

THE STRUCTURE OF SELF-REPAIRS IN THE SPEECH OF HUNGARIAN LEARNERS OF ENGLISH

JUDIT KORMOS

Department of English Applied Linguistics
Eötvös Loránd University
19–21 Ajtósi Dürer sor
H–1146 Budapest
Hungary
kormos.j@chello.hu

Abstract: The study reported in the paper investigates the structure of L2 self-corrections in the speech of 30 Hungarian learners of English and 10 Hungarian native speakers. The aim of the research is to examine what the well-formedness of the corrections, the use of editing terms, the placement of cut-off points and the effect of the participants' level of proficiency on the structure of self-repairs reveal about the psycholinguistic processes of speech production. The results of the study lend additional support for modular models of speech production (e.g., Levelt 1983, 1989; Levelt et al. 1999) and reveal an important role of pragmatic constraints in psycholinguistic processing.

Keywords: speech production, self-correction, monitoring, editing, second language

1. Introduction

The self-correction behaviour of second language learners has already been explored from a number of aspects. In earlier studies it has been argued that the investigation of self-repairs in L2 speech can yield insights into mechanisms of speech production (e.g., Kormos 1999, 2000a,b; Levelt 1983, 1989; Poullisse 1999; Poullisse–Bongaerts 1994; van Hest 1996), processes of speech automatization (Kormos 2000b) and the allocation of attention (Kormos 2000a). In this regard, the distribution of self-repairs and its relation to the frequency of errors and to the development of proficiency as well as the timing of self-corrections was studied

by a number of researchers (for a comprehensive review of self-repair research see Kormos 1999). A less frequently investigated aspect of self-corrections in L2 is the structure of self-repairs. The existing studies in this field mainly used either native speakers of Dutch or Dutch learners of English and were descriptive in nature (with the exception of Brédart's (1991) study, whose participants were French). No inferential statistical analyses to support the claims were carried out in earlier research, and the effect of L2 proficiency on the structure of self-corrections was not investigated either. The present research aimed to fill in this gap and used Hungarian native speakers as well as Hungarian learners of English at three different levels of proficiency as participants (for an overview of Hungarian speakers' self-repair behaviour see Gósy 2004; Huszár 2005). The objectives of the research were to test whether the hypotheses set up in earlier research concerning native speakers of an Indo-European language are also tenable for native speakers of Hungarian as learners of English. The study also intended to investigate whether English language competence has any marked effect on the structure of self-corrections. The results of the study can help further refine or support existing theories of speech production and can shed new light on certain pragmatic aspects of self-corrections.

2. Review of literature

2.1. A brief overview of theories of speech production

Speech production researchers all agree that language production has four important components: (1) **conceptualization**, that is, planning what one wants to say, (2) **formulation**, which includes the grammatical, lexical and phonological encoding of the message, (3) **articulation**, in other words, the production of speech sounds and (4) **self-monitoring**, which involves checking the correctness and appropriateness of the produced output. There is also agreement on the questions that conceptualization, formulation and articulation follow each other in this order, and that in L1 production planning the message requires attention, while formulation and articulation are automatic, and therefore processing mechanisms can work in parallel, which makes L1 speech generally smooth and fast. Researchers also share the view that one of the basic mechanisms involved in producing speech is **activation spreading**. Activation spreading is a metaphor adapted from brain research, which is based on the finding of

neurological studies that neural networks consist of interconnected cells (neurons) that exchange simple signals called activations via the connections they have with each other (Hebb 1949). The speech processing system is assumed to consist of hierarchical levels (conceptualization, formulation, articulation), among which information is transmitted in terms of activation spreading, and of knowledge stores such as the lexicon and conceptual memory store, within which activation can also spread from one item to related items. Decisions are made on the basis of the activation levels of the so-called **nodes** that represent various units such as concepts, word forms, phonemes etc.

