in the case of hesitations, but not in the case of fresh starts. The distinction would be similar to what Goffman (1981) calls 'flat' versus 'strident' corrections. Cutler (1983) found characteristic intonational differences between these two types. Prosodic cues should especially be important to distinguish fresh starts without editing term from 'unfilled' hesitations: our corpus does not include hesitations without editing term or repetition, but they exist, and the listener would need a cue to distinguish these from fresh starts, so that they can be interpreted as normal uninterrupted speech. For further analyses see Levelt and Cutler (1983).

Taken together, one can conclude that, for the repairs in our corpus, a listener *could* decide on the continuation problem no later than at r_1 , the first word of the repair proper. If the listener would capitalize on the constraints and cues discussed in this section, the decision would be correct in almost all cases. It is an empirical issue whether listeners in fact do decide so early in the process. Or, in other words, is the 'optimal efficiency' of listeners claimed

Table 9. *Editing terms for hesitations without repeats, and for fresh starts*

| | 'Uh' | 'Of' | 'Dus' | 'Nee' | 'Sorry' | Other | None | Total |
|--------------|-----------|--------|---------|--------|---------|----------|-----------|------------|
| Hesitations | 154 (92%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 13 (8%) | 0 (0%) | 167 (100%) |
| Fresh starts | 10 (6%) | 3 (2%) | 15 (9%) | 3 (2%) | 1 (1%) | 22 (13%) | 111 (67%) | 165 (100%) |

by Marslen-Wilson and Tyler (1981) for the processing of single words, and conjectured for more extensive stretches of speech, also valid for the processing of repairs?

Whether or not this turns out to be so, the finding that speakers *do* apparently make the critical information available no later than the first word of the repair proper highlights a remarkable feature of self-repairs.

As a summary of the present section it can be said that in making a repair, the speaker respects certain constraints which facilitate the hearer's task to relate the new formulation to what was said before. There is, firstly, a well-formedness rule which specifies the class of possible repairs that can go with an interrupted original utterance. This rule replaces the traditional constituent boundary rule, which is both trivial and still not correct because it is blind to the structure of the original utterance. Secondly, the way of restarting often reflects the intention of the repair: if a speaker wants to correct an error, the repair is 'conservative', keeping close to the wording of the original utterance; if, however, the original utterance was correct but not fully appropriate, the speaker tends to give a newly constructed qualification which may or may not be followed by a citation of part of the original utterance. Thirdly, there are strong constraints on how the first word of the repair proper can be related to previous elements in the original utterance. These constraints hinge on the identity and syntactic category of that first word, given the OU. Together with additional cues from editing terms, these features of repairs make it in principle possible for the listener to predict the insertion (or non-insertion) relation between the repair proper and the original utterance no later than upon the first word after restart.

7. General discussion

In the previous sections analyses were made of the three major phases of self-repairs: monitoring and interrupting the flow of speech, the use of editing terms, and, finally, the ways in which a repair proper is made.

The present section will return to the issues raised in Section 2, the relations between self-repair and a theory of language production. There are, in

particular, two aspects of self-repairs that are of quite general relevance for a theory of speaking. The first one is the monitoring process, which is, obviously, not limited to cases of repair. The second one is the structural relation between repair proper and original utterance, a relation which echoes similar relations between members of a coordinate construction, and, as will be shown, between question and answer. There are, furthermore, relations between these two aspects of self-repairing, since both seem to involve the speaker as his own listener.

