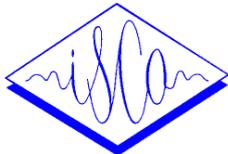


Phonetic correlates of self-repair involving word repetition in German spontaneous speech

ISCA Archive



<http://www.isca-speech.org/archive>

Ramona Benkenstein & Adrian P. Simpson

University of Jena, Jena, Germany

Abstract

A phonetic description of self-initiated self-repair sequences involving the repetition of words in German spontaneous speech is presented. Data are drawn from the *Kiel Corpus of Spontaneous Speech*. The description is primarily impressionistic auditory, but it also employs acoustic records to verify and objectify the impressionistic findings. A number of different patterns around cut-off are identified. The comparison of phonetic differences between reparandum and repair tokens is used to argue that repair sequences can also provide an interesting insight into the way in which fluent stretches of spontaneous speech are phonetically organized.

1. Introduction

One of the most important aspects of disfluency in spontaneous speech is the phonetic management of self-initiated self-repair. Previous work has investigated aspects of pitch, tempo, duration, coarticulation, and phonatory patterns prior to cut-off, in pausal behaviour between reparandum and repair, and at the beginning of the repair proper (e.g. [11, 12, 14, 15]).

From an analytical point of view, self-repairs involving word-repetition represent a set of repair sequences of particular interest as they allow for a direct analysis of the phonatory and articulatory differences in the production of identical lexical material in reparandum and repair by the same speaker.

While the detailed phonetic analysis of repair sequences involving word repetition is required to give us a better understanding of the strategies speakers employ when dealing with repair, this subset of disfluencies is also of more general phonetic interest. Analyses of phonetic change in lexical material across different styles or at different tempos is generally elicited using spoken prose with speakers being instructed to speak faster or slower. Leaving aside the problems of the elicitation procedures themselves, it is always difficult to know the extent to which the results of a study done on prose can be transferred to spontaneous talk. An insufficient number of tokens of the same lexical word or phrase in comparable contexts is undoubtedly one of the main reasons why phoneticians have only slowly been drawn away from the controlled elicitation of prose to the analysis of spontaneous speech.

Word repetition in repair sequences seems to provide an interesting place for solving some of these problems. Within a very short time a speaker repeats the same lexical material, but does so using different phonetics. Although some of the differences can be accounted for in terms of the typical variance inherent in such a complex system, we must assume that any auditory impressionistic differences we can observe and verify acoustically are due to immediate or longer domain contextual differences. Regularities arising from the differences observed in a number of cases can begin to provide an account of what the different factors are.

This study describes some of the detailed impressionistic and acoustic phonetic patterns observed in word-repetition repair sequences in German, and attempts to account for the patterns we describe.

We will show, in common with work carried out on English and Dutch cited above, that speakers of German exhibit a number of articulatory and phonatory patterns which, even in the absence of explicit editing terms, serve to demarcate the reparandum and indicate the initiation of the repair proper. Furthermore, by looking in detail at individual repair sequences we will begin to substantiate the claim made above that the subset of repairs containing word repetition is of more general phonetic interest.

Besides being a contribution to the analysis of repair in general, it is also a contribution to repair in German, something which has received relatively little attention ([9, 13]).

2. Method

2.1. Data collection

Volumes 1–3 of the *Kiel Corpus of Spontaneous Speech* [5] contain mixed and same sex dialogues from 18 female and 24 male speakers with a North German linguistic background. The data were collected as part of the *Verbmobil* project [6].

One condition imposed on the recordings was maximum channel and turn separation. Channel separation was achieved by seating subjects in separate sound-treated rooms and having them communicate via headsets. Turn separation was achieved using a technical setup akin to an intercom. Subjects could only speak and be heard by their partner when pressing a button. This simultaneously blocked the channel for the other speaker. The state of the buttons was indicated to each of the speakers by means of lights.

The dialogues were elicited using an appointment-making scenario in which speakers had to arrange fictitious meetings and trips over a two month period. Each speaker had a sheet covering the same two month period, but with different days shaded in. Speakers were instructed not to make appointments on the shaded days on their calendar sheets. Six of these dialogues, one for each two month frame, were collected from each speaker group. One further dialogue was elicited using a more immediate time frame (the following seven days). This was designed to get speakers using deictic time expressions.

