CORRECTING ERRONEOUS INFORMATION IN SPONTANEOUS SPEECH: CUES FOR A.S.R.

P. Howell

University College London, England

ABSTRACT

Previous research has shown that when speakers make a mistake and repair it, they signal that the forward flow of speech has been altered by insertion of a pause, and highlight what has been altered by adding stress on the first word of the alteration. The question of whether these prosodic patterns are specific to repairs or whether they are shared by related grammatical structures was not addressed. In this contribution, the prosody of phrase and word repetitions and conjunctions, taken from a corpus of unrestricted speech, were analyzed. These constructions were chosen because they have certain similarities with repairs and the prosodic analysis procedure can be applied to them. Repetitions behave like repairs whereas conjunctions are dissimilar. It is concluded that the prosody of repairs is a reliable indication that errors are being corrected, and the implications for A.S.R. are discussed.

1. INTRODUCTION

When we listen to speech it is not usually difficult to determine speakers' intentions, even when they make alterations or start a sentence, break off, and recommence. These processes are termed "speech repair" [1]. An example of a repair is "Go left at the, I mean, go right at the crossroads". There are typically three identifiable parts in a repair - the original utterance (OU), the editing phase and the repair proper. In the example, "Go left at the" is the OU, and "go right at the" is the repair. The OU contains the word or words to be repaired, termed the reparandum, ("left" here). The speaker has gone past the erroneous word and so the repair is said to have an overshoot. The editing phase here is the phrase "I mean". Levelt counts "er" as an editing term, but it is considered here as a form of pause. With "er" excluded, the editing phase is rare in our corpus of repairs from unrestricted speech (4.2%) and is not discussed further. The final phase is the repair which includes the alteration (here the word "right"). Note also that the speaker has backed up to a point prior to where s/he wants to make the alteration and, so, the repair contains a retrace.

Levelt [1] has identified several categories of repair, but attention here is restricted to repairs in which erroneous information has been altered (termed error repairs). Though these may be sub-divided into repairs which occur on different grammatical units, the analysis reported here applies to all categories of error repair so the sub-categories are not discussed further. The only obligatory parts of error repairs are the alteration and reparandum.

Automatic speech recognition (A.S.R.) systems for timetable or directory enquiries will have to be able to identify when an error has been made in voiced input and what information is being altered. Incorrect recognition could result in erroneous information being generated by the system in response to an enquiry. Though analysis of speech can indicate what speakers do, it is important that these be assessed perceptually to check that listeners use these cues. Clearly information about how human listeners are able to understand (1) that a repair is being made (the forward flow of speech has been stopped) and (2) what erroneous information is being altered may have important implications for A.S.R.

2. PROSODY IN SPEECH REPAIRS

Prosody helps listeners both detect that the forward flow of speech has stopped and locate where the altered information starts [2]. Prosody refers to changes in timing, loudness and pitch movements over groups of segments. The two aspects of prosody that are known to signal information about repairs are pauses and stress. Pauses, as noted previously, include filled pauses as well as periods of silence. Stressed syllables tend to be longer and louder than their unstressed counterparts. Primary and secondary stress are marked in the transcriptions, primary stress indicating a higher level than secondary stress.

Howell and Young [2] analyzed a corpus of 272 repairs drawn from the Survey of English usage (SEU) [3]. This corpus is of unrestricted speech and has pauses and stresses transcribed. The speech was parsed into the retraced section and the corresponding section that occurred prior to it. There was a marked tendency for sections of pauses to be added before the first word of the retraced section when such sections occurred but no systematic tendency to increase stress on the first word of the retrace compared with its first occurrence. This shows that pauses are used to mark the interruption to the forward flow of speech. The highest degree of stress that occurred on any syllable of the first word of the alteration was compared with this same measure on the reparandum. This analysis showed that stress was added on the first word of the alteration, and this was interpreted as showing that stress is used to highlight the altered information. There was no tendency to add pauses before the alteration in comparison with the reparandum except when no retrace occurred. In the latter cases, the alteration starts immediately after the forward flow of speech has been interrupted, and is consistent with pauses being used to mark such locations.