It was suggested in Section 2 that the formulating processes are, normally, opaque to the speaker, and that monitoring should, rather, be regarded as based on the parsing of inner or overt speech. We called this a perceptual theory of monitoring, and presented some arguments to support it. The analysis of the first phase of repairing in Section 4 added some further evidence for a perceptual theory. Though none of the findings are remotely decisive in this respect, the following results are of particular importance. The Main Interruption Rule, which says that the flow of speech is immediately interrupted upon detection of trouble, could be confirmed in various ways; some allowance had only to be made for preservation of the integrity of words that are themselves not erroneous. It was then possible to show, firstly, that the *detection* of trouble is often much delayed with respect to the trouble spot or reparandum. This is not an attractive result for a production theory of monitoring. Take for example the case of repairs of lexical errors. If monitoring is to be located inside the formulating mechanism itself, it can directly observe the output (if not the workings) of the individual subcomponents, in this case the lexical retrieval mechanism. If that component produces an erroneous item, there is no clear reason why the monitor would detect it only after several more words have been retrieved. If one would argue that the monitor relates the word to other and later parts of the utterance, that would be tantamount to saying that the input to the monitoring device is at the level we called 'inner speech'. Such a theory is then indistinguishable from the perceptual theory, since there is no advantage any more to having direct access to the subcomponents of production. The second result of relevance here is that the detection process in monitoring apparently modulates with phrase structure: detection chance increases sharply towards the end of surface constituents. Bock, in her recent review paper (1982), cites evidence in support of the notion that instructions to the motor programming of speech, i.e., to what was called the Articulating component in Section 2, occur roughly phrase by phrase. If this 'inner speech' is the major input to self-monitoring as surmised, the occurrence of a similar phrase-by-phrase organization of the monitoring process should not be surprising.

The great advantage of a perceptual theory is that controlling one's own

speech is like attending to somebody else's talk. This makes it natural for the speaker to apply the same parsing procedures and sources of knowledge to his own speech as to other people's speech. More particularly, the speaker will try and interpret his own speech in the context of what was previously said by himself or by another person. He may thus become aware of ambiguity, vagueness, indeterminacy of reference, incoherence, etc. The first function of monitoring, *matching* (see Section 2, sub-section E Monitoring) can for the large majority of cases in our corpus be performed by comparing the result of parsing one's own (inner) speech, i.e., the derived message, to one's original intentions. It was argued in Section 2 that the speaker has access to both the intended and the derived message; the process of comparison can thus take place 'in' working memory. It should be stressed that a perceptual theory of monitoring is perceptual only in that the same parser is involved in understanding an interlocutor's speech and in deriving the message from one's own (inner) speech. What is *done* with the derived message is quite different in the two cases. In listening to somebody else, one normally matches it to the current discourse model, in order to modify or extend the latter. In listening to oneself the matching is with the intended message, and the criterion is identity of intention. In the event of substantial mismatch, one will have to add something to the utterance.

This brings us to the second function of monitoring mentioned in Section 2, namely that of *creating instructions for adjustment*. It was found that methods of adjusting were quite different for different classes of mismatches. In the case of real errors (e.g., *horizontal* for *vertical*, or *red* for *blue*) adjustments were highly conservative in that they closely resembled the original utterance, and this strategy resulted in rather minimal changes. In the case of appropriateness repairs, however, speakers afforded themselves a high degree of freedom in shaping the adjustment. This difference can easily be interpreted in terms of the theory of Section 2: In the event of error essentially the same message (or part of it) is again sent to the formulator. But correcting an inappropriateness often requires the construction of an additional concept or message to be formulated, which then appears as a pre-specification or fresh start. So much for the monitoring process in repairing.

The next major issue for a theory of language production is that of the structural relations which hold between the repair proper and the original utterance. The original utterance restricts the options available to the speaker concerning the way he formulates the repair. Such restrictions showed up in both the well-formedness rule, and in the two 'conversational' constraints (category and identity) to which speakers apparently adhere in making repairs. The character of these restrictions allows one, firstly, to take a further step towards a unified theory of the speaker, involving repairing, coordinat-

ing, as well as question answering. Secondly, it gives a further indication for the existence of a 'perceptual loop' in the production process.

If the aim is to produce a unified theory, one should observe that the well-formedness rule in Section 6 carries the implication that well-formedness restrictions on repairs are essentially the same as those involved in producing coordinations. This observation suggests that a speaker, in producing coordinations, feeds some of the first conjunct's structural properties into the formulator jointly with the content-information. In this way some of the formulator's procedures may become directly activated by *structural* properties of previous speech, over and above activation by conceptual input ('the message'). The precise definition of these structural properties is not an easy matter. An extensive analysis of semantic and syntactic parallels to be preserved between two conjuncts of a coordination is given in Lang (1982). Klein (1981) discusses several rules for ellipsis in coordination. A general formal treatment of syntactic constraints on coordination is given by Gazdar (1980, see Reference Note 3). Here it must suffice to make the general point that both in making a repair and in making a conjunction the speaker transfers the structural commitments established by the original utterance to the second one (though in different ways, as expressed by the well-formedness rule), and that these commitments are, apparently, highly similar in the two cases.

