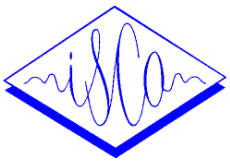

ITRW on Disfluency in
Spontaneous Speech (DiSS'01)
Edinburgh, Scotland, UK
August 29-31, 2001

# Prosodic Marking of Self-repairs

*Tapio Hokkanen*

Linguistics
University of Joensuu, Finland
tapio.hokkanen@joensuu.fi

## Abstract

Slip studies predominantly focus on either structural or semantic properties of the errors. Since most analyses have been based on pen-and-paper collections, i.e., on-line notes, it is quite understandable that suprasegmental of errors have remained a neglected area.

The present prosodic analysis is based on acoustical measurements of 307 self-repairs. Each repair has been measured with the Praat program. In order to make the measurements psychoacoustically relevant and comparable across speakers, the changes in $F_0$ are expressed in terms of semitones.

In general, speakers repair slightly less than three quarters of the errors they commit whereas one quarter remains either totally undetected or at least without a repair. With respect to prosodic marking, it appears that the proportion of marked repairs in the present data is significantly larger than in previous studies: approximately two thirds of self-repairs are marked with remarkably higher pitch (>+3ST), and a total of 96.7 per cent with a somewhat heigthened pitch. It is concluded that alternations of fundamental frequency are utilized in marking self-initiated repairs.

## 1. Introduction

When an error occurs, speakers tend to interrupt the flow of speech and initiate a self-repair. From a communicative point of view, the abrupt interruption and the following attempts to repair the troublesome utterance are a special form of disfluency: The listener has to decide how much of the context must be rejected and where the repair will start. Levelt [1] refers to this as continuation problem. The speaker, in turn, has to signal to the listener that a repair will follow.

Here it is hypothesized that speakers need not restrict to structural matters such as editing expressions, retracings to the previous context, or fresh starts in indicating the initiation of a repair. Instead, they may also provide the listener with certain prosodic cues, such as pausing, higher pitch, loudness, and variations in speech tempo [2] that supposedly play an important role in self-repairs. The purpose of this study is to analyse the micro-prosody of self-repairs in naturally occurring slips of the tongue.

Previous studies (e.g., [3] and [4]) of prosodic marking in self-repairs have relied on auditive impressions and interpretations by the researchers themselves. Cutler [3] argues that there are two prosodic strategies for the speakers to follow. In the first one, which she refers to as **unmarked**, the speaker wants to minimize the disruptive effect of the repair to the message by keeping the pitch of the repair as close as possible to the original troublesome item. The **marked** alternative, in turn, takes advantage of noticeable up- or downward changes in the pitch.

Cutler [3] found that speakers did not use prosodic marking in repairs that follow errors at the phonetic level but that 38 per cent of lexical error corrections were marked with remarkably higher pitch. Levelt and Cutler [4], in turn, report a 45 per cent proportion in a pattern description task. Levelt [1] reminds us that there are also personal stylistic factors of the speakers when he writes, "certain speakers [- -] would, so to say, cry out the corrections".

The present study follows the argumentation of Hokkanen [5], where it was claimed that self-repairs are predominantly marked with a remarkably higher pitch.

## 2. The Data

The following analyses are based on a corpus of tape-recorded, naturally occurring slips of the tongue (see [5]) that have been collected from radio interviews, sportscasts, and archived samples of spoken Finnish. There are 2,202 errors in the entire corpus, approximately three fourths (N = 1,683) of which are repaired by the speaker. In this study, a sample of 307 self-initiated repairs will be analyzed prosodically. All repairs were digitized with a 22.05 kHz sample rate.

## 3. Method and hypotheses

In order to avoid subjective interpretations, the markedness analysis should not be based on auditive impressions but on acoustic measurements, the results of which are then interpreted with certain psychoacoustic criteria. Therefore, pitch is here acoustically correlated to musical semitones (for the advantages of expressing $F_0$ changes in semitones, see [6] and [7]). This method also allows us for a reliable comparison of fundamental frequency changes across various speakers and occasions.

For the purposes of this analysis, fundamental frequency has been measured at the following points in each repair: first, at the onset of the troublesome item, secondly, at the end of the troublesome item before the interruption, thirdly, at the onset of possible editing expressions, and, finally, at the onset of the repair. These points supposedly provide a rough, yet an objective idea on whether speakers mark their self-repairs prosodically, and if so, what is marked and how. The measurements were conducted with the Praat synthesis and analysis program [8]. In this program one can also convert the actual speech signal into a human-like humming, which enables the researcher to listen to the flow of $F_0$ pulses without the actual phones. The prosodic patterns of all repairs can thus be separated from the segmental and semantically motivated means of marking a repair.

The question now arises, what is the differential threshold expressed in terms of semitones in normal communicative situations? The criterion for markedness in this study can be found in 't Hart [9]. He argues that in order to become perceived, the changes in fundamental frequency should

exceed the limit of ±3 semitones (henceforth ST). If we adopt this criterion to the present errors, all repairs indicating an $F_0$ change larger than ±3 ST will be interpreted as marked with either a significanlty higher or lower pitch.