Although the dialogues are spontaneous the technical imposition of turn separation makes them unsuitable for the investigation of many aspects of conversational interaction. However, the dialogues are a rich source of the phenomenon under investigation here, namely self-initiated self-repair.

Turns from the dialogues will be referred to using the same nomenclature as they have in the *Kiel Corpus*, e.g. g125a004, which refers to the fourth turn from the fifth task in dialogue g12.

2.2. Segmentation and annotation

The dialogues were transliterated and manually segmented and phonetically annotated. A number of disfluencies were marked in the transliteration, including pauses, editing terms (“uh(m)”), truncations, etc. The reparanda of overt self-repair sequences involving repetition with and without insertions were parenthesized in the transliteration.

For a complete description of the elicitation, transliteration, segmentation and annotation of the dialogues in the *Kiel Corpus* see Kohler et al. [8].

2.3. Analysis

The dialogues contain some 400 self-initiated, overt self-repair sequences. Of these a subset involving clear cases of word and phrase fragment repetition was selected for impressionistic phonetic and acoustic analysis. Typical examples of such repair sequences are given in Table 1.

Table 1: Examples of self-initiated self-repairs involving word repetition from the corpus.

Example	Gloss
<i>ja das ist- das ist ideal</i>	‘yes that’s- that’s ideal’
<i>auf- auf fünf Tage</i>	‘to- to five days’
<i>nehmen wir doch den- gleich den ersten</i>	‘let’s take the- ADV the first’

In the impressionistic analysis the same word sequences in reparandum and repair were compared. Attention was paid to a number of segmental and suprasegmental parameters. Our principle assumption, following Kelly & Local [7], is that if we can hear a phonetic difference we must assume that it has relevance for speaker and hearer.¹

The main objective of the acoustic analysis was to illustrate and attempt to quantify the impressionistic auditory observations.

3. Description of repair

3.1. General

Table 2 shows the distribution of the items at cut-off across different word classes. The most significant finding which emerges from the table is that in the vast majority of the cases the item directly adjacent to cut-off in the subset of repair sequences analyzed here belongs to the class of function words,² i.e. prepositions, articles, pronouns, modals. This seems to be at odds with the findings from English, in studies such as Fox & Jaspersen [3], who present many examples with content words at cut-off.

Table 2: Frequency of occurrence of items at cut-off across different word classes.

Functional Words	N	Content Words	N
Preposition	26	Noun	2
Pronoun	11	Adjective and Numeral	3
Article	10	Verb	1
Auxiliary	2	Adverb	6
Conjunction	2		
Σ	51		12

¹ Phoneticians have long been aware of our acuteness to relative detail: “The difference between vowel sounds separately pronounced, will sometimes appear so slight that the ear may be perplexed to discriminate them; but in the compounds of speech the minutest shades of elementary variety create unmistakable distinctions.” [1].

² At least from a traditional perspective, since most contemporary theories treat some prepositions as content words.

The first observation one can make when comparing the shape of the same word material in repair and reparandum appears at first to be almost trivial: phonetic identity in reparandum and repair is seldom.

In comparing the phonetic shapes of the same item in reparandum and repair we observed differences in vowel quality, vocalic quality (secondary resonance) of consonantal portions, tension, voice quality, length of vocalic and consonantal strictures, tempo and pitch. Furthermore, there were also differences in the ways in which the cut-off and subsequent pause was produced.

The differences observed in repetitions of the same item in reparandum and repair can be summarized in general terms as follows. The phonetic shape of the repair item is appropriate to the phonetics of the stretch it is situated in; the phonetic shapes of item in the reparandum, are not consistent with the same stretch as that in the repair and hence are different.

This leaves open two interpretations for the phonetics of the reparandum:

- they are appropriate to a stretch which was not completed, but was broken off when an error had been detected or a change in plan occurred.
- they are designed to initiate the repair sequence and project the cut-off itself.

3.2. Repair types

On the basis of a bundle of phonetic features located primarily around the location of the cut-off we make a tentative grouping into three different types of repair sequence, the first two sharing many of features found in different pause types by Local & Kelly [12].