Turning now to the conjecture that these structural constraints are derived by means of a 'perceptual loop', i.e., by parsing one's own original utterance or first conjunct, a third domain of phenomena in language production may also fall in place. There are striking correspondences between repairing and question answering. Take, for instance, examples (71) and (72) which are examples of well-formed and ill-formed repairs (the latter is identical to the previous example (32)):

- (71) *With his sister he talked frequently, uh with his mother he talked frequently*
 (72) *With his sister he talked frequently, uh his mother he talked frequently*

Notice that the repairs proper maintain the same well-formedness relations as answers to question (73):

- (73) *With whom did he talk frequently?*

The answer *With his mother he talked frequently* is all right in this situation, but the preposition *with* cannot be deleted. There is, moreover, no question for which the inverse relation would hold. Another pair of examples is (74), (75):

- (74) *From the purple node you go left, uh you go right*
 (75) *From the purple node you go left, uh purple node you go right*

The latter repair (75) is ill-formed, whereas (74) is all right (in conformity with the well-formedness rule). The corresponding question would be (76):

- (76) *Where do you (have to) go from the purple node?*

Clearly, *you go right* would be all right as an answer, but *purple node you go right* is very odd.

Such examples can be generated *ad libitum* by replacing the repaired element by *Wh*, and transforming the original utterance accordingly into a question. An ill-formed repair is also an ill-formed answer to such a question. If the OU is incomplete itself, it has to be completed as specified in the well-formedness rule before being transformed into a question. An example is the ill-formed (but real) repair (77) (identical to (45)):

- (77) Links daarvan een, dat zwarte een roze rondje
Left thereof a , that black a pink disc
 'Left thereof a that black one a pink disc'

The corresponding question would be (78):

- (78) Waarvan links een roze rondje?
Whereof left a pink disc?
 'Left of what is a pink disc?'

and the answer 'dat zwarte een roze rondje' (*that black one a pink disc*) would be as ill-formed as the repair. (This works exactly the same in real—not transliterated—English; the corresponding examples were added as third lines in (77) and (78).) For a recent discussion of determinants of deletion in question-answering see Kuno (1982).

If the 'structural transfer' in question-answering is indeed highly similar to the transfer in coordinating and repairing we have an additional argument for the claim that the properties-to-be-transferred are derived in a parsing procedure, where previous speech (overt or inner) is the input. In question-answering this previous speech is the interlocutor's, in self-repair and coordination it is one's own.

Levelt and Kelter (1982) showed experimentally that there is transfer of lexical items from question to answer, even in case the item plays no particular semantic or pragmatic role. Example (79) is an English translation of such a case in Dutch:

- (79) Q: (At) *what time do you close?*
 A: (At) *five o'clock.*

The answer tends to agree with the question in the use of the preposition. Important is the further finding that working memory plays a crucial role in this transfer. Loading memory with additional speech diminished or fully obliterated the transfer. But the same can be observed in self-produced coordination. Example (80) is a natural form of structural transfer (called 'gapping'):

- (80) *Mary visited the Rijksmuseum, and John the zoo.*

But if the speaker loads his own memory with additional speech, he will tend to lose the structural properties obtained by parsing the first clause, and transfer will be less likely. As a consequence, (81) sounds unnatural:

- (81) *Mary visited the Rijksmuseum, you know I told you about this beautiful exhibition on expressionist art which is running there till the end of May, and John the zoo.*

There is one additional consideration relating to the general 'parsing loop' hypothesis, i.e., the idea that working memory has access to perceptual parsing results and may feed these into the formulator which produces a repair, a second conjunct, or an answer. One would like to see a similar explanation of the additional, and highly unexpected finding of the present data analysis, namely that speakers adhere to the category and identity constraints in making their repairs. The issue is whether *these* constraints are also derived through parsing one's own original utterance, and transmitting the relevant category and word identity information to the formulator via working memory. In other words, is it the case that the speaker solves the listener's continuation problem by solving his own continuation problem, rather than by keeping a running model of the listener's state of knowledge? This would surely be a parsimonious result. Nevertheless, it should be again emphasized that, on the present evidence, it would be premature to reject all alternatives to the perceptual theory of monitoring and repair.

