

PROSODIC MARKING IN SPEECH REPAIR

Willem J.M. Levelt and Anne Cutler

Abstract

Spontaneous self-corrections in speech pose a communication problem; the speaker must make clear to the listener not only that the original utterance was faulty, but where it was faulty and how the fault is to be corrected. Prosodic marking of corrections - making the prosody of the repair noticeably different from that of the original utterance - offers a resource which the speaker can exploit to provide the listener with such information. A corpus of more than 400 spontaneous speech repairs was analysed, and the prosodic characteristics compared with the syntactic and semantic characteristics of each repair. Prosodic marking showed no relationship at all with the syntactic characteristics of repairs. Instead, marking was associated with certain semantic factors: repairs were marked when the original utterance had been actually erroneous, rather than simply less appropriate than the repair; and repairs tended to be marked more often when the set of items encompassing the error and the repair was small rather than when it was large. These findings lend further weight to the characterization of accent as essentially semantic in function.

1. Some determinants of intonational marking in self-corrections

At least two people are in trouble when a speaker interrupts the flow of speech in order to make a self-correction. The first person is the speaker himself* who apparently became aware of some unclarity or error in what he just said. The second person is a listener who is confronted with an abrupt break, and with the task to find out whether what is going to follow is just a continuation, as after a mere hesitation, or whether it is a repair of something said previously. In the latter case, moreover, she has to find out what the reparandum is, and to replace it by the appropriate items in the correction. This will be called the listener's *continuation problem*.

* For ease of reference we will in the following treat the speaker, i.e. the trouble maker, as male and the listener, i.e. the victim, as female.

How does the speaker deal with the trouble he created for himself and for the listener? Different aspects of this question were treated in previous papers by the present authors. Levelt (1983) analysed the different sources of trouble, or occasions for making a correction, and related them to the ways in which the speaker restarts. It appeared that speakers put severe constraints on the ways in which they make the correction. They not only signal to the listener what sort of trouble they had encountered, but they also give unambiguous cues for the listener to solve her continuation problem. The cues analyzed in that paper were of a syntactic and lexical character. Syntactically, it turned out, the original (interrupted) utterance and the repair relate to one another very much like two conjuncts in a coordination. This guarantees semantic interpretability of the repair, given the original utterance. With respect to lexical cues, they play a significant role in relating the first word of the repair proper to a corresponding place in the original utterance, the place where the repair has to be "inserted". The paper, however, did not analyze potential intonational cues, in spite of the fact that the 957 repairs in the corpus were tape recorded.

Cutler (1983), on the other hand, dealt almost exclusively with prosodic aspects of spontaneous self-corrections. Following a suggestion of Goffman (1981), Cutler drew a major distinction between repairs that are prosodically marked and those that are unmarked. In an unmarked repair "the speaker utters the correction on, as far as possible, the same pitch as the originally uttered error" or trouble item. Amplitude and relative duration of the repair item also closely mimic the trouble item. A correction is marked when the prosody of repair item and trouble item differ. Hence, the notion is a relational one; it is not necessarily the case that a high-pitched correction is marked, or that a low-pitched one is unmarked. Marking can be accomplished by a noticeable increase or decrease in pitch, in amplitude, or in relative duration.

Cutler's analyses showed that, in her corpus of repairs, corrections of phonetic errors are always unmarked, only lexical errors are frequently marked. However, even lexical errors are unmarked in 62% of the data. What, then, determines whether a lexical correction will be marked or not?

There are two possible sets of determinants. The first set will be called *syntactic*. These are properties of the repair such as the extent to which the interruption is delayed, and the amount of previously uttered material which is repeated in the repair. Interruptions may occur early, i.e. within the trouble item or immediately after it, as in (1), or they may be delayed by one or more syllables, as in (2):

- (1) Well, let me write it back - er, down, so that ...
- (2) ... what things are this kid - is this kid going to say incorrectly?