There exist two major theories of L1 speech production: **spreading activation** (the name is somewhat misleading because as just mentioned, both models assume that the way information is transmitted in the speech processing system is activation spreading) (e.g., Dell 1986; Dell–O’Seaghdha 1991; Stemberger 1985) and **modular theories** (e.g., Fry 1969; Garrett 1976; Laver 1980; Levelt 1989; Levelt et al. 1999; Nootboom 1980), and there are two major differences between them. The first main difference between these two theories is whether they allow for feedback between the various levels of encoding. Spreading activation models allow for the backward flow of activation from a subordinate level to the superordinate level, while in modular theories activation can only spread forwards. This means that in spreading activation theory, if an error occurs in one specific process, a warning signal is immediately issued, and activation flows upwards to the superordinate level. Processing starts again from this superordinate level. In modular models the error is not noticed at the level it is made, but only once the erroneous fragment of speech has been phonologically encoded or later when it is articulated. Therefore, in this view, bits of message that contain an error need to be encoded again from the level of conceptualization. Researchers working with modular theories argue that the processing components in the speech production system are autonomous, that is, have their own characteristic input, and they process this input independently of other components. Hence the name ‘modular theory of speech production’. The second major difference between these theories concerns syntactic and phonological encoding. In spreading activation theories it is assumed that speakers first construct frames for sentences and for phonetic representations and then select the appropriate words or phonetic features for the slots in the frame. Modular models are **lexically driven**, which means that words activate syntactic building procedures, and they postulate that lexical en-

coding precedes syntactic encoding and that phonological encoding can only start once lexico-syntactic processes are ready.

2.2. Overview of earlier studies on the structure of self-repairs

The structure of self-repairs can be characterised by three variables. The first of these concerns the syntactic relationship of the reparandum and the reparatum, which provides us with information what rules govern how these two constituents of the self-correction are conjoined. The second related characteristic is the use of editing terms, which reveals by what lexical means this conjunction is expressed. Finally, the point where the erroneous or inappropriate utterance is interrupted also lends insight into the structural aspects of the repair.

Several studies have investigated whether the syntactic structure of self-repairs shows any signs of systematicity (e.g., de Smedt–Kempen 1987; Levelt 1983) and have found that the majority of self-corrections follow a specific rule, which was named the Well-formedness Rule by Levelt (1983). (Previous investigations concerning this field were discussed in detail by Kormos (1999), therefore, only a very brief summary of earlier findings will be presented here). According to the rule, “an original utterance ⟨O⟩ plus repair ⟨OR⟩ is well-formed if and only if there is a string of zero or more words ⟨C⟩ to complete the utterance so that the string ⟨OC or R⟩ is well-formed, where C is a completion of the constituent directly dominating the last element of O” (Levelt 1983, 78). In other words, this rule says that the utterance and the repair have to follow the rules of syntactic coordination. Examples (1) and (2) illustrate well-formed repairs, while examples (3) and (4) ill-formed ones.

- (1) all chairs *have handles*. *And er sorry arms*
- (2) how many people er would will *come there or here* yes?
- (3) you *can this er reserve this* er er room
- (4) it’s almost er er the number we are *the rules er which i... which is written in our rules*

The only study that investigated the well-formedness of L2 self-repairs was conducted by van Hest (1996), who found little difference between L1 and L2 self-repairs in this respect. Following Levelt’s (1983) rule of classification, 70% of the L1 self-repairs and 80% of the L2 self-repairs were

well-formed in Van Hest's corpus. On the basis of van Hest's results, one might claim that the self-repair behaviour of L2 learners also follows the well-formedness rule. Her results, however, might only be due to the fact that the syntactic structure of Dutch and that of English are similar in a number of aspects. Nevertheless, Kormos (1999) argued that van Hest's results can be taken as a proof for modular models of speech production in L2 (e.g., de Bot 1992; Levelt 1989; Levelt et al. 1999) as the findings indicate that L2 speakers re-process the whole speech plan when making a correction and do not restart the utterance from an intermediary level of production. The present research aims to substantiate this claim with data collected from Hungarian speakers.

Earlier studies on the use of editing terms concluded that the default repair structure both in L1 and L2 contains no editing term (e.g., words and expressions such as *I mean, well, or*), but if editing expressions are used, they are most likely to be filled non-lexicalised pauses (Levelt 1983; van Hest 1996). Levelt (1983) found that error-repairs were more frequently accompanied by editing terms than repairs which were concerned with the appropriacy of information. In van Hest's (1996) corpus, however, there was little difference between these two types of self-corrections in this respect. Similarly to Poulisse and Bongaerts' (1994) study, van Hest's data also suggest that L2 speakers frequently use L1 editing terms when speaking an L2 which, in most of the cases, can be regarded as unintentional code-switching. It has to be noted that these studies only used descriptive statistics and correlational analyses on the cooccurrence of editing terms and different types of self-repairs were not performed. In a recent study, Clark and Fox Tree (2002) argued that it is not only the choice of certain lexical editing terms, such as *well, I mean* etc., that is systematic, but speakers also make conscious decisions when using *uh* and *um* as delay signals. On the basis of the analysis of the spoken data in the London-Lund corpus, they claimed that *uh* signals minor delay in the speech production process, whereas *um* indicates a major delay.