Whatever the source of the structural restrictions to which the formulator apparently adheres, a theory of the speaker will have to explain the mechanisms by which these restrictions are realized. This is still an enigmatic issue. Most production models whether formulated in psychological (see Bock, 1982 for a review) or AI terms, generate their output exclusively, or almost exclusively, from a semantic/conceptual base and, accordingly, have great problems with the production of most forms of ellipsis in coordination. An interesting development is the ellipsis component of the HAM-ANS pro-

ject, documented in Jameson (1981, see Reference Note 5). This component indeed involves a parsing loop in the generation of ellipsis: 'An elliptical utterance is assumed to be derived from a complete formula relative to a previous utterance by finding the smallest admissible reduction of the formula which will be correctly reconstructed by the hearer on the basis of a comparison with the previous utterance'. Apart from parsing the previous utterance, it is thus also necessary to parse a candidate utterance before it is produced. One wonders whether this result cannot be realized in a more direct way. It would be preferable for the structural properties of the original utterance, the first conjunct, or the question to affect certain parameters of the formulator's operations in such a way that the elliptical form is *directly* generated, i.e., without mediation of any 'full form'. Ideally, there would then be less rather than more work for the formulator to do, insofar as certain syntactic procedures can either be completely omitted (as in producing gapping or other elliptical constructions), or reduced (for example, as in replacing a full NP by a pronoun). The functional significance of this is clear: the transfer of these sorts of restrictions from one utterance to the next will at the same time increase the fluency of speech (by reducing the number or size of the formulator's operations), and the coherence of discourse (by establishing structural relations between present and previous speech).

In conclusion, the apparently close relationship between repairing, on the one hand, and coordinating and question answering, on the other suggests that the ways in which 'natural language handles its intrinsic troubles' (Schegloff *et al.*, *op. cit.*) may, after all, not be so very different from the ways in which it generally handles coherence and fluency in discourse. The study of self-repairs can thus add to an understanding of these basic properties of language use in context.

References

- Bock, J.K. (1982) Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychol. Rev.*, 89, 1-47.
- Bresnan, J. (1978) A realistic transformational grammar. In M. Halle, J. Bresnan, and G.A. Miller (eds.), *Linguistic Theory and Psychological Reality*. Cambridge, Mass., MIT Press.
- Butterworth, B. (1980) Evidence from pauses. In B. Butterworth (Ed.), *Language Production*. Vol. I. London, Academic Press.
- Cohen, A. (1980) Correcting of speech errors in a shadowing task. In V.A. Fromkin (Ed.), *Errors in Linguistic Performance*. New York, Academic Press.
- Cole, R.A. (1973) Listening for mispronunciations: A measure of what we hear during speech. *Percep. Psychophys.*, 13, 153-156.
- Cole, R.A., and Jakimik, J. (1980) A model of speech perception. In R.A. Cole (Ed.), *Perception and production of Fluent Speech*. Hillsdale, NJ, Erlbaum.

