TEMPORAL CHARACTERISTICS OF LEXICAL ERROR AND APPROPRIATENESS REPAIRS IN SPONTANEOUS DUTCH SPEECH

Leendert Plug & Paul Carter

University of Leeds, UK
l.plug@leeds.ac.uk; p.g.carter@leeds.ac.uk

ABSTRACT

This paper presents a phonetic analysis of lexical repairs taken from a corpus of spontaneous Dutch speech. The analysis focuses on the temporal relationship between the reparandum and repair components. Two predictors are tested, alongside several control variables: the pragmatic type of repair (*error* or *appropriateness*) and the structural type (*interrupted* or *completed*). Results suggest that pragmatic type has no consistent effect on the temporal organization of repairs, while structural type has some effect. Moreover, a significant effect is found for a measure of the relative phonological complexity of the reparandum and repair.

Keywords: prosodic marking, articulation rate, self-repair, Dutch

1. INTRODUCTION

Self-repair, in which a speaker interrupts the flow of speech to correct something said before, as in *Thursd- Friday*, is a common feature of normal speech. Yet, relatively little is known about the way in which the correct information is delivered. We might expect this to be marked as 'different' from what comes before, to make it clear that it replaces what has already been said, but previous research suggests that the nature of this marking is complex and subject to a number of constraints.

First, the semantics of the repair constrain its phonetics, in that repairs of linguistic or factual errors (*error* repairs) are less frequently 'prosodically marked' than repairs of pragmatic infelicity (*appropriateness* repairs) [2, 5, 11]. The details of 'prosodic marking' are elusive: according to Levelt and Cutler [5] p.206, it can be accomplished by 'a noticeable increase or decrease in pitch, in amplitude, or in relative duration'. An obvious question is whether further patterns can be discovered in the clustering of these parameters.

Second, the phonetics of repair are constrained by the timing of the repair relative to the onset of the reparandum [7]. In a study of elicited speech error repairs, Nooteboom [7] shows that repairs in which the erroneous production is interrupted, as in *ba- dark boat*, are phonetically different from those in which the erroneous production is completed, as in *bark… dark boat*: *interrupted* error repairs are done without delay, with high pitch and intensity, while *completed* error repairs are done with more delay, with low pitch and intensity. The difference can be attributed to interactions between pre-articulatory and post-articulatory self-monitoring. An interesting question is how this pattern interacts with that observed by Levelt and Cutler [5].

The present study provides a starting point for addressing these questions by considering a single phonetic parameter – articulation rate – across the sub-types of repair mentioned above in spontaneous Dutch speech. Recent work on similar data [9] points towards a preponderance of temporal compression – that is, a local increase in articulation rate – following the initiation of repair, and reports no evidence of distinction between error and appropriateness repairs. However, the database used in [9] appears relatively hetero-geneous, comprising lexical, phonological and syntactic errors repairs, lexical appropriateness repairs and 'different repairs' [4]. Moreover, the study does not consider the timing of the repair relative to the onset of the reparandum as a possible interfering factor. The present study considers a larger and more homogeneous collection of repairs, focusing on lexical error and appropriateness repairs, and tests, among other things, whether differences between interrupted and completed repairs can be observed.

2. METHOD

2.1. Data selection

Instances were selected from the Spontaneous Face-to-face and Broadcast sub-corpora of the Spoken Dutch Corpus [8], through a corpus search for utterances marked as incomplete. Only instances involving the correction of one (accurately pronounced) word for another were included, with a small number of exceptions (<5%) in which one word replaces or is replaced by a short phrase. Compounds and phrasal verbs were counted as single words. In total, 145 instances were selected. Representative examples are given below, with the core repair components in bold.