On the basis of previous findings by Cutler [3] and Levelt and Cutler [4], it is hypothesized, firstly, that only less than half of lexical errors are marked with a higher pitch, although errors may sometimes also be marked with a remarkably lower pitch. Secondly, the previous studies also suggest that speakers should mark only repairs following lexical errors whereas repairs of phonetic errors remain prosodically unmarked.

It is furthermore intuitively hypothesized that prosodic marking would be dependent on detection latency, *i.e.*, on the interval between the error occurrence and the moment of interruption. The further the interruption appears from the troublesome item, the greater the change in fundamental frequency. Moreover, one may also hypothesize that the longer the pause between the error and its repair, the greater the change in $F_0$. The latter hypothesis is based on physiological grounds: a long pause could be used for breathing, which, in turn, might result in a somewhat higher pitch.

## 4. The analysis

In general, it appears that 66.1% (N = 203/307), *i.e.*, approximately two thirds, of all repairs exceed the ±3 ST differential threshold level and can, thus, be regarded as prosodically marked. This figure is fairly high when contrasted to the 38 and 45 per cent proportions reported previously by Cutler [3] and Levelt and Cutler [4]. This difference can not solely be attributed to the language analyzed (*viz.* Finnish). Rather, it probably depends on the method applied: as opposed to the previous studies, the present distribution has been obtained by virtue of acoustical measurements of naturally occurring errors and their repairs.

To test the second hypothesis, namely that only repairs following lexical slips would be marked with a remarkably higher pitch [3], marking was analyzed at various linguistic levels. This comparison indicated that there arose no statistically significant differences across the rates of prosodic marking between repairs of lexical and phonological errors: the rates were 67.9 (N = 66/103) and 64.1 per cent (N = 93/137), respectively ($\chi^2$ = 0.609, 1 d.f., not significant). This result does not give support to the first hypothesis. On the contrary, it may be concluded that prosodic marking is not restricted to the lexical level or levels above it. Instead, speakers use prosodic marking in their self-repairs regardless of the linguistic level of the error. This finding has also been interpreted [5] to indicate that it is neither structural well-formedness (phonological errors) nor communicative aspects (lexical ones) that primarily determine prosodic marking in self-repairs.

It is worth noting that none of the repairs in the present data was marked with a significantly lower pitch: the maximum downward change in $F_0$ was –0.66 ST which does not meet the current criterion of prosodic marking. For the sake of comparison, the maximum upward change was +19.28 ST. Since the downward changes were minimal and since there were only nine of them in the entire data, they may interpreted as falling within the limits of normal $F_0$ downdrift. Therefore, they become classified here as indications of prosodically unmarked repairs. The current results refer to the possibility that, as opposed to Cutler's [3] findings, speakers of the Finnish language do not mark their repairs with a remarkably lower pitch at all or do so only seldom.

With respect to the hypothesized relation between the latency of interruption and the change in $F_0$, no statistically significant correlation can be established here (Pearson correlation r = 0.109, n.s.). Therefore, it is evident that speakers do not use prosodic marking to signal their delayed repair initiations. An identical result is obtained in the correlation between the duration of the pause and the change in $F_0$ as well. It may be concluded that the purpose of the pause is something else than to provide the speaker with time to readjust the fundamental frequency. It has been suggested [5] that the duration of the pause is dependent on the linguistic level the error has arisen from.

In fact, none of the hypotheses presented so far receive direct support from the present data. Instead, it appears that the current repairs are in favor of the proposition by Levelt [1], namely, that prosodic marking of repairs is primarily semantically and communicatively motivated: speakers repair what they find necessary to repair, and prosodic marking means, first of all, marking of contrast between the troublesome item and the repair.

The suggested kind of contrast can best be seen in those instances where the prosodic marking runs counter to the normal prosodic patterns of the Finnish language. For instance, in (1) the error regards an inflectional suffix:

(1) *tulona, tules-, tulossa*
'coming'

In (1), the targeted inessive form (*tulossa*) is first substituted by the elative case (*tulona*) and then followed by a word search (*tules-*). What is essential here is that the speaker marks the repaired suffix (*i.e.., -ssa*) of the repair with an extra $F_0$ peak: the suffix is articulated 6.8 semitones higher (change from 101.3 Hz to 149.7 Hz) than the corresponding syllable in the troublesome word. The exceptional intonation contour of this particular repair can be best viewed in Figure 1 (see [5]):

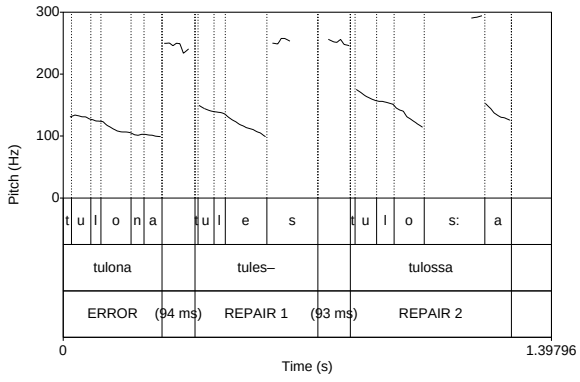


*Figure 1:* $F_0$ contour of error (1).

