Chapter 12

Self-Monitoring and Self-Repair

Speakers monitor what they are saying and how they are saying it. When they make a mistake, or express something in a less felicitous way, they may interrupt themselves and make a repair. This is apparent not only in spontaneous conversations, but in all kinds of discourse. Here are three examples. The first was reported by Schegloff, Jefferson, and Sacks (1977):

(1)
A: And he's going to make his own paintings.
B: Mm hm,
A: And – or I mean his own frames
B: Yeah

In spite of B's *mm hm*, a sign of acceptance, A became aware that she had said *paintings* instead of *frames*, and corrected this on her next turn.

The second and third examples are pattern descriptions obtained in an experiment reported in Levelt 1982a,b. The subjects were asked to describe patterns such as the one shown in figure 12.1. They were told that their descriptions would be tape recorded and given to other subjects, who were to draw the patterns on the basis of these recordings (see also subsection 4.4.2 and figure 4.5). One subject was in the process of indicating the connection between the yellow node and the pink node and said

(2) And above that a horizon –, no a vertical line to a pink ball

This is much like the repair of an error in example 1, but here the speaker was very quick in effectuating the repair; the trouble item (*horizontal*) was not even completed. Another subject, going from the yellow node to the blue one, said

(3) To the right is blue – is a blue point

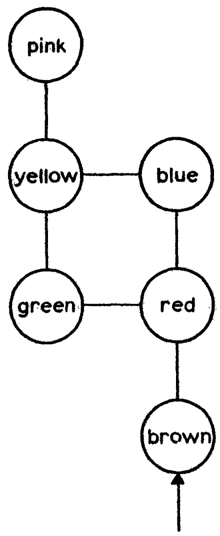

Figure 12.1
One of a set of patterns to be described. The nodes had real colors in the experiment. (After Levelt 1982a.)

This is, clearly, not the correction of an error. Rather, the speaker has made the description of the blue node more precise—more appropriate to the task. The distinction between error repairs (such as in examples 1 and 2) and appropriateness repairs (example 3) is an important one, as will become apparent in this chapter. The three examples show that speakers can monitor and repair their own speech. And this self-monitoring doesn't need the interlocutor's feedback. Speaker A in example 1 discovered her error even though B ignored it. The other two repairs were also made spontaneously; there was no interlocutor to initiate the repair by asking *What?* or *What do you mean?*.

The first section of this chapter will deal with this process of self-monitoring. How dependent is fluent speech on self-monitoring? What does a speaker monitor for? How does a speaker become aware of trouble in his own speech? We will, in particular, consider two theoretical accounts of monitoring: the editor theories and the connectionist models. In section 12.2 we will turn to what the speaker does upon detecting trouble. He may, sooner or later, interrupt his speech to initiate a repair. This usually creates a problem for the interlocutor or the listener. The normal flow of speech is interrupted, and the interpretation of the ongoing speech has to be changed or qualified. Speakers often signal this to their listeners by sudden pausing and by the use of editing expressions, such as *er*, *rather*, *no*, and *I mean.* We will analyze this process of self-interruption and the systematicity in the employment of editing expressions. Finally, the speaker will initiate the repair proper. Section 12.3 will discuss the way in which speakers construct their repairs. Contrary to certain claims in the linguistic literature, repairs are constructed in a highly regular fashion. They are subject to strict linguistic rules and strict conversational rules.

12.1 Self-Monitoring

12.1.1 What Do Speakers Monitor for?

That speakers can attend to various aspects of the action they are performing is apparent from the kinds of spontaneous self-repairs speakers make. Some major targets of monitoring seem to be the following.

Is this the message/concept I want to express now?

An example from the above-mentioned set of pattern descriptions (Levelt 1983) is this:

(4) We go straight on, or – we enter via red, then go straight on to green.

Here the speaker realized that his linearization was not adequate; he should have mentioned the brown entry node before moving to the green one. He stopped and restarted the right way.

Also, a particular message may, on closer inspection, not be correct or adequate with respect to the intention. This seems to have been the case in the following utterance (from Schegloff 1979):

(5) Tell me, uh what – d'you need a hot sauce?

Here the speaker probably started to say *what do you need?*, but it was apparently more adequate to issue a different directive, a Yes/No question. The original utterance was interrupted on the fly, and a somewhat different speech act was performed.

In these and similar cases the speaker's trouble is at the conceptual level. The speaker can directly monitor the messages he prepares for expression, and he may reject a message before or after its formulation has started. In the former case no overt repair will result, though there may be some hesitation. In the latter case the original utterance will sooner or later be replaced by a formulation of the alternative message.

Is this the way I want to say it?

Even if the speaker is sure about the information to be conveyed, he may get second thoughts about the way it should be expressed, given the discourse record—i.e., given the topic and content of previous discourse, given what was literally said earlier, and so on (see section 4.2). Consider the following example from the pattern descriptions:

(6) To the right is yellow, and to the right – further to the right is blue

Here the speaker started expressing the intended concept, but then realized that the move described previously had also been to the right. By adding *further* the speaker made his utterance more cohesive with previous discourse.

The speaker may also realize that what he is saying involves a potential ambiguity for the listener. Again, the intended concept was expressed, but it was, in retrospect, not sufficiently contextualized. The following example from Schegloff et al. 1977 is a case in point:

(7) Hey, why didn't you show up last week. Either of you two.

Here the speaker addressed the interlocutors with a question about not showing up last week. But then he realized that the communicative situation could be ambiguous. He had intended to address the interlocutors not as a group but as individuals. That ambiguity was taken away by the repair.

At this level, in short, the speaker's monitoring concerns the contextual appropriateness and sufficiency of the information expressed. Closely related to this focus of monitoring is the next one.

Is what I am saying up to social standards?

One's choice of words will, normally, depend on the level of formality required by the context of discourse. In court one will say *policeman* rather than *cop*; this is a choice of *register*. There is some evidence that speakers monitor their speech for unacceptable deviations from standards of formality and decency. In particular, Motley, Camden, and Baars (1982) have shown experimentally that speakers are very good at catching taboo words before they are uttered.

Am I making a lexical error?

The speaker's message may be as intended and contextually appropriate; still, flaws of formulation may appear, and speakers do monitor their speech for these errors. The most frequently caught error of this kind is the *lexical error*. Consider an example from the pattern descriptions; it appeared in one subject's description of the pattern shown in figure 12.1:

(8) Left to pink – er straight to pink

Here the speaker almost certainly intended to express the concept "straight". The previous utterance had introduced the yellow node, and it is even impossible to go left from there. Still, the wrong lemma, *left*, became activated (owing to causes discussed in chapter 6). The speaker caught the error and corrected it. Here is another case, from Levelt and Cutler 1983:

(9) Well, let me write it back–er, down, so that ...

Speakers monitor for lexical errors of any grammatical category, but not all lexical errors are caught.

Are my syntax and my morphology all right?
Certain formulating errors are due not so much to lexical access as to other trouble in grammatical encoding; and sometimes speakers do become aware of deviant syntax or morphology, as is evident from their repairs. Note the following instances:

(10) What things are this kid – is this kid going to say incorrectly?
(from Levelt and Cutler 1983)

Here the speaker noticed an error of agreement and corrected it.

(11) Why it is – why is it that nobody makes a decent toilet seat?
(from Fay 1980b)

Here an ordering error, either an error of syntax or a shift, was caught and immediately corrected.

Am I making a sound-form error?
Trouble in phonological encoding is often recognized by speakers, as is apparent from spontaneous repairs. Cases 12 and 13 are examples of segmental and suprasegmental phonological trouble that was apparently quickly noticed by the speaker.

(12) A unut – unit from the yellow dot (from Levelt 1983)

(13) ... from my prOsodic – prosOdic colleagues (from Culter 1983b)

Below we will discuss experimental evidence presented by Baars, Motley, and MacKay (1975) which demonstrates that speakers can notice and intercept an imminent nonword before uttering it.

Has my articulation the right speed, loudness, precision, fluency?
There is some minimal evidence that speakers monitor their speech delivery for parameters of this sort, but it does not stem from spontaneous self-corrections. It is exceptional indeed for a speaker to spontaneously repeat a word with more precision, or more slowly, or more loudly. Such corrections are typically induced by the interlocutor, who says *what?*, knits his brows, or otherwise signals that the speaker's delivery was not optimal. There is some experimental evidence for self-monitoring of loudness. Speakers immediately increase the loudness of their speech when it becomes masked by loud noise—this happens naturally at cock-

tail parties, but it can also be provoked experimentally (Siegel and Pick 1976).

It is surely possible to make more fine-grained distinctions between foci of self-monitoring. Schegloff et al. (1977) mention, for instance, the monitoring of word selection, of person reference, and of next-speaker selection (example 7 is a case of the latter). Cutler (1983b) reports interesting findings about speaker's monitoring of pitch-accent placement. The main conclusion here can, even without these finer distinctions, be straightforward: Speakers can monitor for almost any aspect of their own speech.

12.1.2 Selective Attention in Self-Monitoring

Do speakers actually attend simultaneously to all these aspects of their speech? This is most unlikely, and there are data to support the view that (i) much production trouble is not noticed by the speaker, that (ii) monitoring is context-sensitive, i.e., contextual factors determine which aspects of speech will be given most scrutiny by the speaker, and (iii) a speaker's degree of attention for self-generated trouble fluctuates in the course of an utterance.

There is both indirect and direct evidence that the meshes of a speaker's trouble net are too wide to catch all queer fish in his own speech. Nooteboom (1980) analyzed Meringer's (1908) corpus of speech errors and found that 75 percent of the registered phonological errors and 53 percent of the lexical errors were repaired by speakers. This is indirect evidence, because a speaker may detect all errors but still not bother to correct each and every one of them.

More direct evidence can be found in Levelt 1983, which reports on color-naming errors in a pattern-descriptions task. Remember that the subject's task was to give a description that would allow another subject to draw the pattern. It was, therefore, essential for a speaker to give correct color names in all cases. All 2,809 pattern descriptions, produced by a total of 53 subjects, were checked for errors in color naming. There were 472 such errors. A speaker would occasionally say *yellow* instead of *green*, *orange* instead of *pink*, *green* instead of *blue*, and so forth. Of these errors, only 218 were repaired by the speaker. That is 46 percent, which corresponds well to Nooteboom's 53 percent for lexical errors. So, even where it is a speaker's given task to produce the correct color name, only about half of the errors are caught. This is most probably due to failures in detection.