Independently of this, after interruption the speaker may instantly introduce the repaired element, as in (1) and (2), or may retrace to an earlier element, as in (3):

- (3) I cannot work out where I ran over - ran across that other name

Prosodic marking may, then, serve as a way in which the speaker can indicate to the listener that he has delayed his interruption, or that he is retracing, or that he is making a fresh start, etc. The listener, in her turn, may use such cues to solve her continuation problem.

The second set of potential determinants for marking can be called *semantic*. Marking could be used by the speaker to express a semantic relation between the repair and the reparandum. The most obvious semantic dimension on which repairs differ is whether or not the trouble item and the repair are compatible or incompatible; that is, was the trouble item actually an error, which must be replaced by a correct version of the intended message, or was the trouble item simply not the most appropriate possible word for the context, so that the repair does not so much replace it as further elaborate upon it? This latter type of repair will be referred to as an *appropriateness* repair; an example is given in (4):

- (4) ... to a dark brown crossing - T-crossing

There were different types of crossing in the domain of discourse, and "T-crossing" is a further specification. There are other forms of correction for appropriateness: a demonstrative can be replaced by a definite description ("from there, from the blue node ..."), a definite article by an indefinite one ("a line to the yellow disc, to a yellow disc"), etc. We might conjecture that the speaker would be more concerned to draw the listener's attention to a repair replacing an error, i.e. completely wiping out the previous version of the utterance, than to a repair which merely elaborates or expands upon the previous version. That is, if marking is a way to signal *rejection*, we would expect corrections for error to be more marked than corrections for appropriateness.

Within the category of error repairs, there is at least one further dimension which might be relevant to the speaker's marking decision, namely the size of the semantic domain in which error and repair contrast. This can be conceived of as being at a minimum when error and repair are antonyms, as in (5):

- (5) Left to green - er, right to green

Other such pairs in our corpus are "horizontal/vertical", "up/down", etc. However, the corpus also includes many cases in which the error is replaced by a word from a more general semantic field, as in (6):

(6) Right of that is green - oh, blue

In this case the speaker was describing patterns consisting of colored nodes, which were connected by either vertical or horizontal black lines. There were 11 different colors involved. It is possible that speaker and listener are mutually aware of the number of alternatives to the trouble item in the domain of discourse, and the larger the number of alternatives, the smaller the sensed degree of opposition, hence the less contrastive it is to mark the repair. We would then expect to find more marking in cases like (5) than in cases like (6).

In summary, there are two semantic dimensions which may be of relevance for the analysis of intonational marking in spontaneous self-corrections: is the intended correction for error or for inappropriateness, and, if it is for error, is the replaced element one of a smaller or a larger set in the domain of discourse?

In her paper, Cutler could not find a systematic relation between marking and syntactic factors. There was, moreover, no clear indication that prosodic marking of lexical repairs was due to semantic determinants. Cutler suggested, however, that analysis of a more extended sample of corrections might reveal effects which could not be discerned in her data. The present paper provides such an analysis. It is based on the sub-sample of 412 lexical corrections in Levelt's repair corpus for which the tape quality was good enough to make a judgment of intonational marking. This sample is indeed large enough to reach more definite conclusions with respect to the determinants of marking in spontaneous self-repairs.

2. *Corpus, judgments of marking, and ways of analysis*

The corpus of self-repairs is extensively described in Levelt (1983), to which the reader is referred. Here it suffices to say that the repairs were obtained in an experiment where each of 53 native adult speakers of Dutch described 53 visual patterns, consisting of colored nodes, connected by black arcs (see above). The average number of repairs per subject was 18.1, with a standard deviation of 10.3. The lexical repairs in the corpus were called "lexical" because the trouble item was a single lexical item. Examples (4), (5) and (6) are English translations of lexical repairs from the corpus.

The two authors of the present paper independently judged each of the 412 lexical repairs for intonational (un)marking. The criterion was as described above: is the prosody of the trouble word roughly the same as the prosody of its correction, or is it different? Here "prosody" refers to pitch, amplitude and duration, since variation in any of these can constitute marking (usually, of course, they vary together). After the judgments were completed, they were compared

between the authors, and it turned out that there was agreement on 299 items, i.e. 73%. This is reasonable, given the fact that one of the authors is not experienced in making prosodic judgments, and the other one is not a native speaker of Dutch. We decided to be ruthless, and to limit the further analyses to the 299 cases where we agreed.