3. ASSOCIATION OF PROSODIC PATTERNS WITH REPAIR

No data has yet been provided concerning whether the prosodic patterns described are specific to repairs or whether they are shared with related structures. In this presentation two structural types which are closely related to repairs are analyzed. These two types are repetitions and conjunctions. Repetitions include word and phrase repetitions, and the repeated sections can be analyzed in the same way as retraces. A total of 364 word and 168 phrase repetitions were located in the SEU and the results of this

analysis are shown in Table I.

Table I. Analysis of pauses and stresses in repetitions

		Word repetitions	Phrase repetitions
Pauses	Added	107	56
	Dropped	21	4
	N	364	168
		sig.	sig.
Stresses	Added	19	12
	Dropped	13	15
	N	364	168
		ns.	ns.

Basically, the table shows that speakers introduce pauses before the first word of a repetition but there is no increase in stress. This parallels the findings with retraced sections of repairs and offers some support for terming these "covert repairs" [1].

The inclusion of conjunctions in the analysis depends upon a rule described by Levelt for ascertaining whether a repair is well-formed or not (WFR). Levelt's WFR suggests that the two main parts of a repair (original utterance and repair proper) are related in the same manner as the two constituents of a coordination. According to Levelt "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 a connective such as "or" or "and")." [1, p.486]. Thus, in "There you can park at the left-hand side of the, the right-hand side of the road," the original utterance is "park at the left-hand side of the". The VP "park at the left-hand side of the" can be completed with "road" (=C).

The repair (R) is "park at the right-hand side of the road". The coordination thus becomes: "There you can park at the left-hand side of the road or park at the right-hand side of the road."

Given a co-ordination, this process can be reversed. Thus the co-ordination "a comparative graphology paper or a historical graphology paper" from the SEU could have been the repair "a comparative graphology, a historical graphology paper". Levelt has pointed to the syntactic relationship between co-ordinations and repairs, so it might validly be asked whether this extends as far as the two structures having similar prosodic properties. Using Levelt's WFR, the retrace in a co-ordination starts after the conjunction and the first non-matching word constitutes the alteration ("historical" in the example, which is to be compared with "comparative").

A total of 244 conjunctions were collected from the SEU. These were single words or phrases which were joined by any conjunction. The constraints placed upon conjunction selection was that the word or phrase before the conjunction started at a constituent boundary and that there was a constituent boundary after the word or phrase after the conjunction. Also, the grammatical category of the word or phrase before the conjunction had to match in type with the grammatical class of the word after the conjunction. No other constraints were applied. As illustrated in the example, this generated material which had sections with retraces and "reparandum/alteration" equivalent pairs and analysis proceeded as described for the repairs. The data are summarised in Table II.

Table II. Analysis of prosodic factors around conjunctions

a) Pauses and stresses on the first word of the "reparandum/aleration" equivalent

		Had no retrace	Had retrace
Pauses	Added	10	2
	Dropped	2	4
	N	198	46
		sig.	ns.
Stresses	Added	61	20
	Dropped	50	4
	N	198	46
		ns.	sig.

b) Pauses and stresses on retraced sections

Pauses	Added	1
	Dropped	2
	N	46
		ns.
Stresses	Added	1
	Dropped	4
	N	46
		ns.

The prosody around conjunctions differs from that around repairs. For the conjunctions that had a retrace, there is a significant tendency to stress the "alteration", as with repairs, but no significant tendency to add pauses prior to the "retrace", unlike with repairs. For the conjunctions that did not have a retrace, there is a significant tendency to add pauses, as with repairs, but no significant tendency to add stresses, unlike with repairs. Thus, it appears that the different types of constructions, though syntactically related are prosodically dissimilar. The prosody in repairs is dissimilar to that in related grammatical structures such as around conjunctions.

REFERENCES

[1] LEVELT, W.J.M. (1983), *"Monitoring and self-repair in speech", Cognition, 14,* 41-104.
[2] HOWELL, P. & YOUNG, K., *"The use of prosody in highlighting alterations in repairs from unrestricted speech", Quarterly Journal of Experimental Psychology,* in press.
[3] SVARTVIK, J. & QUIRK, R. (1980), *"A corpus of English conversation"*, Lund: Gleerup.