Not all sources of trouble are given equal attention. The context of discourse is an important determinant of the kind of flaws a speaker will try to prevent, intercept, or correct. One would expect a speaker to attend most carefully to trouble that is potentially disruptive for the ongoing discourse. Evidence pointing in this direction was provided by Cutler (1983b), who analyzed a corpus of lexical-stress errors (such as the one in example 13 above). About 50 percent of these errors were spontaneously repaired. The likelihood of repair depended on how disruptive the error might be for the listener. A potentially disruptive stress error is one in which there is also a segmental change, in particular a change in vowel quality. That holds for *prOsodic* in example 13, where the second vowel is reduced to [ə] and is no longer perceivable as [ɒ]. It does not hold for the stress error *sarcAsm*, where both vowels keep their intended vowel quality. That the former kind of stress error is in fact far more disruptive for the listener than the latter was experimentally confirmed by Cutler and Clifton (1984). They found that potentially disruptive stress errors were spontaneously repaired in 63 percent of the cases, whereas no more than 23 percent of nondisruptive stress errors were repaired.

It should, of course, be noted that this evidence is indirect. The speaker may have noticed a nondisruptive stress error without bothering to repair it. A disquieting finding in this connection is Nooteboom's, mentioned above. One would expect lexical errors to be more disruptive for discourse understanding than phonological errors; but only 53 percent of the lexical errors were spontaneously repaired, as against 75 percent of the phonological ones. Did Meringer's speakers not bother to correct their lexical errors? This is most unlikely. They probably didn't notice them to start with. But then there is a problem with the notion that speakers will attend most to flaws that are most disruptive for their listeners. Did Meringer carefully register all the repairs? His research target was speech errors, not repairs. The case is, clearly, undecided, but we will shortly return to it when we consider theories of self-monitoring.

Much more direct evidence for the context sensitivity of monitoring proceeds from various experimental studies by Baars, Motley, and others. All these studies used, in one way or another, the speech-error-inducing technique discussed in subsection 9.5.2. A subject is presented with a series of biasing word pairs, such as *ball dome*, and is asked to read these pairs silently. They are then followed by the target pair, for instance *darn bore*, which is to be read aloud. When the word-initial phonemes of the target pair are the same as those of the bias pairs but the order is reversed, speakers make occasional errors on the target pair, saying (e.g.) *barn door* instead of *darn bore*. The technique has been used not only to study the generation of speech errors but also to study the selective attention of speakers in monitoring for such errors.

A good instance of such a study is the original one, in which Baars, Motley, and MacKay (1975) studied the lexical-bias effect (the finding that errors that are real words are more likely to occur than errors that are nonwords). In subsection 9.5.2 above, we considered the activation-spreading account for this finding: Backward spreading of activation can only affect lexical nodes; there are no such nodes for nonwords. But the original interpretation of the lexical-bias effect (Baars et al. 1975) was a different one. They concluded that speakers edit their speech for lexical status before it is uttered. A nonlexical slip is intercepted just before it is uttered, but a lexical slip may pass through the sieve. We will return to this editing theory shortly; the important point here is that Baars et al. were able to show that this monitoring of nonlexical slips is subject to selective attention. Lexical bias is not a necessary effect; it depends on the contextual setting. When *all* pairs of items in the list were nonwords, there was no bias against nonword slips. Word and nonword slips appeared at equal rates. But as soon as some pairs of real words were added to the list (the target pairs were left unchanged; i.e., they consisted of nonwords), the lexical bias appeared again—the slips tended to be words. If a speaker's task deals exclusively with nonsense words, he apparently doesn't bother to attend to the lexical status of his output; however, this changes drastically when normal words appear in the testing materials. This finding is hard to reconcile with the activation-spreading account of the lexical-bias effect, which applies equally to both conditions.

The same technique was also helpful in establishing other context sensitivities in a speaker's selective attention. Some of these findings can be accounted for by activation spreading, but for others this is less easy. Here we will take them on their face value as contextual effects on a speaker's self-monitoring. The alternative theories of monitoring will be taken up in subsequent paragraphs.

Motley (1980) used interference items of a sort that might introduce a certain *semantic* bias in a speaker's editing. For instance, the target pair *bad-mug* might be preceded by semantically biasing pairs such as *irate-wasp* and *angry-insect*. These may weaken the resistance against a slip such as *mad-bug*. And that is what Motley found. Such slips were edited out less often when the context was semantically biasing than when it was semantically neutral. Similarly, *syntactic* biases could be induced—i.e., biases for particular syntactic constructions, such as adjective-noun (Motley, Baars, and Camden 1981).

It is also interesting that Motley (1980) was able to create a biasing *conversational setting*. In this experiment he used target pairs preceded by "standard" phonological interference (biasing) items. The target items were of two kinds. One kind is exemplified by *shad-bock*, which when preceded by appropriate interference items may lead to the slip *bad-shock*; this was the "electrical" kind of target. The other targets were "sexy" ones, such as *lood-gegs* and *goxi-firl* (the intended slips are obvious). The two target types were mixed in the list. Half the subjects were attached to fake electrodes and told that mild shocks could be given. The luckier half of subjects underwent no such threat, but had an attractive and provocatively attired female experimenter. The resulting slips corresponded to the condition of treatment. In the electrical condition, "electrical" speech errors prevailed; in the sexy condition, "sexy" speech errors were dominant. Because all these speech errors were induced by *phonological* interference items, Motley concluded that the difference was an editing effect. When one expects things electrical, a phonological slip that produces such an item will not be filtered out by the editor, and similarly for sexy items. There is an attentional bias in the subject.

In conclusion: A speaker can attend to different aspects of his speech output. In this way, potential flaws can be intercepted before they are overtly uttered. This is called *prearticulatory editing*.

There is, further, evidence that a speaker's attention to his own output fluctuates in the course of an utterance. The evidence proceeds from an analysis of the 472 color-name errors mentioned above (Levelt 1983). For each of these errors it was determined where it occurred in the ongoing phrase. More precisely, it was determined how many syllables separated the erroneous color word from the end of the syntactic constituent (usually a noun phrase) to which it belonged. For example, the erroneous color words in examples 14, 15, 16, and 17 are zero, one, two, and three syllables away from the end of the phrase (which is marked by a slash).

(14) And then you come to *blue* / – I mean green

(15) There is a *yellow* node / to the right of the red one

(16) To the right is a *black* crossing / from which you can go up or down

(17) You enter at a *green* nodal point /

It was then determined how many color-name errors in these different positions were noticed by the speaker and repaired (as in example 14). The results are given in figure 12.2, which shows clearly that the error-detection rate increases sharply toward the end of the phrase. Of the phrase-final color-name errors, 57 percent were detected and repaired;

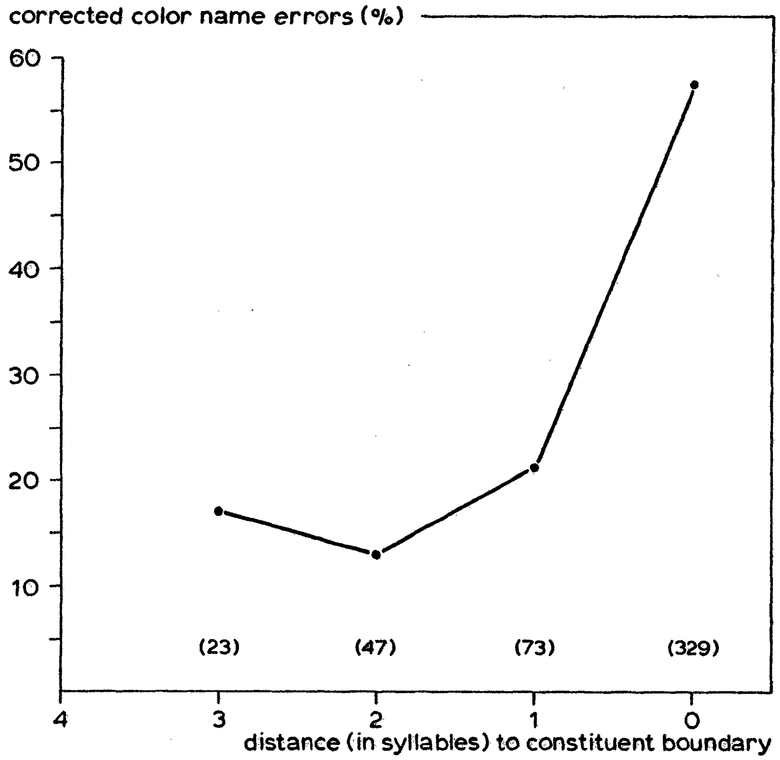

Figure 12.2
Proportion of color name errors detected and repaired by speakers for different positions in the current phrase. Numbers in brackets depict the total number of color name errors for that position. (After Levelt 1983.)

for non-phrase-final errors, the percentage was no greater than about 15. In other words, a speaker's selective attention to his own output increases toward the ends of phrases. During speech a speaker's attentional resources are mainly used for message planning, but by the ends of phrases attention can momentarily shift in order to evaluate the current speech output.

I have dealt with selective attention in monitoring, but I have delayed the discussion of theories of monitoring. There is no one generally accepted theory of self-monitoring in speech, but the main classes of theories have been foreshadowed here. There are *editor theories* of monitoring, and there are *connectionist theories* such as activation-spreading accounts. Let us take these up in turn.

12.1.3 Editor Theories of Monitoring

The major feature of editor theories is that production results are fed back through a device that is *external* to the production system. Such a device is called an *editor* or a *monitor*. This device can be distributed in the sense that it can check in-between results at different levels of processing. The editor may, for instance, monitor the construction of the preverbal message, the appropriateness of lexical access, the well-formedness of syntax, or the flawlessness of phonological-form access. There is, so to speak, a watchful little homunculus connected to each processor. Distributed editing has been proposed by Laver (1973, 1980), De Smedt and Kempen (1987), and van Wijk and Kempen (1987).

A major problem with distributed editing is reduplication. The editor that evaluates the output of a particular processing component must incorporate the same kind of knowledge as the component it monitors; how else could it evaluate the component's output? Hence, for each level of processing there is a reduplication of knowledge: the processor's and the monitor's. Distributed editors are, moreover, not on speaking terms with the notion of components as autonomous specialists. Section 1.3 proposed a partitioning of the system such that a component's mode of operation is minimally affected by the output of other components. This principle is undermined when each processing component is controlled by some monitoring agent.

A more restricted editing device was proposed by Motley, Camden, and Baars (1982). It cannot inspect all intermediary output in the generation of speech, but only the prearticulatory output. Editing follows phonological encoding, according to these authors. The editor can intercept or veto troublesome output before it becomes articulated—hence the notion of "prearticulatory editing."

Some evidence for the existence of such an editor was reviewed in the previous subsection, but particularly important in this respect is the study by Motley et al. (1982). In an experiment very much like the abovementioned one by Motley (1980), they induced socially less appropriate speech errors (e.g., *tool kits* → *cool tits*). In these cases a speaker could either (i) not make the error, or (ii) make the error, or (iii) make a partial error. A partial error is one in which there is no full exchange, but rather an anticipation or perseveration, such as *cool kits* or *tool tits*. Of these two, the former is socially more appropriate than the latter. Motley et al. found that the appropriate partial speech error was made more frequently than the inappropriate one, and they explained this by a mechanism of prearticulatory editing. The speaker, they conjectured, would internally generate the full taboo error (*cool tits*) and start articulating it. If there were enough time to recognize the taboo word, they might be able to intercept it and correct it before uttering it. The result would be the innocent partial error (*cool kits*).