The marking values were added to the (computerized) codes which were already available for these repairs (cf Levelt, *op. cit.*). These involved various syntactic and semantic aspects of the corrections, among them those mentioned in the previous section. It was, finally, easy to compute the distribution of intonational marking for different levels of the hypothesized determinants. The next two sections will discuss the results for syntactic and semantic determinants, respectively.

3. Syntax and marking

3.1 Delay of interruption

Examples (1) and (3) above were cases where the speaker interrupted the flow of speech immediately after the trouble item; examples (4) and (6) are also in this category. There is also an even more immediate way of interrupting, namely immediately within the trouble item. An example is given in (7):

(7) First a brow - er, yellow and a green disc

Here the final /n/ of "brown" (in fact of Dutch "bruin") is not articulated. In all other cases, such as in Example (2) above, the interruption is more or less delayed. Levelt (1983) found that delay of interruption is mainly caused by delay on the part of the speaker of detecting the trouble; interruption occurs, in effect, immediately after detection. Would a speaker mark the correction of the trouble word more after a delayed than after an immediate interruption? The "default" case for the listener could be that it is the last word spoken which needs repair (this is, in fact, the most frequent case in the data). Prosodic marking might help her to consider an earlier element for repair.

Table 1 Intonational marking in repairs with different moments of interruption

moment of interruption	within trouble item	immediately after trouble item	delayed	total
marked correction	26	64	44	134
unmarked correction	23	87	55	165

Table 1 gives the distribution of intonational marking over the three categories of interruption, described above. A chi-square test shows no significant differences between these categories. A slight tendency for corrections after interruption within the trouble item to be more marked than those after interruptions following the trouble item - either immediately or delayed - has an obvious semantic explanation, to which we shall return in the next section. Here, one can safely conclude that speakers do not use intonational marking to tell the listener whether or not the trouble item occurred just before interruption of the flow of speech, or earlier.

3.2 Retracing

There are different ways for a speaker to restart after interruption. Examples (1) and (2) above were cases where the speaker introduced a replacement for the trouble item instantly, as the first word of the correction. The same is true for examples (4) through (7). Example (3) was a case where the speaker restarted at an element which preceded the trouble item in the original interrupted utterance. Such retracings are quite frequent in the corpus; another example is (8):

(8) ... and it ends then in a black - rather, in a purple ball

Here the speaker prepares for the trouble element ("black") by retracing to the beginning of the prepositional phrase in which it occurred. There are also other ways for a speaker to restart (cf. Levelt, *op. cit.* for details), but they are so infrequent in the present sample that we can refrain from discussing them, and classify them as "other". It should be noticed that this categorization ignores such interjections as "er", "rather", etc. The repair proper is often preceded by "editing expressions" of this sort. We will return to them in the next section.

Do speakers use intonational marking to inform the listener about the type of restart they are making? One might conjecture, for instance, that instant repairing is the default case: the listener assumes that the first word of the repair proper is the replacement for the trouble item. If the speaker retraces, however, it would be helpful to mark the focussed element which is to replace the trouble item.

Table 2 Intonational marking in repairs with different ways of restarting

way of restarting	instant	retraced	other	total
marked correction	75	53	6	134
unmarked correction	92	57	16	165

The relevant data for answering this question are presented in Table 2. It gives the distribution of marked and unmarked corrections over the categories of instant repair, retraced repair and "other". Here again, a chi-square test revealed no significant differences between the categories. Speakers do not use intonational marking to tell the listener what sort of restart they have chosen to make. It should, finally, be added that there is nothing in the data to suggest that particular ways of restarting are more marked under particular conditions of delay, neither is there any interaction between delay, re-starting, and semantic type of correction (error versus appropriateness) with respect to prosodic marking.