- Cutler, A. (1983) Speakers' conception of the function of prosody. In A. Cutler and D.R. Ladd (Eds.), *Prosody, Models and Measurements*. Heidelberg, Springer.
- Dell, G.E., and Reich, P.A. (1981) Stages in sentence production: an analysis of speech error data. *J. verb. Learn. verb. Behav.*, 20, 611-629.
- Dietrich, R. (1982) Selbstkorrekturen. Fallstudien zum mündlichen Gebrauch des Deutschen als Fremdsprache durch Erwachsene. *Zeitschrift für Literaturwissenschaft und Linguistik*, 45, 120-149.
- Du Bois, J.W. (1974) Syntax in mid-sentence. In *Berkeley Studies in Syntax and Semantics, Vol. 1*. Berkeley, Calif., Institute of Human Learning and Department of Linguistics, University of California, III-1 - III-25.
- Foss, D.J. (1967) Decision processes during sentence comprehension: Effects of lexical item difficulty and position upon decision times. *J. verb. Learn. verb. Behav.*, 8, 457-462.
- Garrett, M.F. (1975) The analysis of sentence production. In G. Bower (Ed.), *Psychology of Learning and Motivation*. Vol IX. New York, Academic Press.
- Garrett, M.F. (1980) Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production*. Vol. I. London, Academic Press.
- Goffman, E. (1981) Radio talk. In E. Goffman (Ed.), *Forms of Talk*. Oxford, Blackwell.
- Herriot, P. (1970) *An Introduction to the Psychology of Language*. London, Methuen & Co.
- Hockett, C.F. (1967) Where the tongue slips, there slip I. In *To Honor Roman Jakobson*, Vol. II. The Hague, Mouton.
- James, D. (1972) Some aspects of the syntax and semantics of interjections. In *Papers from the Eighth Regional Meeting, Chicago Linguistic Society*, 162-172.
- James, D. (1973) Another look at, say, some grammatical constraints on, ok, interjections and hesitations. In *Papers from the Ninth Regional Meeting, Chicago Linguistic Society*, 242-251.
- Kempen, G., and Huybers, P. (in press) The lexicalization process in sentence production and naming: indirect election of words. *Cog.*
- Klein, W. (1981) Some rules for regular ellipsis in German. In W. Klein and W.J.M. Levelt (Eds.), *Crossing the Boundaries in Linguistics. Studies Presented to Manfred Bierwisch*. Dordrecht, Reidel.
- Kuno, S. (1982) Principles of discourse deletion—case studies from English, Russian and Japanese. *J. Seman.*, 1, 61-93.
- Labov, W. (1970) The study of language in its social context. *Studium Generale*, 23, 30-87.
- Lang, E. (1982) *The Semantics of Coordination*. Amsterdam, Benjamin.
- Laver, J.D.M. (1973) The detection and correction of slips of the tongue. In V.A. Fromkin (Ed.), *Speech Errors as Linguistic Evidence*. The Hague, Mouton.
- Laver, J.D.M. (1980) Monitoring systems in the neurolinguistic control of speech production. In V.A. Fromkin (Ed.), *Errors in Linguistic Performance*. New York, Academic Press.
- Levelt, W.J.M. (1972) Some psychological aspects of linguistic data. *Ling. Berichte*, 17, 18-20.
- Levelt, W.J.M. (1974) *Formal Grammars in Linguistics and Psycholinguistics*, Vol. III. The Hague, Mouton.
- Levelt, W.J.M. (1978) Skill theory and language teaching. In A. Valdman (Ed.), *Studies in Second Language Acquisition*. Proc. of the 4th Annual Neuchâtel Colloquium on Theoretical Issues in Applied Linguistics. Vol. I, 1, 53-70.
- Levelt, W.J.M. (1981) The speaker's linearization problem. *Phil. Trans. Roy. Soc. London B295*, 305-315.
- Levelt, W.J.M. (1982a) Linearization in describing spatial networks. In S. Peters and E. Saarinen (Eds.), *Processes, Beliefs, and Questions*. Dordrecht, Reidel.
- Levelt, W.J.M. (1982b) Cognitive styles in the use of spatial direction terms. In R.J. Jarvella and W. Klein (Eds.), *Speech, Place, and Action*. Chichester, Wiley.
- Levelt, W.J.M. (1982c) The speaker's organization of discourse, Report to Plenary Session on Psycholinguistics XIIIth International Congress of Linguists, Tokyo. Reprints of the plenary Session Papers, 241-259.
- Levelt, W.J.M., and Cutler, A. (in press) Prosodic marking in speech repair. *J. Seman.*
- Levelt, W.J.M., and Kelter, S. (1982) Surface form and memory in question answering. *Cog. Psychol.*, 14, 78-106.