This conjecture predicts two nontrivial results. The first one is that speakers will be more prone to manufacturing innocent partial errors when the potential taboo word comes second than when it comes first. It comes second in the just-mentioned example; it would come first in a case like *fits tall* → *tits fall*. In the latter case there will hardly be enough time for the editor to recognize and intercept the taboo word. And indeed, this is what was found in the reported experiment. The second result predicted is that, because a *potential* taboo word is recognized by the speaker, the speaker should give the corresponding emotional reaction. The authors measured their speakers' galvanic skin response, which is known to reflect emotional arousal. This prediction is quite nontrivial for the following reason: An innocent partial speech error, such as *cool kits*, could have been produced as a simple error of anticipation without any intervention of an editor. In that case no taboo word would have been generated in phonological encoding, and the speaker would have had no reason to show an emotional response. The prearticulatory-editing explanation, however, presupposes that the full error (*cool tits*) was internally generated, and that it would have produced the corresponding galvanic skin response despite the fact that no overt taboo word ever appeared. Again, this is what Motley et al. found. There is also good clinical evidence for the existence of prearticulatory editing in agrammatic patients (Garnsey and Dell 1984).

Unlike distributed editors, the prearticulatory editor is not omnipresent in the language-production system. It cannot evaluate intermediary results, and it leaves the operation of most processing components unaffected. Whether this editor involves reduplication is not resolved. It must, of course, be able to evaluate semantic, syntactic, phonological, and other aspects of prearticulatory or internal speech. How can this be conceived without implicating reduplication?

An obvious solution is to identify the editor with the language-understanding system. A speaker can attend to his own speech in just the same way as he can attend to the speech of others; the same devices for understanding language are involved. In Levelt 1983 I elaborated this proposal by supposing that there is a double "perceptual loop" in the system—that a speaker can attend to his own *internal* speech before it is uttered and can also attend to his self-produced *overt* speech. In both cases the speech is perceived and parsed by the normal language-understanding system. This view of self-monitoring is depicted in figure 12.3, which is nothing but a condensed version of figure 1.1. It should be

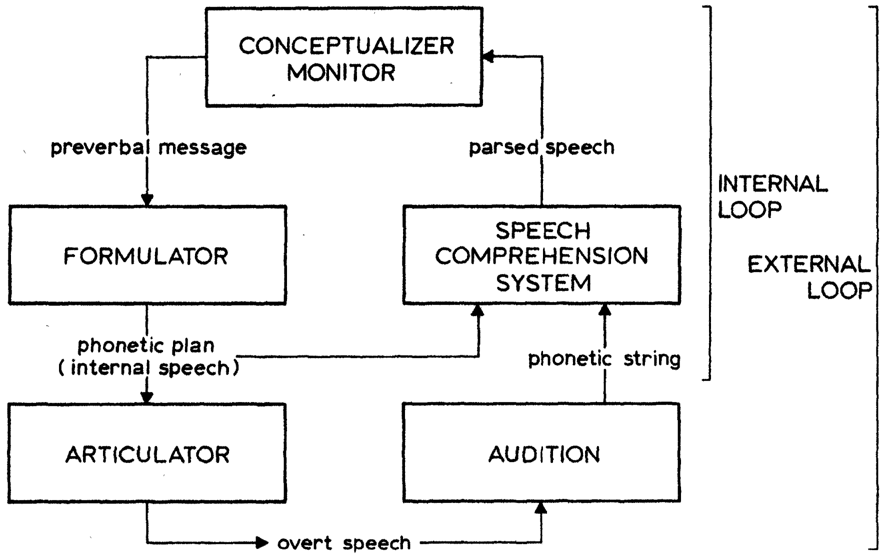

Figure 12.3
The perceptual loop theory of self-monitoring.

noted that the language-understanding system is not only able to derive a message from its speech input, it is also able to detect deviations from linguistic standards. When we listen to the speech of others we can discern deviant sound form, deviant morphology, and deviant syntax. According to the perceptual-loop theory, the same mechanism is involved in monitoring one's own internal or overt speech.

The major advantage of this approach is that no *additional* editing devices have to be conjectured. There are no special-purpose editors to check the outputs of lemma access, of grammatical encoding, of segmental spellout, and so forth. Only the final (prearticulatory) phonetic plan or internal speech and the overt speech produced can be attended to by the language-understanding system. The aspects of the self-produced speech to which a speaker will be especially sensitive will then depend on the context, in agreement with the findings of Motley and colleagues.

This is important, because MacKay (1987) has argued against editor theories on the ground that the errors one detects in one's own speech are different *in kind* from those one detects in the speech of others. MacKay compared Nooteboom's (1980) counts in the Meringer data on self-correction and in Tent and Clark's (1980) data on error detection in other-produced speech. As mentioned, Nooteboom found a 75-percent correction rate for phonological errors and a 53-percent rate for lexical errors. Tent and Clark, who asked their subjects to exactly transcribe the sentences they heard (and these sentences could contain various kinds of errors), found that a phonemic mispronunciation was noticed far less often than a morphemic or a syllabic error.

But this argument is invalid. Apart from the fact that the two data sets are highly incomparable (for instance, white noise was added to Tent and Clark's experimental sentences), it is obvious that the attentional criteria are quite different in listening and speaking. A listener's first concern is to extract meaning from the speech signal, even in spite of ill-formedness. A speaker, however, must guard not only the meaningfulness of his speech but also its satisfaction of grammatical standards. His attention will be tuned accordingly.

Moreover, the monitoring of other-produced speech is itself highly dependent on context. The detection rate for errors of different types is highly dependent on the kind of task assigned to the listener (Bond and Small 1984). When a listener is instructed to monitor for errors of a particular kind, the chances of detection rise sharply (Cooper, Tye-Murray, and Nelson 1987). In short: A difference between the detection distribution for self-monitoring and that for the monitoring of others can hardly be used as an argument against the perceptual-loop model. Both distributions depend on context and on task, and in normal conversational situations these contexts are systematically different for self-monitoring and other-monitoring.

The perceptual loop consists, according to figure 12.3, of two pathways, one involving perception of self-produced overt speech and one involving internal speech. Can these two routes be distinguished experimentally? Are there, in particular, qualitative differences in the features picked up by these two systems, and are there differences in speed of feedback?

An inventive study by Lackner and Tuller (1979) produced affirmative answers to both of these questions. Subjects were given a string of four syllables (for instance, *pi-di-ti-gi*) and asked to repeat this sequence again and again over a period of 30 seconds. The speaking of these syllables was paced by means of a light flashing once per second. This is a tongue-twister-like situation, and speakers occasionally made errors in pronouncing the syllables. They were asked to push a button every time they noticed an error. In this way two aspects of monitoring could be registered: the detection or nondetection of a self-produced error, and the latency of detection (measured as the time interval from speech onset to button press). The experiment was designed to compare a condition where there was overt feedback with one where there wasn't. The latter condition was realized by using earphones to present the speakers with strong white noise, which largely or fully masked the self-produced overt speech.

Are there qualitative differences in self-monitoring between these two conditions? Lackner and Tuller used three kinds of tongue-twisting strings. Apart from the type given above (involving CV syllables differing in onset consonants), they gave strings where CV and V syllables alternated (e.g., *pi-æ-ti-o*) and strings consisting of V syllables only (e.g., *æ-i-o-u*). For the first two kinds of strings, errors could be categorized as *place-of-articulation* errors (e.g., *di* for *gi*), as *voicing errors* (e.g., *di* for *ti*), or both (e.g., *di* for *pi*). Did the detection rates for these three kinds of error differ between the two testing conditions? They did. The detection rate of voicing errors was markedly lower in the auditory masking condition than in the unmasked condition. There was, however, hardly any difference in detection rate for place-of-articulation errors or for combined errors. Trouble in voicing is, apparently, far better perceived via overt speech than via internal speech. This result is not surprising from the perceptual-loop point of view. In the articulatory plan, the distinction between voiced and unvoiced plosive consonants hinges on a tiny difference in the moment of voice onset after the release of the obstruction in the vocal tract. This, however, creates a substantial acoustic effect, which can be picked up by the external loop. Similarly, Lackner and Tuller found that errors in vowel quality were detected much better when there was no masking. In the masking condition, vowel errors that involved small deviations from the target vowel (e.g., [ɛ] for [ɪ]) often went unnoticed.

There were also differences in detection latencies between the two conditions. Although one would expect the masking condition to be more difficult and therefore to create longer latencies, the reverse turned out to be the case. Detection latencies were shorter in the masked condition. And this is as it should be. The overt-speech loop is longer; extracting features from overt speech requires both articulation and acoustic analysis. These processes are not involved in the internal loop. Masking noise suspends the longer external loop.

Lackner and Tuller argued that, in normal speaking situations, both loops are involved. This should predict that the detection latencies for self-produced errors will be shorter than those for other produced errors, since some self-produced errors will be detected via the internal loop. An experiment comparing the detection latencies of self-produced and other-produced errors showed that this was indeed the case. For the kinds of materials used in the above experiment, detection latencies were more than 100 milliseconds shorter in self-monitoring than in other-monitoring.

Is the perceptual loop short enough to intercept an error? The external loop isn't, of course; a word is evaluated only when it is articulated. The effectiveness of the internal loop, however, depends on how far phonetic planning is ahead of articulation. We have seen that it can be one or even a few phonological phrases ahead. The more articulatory buffering, the more opportunity for self-monitoring. But what about fast-running speech, where buffering is presumably minimal?

Let us assume that the internal loop and the Articulator can simultaneously start working on a phonological word delivered by the Phonological Encoder. How much time does the Articulator use to unpack the phonetic program and to initiate articulation? The only measures we have are the latencies in the fourth column of table 11.1. These were the simple reaction latencies in the experiments of Klapp et al. They were about 310 msec. But these latencies included perceiving the Go signal (say, 80 msec). The Articulator's own latency can therefore be estimated at about 200–250 msec. This means that there is at least this much time between the delivery of the phonetic plan and the initiation of articulation. Is this long enough for the internal loop to recognize the word and, in case of trouble, to send an interruption signal to the Articulator?

How long does it take to recognize a word in running speech? The results of Marslen-Wilson and Tyler (1981, 1986) indicate that, on the average, recognition is possible about 200 msec after word onset. This corresponds roughly to the length of a syllable or a monosyllabic word. This short time holds for listening to *overt* running speech; i.e., it includes the time taken by the auditory process. Recognition of one's own *internal* speech can, therefore, be even speedier, say with a latency of 150–200 msec. This would leave the internal loop with 0–100 msec to send an interrupt signal to the Articulator in case of trouble. In other words, the internal loop may in many cases be short enough for effective word-level monitoring in fluent speech—i.e., short enough to interrupt articulation before the troublesome word is spoken. And this is the extreme case where there is no buffering. If there is more phonetic plan available in the buffer, there should be no timing problem whatsoever.