The third aspect of the structure of self-corrections concerns the point of interruption of the erroneous or inappropriate utterance. The Main Interruption Rule, which was first proposed by Nootboom (1980) and later elaborated by Levelt (1983) says that speakers stop the flow of speech immediately upon the detection of trouble. This rule suggests that regardless of the nature of trouble in the utterance, the reparandum will be interrupted immediately after the problem is perceived, and word or syllable integrity will not be respected in these cases. Levelt (1983) and

Brédart (1991), however, observed that in their study speakers tended to finish inappropriate words, while error-repairs adhered to the Main Interruption Rule. Van Hest's (1996) study with L2 learners reproduced the same result. This finding also needs to be substantiated with learners of English and speakers from a different native language background.

3. Research questions

The study reported in this paper investigated the following issues:

1. Do self-repairs in the speech of Hungarian speakers and learners of English also follow the well-formedness rule?
2. To what extent does well-formedness depend on the type of repair in the corpus?
3. How does the language proficiency of the participants affect the rate of well-formed repairs?
4. What is the relationship between the types of self-correction and the editing terms used by Hungarian speakers and learners of English?
5. What is the effect of language proficiency on the frequency and type of editing terms used by the participants?
6. Where is the point of interruption in the reparandum in the speech of Hungarian speakers and learners of English?
7. Does the language proficiency of the participants influence the placement of cut-off points?

4. Method

4.1. Settings, procedures and participants

The data for this study came from a corpus of speech samples collected from 30 Hungarian learners of English and 10 Hungarian native speakers. The database and the data collection procedures were discussed in detail in Kormos (2000a,b), therefore, the research design will be described here only briefly. Participants of the study performed a role-play task with the researcher being the interlocutor. The task was recorded and an immediate retrospective interview followed the role-play activity. The interview was conducted in accordance with the guidelines drawn up by Ericsson and Simon (1980; 1993). A C-test,¹ which had been validated

¹ In a C-test, the second half of every second word is missing.

by Dörnyei and Katona (1992) was administered to all the participants to measure their level of proficiency.

Participants of the study were all native speakers of Hungarian aged between 16 and 35. Ten students were learners in a language school and their language proficiency was at the intermediate level. Twenty participants were English majors at Eötvös Loránd University, Budapest, whose English competence was advanced. Half of the English majors performed the task in Hungarian and half of them in English. (It was not possible to collect Hungarian and English data from the same participants because of the strong effect task-repetition would have had on performance). Ten students were participants of an evening training course at Eötvös Loránd University, and their proficiency was upper-intermediate. The one-way analysis of variance of the C-test scores showed that the level of proficiency of the three groups was significantly different ($F = 130.45$; $p < 0.001$).

Both the speech produced while performing the task and the text of the retrospective interview were recorded and transcribed by trained research assistants. Transcriptions were checked by the researcher.

4.2. Analysis

The self-repairs were identified and classified in the texts with the help of the retrospective interviews. Four major types of self-repairs were established with sub-categories within each group. **Different information (D-) repairs** involve the decoding of different information than the speaker is currently formulating. **Appropriacy (A-) repairs** are used when the speaker modifies the originally intended information. **Error repairs** involve corrections of accidental lapses in accessing words (**lexical error repairs**), reprocessing the grammatical structure of the utterance (**grammatical error repairs**) and revised phonological encoding mechanisms (**phonological error repairs**). **Rephrasing-repairs** are used by L2 speakers when they are uncertain about the correctness of the original utterance, and in this case they encode their original message in a different form. (For subtypes of D- and A-repairs as well as for examples see the appendix).

The analysis of well-formedness employed in the study differed slightly from previous research in this field. If one considers Levelt's (1983) rule presented above, it does not specify how one should proceed in the case of within-word interruptions, as it only says that there should be a

string of zero or more words to complete the reparandum. In the case of within-word interruptions, however, it is often a fragment of a word (and frequently not even a well-formed syllable) that is needed to finish the original utterance. Due to the fact that in these cases it can only be speculated how speakers intended to utter their message, no accurate judgements can be made concerning the grammaticality and well-formedness of repairs with within-word interruptions. Therefore, these types of repairs (N = 33, 11.4% of all the repairs in the corpus in the L2 corpus and N = 8, 16.3% in the L1 corpus) were excluded from the analysis of well-formedness in this study. The L1 corpus consisted of 59 and the L2 corpus of 289 self-repairs altogether. It has to be noted here that the number of L1 self-repairs was not high enough to make overall generalizations concerning the structure of self-corrections in Hungarian, but it was sufficient as background to the L2 data.