- Levelt, W.J.M., and Maassen, B. (1981) Lexical search and order of mention in sentence production. In W. Klein and W.J.M. Levelt (Eds.), *Crossing the boundaries in linguistics. Studies presented to Manfred Bierwisch*. Dordrecht, Reidel.
- Maclay, H. and Osgood, C.E. (1959) Hesitation phenomena in spontaneous English speech. *Word*, 15, 19–44.
- Marslen-Wilson, W., and Tyler, L. (1980) The temporal structure of spoken language understanding. *Cog.*, 8, 1–71.
- Marslen-Wilson, W., and Tyler, L. (1981) Central processes in speech understanding. *Phil. Trans. Roy. Soc. Lond.* B295, 317–332.
- Marslen-Wilson, W., and Welsh, A. (1978) Processing interactions and lexical access during word-recognition in continuous speech. *Cog. Psychol.*, 10, 29–63.
- Meringer, R. (1908) *Aus dem Leben der Sprache*. Berlin: Behr.
- Nisbett, R.E., and Wilson, T.D. (1977) Telling more than we can know: Verbal reports on mental processes. *Psychol. Rev.*, 84, 231–259.
- Nooteboom, S.G. (1980) Speaking and unspeaking: detection and correction of phonological and lexical errors in spontaneous speech. In V.A. Fromkin (Ed.), *Errors in linguistic performance*. New York, Academic Press.
- Schegloff, E.A., Jefferson, C., and Sacks, H. (1977) The preference for self-correction in the organization of repair in conversation. *Lang.*, 2, 361–382.
- Seuren, P. (1978) Grammar as an underground process. In A. Sinclair, R.J. Jarvella, and W.J.M. Levelt (Eds.), *The Child's Conception of Language*. Heidelberg, Springer.
- Uit den Boogaart, P.C. (1975) *Woordfrequenties in geschreven en gesproken Nederlands*. Utrecht, Oosthoek, Schellema & Hoikema.
- Wernicke, C. (1874) *Der aphasische Symptomencomplex*. Breslau, Max Cohn & Weigert.
- Yngve, W.H. (1961) The depth hypothesis. In R. Jakobson (Ed.), *Structure of Language and its Mathematical Aspects*. Proc. 12th Symposium in Applied Mathematics. Providence, RI, American Mathematical Society.
- Yngve, V.H. (1973) I forget what I was going to say. In *Papers from the Ninth Regional Meeting of the Chicago Linguistics Society*, 688–699.

Reference Notes

1. Dell, G.S. (1980) Phonological and lexical encoding in speech production: An analysis of naturally occurring and experimentally elicited speech errors. Unpublished Ph.D. thesis, University of Toronto.
2. Garrett, M.F. (1981) Production of speech: Observations from normal and pathological language use. Unpublished manuscript.
3. Gazdar, G. (1980) Unbounded dependencies and coordinate structure. Unpublished manuscript. Cognitive Studies Programme, School of Social Science, University of Sussex.
4. Hoenkamp, E. (1980) Spontaneous speech as a feedback process. Internal Report 80 FU 09, Department of Psychology, Nijmegen University.
5. Jameson, E. (1981) Documentation for three HAM-ANS components: Ellipsis, Normalize, and Normalize-1. Memo ANS-4, Hamburg: Research Unit Information Science and Artificial Intelligence.

Résumé

L'auto-correction dans le discours se fait typiquement en trois temps. Dans un premier temps, le locuteur contrôle sa propre parole et l'interrompt lorsqu'il rencontre un problème. Une analyse de 959 corrections spontanées indique que l'interruption suit de très près la perception du problème, à l'exception près que le

locuteur a tendance à finir les mots corrects. Les résultats de cette analyse indiquent d'autre part que la perception du problème s'améliore vers la fin des constituants. Le deuxième temps se caractérise par des hésitations, des pauses, mais surtout par l'utilisation de ce qu'on peut appeler les 'commentaires rédactionnels'. Ceux-ci sont liés de façon suffisamment régulière à la correction particulière qui est faite: ils sont différents lorsqu'il s'agit d'une véritable erreur et lorsqu'il s'agit simplement d'une mauvaise tournure de phrase. La présence immédiate du problème est signalée par l'utilisation de 'uh'. Dans le troisième temps a lieu la correction elle-même. La bonne-formation des corrections ne dépend pas de ce que le locuteur respecte l'intégrité des constituants, mais plutôt de la relation structurelle qui existe entre le premier énoncé et la correction. Cette relation est liée à la relation correspondante entre les éléments conjoints d'une coordination par une règle de bonne formation bi-conditionnelle. On peut également suggérer qu'il existe une relation semblable entre questions et réponses. Dans ces trois cas, le locuteur respecte les contraintes structurelles de son premier énoncé. Enfin, l'analyse démontre que l'ensemble formé par le 'commentaire rédactionnel' et le premier mot de la correction elle-même contient presque toujours des éléments d'information permettant à l'interlocuteur de décider comment il faut relier la correction au premier énoncé. De ce point de vue, les locuteurs ne produisent presque jamais d'énoncés qui pourraient induire leur interlocuteur en erreur.

Ces résultats indiquent que le locuteur a peu ou pas du tout d'accès au processus de production d'énoncés; l'auto-contrôle se fait plutôt à partir de la compréhension de sa propre parole intérieure ou extérieure.