This, however, requires optimal attentional conditions. These are not always given, as we saw in the previous subsection. The speaker may then still be too late in intercepting the trouble item. This can lead to repairs such as the following:

(18) To the left side of the purple disk is a *v* – a horizontal line

Here execution (of the word *vertical*) had just set in before speech could be halted. Such a repair shows that self-monitoring indeed need not be based on overt speech. The erroneous word had been planned, as is evident from the initial [v] (and the context of pattern descriptions), but too little of the word was pronounced to make recognition via the external loop possible.

Summing up: We have seen good evidence for both prearticulatory and postarticulatory editing. That there is prearticulatory editing is apparent from partially intercepted errors such as example 18, but especially from the experiments of Motley et al., whose nontrivial predictions about partial repairs and psychophysiological responses were fully confirmed. That there is, in addition, monitoring of *overt* speech is apparent from the results of Lackner and Tuller, which show a differential sensitivity to self-produced errors under conditions of auditory feedback versus impeded feedback. There is, in short, evidence for the existence of both the external and the internal feedback loop.

Before we turn to the connectionist accounts of self-monitoring, one additional remark should be made: Not all self-monitoring during speech is mediated by the perceptual loops. Speakers can also monitor their messages before they are formulated. They can attend to the appropriateness, the instrumentality, and the politeness of the speech act they are planning. This is an entirely conceptual activity; it need not involve any formulation or parsing.

12.1.4 Connectionist Theories of Monitoring

It is characteristic of connectionist accounts that there are no mechanisms external to the speech-production apparatus involved in the control of one's own speech. In positive terms: The system's self-control is due to the same inherent feedback that is presumably operative anyhow in the generation of speech. That is, the bottom-up priming from lower-level nodes to higher-level nodes in the network (see subsection 9.5.1). There is no editor watching the production system's output.

Even strong proponents of an activation-spreading account, such as Dell (1986) and Stemberger (1984), do not dogmatically exclude other kinds of feedback—e.g., through the language-understanding system. Still, one naturally tries to push the notion as far as possible. The most far-reaching network account of monitoring is that of MacKay (1987). His *node structure theory* explains error detection along the following lines.

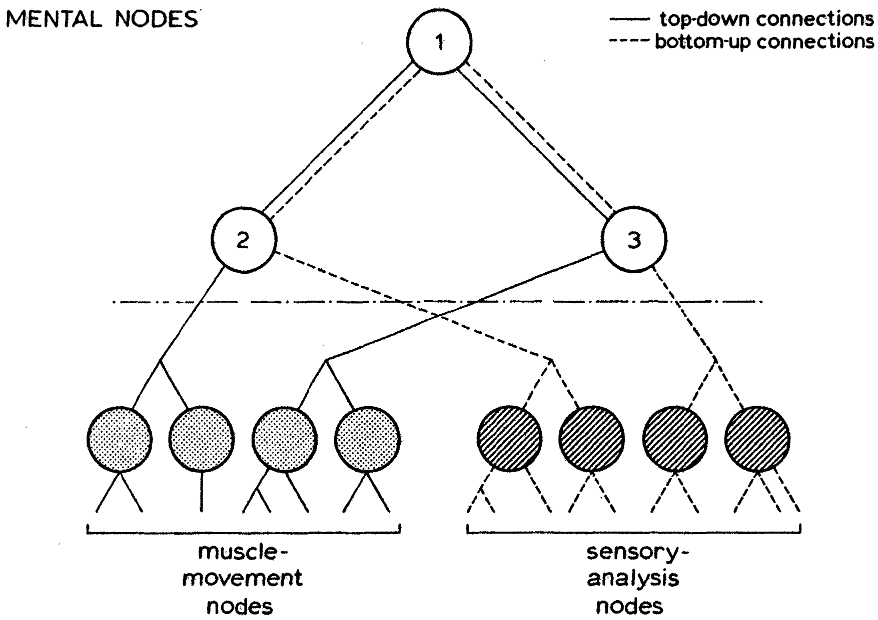

Figure 12.4
Common and specific nodes in MacKay's (1987) node structure theory.

Language production and language understanding are largely subserved by one and the same *node structure*, a layered network of mental nodes. This is schematically represented in figure 12.4, which pictures a very small part of the network. The three mental nodes represented in the upper portion of the figure are common to production and perception. The two node types in the bottom part of the figure are specific. They are, respectively, "muscle-movement nodes" and "sensory-analysis nodes". These nodes are involved in articulation and audition, respectively.

The network for language has several layers of mental nodes. There are layers of "propositional nodes," of "conceptual nodes," of "lexical nodes," of "syllable nodes," of "phonological nodes," and of (distinctive) "feature nodes." All these layers of mental nodes are shared by the production system and the comprehension system. An activated mental node primes all nodes connected to it. These, in turn, prime—to some lesser extent—the nodes connected to them, and so on. The theory is quite similar to Dell's (1986) in that it assumes bidirectional priming between layers of mental nodes. It is also similar in another respect. In Dell's theory, a node could become "current node," which would boost its activation level. When, for instance, a syllable-onset phoneme was structurally required at some moment in time, the most strongly primed onset phoneme would become the current node. The same principle is used in the node-structure theory. The most primed node in a particular domain (e.g., the domain of onset phonemes) will become the "activated" node at the moment it is structurally required. An "activated" node in the node-structure theory is equivalent to a "current" node in Dell's spreading activation theory.

Also, errors arise in the same way in the two theories: The "wrong" node becomes "current" or "activated." This happens when the "wrong" node is the most strongly primed one of its kind at the moment such a node's activation is needed. This can have several causes. The speaker may want to say *the green light* but may happen to say *the red light.* When the speaker prepares his proposition, mental nodes for the modifier and entity concepts GREEN and TRAFFIC LIGHT are, at some moment, activated. They prime nodes connected to them—for instance, COLOR (from GREEN) and RED (from TRAFFIC LIGHT). They also prime nodes at the next, lexical level. The node for *green* is primed by GREEN and by COLOR. The node for *red* is primed by RED and by COLOR. It may happen that *red*'s priming just exceeds *green*'s at the moment an adjective is needed. In that case *red* will become the current or activated adjective node, though GREEN had been the only "current" concept node of type "modifier." Bottom-up priming may be another cause of erroneous activation.

How is an error detected in the node-structure theory? Through backward priming. When *red* is erroneously activated, it spreads its activation back to the concept node RED, which undergoes a sudden increase of priming. And this precisely is *perception* in the node-structure theory: the bottom-up priming of mental nodes. In this way, MacKay claims, an error is perceived almost immediately, and a corrective action can start before articulation of the error has been initiated. The mechanism of such a corrective action is, as yet, not very well defined in the node-structure theory.

When we compare the editing account and the connectionist account, the most worrisome diagnosis is that, as they stand, both are very hard to disconfirm. We saw that the initial editor theories suffered from reduplication of knowledge. This was repaired in the perceptual-loop account, where the monitoring device is just the speaker's language-understanding system. In MacKay's node-structure theory, even more economy is attained by equating the mental networks for the production and the understanding of language. This is a bold step not followed in the present book. There is, on my view, too much empirical evidence, both experimental and clinical, for a substantial separation of processors involved in the producing and the understanding of language.

Still, there are no doubt intimate connections between the speaking system and the listening system (Cutler 1987). One showpiece is the phenomenon of *delayed auditory feedback*, extensively studied and discussed by MacKay (for a review of this work, see MacKay 1987). The basic phenomenon is this (Lee 1950): A speaker listens to his own speech over earphones. The speech is amplified and delayed. When the delay is of the order of 200 milliseconds, serious dysfluencies arise in the speaker's speech delivery. It becomes a kind of stuttering, with repetitions and delays of speech sounds and drawls in the articulation of vowels. Here speech perception immediately interferes with speech production, and this wouldn't be possible without short-circuited connections between the two systems.

MacKay's account of this effect is, in essence, this: It is an essential property of the node-structure theory that an activated node is completely inhibited shortly after its activation. As a consequence, it will not be the most active node at the next moment in time. This opens the way for another node in the same domain to become the next activated or current node. The phase of inhibition is, however, short. Nodes representing phonemes or distinctive features can be reactivated some 150 or 200 msec after inhibition. And that is what happens systematically in a situation of delayed auditory feedback. The sensory nodes (figure 12.4) that react to the speech that is fed back prime precisely the same mental nodes that had been active some 200 msec earlier. This priming is substantial because of the amplification of the speech. In many cases, the same node will then again be selected as the current node, because it has again become the most primed of its kind. This then leads to repetition of the same motor activity–i.e., to stuttering.

Other things are more problematic for the node-structure theory's account of self-monitoring. The theory implies that every error is detected almost immediately; bottom-up priming is always direct and automatic. This, however, is not in agreement with the findings in Levelt 1983 that error detection can be substantially delayed. This will be the first issue taken up in the next section.

Also difficult to handle are some of the set effects Motley and his colleagues established. One of the hardest is the finding by Baars, Motley, and MacKay (1975) that the lexical-bias effect disappeared when only nonwords were used in the error-eliciting procedure (see subsection 12.1.2). There is as much bottom-up priming of lexical nodes in the latter case as when real words are used, but no lexical bias arises. Self-monitoring may not be as automatic as the node-structure theory suggests. A subject may have reasons to bother about lexical status, or not to bother. And the same holds for his attitude with respect to other features of his speech.

12.2 Interrupting and the Use of Editing Expressions

When a speaker detects trouble that is sufficiently alarming according to the speaker's current standards, the decision will be taken to interrupt speech and to initiate a repair. There are two issues to be considered here. The first one is the temporal relation between detection and interruption: Does a speaker complete certain parts of speech before halting, or is interruption immediate? The second is whether and how the speaker signals to the addressee that trouble is at hand and that some repair is about to be made.

12.2.1 Interrupting the Utterance

The present evidence on spontaneous self-interruptions allows us to maintain the following *Main Interruption Rule.* (One minor but interesting exception to the rule will be discussed below.)

Main Interruption Rule:
Stop the flow of speech immediately upon detecting trouble.

This rule was first suggested and discussed by Nooteboom (1980) in his analysis of the repairs in the Meringer (1908) corpus. A detailed empirical analysis of the rule on the basis of almost 1,000 tape-recorded spontaneous self-repairs in the visual-pattern descriptions discussed above was presented in Levelt 1983. Some of the main findings of that analysis will be summarized here.