So far, the present analysis confirms the findings of Cutler (op. cit.): there is no indication that the difference between marked and unmarked lexical corrections has anything to do with the interruption-and-restart structure of the repair. Let us now turn to the second possibility, semantic determinants of intonational marking.

4. *Semantics and marking*

4.1 *Error or appropriateness*

It was discussed above that there are two major classes of reasons for a speaker to interrupt and repair his utterance. The utterance can, in the first place, contain a straightforward error. This is the case for examples (1) through (5), (7) and (8) above. The error can, still, be of different sorts. Lexical errors are often referential misnomers, such as "green" for blue (cf. (4)), "left" for right (cf. (5)), or "over" for across (cf. (3)). In these cases the substituted word has an obvious semantic relation to the intended word. Other types of lexical error are also possible - for instance, where the relation between the error and the intended word is one of form rather than of meaning; but in the present corpus of lexical corrections, almost all cases of error are of a referential sort. (Further kinds of error - syntactic, morphological, phonetic, prosodic - are beyond the scope of this paper.)

The second main reason for making a repair is that the utterance is not fully appropriate, given the context in which it occurs. In (6) the word used is too vague, given the set of contextual alternatives. This is especially often the case when demonstratives are used, as in (9):

(9) And right of that one - of that purple ...

Also, an otherwise correct word is sometimes replaced because it does not match previously used terminology. A speaker may, for instance, decide to replace a static verb by a verb of motion, because he is

LEVELT & CUTLER

giving a dynamic description of the spatial network, i.e. in terms of an imaginary tour. An example of such an appropriateness repair is given in (10):

(10) If you go up one, there's - er, you come to yellow

Here the static "there is" is replaced by the dynamic "come", though the speaker could have completed the original static utterance.

Only corrections for error involve *rejection* of the reparandum, and this is often marked by the editing expression the speaker uses before making the repair proper. The explicit denial "no" (*nee*), for instance, occurs almost exclusively in corrections for error (Levelt, op. cit.). The repair is therefore an act of *contrasting*. This is not so in the case of correcting for appropriateness. There is no rejection, but rather specification of the reparandum. Here, what was said is *confirmed*, and this is often apparent from the editing terms speakers use as interjections. In the Dutch repair corpus, "dus" (literally "thus", "therefore"; the English contextual equivalent for the present repairs is "that is") is frequently used in corrections for appropriateness, but never for error repairs. Repairing for appropriateness is an act of *elaboration*.

It should be remembered that a correction was defined as marked when the repair differed prosodically from the reparandum. Do speakers apply prosodic differentiation when they are in the act of contrasting, rather than when they are in the act of elaborating? This can be tested by analysing the marking distributions for appropriateness and error repairs. Table 3 gives the results.

Table 3 Intonational marking in repairs for error and in repairs for appropriateness

correction for	error	appropriateness	total
marked correction	121	13	134
unmarked correction	108	57	165

It shows a highly significant ($p < .001$ by chi-square test) difference in marking between the two types of repair. Of the corrections for error 53% are marked, whereas corrections for appropriateness receive marking in only 19% of the cases. Hence it may be concluded that a main function of intonational marking in spontaneous self-repairs is to *reject* by establishing contrast.

The higher rate of marking in error corrections also explains the slightly higher occurrence of marking in repairs with interruptions within the trouble item which was observed in the previous section (cf. Table 1). Such interruptions occur almost exclusively in correction for errors, not in repairs for appropriateness (for reasons explained in Levelt, op. cit.).

Two points are left to be explained. The first one is why only 53% of the corrections for error are marked, given the fact that all of them presumably involve rejection of an item in the original utterance. This issue will be dealt with in the next section. The second point is why there is still 19% marking in appropriateness repairs. This will be taken up first.