There are also examples regarding constituents of compounds in which an extra $F_0$ peak is assigned precisely to the first

syllable of the repaired constituent. These kinds of prosodic marking suggest that speakers want to make the contrast between the troublesome item and its repair more prominent by virtue of a substantially higher pitch and that this marking can be aligned exactly with the troublesome point in the utterance. It is also worth noting that the strive for marking a repair is so strong that it overrules the basic prosodic patterns characteristic of, *e.g.*, world-final syllables in polysyllabic words as well as right-hand constituents of compounds [5].

There is another strategy of prosodic marking as well, in which the $F_0$ level of the repair is matched with the $F_0$ level of the error. The technical criterion for this strategy is that the rise of the fundamental frequency contour is less than the critical +3 ST limit, yet positive. This definition distinguishes the prosodic marking from the normal $F_0$ downdrift. With respect to the current data, 30.6% (N = 94/307) of all repairs follow this strategy. It can be interpreted that speakers take advantage of prosodic marking not only to contrast the troublesome item and its repair but also to signal the start of the repair by readjusting the $F_0$ level. Alongside with structural means, such as, *e.g.*, repetitions and fresh starts, this $F_0$ readjustment presumably helps the listeners to easier adjust the repair to the original utterance.

The third alternative, *i.e.*, unmarked repairs, corresponds to the normal $F_0$ downdrift and to the speaker's choice of not marking a repair prosodically. All negative semitone values that do not exceed the –3 ST limit fall into this category. Nevertheless, these kinds of repairs are marginal, since they only account for 3.3 per cent (N = 10/307) of all repairs in the current data. They can be defended from a communicative point of view: despite the error, this strategy guarantees an undisturbed prosodic pattern and supports the continuation of the utterance. None of these repairs exceeded the –0.66 ST limit mentioned above.

Finally, there were no instances in which the repair would have been marked with a significantly (>–3ST) lower pitch. Consequently, the findings reported by Cutler [3] do not receive support from the present self-repairs.

## 5. Results and concluding remarks

It was argued above that speakers take advantage of changes in $F_0$ to mark the contrast between the troublesome item and the actual repair. A 66.1 per cent proportion of all repairs in the present data exceeded the +3 semitone limit set as a psychoacoustically motivated criterion of markedness. Instances of $F_0$ readjustment, in turn, account for slightly less than one third (30.6 per cent) of all repairs. Since these two strategies both deviate significantly from the normal $F_0$ downdrift, *i.e.*, from the third strategy, they can be treated as marked. Together the two marked strategies account for a vast 96.7 per cent majority of self-repairs, which allows for the conclusion that speakers predominantly use prosodic means in marking their repairs. In general, the proportion of prosodically marked repairs in the current data is substantially larger than those reported in previous studies.

On the basis of the present findings it is evident that speakers combine structural means with suprasegmental cues to help the listener overcome what is referred to as the continuation problem, *i.e.*, the fact that the flow of speech is abruptly interrupted and the troublesome item replaced with some other material.

It also appeared that prosodic marking is not restricted solely to the repairs following lexical errors. Instead, there arose no statistically significant differences between repairs of lower-level errors, which suggests that prosodic marking is not dependent of the linguistic level. No correlation was found between the duration of the pause after the interruption of the speech flow. This finding was interpreted here as an indication for the fact that the pause does not serve as a point for resetting the fundamental frequency.

## 6. References


[1] Levelt, W.J.M., *Speaking: From Intention to Articulation*. Foris, Dordrecht, 1989.

[2] Nakatani, C.H. and Hirschberg, J., A corpus-based study of repair cues in spontaneous speech. *J. Acoust. Soc. Amer., Vol. 95, 1994, pp. 1603-1616.*

[3] Cutler, A., Speakers' conceptions of the function of prosody. In: A. Cutler & D.R. Ladd (eds.), *Prosody: Models and Measurements*. Springer Verlag, Berlin, 1983. pp. 79-92.

[4] Levelt, W.J.M. and Cutler, A., Prosodic marking in speech repair. *J. Sem., Vol. 2, pp. 205-217.*

[5] Hokkanen, T., *Slips of the Tongue. Errors, Repairs, and a Model*. Finnish Literature Society, Helsinki, 2001.

[6] Hayward, K., *Experimental Phonetics.* Longman, London, 2000.

[7] 't Hart, J., Collier, R. and Cohen, A., *A Perceptual Study of Intonation. An Experimental-phonetic Approach to Speech Melody*. Cambridge University Press, Cambridge, 1990.

[8] See http://www.fon.hum.uva.nl/praat/

[9] 't Hart, J., Differential sensitivity to pitch distance, particularly in speech. *J. Acoust. Soc. Amer., Vol. 69, 1981, pp. 811-821.*



**Acknowledgement**

I am grateful to Professor Jussi Niemi (Joensuu, Finland) for his encouraging words and comments on an earlier version of this article.