There is a variable distance between the troublesome item and the point of self-interruption. The speaker may discover trouble and interrupt himself before the trouble item is uttered. That is probably the case in a *covert* repair, such as the following:

(19) Here is a – er a vertical line

We do not know in this case what the troublesome item was. Maybe the speaker was about to say *horizontal.* At any rate, there was some reason to interrupt and restart. The repair is called "covert" because we don't know what was being repaired; 25 percent of the repairs in the corpus were of this kind. Not knowing the source of trouble, we cannot be sure about the delay between the source of the trouble and the moment of

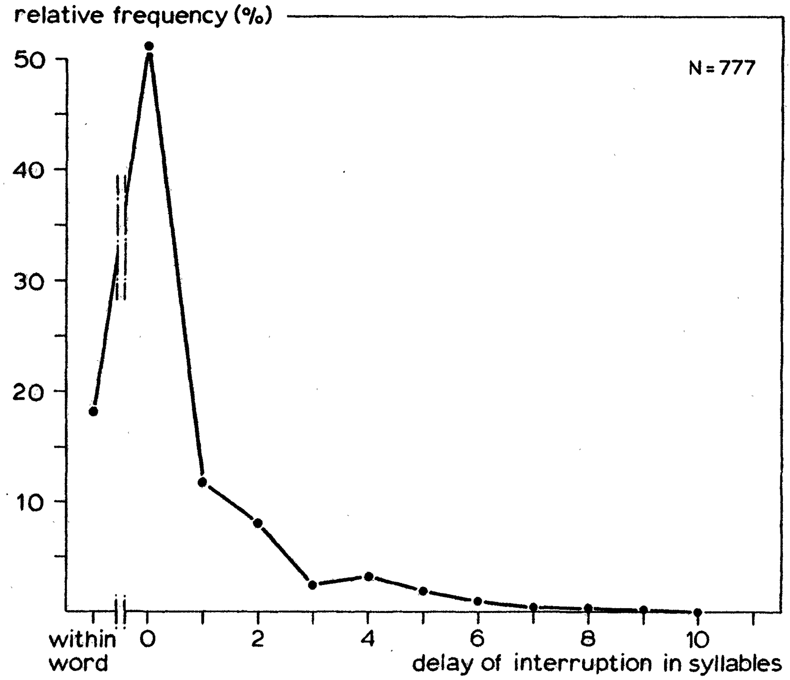

Figure 12.5
Distribution of interruption moments in number of syllables after trouble item. (After Levelt 1983.)

interruption. In the following we will, therefore, ignore these covert repairs and limit ourselves to *overt* repairs. The troublesome items will be italicized.

In overt repairs, interruption can take place during the utterance of the troublesome item, or right after it, or one or more syllables later. Figure 12.5 shows the distribution of interruption moments in the overt repairs of the pattern-description data. If indeed interruption follows on the heels of detection, the curve in figure 12.5 also reflects the distribution of error detection.

Let us begin with the immediate within-word interruptions. The following repair is an example:

(20) We can go straight on to the *ye-*, to the orange node

About 18 percent of the overt repairs in the corpus were of this type. Interruption can also occur just after the troublesome item, as in the following:

(21) Straight on to *green* – to red

This is, in fact, very common. It is the most frequent place of interruption, occurring in 51 percent of all self-repairs.

But interruption can also be delayed by one or more words. This happens in the remaining 31 percent of the overt repairs. In example 22, the troublesome element was *green.*

(22) And from *green* left to pink – er from blue left to pink

Interruption occurred only three words later. An infrequent case of such delayed interruptions (4 percent out of the 31 percent) was one in which the interruption occurred within a word, as in the following:

(23) And *over* the gray ball a pur- or er right of the gray ball a purple ball

The Main Interruption Rule says that, in all these cases, the speaker halted immediately upon becoming aware of trouble. Delays in interruption, therefore, are ascribed to delays in *detection.* An alternative to the rule would be that speakers do not interrupt immediately, but prefer to complete the current word, phrase, clause, or whatever constituent. The evidence does not favor such an alternative. Any smaller or larger constituent can be interrupted for repair. This is the case in example 20, where there is not only interruption of a word but also interruption of a noun phrase; there is no delay. And when there is delay, there need be no constituent completion. Example 23 demonstrates this. There is delay, but neither the integrity of the word nor that of the noun phrase (*a purple ball*) is respected. Delayed interruptions with incomplete constituents are quite common in spontaneous self-repairs.

Still, one could make a statistical argument. Speakers may have *some* tendency to delay interruption till after completion of the current phrase. If there is such a tendency, delayed interruptions in repairs will respect phrase boundaries more often than immediate interruptions. However, the reverse turned out to be the case. Interruptions immediately after the troublesome item respected phrase boundaries in 74 percent of the cases, whereas delayed interruptions coincided with phrase boundaries in 66 percent of the self-interruptions. Hence, there is no reason to suppose that a speaker delays interruption in order to complete the current phrase.

But do speakers respect the integrity of *words* when they interrupt? Nooteboom (1980), in his analysis of the Meringer data, found that words are almost never broken up; speakers tend to delay interruption until at least the current word has been completed. It should be remembered, however, that Meringer was interested in speech errors rather than in the way speakers self-interrupt. It is, in particular, very hard for even an attentive listener to register exactly where a speaker interrupts his own speech. The repairs in my pattern-description data were all tape-recorded; this made it possible to determine precisely where the self-interruption took place. It turned out that 22 percent of interruptions occurred within words. Words are not sacred cows in self-repair.

Still there is something special about within-word interruptions. Words that are real errors are much more prone to interruption than words that are merely not fully appropriate or words that are correct. Compare the following three examples:

(24) First a *bro-* – er a yellow and a green disk

(25) To the left of it a *blanc*, or a white crossing point

(26) *Left* to the pink disk, or right to the pink disk

In example 24, the word *brown* is wrong and the speaker interrupts it for repair. In example 25, the word *blanc* is not an error. It is, rather, somewhat inappropriate because the speaker always used *white* to describe white nodes. Note that the speaker does not interrupt *blanc*. Erroneous words were interrupted more than three times as often in the pattern-description data as merely inappropriate words. Example 26 is still different. There is an error word (*left*), but the speaker was late in detecting the trouble. If detection occurred when the word *disk* was being uttered, the speaker should have interrupted that word, according to the Main Interruption Rule. The data show, however, that words are seldom broken up in such cases of delayed interruption. The bulk of within-word interruptions are cases where the broken-up word is an error itself.

Can this be adduced to the *detection* of trouble? Hardly. There is no reason to assume that the detection of error (as in example 24) occurs more frequently within the troublesome word than the detection of inappropriateness (as in example 25). And it is even less likely that when detection of trouble is late, as in example 26, it won't occur within a word. That is, we have a real exception to the Main Interruption Rule: Words that are not errors themselves tend to be completed before interruption. One would be inclined to a pragmatic interpretation of this finding. By interrupting a word, a speaker signals to the addressee that that word is an error. If a word is completed, the speaker intends the listener to interpret it as correctly delivered.

Are syllable boundaries respected when a speaker halts within a word? Compare the following repairs:

(27) Then follows a *horizon-*, no, a vertical line

(28) Over the *gree-*, no I am wrong, left of the green disk . . .

(29) I go *f-*, all the way straight

In example 27, the word is broken up but the syllable boundary is respected. In examples 28 and 29, syllable boundaries are not respected; this is quite common in self-interruptions. But there is a difference between the latter two examples. In example 28 the pronounced part of the word is itself a *possible* syllable. In that sense the interruption respects the phonotactic constraints of the language. This is not so for example 29, where the speaker produced just a fricative (probably the first consonant of *further*) that was not a possible syllable. A careful analysis by ear of all 172 within-word interruptions in the data revealed that 39 percent were of this latter kind, i.e., the production of a sound that was not a well-formed syllable in the language. When speakers interrupt an erroneous word, they are quite inconsiderate with respect to its phonological integrity. Here the Main Interruption Rule—stop as soon as trouble is detected—appears to be respected.

The Main Interruption Rule, in short, can be maintained with one exception: Speakers tend to complete the current *correct* word upon detection of trouble. An important consequence of the rule is that delayed interruption signals delayed detection of trouble. But delayed detection of trouble cannot be accounted for by the node-structure theory as it stands.

12.2.2 The Use of Editing Expressions

The speaker's self-interruption is usually followed by a short pause or by what is called an *editing expression* (Hockett 1967). Editing expressions—simple ones such as *er*, *that is*, *sorry*, *I mean* and elaborate ones such as *oh, that's impossible; I will start again, OK?*—play a significant role in signaling to the addressee that there is trouble, and what kind of trouble it is. James (1972, 1973), for instance, analyzed the uses of *uh*, *oh*, and *ah*, and how they differ semantically. The interjection *uh* or *er* signals, according to James, that something has been temporarily forgotten. DuBois (1974) argued that the phrase *that is* is used to specify a referent in a repair, as in the following:

(30) He hit Mary – that is, Bill did.

The use of *rather*, according to DuBois, is for "nuance editing," as in

(31) I am trying to lease, or rather, sublease my apartment.

And *I mean* is used when there is an out-and-out mistake:

(32) I really like to – I mean – hate to get up in the morning.

These are constructed cases, not observed ones, but they strongly suggest that there are special editing expressions for signaling straight mistakes or errors (as in example 32) and others for signaling appropriateness problems (as in examples 30 and 31). What do the repair data tell us?

It is indeed the case that a speaker's use of editing expressions is quite different after an error than after a mere inappropriateness. The most remarkable difference in my data (Levelt 1983) is that errors were followed by editing expressions in 62 percent of the cases, whereas for appropriateness repairs the percentage was only 28. And this makes communicative sense. In the case of an error (as in example 32), the speaker may want to warn the addressee that the current message is to be replaced. But in the case of an inappropriateness (for instance, an underspecification, as in examples 30 and 31), no such drastic steps are imminent. The addressee will only be given an additional specification. This distinction was already apparent in the speaker's tendency to interrupt erroneous words but not inappropriate words.

The repair data also showed a systematic difference in the kinds of editing expressions used. Errors were mostly followed by *er*, *or*, and *no*. The latter two tell the addressee that an alternative is about to come, and that something is being rejected. Such oppositions are not often created in appropriateness repairs. The most frequent appropriateness-editing term in the Dutch data was *dus* (literally *thus*)—a connective that normally presupposes the correctness of the previous propositions and introduces some consequence or state of affairs that is compatible with it. In other words, the speaker self-interrupts but tells the listener to continue the same line of interpretation. This editing term was never used when the occasion for repair was an error. Also, speakers excused themselves (by *sorry* or the like) after errors, but almost never after cases of inappropriateness.

A very special editing term is *er*. It is the most frequently used editing expression, used in 30 percent of all repairs. It is also the only editing expression that is practically universal; it exists, with only minor phonetic variations, in many if not all languages. The latter should make one suspicious; *er* would be the only universal word. But is *er* a word, or is it rather a neutral sound that is likely to occur under certain speaking conditions?