The 13 marked corrections for appropriateness in our sample are very heterogeneous in character, and for most cases we have not been able to find an explanation for the marking that occurred. One subject marked every repair of either type. Another subject marked the same correction in one case but not in a second case (these were repairs where "blanco", i.e. blank, was replaced by "wit", i.e. white). Other cases in this set were, for instance, "door" - "rechtdoor" ("on" - "straight on"), "vanuit" - "door" ("from" - "through"), and the unusual case "het" - "een" ("the" - "a"). No uniform pattern emerges from these cases.

4.2. *Number of alternatives*

In order to explain why not *all* repairs for error are intonationally marked, a further partitioning of these errors should be considered. Earlier we suggested that, dependent on the context of discourse, speaker and listener may be mutually aware of the set of alternatives to the trouble item that caused the speaker to interrupt speech. The sense of contrast should be higher if this set is small, such as in case of antonyms and the like. The conjecture can be made that these cases especially will induce a speaker to mark the contrast by intonation.

It is possible to test this conjecture for the present corpus of repairs. A comparison can be made between two classes of error repairs. The first class consists of color name repairs; there are 119 of them among the 229 corrections for error. For these trouble items the set of alternatives is known: speaker and hearer knew that there were 11 different colors in the patterns described. The second class contains the repairs for directional terms. There are 61 of these in the sample. The directional terms almost always came in pairs: "left" - "right", "up" - "down", "horizontal" - "vertical". Since there were only four possible directions in the patterns, the maximum number of contextual alternatives at a particular choice point was four. The set of alternatives is, therefore,

LEVELT & CUTLER

substantially smaller for directional expressions than for color names. Does this correspond to a difference in the amount of marking?

Table 4 Intonational marking in repairs for color and in repairs for direction

correction for	color	direction	total
marked correction	59	44	103
unmarked correction	60	17	77

Table 4 presents the marking data for these two classes of error repairs. The difference is in the expected direction and significant ($p < .01$ by chi-square test): only half of the color word repairs are marked, but 72% of the direction term repairs. This supports the notion that there is more intonational marking for smaller sets of contextual alternatives to the trouble item.

5. Discussion

How far have we proceeded in finding an answer to the question why some lexical repairs are intonationally marked and others are not? Cutler's (op. cit.) data showed a "marking rate" of 38% for lexical repairs. The sample analyzed in the present paper has a rate of 45%. It was shown first, that syntactic factors i.e. the interrupt-and-restart structure of the repair played no role in marking. But a word of caution is in place here. Though marking, in the sense of prosodic contrast, is apparently not used for this purpose, it is possible and even likely that intonation does play a role in the solution of the continuation problem. If, for instance, a speaker makes an unmarked retracing, i.e. repeating elements that occurred before the trouble item, the listener could use the correspondence in intonation contour for identifying the part of the original utterance with which the repair proper overlaps. The other obvious cue here is the lexical identity of the repeated words (cf. Levelt op. cit.). Such lexical identity is not present in instant repairs, where the first word of the repair proper replaces the trouble item. In the absence of such a lexical joint between repair and original utterance, the listener may very well use intonational cues to match the repair to the trouble item. But notice that such a match exists only for unmarked repairs; in the marked case the prosody of the items to be matched is different. In other words, if intonation is used in this way for solving the continuation problem, intonational marking would be likely to interfere.

A second type of factor, however, showed a clear relationship with intonational marking. Corrections for error were marked in 53% of the cases, whereas corrections for appropriateness reached a mere 19%. The first, but not the latter of these repair types involves *rejection* of what was said before. Marking, it was argued, is used to express rejection. It was further found that marking is even more frequent if the number of contextual alternatives to the rejected item is small, i.e. if the contrast acquires the character of opposition. Corrections for directional terms ("left" versus "right", etc.) show a marking rate of 72% in the present sample. One might, for the sake of theoretical clarity, wish to distinguish between degree of opposition and number of contextual alternatives. The degree of opposition is the exclusiveness of the repair vis-a-vis the trouble item. If the task of the speakers involved just four different colors (instead of eleven), and these colors had been purple, pink, orange, and yellow, the number of color alternatives would have been the same as the number of directional alternatives. Nevertheless, the degree of opposition might still have been less, since the colors are sensed as fairly similar, whereas the four directions are highly exclusive. The present data do not allow us to make a choice between these two notions, but we would conjecture that it is the sensed degree, of opposition or exclusiveness, rather than the size of the set of contextual alternatives *per se*, that primarily underlies intonational marking.