The first, most common, type can loosely be described as lax. The item at cut-off and in particular the final vocalic or consonantal portion is long, there may be a drop in pitch, voicing is breathy, and the glottis opens into the subsequent pause, which may contain an audible inbreath. Friction occurring at cut-off, although longer, is often laxer than the corresponding portion in the repair token. Acoustically, the laxer friction in the reparandum token is lower in intensity, the spectrum more diffuse. Articulatorily, this could be due to a wider stricture of close approximation. Interestingly, increased nasality was not a feature found at cut-off. In fact, the vocalic portion preceding a nasal at cut-off (e.g. article *den*) was generally found to be less nasalized than the vocalic portion of the corresponding item in the reparandum. This was also the case when there was little durational difference between the vocalic portions. [Example: g072a008, *das ist- (P) das ist leider*].

In the second type, which can loosely be described as tense, the cut-off was initiated with glottal closure, possibly accompanied by creak.³ Glottal closure was maintained for the duration of the subsequent pause. There was no pitch drop. Minimally, if there was no pause following cut-off, creak could accompany the final portion of the reparandum item before the subsequent start of the repair item. The reparandum item at cut-off could be longer. There were no noticeable tension differences between corresponding fricatives in reparandum and repair items. This type of repair seems to be similar to Local & Kelly’s [12] turning holding pauses in English. [Example: g375a002, *die- die Zeit*]

³ Glottalization element at cut-off in the same data set is described acoustically by Kohler et al. [9].

In the third type, the repair proper followed directly on from the reparandum without there being any break or change in phonation, or pitch. The reparandum item could be longer. [Example: g376a006, *der- der*]

In an attempt to find a systematic qualitative differences between the vowel qualities of the reparandum and repair tokens, such as centralization, F2 at vowel midpoint for a number of front vowel categories was estimated. It was hypothesized that if the vowel tokens of the reparandum were consistently more centralized, F2 should be lower. As can clearly be seen from Figure 1 this was not the case. We will offer a possible account for this in the next section.

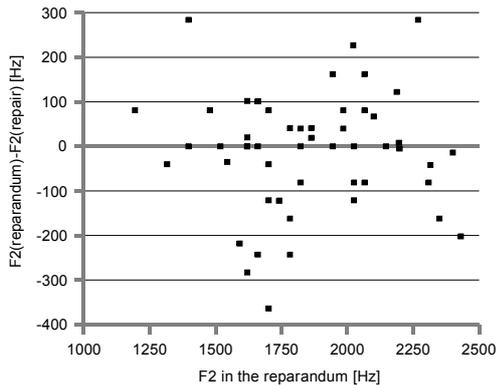


Figure 1: Difference between F2 of reparandum and F2 of repair as a function of the F2 of reparandum for front vowels.

3.3. Observations on individual repair sequences

Figure 2 contains the spectrogram of a portion from a repair sequence containing repetition of the definite article *den*. From an impressionistic point of view the token in the reparandum has a slightly fronter vocalic portion, the vocalic quality of the plosive release is clearer.

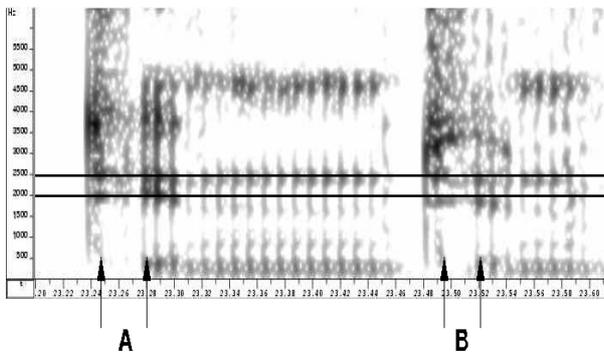


Figure 2: Repeated tokens of the definite article *den*, from the repair sequence *den- den Wochenplan*. Arrows at A and B are placed during plosive release and at the first period of the vocalic portion. [Ref: g215a004]

These auditory differences can also be seen in the acoustic representation. Two horizontal lines intersecting the frequency axis at 2 kHz and 2.5 kHz facilitate a comparison of the positions of F2 and F3 during plosive release, and during the subsequent vocalic portion. As we can see, both F2 and F3 in the reparandum token are higher than in the repair token.