Analysis of the pattern-description repair data reveals that there is one dominant factor determining the use of *er*: the actuality or recency of trouble. The most frequent use of *er* was in covert repairs, such as

(33) We start with a green – er – green point.

Here trouble is clearly present, though we cannot tell precisely what kind of trouble. Of the covert repairs in the corpus, two-thirds contained an editing expression. Of these, *er* was by far the most frequent one, occurring in 84 percent of the cases. The next most frequent use of *er* was right upon the interruption of a word, as in

(34) Right to *lef* – er to yellow.

If an editing term was used in these cases, it was *er* in 80 percent of the cases. The use of *er* declined when the delay between the trouble spot and the interruption increased. When the troublesome word has just been completed, as in example 35, *er* was no longer the most frequent editing term; still, 35 percent of the editing expressions were *er*.

(35) Left – er – right in front of me

When the interruption was delayed by one or more words, as in example 36, the use of *er* dropped to 21 percent.

(36) From white I go *straight* to – er – right to blue.

Another way of expressing the same result is that the difference between the use of *er* and the use of other editing expressions is that *er* is used when the delay between the troublesome word and the interruption is short (in fact, the average delay for *er* is 1.7 syllables); the other expressions are used at larger delays (average: 4.3 syllables). The interjection *er* apparently signals that at the moment when trouble is detected, the source of the trouble is still actual or quite recent. But otherwise, *er* doesn't seem to mean anything. It is a symptom, not a sign.

12.3 Making the Repair

After self-interruption there is a moment of suspense. The speaker's *er* may help him to keep the floor, and various editing expressions may be used for the same purpose or to indicate the kind of trouble at hand. This moment of suspense will be used to prepare the correction. It is, at the same time, a moment of trouble for the listener. The running interpretation of the flow of speech is suddenly interrupted, and the listener has a *continuation problem*. How much of the interrupted utterance is to be reinterpreted or qualified? Or is it just a moment of hesitation after which speech will continue normally?

It will be argued in this section that speakers make repairs in such a way that their listeners' continuation problems are maximally accommodated. Speakers repair, first, in a linguistically principled way. The repair entertains a simple and systematic syntactic relation to the interrupted utterance. This makes it possible for the listener to derive an interpretation for the *pair* of original utterance and correction. Second, speakers include more or less of the original utterance in their correction, which is an important way of letting the listener know whether the occasion for repair is an error or an inappropriateness. Third, the very first word of the repair is usually enough to allow the listener to determine where the repair is to be inserted in the original utterance. Fourth, the speaker can focus the replacement by assigning it pitch accent. The next four subsections will deal with these four systematic properties of repairs. In a final section some attention will be given to less conventional ways of self-repairing.

12.3.1 The Syntactic Structure of Repairs

It has often been observed that speakers restart their speech at a phrase boundary or a constituent boundary (DuBois 1974; Nooteboom 1980). This is the *constituent rule*. The rule is not only correct, it is also trivial. A check of the 957 repairs in my corpus (Levelt 1983) shows that there are only minor exceptions, but a further analysis makes it clear that this adherence to the rule is due to the language rather than to the way speakers repair. In a right-branching language, such as English or Dutch, almost any word in a sentence marks the beginning of a syntactic constituent. This is exemplified in the following.

(37) to the red node
  1 2 3 4 5

If a repair were to start at position 1 in example 37, it would initiate a prepositional phrase. Beginning at position 2 or 3, it would introduce a noun phrase. Whether position 4 is also to be taken as the beginning of a noun phrase depends on one's theory. The important thing to observe is that when syntactic branching predominantly proceeds to the right, as in example 37, new phrases are started in subsequent positions. Only phrase *endings* coincide. In example 37, all phrases end in position 5. In the pattern descriptions, 89 percent of all the occurring words were phrase initial. It is, in other words, quite hard for a speaker to violate the constituent rule while making a repair.

What is worse, the rule does not guarantee that the repair is well formed. Just as we have intuitions about the well-formedness of sentences, we can have rather strong feelings about whether a repair "fits" or doesn't "fit." The following (constructed) example may help to demonstrate this point:

(38) Is the doctor seeing – er – the doctor interviewing patients?

This repair sounds rather funny; it is intuitively ill formed. Still, it respects the constituent rule; the repair starts with a noun phrase (*the doctor*). Is there something in the repair itself (*the doctor interviewing patients*) that makes it a problematic case? No; the same repair can be fine in the following context:

(39) Is the nurse – er – the doctor interviewing patients?

This repair is intuitively well formed. Well-formedness of a repair is, apparently, not a property of its intrinsic syntactic structure. It is, rather, dependent on its relation to the (interrupted) original utterance. The constituent rule ignores that relation.

What sort of relation is a repair to maintain with the original utterance? The relation is, in essence, quite simple. Syntactically speaking, an utterance and its repair constitute a kind of *coordination* (Levelt 1983; De Smedt and Kempen 1987), and the syntactic rules of coordination have to be followed. This state of affairs can be captured in a *Well-Formedness Rule* for repairs. The rule is somewhat abstract on first appearance, but is easily explained. In the rule, O means the original utterance, R means the repair proper, and C is a string of zero or more words that is to complete the original utterance. The rule ignores the use of editing expressions.

Well-Formedness Rule for repairs

An original utterance plus repair ⟨OR⟩ is well formed if and only if there is a string C such that the string ⟨OC*or*R⟩ is well formed, where C is a completion of the constituent directly dominating the last element of O (*or* is to be deleted if that last element is itself a connective such as *or* or *and*).

To demonstrate this rule, we will apply it to the just-discussed ill-formed and well-formed cases.

In the ill-formed repair given as example 38, O, the original utterance, is *Is the doctor seeing*. The interrupted constituent in the original utterance is the verb phrase (*seeing* . . .). It can, for instance, be completed by an NP such as *the surgeon*; let us take this as C, the completion. Then follow *or* and the repair R, *the doctor interviewing patients*. The result is this:

(40) Is the doctor seeing the surgeon or the doctor interviewing patients?

This coordination is as ill formed as the repair in example 38.

What about the well-formed case (example 39)? There O is *Is the nurse.* Its last word is *nurse*, which itself completes the NP constituent, so the completion C can be empty. R is the repair *the doctor interviewing patients,* and the coordination specified by the rule will be the following:

(41) Is the nurse or the doctor interviewing patients?

This is all right as a coordination. For these examples, therefore, the biconditional rule appears to work.

How does it work for naturalistic data? An analysis of the pattern-description repairs showed that the rule was adhered to with only minor exceptions—i.e., speakers usually made repairs that were syntactically well formed. Let us apply it to the following observed repair:

(42) From purple up to, straight up to red

This repair is, according to the rule, syntactically well formed because the corresponding coordination is. O is *from purple up to.* Its last word (*to*) introduces a prepositional phrase; let us complete it with *green* (= C). R is *straight up to red.* The coordination thus becomes

(43) From purple up to green or straight up to red

which is well formed.

So far, the conclusion is warranted that, in contrast with what has been suggested in the linguistic literature about repairs, self-repair is a syntactically regular process. In order to repair, the speaker tends to follow the normal rules of syntactic coordination.

This does not mean, of course, that the Well-Formedness Rule cannot be violated in a speaker's repair behavior. The study of speech errors has made it abundantly clear that just about any linguistic rule is occasionally violated by speakers. Such irregularities are especially likely to arise when there are attentional lapses, or when there are high processing demands (such as in fast speech). Repair situations are almost always "loading" moments for a speaker. It is, therefore, quite surprising how *regular* speakers' repairs usually are.

But just as one can experimentally elicit speech errors, one can also experimentally elicit ill-formed repairs. Van Wijk and Kempen (1987) designed an inventive experimental procedure to induce repairs in syntactically well-controlled sentences. The speaker's task was to describe an event depicted in a picture. An example of such a picture is given in figure 12.6a. It can be described as follows:

(44) The bald man with the spectacles pushes the sad clown.

The subject was first made familiar with the possible agents and re-

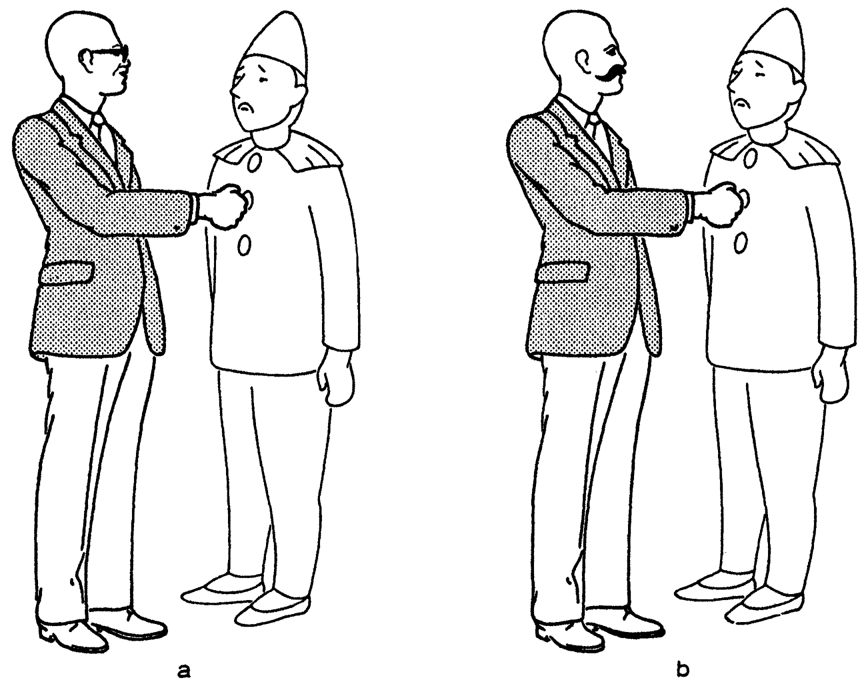

Figure 12.6
Pictures used in van Wijk and Kempen's (1987) repair elicitation experiment.

cipients, and with the kinds of actions depicted in the pictures. They were introduced as *the man*, *the clown*, etc., so that no naming problems would arise during the experiment. In the main experiment, the protagonists in a picture did not appear simultaneously, but one after the other. The picture was built up over time. If the man appeared first in the picture of figure 12.6a, the subjects usually started making an active sentence. When the clown appeared first, followed by the pushing man, the sentence being uttered usually became a passive one (*The sad clown is pushed by the bald man with the spectacles*). In this way different syntactic constructions could be induced. The important variable, however, was this: At some moment during the speaker's utterance the picture was changed. Some attribute of the one protagonist or the other, or the action depicted, was altered. The man with spectacles, for instance, suddenly became a man with a moustache, as in figure 12.6b. When the change came some time after the speaker had uttered *with the spectacles*, he normally halted and repaired to replace *spectacles* by *moustache*. A typical repair was the following:

(45) The bald man with the spectacles – er with the moustache pushes the sad clown.