(1) *met de* ***au****- met de* ***bus*** (‘by ca- by bus’)
(2) *als er met tekst* ***gebrui****-* ***gewerkt*** *wordt* (‘when one use- works with text’)
(3) *een* ***leuke*** *k- een* ***mooie*** *keuken* (‘a nice k- a beautiful kitchen’)
(4) *een* ***paar jaar*** *gel- of uh een* ***tijd*** *geleden* (‘a couple of years a- or er a time ago’)
(5) ***Kerk****plei- of* ***Konings****plein* (‘Church Squa- or King’s Square’)

2.2. Pragmatic and structural analysis

Each instance was classified as appropriateness or error repair as in [4]. Instances involving a factual inaccuracy or linguistic ill-formedness, such as (1), (2) and (5) above, are error repairs; all other instances, including (3) and (4), were classified as appropriateness repairs. Classification was done independently by two Dutch linguists. Their classifications matched for 137 instances (94%). Following reconsideration, agreement was reached on a further 4, and the 4 instances that remained unclear were excluded from subsequent analysis.

In addition, each instance was classified as interrupted or completed depending on whether the lexical item which occasions the repair is cut off or completed before being corrected. Examples (1) and (2) are interrupted; (3) and (4) are completed. In cases of compound items, internal structure was taken into consideration: (5) was classified as completed since *Kerk* ‘church’ is, and *plein* ‘square’ is not subject to correction. Table 1 shows cross-tabulated counts following pragmatic and structural classification, showing a significantly greater tendency for appropriateness repairs to be interrupted ($\chi^2(1)=6.1$, $p=0.01$).

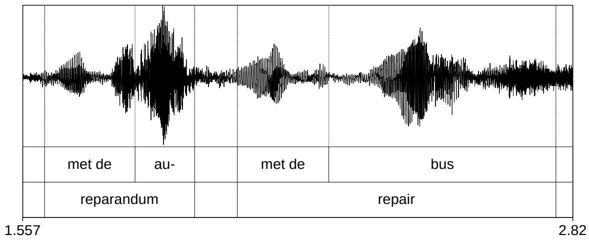
Table 1: Cross-tabulation following pragmatic and structural classification.

Type	N interrupted	N completed	*Total*
Error	47	32	79
Appropriateness	49	13	62
Total	96	45	141

2.3. Temporal analysis

Each instance was segmented into a reparandum and repair portion. Any repeated lexical items and ‘editing’ items, such as *of* ‘or’ in (4) and (5), were labeled separately (see Figure 1).

Figure 1: Segmented waveform for the repair in (1).



Articulation rate figures were calculated for all separately labeled portions in both canonical syllables and canonical segments per second. To facilitate comparison between the reparandum and repair, articulation rate figures associated with the repair were divided by those for corresponding stretches in the reparandum, following [9]. Proportional values above 1 reflect an increase in articulation rate in the repair portion; values below 1 reflect a decrease. Editing items such as *of* ‘or’ are present in a small minority of instances only (12% if ‘uh’ is included) and were not considered in the analysis reported here.

2.4. Statistical analysis

Statistical analysis was done through linear mixed-level modeling [1], with proportional articulation rate as the dependent variable and pragmatic and structural type as crucial predictors. Several additional factors were entered into the analysis. First, it may be expected that longer or more complex target words are produced at a higher articulation rate than shorter or simpler ones [6]. As a crude measure of the difference between core reparandum and repair items in complexity, the number of segments in the (projected) reparandum item – *auto* ‘car’ in (1) – was subtracted from the number of segments in the corresponding repair item. Second, the delay between reparandum offset and repair onset was included to test the hypothesis that repairs which are initiated without delay are produced faster than those with delay [7]. Third, the sub-corpus from which the instance was sampled and speaker nationality (Dutch or Flemish) were entered as random factors.

3. RESULTS

3.1. Dependent variable selection

As indicated above, articulation rate was calculated in syllables and segments per second, across all labeled portions. This yields at least four instantiations of the dependent variable: proportional values in *either* sylls/sec *or* segs/sec, calculated *either* across the entire reparandum and repair *or* across core repair items only – that is, excluding repeated lexical items.

Inspection of these variables suggests, firstly, that excluding repeated lexical items from the calculation of proportional values has little effect on their distribution. Repeated lexical items are present in 76 instances (52%), and these items are not uniformly produced at a higher or lower rate than the core repair items. The results below are for core repair items only. Secondly, while none of the four variables are normally distributed, the values calculated in segs/sec are closer to normal than those calculated in sylls/sec. Therefore, statistical analysis focused on the former.