The backer quality of the repair token is consistent with the phonetic shape of the initial stressed syllable [vɔχ] of the following noun. Notice, however, that the qualitative differences between these two tokens cannot be accounted for in terms of undershoot arising from duration differences. While the final nasal in the reparandum is longer, the vocalic portions are almost identical in duration.

An example of a two-word repetition is shown in Figure 3. As in the previous example, the two *das ist* tokens exhibit subtle phonetic differences which are compatible with the repair token harmonizing with the initial syllable of the following adjective *ideal*. Noticeable auditory differences between the two tokens are the relative openness of the vocalic portion in the reparandum *das*, a more central vocalic portion in *ist* together with a darker vocalic quality of the final alveolar friction.

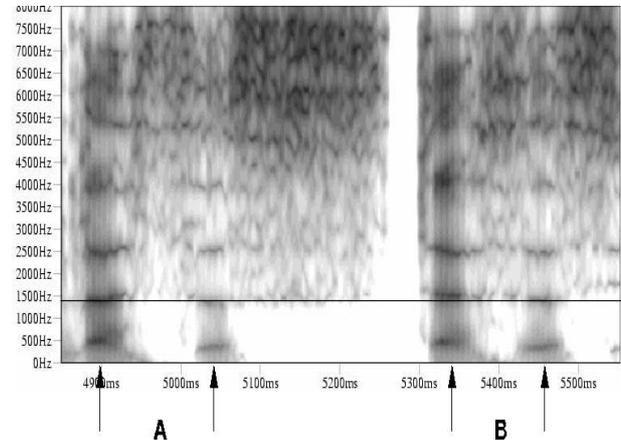


Figure 3: Repeated tokens of the fragment *das ist* from the repair sequence *das ist- das ist ideal*. Arrows at A and B are centred on the vocalic portions of *das* and *ist*. Formant track visibility has been enhanced by using a LPC spectrogram (order: 30, window: 10 ms). [Ref: g117a007].

Of the auditory differences, it is again the height of F2 which is most apparent in the acoustic record. In Figure 3 a horizontal line has been drawn to coincide with the F2 minimum in the vocalic portion in the reparandum *das*. The left arrow at B in Figure 3 indicates the centre of the vocalic portion in the repair *das*, and, as we can see, F2 is greater than the F2 minimum of the reparandum *das* throughout.

Differences in the shape of the formant tracks in diphthongs have typically been related to temporal differences, (e.g. [4]). Repairs containing repeated items with diphthongs indicate that additional factors need to be considered when determining diphthongal quality. The correlate of coda *r* in many varieties of German is a central vowel of variable height between mid and half-open. This gives rise to a range of diphthongal vowel qualities. Figure 4 contains repetition of the prepositional phrase *bei mir* with the pronoun *mir* at cut-off. Comparison of the diphthongs (at A and B) reveal interesting auditory and acoustic differences.

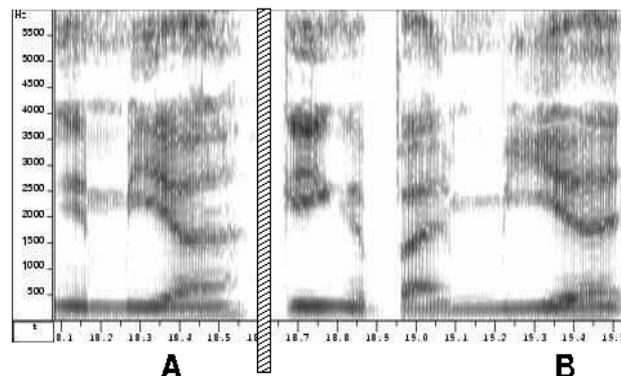


Figure 4: Portion of the repair sequence *bei mir- ginge bei mir*. The diphthongal portion of the repeated pronoun *mir* is at A and B. The vertical block indicates a deleted part of the pause in the repair. [Reference: g421a001]

Auditorily, the diphthong of the reparandum token (A) begins more centralized and ends more open than the repair token (B). From an acoustic point of view, differences in the duration of the *mir* tokens is attributable solely to the longer initial nasal of the repair item; the vowels themselves have approximately the same duration of 280 ms. The qualitatively closer end to the repair token, visible in the higher F2 and lower F1 might be expected as a local coarticulatory effect due to the close vowel in *im* which directly follows. The more open end of the reparandum, in its turn, can be seen as appropriate to the cut-off context. However, the differences in the formant tracks are not restricted to the ‘ends’ alone. In the repair token F2 remains high for at least 100 ms before beginning its fall, at least twice as long as it does in the repair token.