Clearly, example 45 adheres to the Well-Formedness Rule; the corresponding coordination would be *The bald man with the spectacles or with the moustache pushes the sad clown*, which is fine. But van Wijk and Kempen developed this technique especially to induce very delayed repairs. When the pictorial change occurred late in the speaker's utterance, ill-formed repairs might occur, such as the following:

(46) The bald man with the spectacles pushes – er – with the moustache pushes the sad clown.

This is ill formed because the corresponding coordination would be something like *The bald man with the spectacles pushes a child or with the moustache pushes the sad clown*, which is horrible. And indeed cases like example 46 do arise regularly under such experimental conditions.

On the basis of these and related experimental results, van Wijk and Kempen conjectured that speakers have, essentially, two ways of making a repair. The first one they call a *reformulation*; the speaker creates a new syntactic construction. An instance of this is example 42 above, where *up* is reformulated as *straight up*, which involves a syntactic addition. These repairs systematically adhere to the Well-Formedness Rule. Reformulations occur at the level of grammatical encoding, in the realm of syntax.

The second repair strategy is *lemma substitution*. A lemma substitution fully preserves the syntax of the original utterance; only a word is replaced. There is no further grammatical re-encoding. Examples 45 and 46 are of this kind. According to van Wijk and Kempen, they may but need not adhere to the Well-Formedness Rule. And indeed, two such cases also appeared in my naturalistic data: example 47 (presented here) and example 56 (below).

(47) right of *purple* is – er – of white is purple

The Well-Formedness Rule creates a coordination like *right of purple is green or of white is purple*, which is ill formed. Further scrutiny of naturalistic data is needed to substantiate the systematic occurrence of such repairs. And if they occur more than occasionally, one should check whether they are experienced as well formed by speakers of the language; only if they are need a subsidiary well-formedness rule be added to the grammar of repair.

12.3.2 Ways of Restarting

One can adhere to the Well-Formedness Rule and still restart in very different ways after self-interruption. The repair may or may not include parts of the original utterance, and these parts can be smaller or larger. A very major determinant of how speakers restart is the *repair occasion*, i.e., the type of trouble the speaker is dealing with. Earlier we distinguished between two major classes of trouble: out-and-out *error* and *inappropriateness*. In the former case the repair must undo the mistake; in the latter case some further specification is needed in order to say the same thing in a more felicitous way. The way in which speakers begin their repairs is systematically different for these two kinds of repair occasion. To demonstrate this, three ways of restarting should be distinguished.

The first one will be called *instant repairing*. In these cases there is a single troublesome word, and the speaker retraces to just that word and replaces it with a new item. In other words, the repair begins with the replacement, as in example 48.

(48) Again left to the same *blanc* crossing point – white crossing point

Here the very first word of the repair, *white*, replaces the troublesome item, *blanc*. Notice that, in this example, there is delayed interruption (till after *crossing point*). Instant repairs do not require immediate interruption. Of all the error and appropriateness repairs in my corpus, 42 percent are instant repairs.

The second way of restarting is to retrace to and repeat some word prior to the trouble element. This is called *anticipatory retracing*. An example is the following:

(49) And left to the *purple* crossing point – to the red crossing point

The troublesome item is *purple*, but after interruption the speaker retraces to the beginning of the prepositional phrase, thus "leading in" the replacement of *purple* by *red*. This happens in 35 percent of the error and appropriateness repairs.

The third way of restarting is by making a *fresh start*. The speaker neither instantly replaces the trouble element nor retraces to an earlier word. Rather, he starts with fresh material that was not part of the original interrupted utterance, as in example 50.

(50) From yellow down to *brown* – no – that's red

The speaker says neither just *red* or *to red*; he starts with a new construction: *that's*.... Fresh starts appear in 23 percent of the error and appropriateness repairs in the repair corpus. A special kind of fresh start is one in which the speaker starts with fresh material but then still copies part of the original utterance; the following is an example:

(51) The road begins with a – or it goes on and it begins with a green crossing.

Table 12.1
Ways of restarting for appropriateness repairs and error repairs. (Data from Levelt 1983.)

	Instant repairs	Anticipatory retracings	Fresh starts
Appropriateness repairs ($N = 290$)	30%	25%	44%
Error repairs ($N = 399$)	51%	41%	8%

After the editing expression *or*, the speaker introduces fresh material (*it goes on and it*), then repeats part of the original utterance (*begins with a*). This special kind of fresh start is called a *prespecification.*

How does a speaker restart when he has made an error, and how when the trouble has been an inappropriateness? Table 12.1 gives the results for these two classes of repair occasions in the pattern-description data.

The table shows a considerable difference in the ways of restarting for the two categories of repair occasion. Repairs of errors are usually made either instantly or by anticipation. It is exceptional to make a fresh start after a mistake; example 50 is such a case. It should be added that three-fourths of the fresh starts in the category of appropriateness repairs are prespecifications, such as in example 51. Such prespecifications hardly ever occur among error repairs.

The upshot of these findings is that most error repairs leave the original utterance unaffected but for the erroneous element. Nothing is changed or added that is not strictly necessary. Repairs of all-out mistakes are *conservative.* When, however, the trouble is an inappropriateness, such as an ambiguity or an underspecification, the speaker tends to introduce fresh materials. Either he will say the same thing differently by creating a new sentence or phrase, or he will make a prespecification (which involves the insertion of fresh syntactic materials into the original utterance). The latter is, in fact, the dominant procedure for appropriateness repairs. Insertions of fresh materials almost always follow the Well-Formedness Rule. This also holds for van Wijk and Kempen's (1987) experimental data. A different way to put this is that in cases of inappropriateness speakers tend to *reformulate* what they had said. In cases of error they *preserve* their original formulation and only exchange the troublesome element.

One might expect that the way of restarting is also influenced by the amount of delay before interruption. It would, for instance, make sense to give an instant repair after an immediate interruption; the troublesome word would then immediately be followed by its correction. Detailed analyses of the repair data showed, however, that there is no such systematic dependency. The way of restarting is by and large independent of where the original utterance was interrupted.

12.3.3 Restarting and the Listener's Continuation Problem

There is another aspect of restarting that deserves particular attention. Not only the speaker but also the listener is in trouble when an utterance is interrupted. It is well known that listeners interpret speech "on-line" as much as possible (see, for instance, Marslen-Wilson and Tyler 1981). That is, listeners tend to syntactically, semantically, and pragmatically integrate each new word of an utterance into whatever they have understood so far. This integration is done very quickly, even *during* the delivery of the word. When an utterance is interrupted by the speaker, the listener's problem is how to insert the new repair information into the developing interpretation of the original utterance.

It would be ideal if the listener could know how to do this as soon as the first word of the repair appears. In particular, a repair's first word (after whatever editing terms) can, first, be the instant replacement for a trouble item, as in example 48. In that case the listener would want to know that this is the case, as well as which word in the original utterance is to be replaced. The first word can also be the beginning of an anticipatory retracing. How can the listener know that this is so, and which element in the original utterance is the target of retracing? A third alternative is for the first word to be the beginning of a fresh start. That is important to know, because the listener will then interpret it as an addition rather than as a change. Finally, the interruption may be due to a covert repair. In that case the first word after the restart is either just the next word of the current sentence, as in example 52, or an item overlapping with the original utterance, as in examples 19, 33, and 53. For the listener, the latter case is like an anticipatory retracing.

(52) Left of the black dot – er – which was already mentioned

(53) A vertical line to a – to a black disk

Does the way in which speakers restart reveal to their listeners right away which of these four cases is at hand? The answer is, by and large, Yes. In particular, speakers adhere to two conventions of restarting, which will now be discussed. Both conventions relate the first word of the repair to some earlier word in the original utterance.

The Word-Identity Convention

If the first word of the repair is identical to some word *w* of the original utterance, the repair is to be interpreted as a continuation of the original utterance from *w* on. (If there is more than one such word in the original utterance, take the last one.)

The idea of this convention is easily exemplified. Case 54 adheres to the convention; case 55 doesn't.

(54) Right to *yellow* – er – to white

Here the first repair word, *to*, also appears in the interrupted utterance. The intended interpretation, therefore, is *right to white*. The continuation of interpretation means, in particular, that the local syntax at the point of insertion is acknowledged. In the case of

(55) And at the bottom of the *line* a red dot – a vertical line

the convention prescribes the interpretation *and at the bottom of the line a vertical line*, and that is not what the speaker intended. Before discussing whether violating cases such as 55 are just exceptional or rather quite frequent, I will introduce the second convention, which applies to cases where the first word of the repair differs from all the words in the original utterance.

The Category-Identity Convention

If the syntactic category of the first word of the repair is identical to the syntactic category of some word *w* of the original utterance, the repair is to be interpreted as a continuation from *w* on, with the first repair word replacing *w*. (If there is more than one such word in the original utterance, take the last one.)

The latter convention is followed in example 56 but violated in example 57.

(56) Down from white is a *red* node and from – pink node

By the convention, the interpretation should be *down from white is a pink node*, since the adjective *pink* is to replace the adjective *red* in the original utterance. This is indeed the intended interpretation.

(57) And then again a *stretch* to a red patch – stripe

Here the convention prescribes the interpretation *and then again a stretch* *to a red stripe*, the noun *stripe* replacing the last noun (*patch*) in the original utterance. This was not intended by the speaker; *stretch* was to be replaced.

Do speakers follow these conventions? Examples 55 and 57 are real exceptions in the data. The Word-Identity Convention was violated in 13 percent of the cases in which it should have applied. In most of these cases not the last but an earlier identical word had to be taken. The Category-Identity Convention was violated in 9 percent of the cases, and again in most of these cases an earlier category-identical word had to be taken (as in example 57). In quite a number of cases, the Word-Identity Convention was violated but the Category-Identity Convention was respected. In fact, in no more than 7 percent of the cases did speakers violate both constraints at the same time. In other words, in a large majority of cases speakers begin their repair in such a way that the listener can immediately insert the repair in the right slot of the original utterance. This guarantees unhampered on-line interpretation of utterance-plus-correction. And the repair's prosody reveals this fact. When the "bridged" parts in repairs 54–57 are spliced out of the tape, normal fluent speech arises (Levelt 1984b).

The reader may have noticed the similarity of the two conventions. The Word-Identity Convention is probably a special case of the Category-Identity Convention. However, it makes psychological sense to keep them apart. In the case of word identity, the listener knows that the repair is going to be an anticipatory retracing—there will be a lead-in to the correction, and a "leisurely" interpretation is possible. In case of mere category identity, there is instant replacement—which probably requires more alertness on the part of the listener.