Editing terms preceding the reparatum were all identified together with the point of interruption of the reparandum in the transcripts. For the analysis of the effect of proficiency on the structure of self-repairs one-way analysis of variance and Chi-square analysis and for the investigation of the relationship of the type and structure of repair Chi-square analysis was used. The level of significance was set for $p < 0.05$ in the study.

5. Results and discussion

5.1. The well-formedness of repairs

As can be seen in Table 1, 87.3% of the repairs in the L2 data-base were well-formed. The highest frequency of ill-formed repairs could be found within the subtype of grammatical error-repairs (see Example 3 above) and message-replacement repairs (i.e., repairs when the speaker abandons the original speech plan and encodes a completely new message). Due to the high number of cells with values lower than 5, no meaningful Chi-square statistics could be computed with subtypes of repairs. The analysis of the relationship of main types of repairs, however, showed that the type of correction significantly influences whether the correction is well-formed ($\chi = 13.78$, $p = 0.003$). Error-repairs were ill-formed more frequently than expected, whereas appropriacy repairs were more frequently well-formed than expected.

Table 1
The distribution of well-formed and ill-formed repairs in L2

Type of repairs	WELL-FORMED				ILL-FORMED			
	Count	Expected value	Row per-cent. (%)	Residual	Count	Expected value	Row per-cent. (%)	Residual
EL-repairs	33	28.8	100.0	4.2	0	4.2	0	-4.2
EG-repairs	26	38.4	59.1	-12.4	18	5.6	40.9	12.4
EF-repairs	8	7.9	88.9	0.1	1	1.1	11.1	-0.1
AL-repairs	39	34.9	97.5	4.1	1	5.1	2.5	-4.1
AA-repairs	3	2.6	100	0.4	0	0.4	0	-0.4
ALC-repairs	2	1.7	100	0.3	0	0.3	0	-0.3
AP-repairs	10	8.7	100	1.3	0	1.3	0	-1.3
AG-repairs	7	6.1	100	0.9	0	0.9	0	-0.9
DM-repairs	34	34.9	85.0	-0.9	6	5.1	15.0	0.9
DI-repairs	16	14.8	94.1	1.2	1	2.2	5.9	-1.2
DO-repairs	3	2.6	100	0.4	0	0.4	0	-0.4
R-repairs	38	37.5	88.4	0.5	5	5.5	11.6	-0.5
Rest	4	3.5	100.0	0.5	0	0.5	0	-0.5
Total	219		87.3		32		12.7	

EL-repairs = lexical error-repairs, EG-repairs = grammatical error-repairs; EF-repairs = phonological error repairs; AL-repairs = appropriate level of information-repairs; AA-repairs = ambiguous reference-repairs; ALC-repairs = coherent terminology repairs; AP-repairs = pragmatic appropriacy repairs; AG-repairs = repairs for good language; D-repairs = different information repairs; DM-repairs = message replacement-repairs; DI-repairs = inappropriate information repairs; DO-repairs = ordering error-repairs; R-repairs = rephrasing-repairs

The Chi-square analysis of the relationship of proficiency grouping and well-formedness showed a significant effect of foreign language competence ($\chi = 6.69$, $p = 0.03$). In the intermediate group ill-formed repairs occurred with a higher frequency than expected whereas in the speech of advanced learners there were more well-formed repairs than expected (see Table 2, overleaf). As for the effect of language, the results showed that the percentage of well- and ill-formed repairs did not differ significantly in Hungarian and in English as L2 ($\chi = 1.15$, $p = 0.28$).