We can propose position in repair sequence or stress (the repair token is stressed) to account for these differences, but interestingly these clear audible and acoustically visible differences cannot be attributed to duration. Indeed, one possibility for the different patterns is that these pronoun tokens are durationally similar for different reasons. While in the stressed repair item it is the correlate of the vocalic nucleus which is long, in the reparandum it is the correlate of the coda *r*, appropriate to its pre-cut-off position, which is longer. In sum, this gives rise to similar durational, but quite different formant patterns.

4. Discussion

In the previous section we have looked at some of auditory and acoustic phonetic detail that can be observed in a subset of self-repairs involving the repetition of words and phrase fragments. The phonetic detail which can be observed throughout such repair sequences is revealing for the way speakers handle this type of disfluency. In phonetic terms, it is possible in many cases to identify that a speaker is producing phonetics appropriate to an upcoming cut-off [15]. Such information is crucial to our understanding of the cognitive processes which underlie the detection and management of repair ([2, 10]).

However, we have also seen that word repetition across repair sequences can offer us a way of looking at how speakers phonetically organize ‘normal’ stretches of utterance in spontaneous talk. The subtle patterns we have described suggest that coarticulation is long-domain, rather than the local phenomenon, restricted to adjacent segments, which it is often treated as.

5. References

- [1] Bell, Alexander M. 1881. *Sounds and Their Relations*. Salem, Massachusetts: Burbank.
- [2] Blackmer, Elizabeth R. & Janet L. Mitton. 1991. Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition*, vol. 39, pp. 173–194.
- [3] Fox, Barbara A. & Robert Jasperson. 1995. A syntactic exploration of repair in English conversation. In: Philip W. Davis (ed.), *Alternative linguistics. Descriptive and theoretical modes*. Amsterdam/Philadelphia: John Benjamins, pp. 77–134.
- [4] Gay, Thomas. 1968. Effect of speaking rate on diphthong formant movements. *JASA*, vol. 44, pp. 1570–1573.
- [5] IPDS 1995–97. *The Kiel Corpus of Spontaneous Speech*, vols. 1–3. Kiel: Institut für Phonetik und digitale Sprachverarbeitung.
- [6] Karger, Reinhard & Wolfgang Wahlster. 1995. *Verbmobil Handbuch – Version 3. Verbmobil Technisches Dokument 35*. Saarbrücken: DFKI.
- [7] Kelly, John & John K. Local. 1989. *Doing Phonology*. Manchester: Manchester University Press.
- [8] Kohler, Klaus J., Matthias Pätzold & Adrian P. Simpson. 1995. *From scenario to segment: the controlled elicitation, transcription, segmentation and labelling of spontaneous speech*. AIPUK 29. Kiel: IPDS.
- [9] Kohler, Klaus J., Benno Peters & Thomas Wesener. 2001. Interruption glottalization in German spontaneous speech. *Proceedings of the ISCA Workshop Disfluency in Spontaneous Speech*, August 29–31 2001, Edinburgh, pp. 45–49.
- [10] Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. Cambridge, Massachusetts: MIT Press.
- [11] Levelt, Willem J. M. & Anne Cutler. 1983. Prosodic marking in speech repair. *Journal of Semantics*, vol. 2, pp. 205–217.
- [12] Local, John K. & John Kelly. 1986. Projection and ‘silences’: notes on phonetic detail and conversational structure. *Human Studies*, vol. 9, pp. 185–204.
- [13] Selting, Margret. 1995. *Prosodie im Gespräch. Aspekte einer interaktionalen Phonologie der Konversation*. Tübingen: Niemeyer.
- [14] Shriberg, Elisabeth E. 1994. *Preliminaries to a theory of speech disfluencies*. Ph.D. thesis, University of California at Berkeley.
- [15] Shriberg, Elisabeth E. 2001. To ‘errrr’ is human: ecology and acoustics of speech disfluencies. *J. Int. Phon. Assoc.*, vol. 31, pp. 153–169.