Table 2 summarizes the distribution of values. It is positively skewed, with half of the values exceeding 1.10. Assuming a JND for tempo variation of 5-10% [10], these values represent a noticeable temporal compression of the repair relative to the reparandum. Similar proportions of values fall in the 'equivalence' range of 0.90-1.10, and below 0.90, representing noticeable temporal expansion in the repair.

Prior to modeling, values were log-transformed to normalize the distribution (K-S Test, D=0.06, p>0.2). Moreover, three outlier values above 3 were removed, reducing the dataset to N=138.

Table 2: Summary distribution of proportional values (segs/sec) across core repair items.

Statistic	Value
Mean (SD)	1.18 (0.51)
Median	1.10
Range	0.37–3.48
Skewness	1.77
Sub-range	N (%)
<0.90 (expansion)	39 (28%)
0.90–1.10 (equivalence)	32 (23%)
>1.10 (compression)	70 (50%)

3.2. Main effects

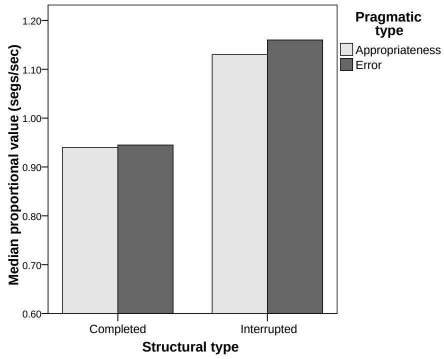
A first fit of a linear mixed-level model with the factors described above revealed main effects for structural type and the difference in number of segments between the (projected) reparandum and repair items. The model was refitted with only these factors, giving the output in Table 3.

Table 3: Mixed-level model for log proportional values (segs/sec) across core repair items.

Factor	F	*p*
Structural type	5.172	0.025
Difference in N segments	9.416	0.003

These results suggest that of the two crucial predictors pragmatic and structural type, only the latter has a significant effect on proportional values. Figure 2 illustrates the interaction between the two predictors. While completed repairs have a median proportional value in the 'equivalence' range, interrupted repairs are on average associated with temporal compression. This pattern is the same across error and appropriateness repairs.

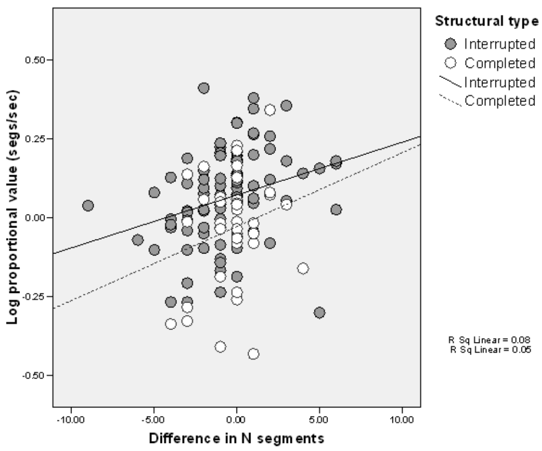
Figure 2: Effects of pragmatic and structural type on proportional values (segs/sec) across core repair items.



3.3. Predictive value

The predictive value of the two main effects can be illustrated in a regression plot. Figure 3 shows that the log proportional values are positively correlated with difference in number of segments between (projected) reparandum and repair. This means that the more complex the repair is relative to the reparandum, the more likely it is to be temporally compressed. The correlation is observed for both completed and interrupted repairs, with the latter showing a higher mean than the former.

Figure 3: Scatter plot of log proportional values (segs/sec) against difference in N segments, with structural type as grouping variable.



A stepwise multiple regression analysis with the factors structural type and difference in N segments between (projected) reparandum and repair gives the output in Table 4. It can be seen that structural type is the strongest predictor, and that together the two predictors account for approximately 13% of the variance of the dependent variable.

Table 4: Stepwise regression with the two predictors structural type and difference in N segments.