The above conventions do not solve all of the listener's continuation problems. If the conditions for neither convention are met, the repair can be either a fresh start or a covert repair of the type shown in example 52. In the former case a new syntactic construction is at hand; in the latter case the syntax of the utterance is simply continued. This is an important difference in on-line interpretation. The problem, however, is smaller than it appears to be. In the large majority of repairs the listener can determine which is the case on the basis of the editing term used. In 92 percent of all covert repairs (i.e. hesitations) of this sort, speakers used *er* as an editing term, whereas only 6 percent of the fresh starts were introduced by *er*. Two-thirds of the fresh starts had no editing term at all, which shouldn't surprise us: Fresh starts are the mark of appropriateness repairs, and in most cases appropriateness repairs go without editing expressions.

In conclusion, therefore, it can be said that, with a low rate of exceptions, speakers solve their addressees' continuation problem when a repair is at hand. No later than upon the first word of the repair the listener can know how to relate the repair to the interrupted utterance. Whether the listener in fact uses that information when dealing with the speaker's trouble is another issue; it is not a concern of this book.

12.3.4 Prosodic Marking in Self-Repair

Following a suggestion by Goffman (1981), Cutler (1983b) drew a distinction between repairs that are prosodically marked and those that are unmarked. A repair is prosodically unmarked when "the speaker utters the correction on, as far as possible, the same pitch as the originally uttered error [or trouble item]." Also, in unmarked repairs, the amplitude and the relative duration of the repair mimic those of the trouble item. A repair is prosodically marked when the prosody of the trouble item and its correction differ substantially. Marking can be done by creating a contrast in pitch contour, in loudness, or duration. Cutler found that repairs of phonetic errors are, as a rule, unmarked. For instance, for the repair

(58) Well it'll all have to be *unsiled* – unsigned

the speaker does not make a prosodic contrast between *unsigned* and the erroneous *unsiled*. Marking typically occurs in lexical repairs, such as

(59) You're *happy* to – wElcome to include it

(where pitch accent is, as usual, indicated by capitalizing the vowel).

Still, in Cutler's corpus only 38 percent of the lexical repairs were prosodically marked. What, then, governs a speaker's decision to mark a repair? To deal with this question, we (Levelt and Cutler 1983) analyzed all of the tape-recorded lexical repairs in the pattern-description corpus. There were 299 such repairs for which the independent markedness judgments of the two authors agreed. Of these, 134 (45 percent) were marked.

A first analysis concerned the interrupt/restart pattern of the repair. Did a speaker mark the replacing word more often when interruption was immediate, i.e., within or right after the trouble word? Or did a speaker do more marking in instant repairs than in anticipatory retracings and fresh starts? It turned out that none of these factors had anything to do with prosodic marking in repairs. In other words, prosodic marking in no way reflects the *syntax* of the repair operation.

Rather, repair prosody is *semantically* motivated. This appears, first, from a comparison of error and appropriateness repairs. Of the error repairs, 53 percent were prosodically marked, whereas only 19 percent of the appropriateness repairs were. It should be remembered that repairs of errors are made to *reject* what was said earlier. This, we saw, is also apparent from the use of editing terms such as *no*, *or*, and *sorry*, and from the tendency to break off the error word. Prosodic marking *accentuates* this contrast. Repairing for inappropriateness, however, does not involve rejection as much as it involves further specification; no contrast is established to what was said before.

Still, the question should be raised why only 53 percent of these error repairs are marked. Don't they all involve a semantic contrast? Yes, they do, but there are degrees of contrast. The speaker and the listener will often be mutually aware of the set of alternatives for a trouble item. The sense of contrast will be highest when this set is small, as in the case of antonyms (*left* – no, *rIght*). It will be less when the number of alternatives to the trouble word is large.

The repair data we studied made it possible to test this notion of degrees of contrast. Among the lexical corrections there were many color-name repairs. The experimental patterns involved eleven different colors; hence, the number of alternatives to a particular color was ten. There were also a substantial number of errors in directional terms. These almost always exchanged *left* and *right*, *horizontal* and *vertical*, or *up* and *down*—i.e., antonyms. In most of these cases, therefore, there was just one alternative. If the degree of contrast matters, one would predict more prosodic marking in the repairs of directional terms than in the color-name repairs. And indeed, this turned out to be the case. Seventy-two percent of the direction-term repairs were marked, versus 50 percent of the color-name repairs. The degree of semantic contrast, therefore, is an important contributing factor to prosodic marking in self-repairs.

These findings place the origins of prosodic marking in repairs at the message level. The speaker intends to give contrastive prominence. This, in turn, induces prosodic focus in grammatical encoding (see subsection 5.2.2), which becomes phonologically realized as pitch accent.

In addition to this genealogy of prosodic marking in repairs, there may well be a personal stylistic factor. Certain speakers became quite upset with themselves when they made a mistake; they would, so to say, cry out the corrections. One speaker even marked every single repair. Others preferred to make their repairs more or less in passing, as if the listener rather than the speaker was to blame for the misunderstanding. What would you do, dear reader?

12.3.5 Repairing on the Fly

There are many more ways of repairing than those reviewed above. Actually, the boundaries between repairing and nonrepairing are quite fuzzy. There are parentheticals and expansions that come, sort of, as afterthoughts. One wouldn't call the following a repair:

(60) He conquered Babylon, the great Alexander.

It is, clearly, a well-formed utterance. If it were a repair, the Well-Formedness Rule would classify it as ill formed. (*He conquered Babylon and the great Alexander* is ill formed if *he* and *the great Alexander* are coreferential.) But other cases are less clear. What about the following?

(61) That's the only thing he does is fight. (from Kroch and Hindle 1982)

(62) It seems to be a good marriage, of her parents. (C. van Wijk, personal communication)

(63) Who did you think else would come? (from Garrett 1980a)

In all these cases the speaker repairs (or seems to repair) something on the fly, without an explicit stop and restart operation (see van Wijk and Kempen 1987 for a review of such cases). Speakers are apparently willing to stretch or even distort their syntax to cope with local trouble, just to maintain fluent delivery of their speech.

How universal are the mechanisms of self-repair that we have reviewed in this chapter? Though the present base of data and analyses is still fairly limited, there is reason to suppose that the organization of repair is quite invariant across languages and cultures. This is, in particular, argued by Schegloff (1987) on the basis of a comparison of repair behavior in Tuvaluan (South Pacific), Quiche (Guatemala), and Thai. According to Schegloff these universal patterns of conduct proceed from "plausibly generic organizational contingencies of interaction."

Summary

Speakers attend to what they are saying and how they say it. They can monitor almost any aspect of their speech, ranging from content to syntax to the choice of words to properties of phonological form and articulation. But they do not continuously attend to all these things simultaneously. Attention is on the one hand selective, and on the other hand fluctuating. Which aspects of speech are attended to is highly dependent on the context and on the task. A speaker can be set to attend to certain kinds of errors or dysfluencies, and to ignore others. Also, the detection rate fluctuates with the developing phrase structure of the current utterance; monitoring is more intense at ends of phrases.

There are, essentially, two classes of monitoring theories: the editing theories and the connectionist theories. The editing theories put the monitor outside the language-production system; the connectionist accounts make it internal to the system. The most parsimonious version of the former type is the perceptual-loop theory, which identifies the monitor with the language user's speech-understanding system. According to this theory, the speaker can monitor both his internal speech and his overt speech (via the internal and the external perceptual loop, respectively). But there is no monitoring access to intermediary results of formulating. The connectionist type of account is most specifically worked out in MacKay's node-structure theory. It assumes that the systems for language production and language understanding are largely coincident networks of connected nodes. Error detection is due to bottom-up priming within this network. Both kinds of theories have their strengths and weaknesses, and both are hard to disconfirm.

After noticing some flaw in content or in well-formedness, a speaker may interrupt himself and begin a repair operation. There is evidence that self-interruption follows immediately upon the detection of trouble. Speakers do not wait till after the completion of a syntactic, a lexical, or a phonetic unit. There is, however, one exception to this Main Interruption Rule: Speakers respect the integrity of words—at least, of words that are themselves correct. Erroneous words, on the other hand, are frequently broken off in self-repairs.

The moment of suspense after self-interruption is often filled with some editing expression, most frequently *er*. This *er* is a symptom of the actuality of trouble. It appears especially when the troublesome item is immediately recognized by the speaker. Other editing terms tend to reveal the *kind* of trouble at hand. When there is an all-out mistake, editing expressions are used that involve rejection or excuse (*no*, *sorry*). When what was said was merely not fully appropriate, editing expressions are not much used, and when used they tend to stress the continuity of interpretation (*that is*).

When the speaker resumes speech after this interlude, he carefully relates the repair to the interrupted original utterance. There is, first, a syntactic relation between these two. The original utterance and the repair are, essentially, delivered as two conjuncts of a coordination. The syntax of repairing is governed by a rule of well-formedness, which acknowledges this coordinative character of repairs. The rule is usually followed in naturally occurring repairs. But it can, as can any linguistic rule, be violated.

A second observation is that the way of repairing depends on the occasion. In correcting all-out errors, speakers tend to preserve the original syntax in the repair. They can cite parts of the original utterance without change, except of course for the item(s) to be replaced. When the occasion of repair is an inappropriateness, speakers often reformulate, either by inserting fresh materials into the original utterance or by starting with a new utterance. Error repairs are conservative; appropriateness repairs are rather innovative reformulations.

A third property of the way speakers restart their speech is that they give cues about how to relate the repair to the original utterance. When the repair is to be understood as the continuation of the original utterance from some word on, the first repair word reveals which word that is. In most other cases the editing term tells the listener how to solve this continuation problem.

Not all repairs are neatly made as stop-and-restart operations. Speakers can preserve fluency by repairing "on the fly," but usually at the expense of syntactic well-formedness.

It is often the interlocutor who makes the speaker aware of some trouble in his speech, by asking *what?* or *what do you mean?* or by giving nonverbal signals. A speaker not only monitors his own speech directly; he also monitors it indirectly by attending to the interlocutor's reactions.

Appendix

Symbols from the International Phonetic Alphabet, with examples

a	stop, cop	m	man, him
ɑ	balm, rather	n	nose, annoy
æ	fat, placid	ŋ	wing, sink
ʌ	but, flood	oʊ	go, soul
aɪ	eye, kite	ɔ	walk, law
aʊ	owl, how	ɒ	want, astronaut
b	boss, hobo	ɔɪ	boil, voice
tʃ	cembalo, pitch	p	pie, nap
d	do, had	r	rim, parrot
dʒ	wedge, rage	s	simon, boss
eɪ	bait, day	ʃ	ship, facial
ə	among, readily	t	town, walked
ɛ	weather, any	θ	thin, myth
f	felt, left	ð	then, weather
g	go, leg	u	roof, flew
h	hat, how	ʊ	look, full
ɦ	Ohio, ahead	ɜr	bird, worst
i	evil, bee	v	voice, give
ɪ	if, remit	w	why, quail
j	yet, canyon	x	(Spanish) proteger, (Dutch) kachel
k	case, back	z	zeal, haze
l	lip, bill	ʒ	leasure, garage