In fact such a conjecture, it will be seen, fits well with what we consider to be the function of the prosodic marking of repairs in the context of prosodic structure in general. We will argue that marking a repair is, in effect, *accenting* it. In prosodic theory, accent is defined simply as the assignment of prosodic prominence to one element or part of an utterance; it is not defined in terms of how the prominence is realised. That is to say, accent is an abstraction; in an actual utterance it can be realised in a variety of ways. Accented words are usually longer and louder than unaccented words, higher in pitch or with more pitch movement, but they need not be - in appropriate circumstances accent can be realised by a noticeable decrease in amplitude, in pitch, etc. In other words, the definition of prosodic marking which we gave above is remarkably similar to a definition of accent.

What factors determine the placement of accent in an utterance? Although syntactic rules can be formulated which will correctly predict accent placement in neutral (acontextual) utterances, such rules only account for the default case; semantic factors will always override the syntactic. In actual utterances, in context, the placement of accent overwhelmingly reflects the semantic structure of the utterance (Cutler and Isard, 1980; Ladd, 1980). If marked repairs are accented repairs, it is little wonder that we found marking to be determined primarily by semantic rather than syntactic influences.

Moreover, the case of prosodic marking of *lexical* repairs allows

an even closer comparison with the function of accent. A lexical repair consists in the replacement of a single lexical item by another, virtually without exception one of the same form class, in the same syntactic context. Accenting of two lexical items of the same form class which are embedded in identical syntactic contexts occurs frequently in speech; it is said to express *contrast*, as in (11):

(11) First we **WROTE** it, then we **reVISED** it

Again it seems in this context hardly surprising that when a speaker wishes to emphasize the contrast between a repair item and the original trouble item which occupied its syntactic slot, he would mark it - or accent it.

Finally, the interpretation of marking as a manifestation of accent allows us retrospectively to account for the finding of Cutler (op.cit.) that marking is applied only to errors at the lexical level or above, never to phonetic errors. The smallest unit to which contrastive accent can be applied is a morpheme, as in (12), in which prefixes are contrasted:

(12) An **IN**crease in pitch but a **DE**crease in amplitude

Thus when the element to be repaired is below the morphemic level, as in (13) in which only a single sound is corrected, the appropriate environment for the assignment of accent is not available:

(13) Well it'll all have to be **un**siled - **un**signed

To apply accent to the word as a whole would be to mislead the hearer into thinking that one word was to be contrasted with another, whereas the desired contrast is in fact only between sounds. One sound cannot be contrasted with another by the application of accent; thus phonetic errors cannot be marked. Prosodic marking in speech repair, therefore, conforms to general constraints on the prosodic structure of language.

Acknowledgements

This research is a European Psycholinguistics Association collaborative project. The second author thanks the Max-Planck-Institut für Psycholinguistik, Nijmegen, for a Visiting Scholarship, during the tenure of which the work described in this paper was carried out. The authors are grateful to John Hawkins and Ewald Lang for useful discussions.

References

- Cutler, A., 1983: Speaker's Conceptions of the Functions of Prosody. In: A. Cutler and D.R. Ladd, (ed.), *Prosody: Models and measurements*. Springer, Heidelberg.
- Cutler, A. and Isard, S.D., 1980: The Production of Prosody. In: B. Butterworth (ed.), *Language Production*. Academic Press, London.
- Goffman, E., 1981: Radio Talk. In: E. Goffman (ed.), *Forms of Talk*. Blackwell, Oxford.
- Ladd, D.R., 1980: *The Structure of Intonational Meaning*. Indiana University Press, Bloomington.
- Levelt, W.J.M., 1983: Monitoring and Self-repair in Speech. *Cognition*. forthcoming.