Model	Adjusted R^2	F change	*p*
Structural type	0.07	11.67	0.001
Structural type * Difference in N segments	0.13	9.08	0.003

4. DISCUSSION

The results of this study confirm that lexical error and appropriateness repairs are similar in temporal organization, and that repeated lexical items in self-repair are not invariably compressed relative to core repair items [9]. Notably, the study reveals a main effect of whether the reparandum is completed or not before the onset of repair: repairs following an interrupted reparandum are done relatively fast, while repairs following a completed reparandum are done at a similar articulation rate to that of the reparandum. This is consistent with the argument that repairs with an interrupted reparandum are likely to have been initiated at the stage of pre-articulatory self-monitoring, and that those that are initiated at this stage are more likely to be executed under time pressure than those initiated at the post-articulatory stage [7].

If the main effect of structural type is indeed due to a difference between repairs initiated in pre-articulatory vs post-articulatory monitoring, the observed greater tendency for appropriateness repairs to be interrupted can be interpreted as a tendency for pragmatic infelicities to be detected earlier than linguistic and factual errors – although this tendency is not strong enough to cause a main effect of pragmatic type on the temporal organization of the repairs. Further work is needed to establish the robustness of this observation. Moreover, it should be noted that this study found no difference between interrupted and completed repairs in the delay between the end of the reparandum and the onset of repair, and no correlation between delay duration and articulation rate values. Therefore, the interpretation of the findings in relation to those of [7] requires careful consideration. One avenue for further work is the implementation of a more fine-grained measure of the degree of completeness of the reparandum.

Interestingly, the results also suggest that the phonological complexity of the lexical items involved in the repair constrains its temporal organization. Again, the direction of the pattern is as might be expected, in accordance with observations that vowel duration decreases as the length of host words increases [6], and that word duration decreases with an increase in host utterance length [3]. The pattern is particularly intriguing in the context of interrupted repair: it suggests that the rate of articulation of a partial phonetic form may be constrained by the complexity of the form as a whole, even if repair is likely to have been initiated *before* the onset of articulation. Still, the measure of complexity implemented in this study was crude, and future work should address this weakness.

Finally, given that the two significant factors account for a small proportion of the variance in one phonetic dimension only, the question of the clustering of phonetic parameters in self-repair raised above remains an important one.

5. REFERENCES

[1] Baayen, R.H. 2008. *Analyzing Linguistic Data*. Cambridge: Cambridge University Press.
[2] Cutler, A., 1983. Speakers' conceptions of the function of prosody. In Cutler, A., Ladd, D.R. (eds.), *Prosody: Models and Measurements*. Heidelberg: Springer, 79-91.
[3] Jacewicz, E., Fox, R.A., Wei, L. 2010. Between-speaker and within-speaker variation in speech tempo of American English. *J. Acoust. Soc. Am.* 128, 839-850.
[4] Levelt, W.J.M. 1983. Monitoring and self-repair in speech. *Cognition* 14, 41-104.
[5] Levelt, W.J.M., Cutler, A. 1983. Prosodic marking in speech repair. *J. Semantics* 2, 205-217.
[6] Nooteboom, S. 1972. *Production and Perception of Vowel Duration in Dutch*. PhD Diss., Utrecht University.
[7] Nooteboom, S. 2010. Monitoring for speech errors has different functions in inner and overt speech. In Everaert, M., Lentz, T., De Mulder, H., Nilsen, Ø. (eds.), *The Linguistic Enterprise*. Amsterdam: John Benjamins, 213-233.
[8] Oostdijk, N.J.H. 2002. The design of the Spoken Dutch Corpus. In: Peters, P., Collins, P., Smith, A. (eds.), *New Frontiers of Corpus Research*. Amsterdam: Rodopi, 105-112.
[9] Plug, L. 2010. Phonetic reduction and informational redundancy in self-initiated self-repair in Dutch. *J. Phon.*, doi:10.1016/j.wocn.2010.08.001.
[10] Quené, H. 2007. On the just noticeable difference for tempo in speech. *J. Phon.* 35, 353-362.
[11] Shattuck-Hufnagel, S., Cutler, A. 1999. The prosody of speech error corrections revisited. *Proc. 14th ICPhS* San Francisco. Vol. 3, 1483-1486.