In the Hungarian data-base 73.2% of the repairs were well-formed and 26.8% ill-formed. The Chi-square analysis showed a significant relationship between the main type of repair and well-formedness ($\chi = 11.78$, $p = 0.002$). However, 33.3% of the cells did not have an expected frequency above 5, which shows that the corpus is too small for drawing a firm conclusion on the basis of the findings. The results indicate

Table 2
The relationship of well-formedness
and the level of proficiency in L2

Level of proficiency (count, expected, row %, residual)	Ill-formed	Well-formed
Intermediate	16	64
	10.2	69.8
	20.0%	80.0%
	5.8	-5.8
Upper-intermediate	10	72
	10.5	71.5
	12.2%	87.8%
	-0.5	0.5
Advanced	6	83
	11.3	77.7
	6.7%	93.3%
	-5.3	5.3
Total	32	219
	12.7%	87.3%

$\chi = 6.69, p = 0.03$

that different-information repairs in Hungarian were more frequently ill-formed than expected.

The findings suggest that in most of the cases, the self-repair behaviour of Hungarian learners and native speakers is also governed by the well-formedness rule. Thus well-formedness seems to be a phenomenon independent of language background. This lends additional support to Kormos's hypothesis that both L1 speakers and L2 learners "are able to store the original syntactic structure of their message in working memory" (1999, 329) and adjust the reparatum to it. The hypothesis that both native and non-native speakers implement the correction by reproducing the syntactic environment of the reparandum, that is, "by grammatically encoding the relevant part of the message anew" (*idem.*) also seems to be supported. The findings of the study lend strong support to the assumption that speech production usually does not start from the intermediary levels where the error was made (Kormos 1999), thus the results indirectly support modular models of speech production (e.g., Levelt 1989; Levelt et al. 1999) and Levelt's (1989) model of monitoring. In this the-

ory three monitor loops are assumed to be responsible for inspecting the outcome of the production processes. The first loop involves the comparison of the preverbal plan with the original intentions of the speaker. The second loop concerns the monitoring of the phonetic plan (i.e., ‘internal speech’) before articulation, which is also called ‘covert monitoring’ (see also Postma–Kolk 1992, 1993; Wheeldon–Levelt 1995). Finally, the generated utterance is also checked after articulation, which constitutes the final, external loop of monitoring, involving the acoustic-phonetic processor. Upon perceiving an error or inappropriacy in the output in any of these three loops of control, the monitor issues an alarm signal, which, in turn, triggers the production mechanism for a second time starting from the phase of conceptualization.

The findings of this study also suggest that well-formedness seems to be a universal phenomenon in speech production. This does not necessarily mean that this study lends support to theories claiming that L2 speakers have access to Universal Grammar (Chomsky 1965). It is rather the case that language users in general aim to produce well-formed sentences, otherwise their interlocutors have difficulty understanding them. Therefore, even if the utterance contains a slip, error or inappropriacy, speakers will strive to adjust the syntactic structure of the newly formulated message to the preceding part of the utterance. Since certain maxims of conversation are supposed to be universal such as the maxim of clarity (Grice 1975), speakers—regardless of L1 and independent of the fact whether they speak their L1 or another language—will aim to make their message clear and unambiguous.

In comparison with previous research that investigated the syntactic structure of self-repairs, it can be stated that the rate of well-formed self-corrections in this project is between the percentages in Levelt’s (1983) (L1: 98%) and van Hest’s (1996) (L1: 70% and L2: 80%) studies. Van Hest explained the lower rate of well-formed repairs found in her corpus with reference to the fact that in her study the task to be performed by the students was more complex than in that of Levelt (1983). She argued that, for this reason, participants of her project were required to use syntactically varied sentences, which resulted in a relatively low percentage of well-formed repairs. Accordingly, van Hest assumed that the higher rate of well-formed repairs in L2 was due to the fact that utterances of speakers are more complex in L1 than in L2, therefore, self-corrections in L2 are syntactically easier to implement than those in L1. It has to be noted that van Hest only used descriptive statistics in

her study, therefore it is not known whether the difference in the rate of well-formed repairs between L1 and L2 was statistically significant. In the present study well-formedness was not significantly influenced by whether the participant spoke L1 and L2. In addition, van Hest's line of argumentation would also mean that with increasing proficiency L2 speakers make a decreasing number of well-formed self-repairs, which seems to contradict the results of this study where an opposite effect was found.

The finding that 20% of the self-repairs produced by intermediate learners was ill-formed as opposed to the 6.7% of ill-formed repairs in advanced learners' speech can be explained with reference to the psycholinguistic mechanism of self-correction and theories of attention. From psychological research it is well known that attention is limited due to the constraints of working memory (Gathercole–Baddeley 1994). It is also an accepted fact that the less automatic a mechanism is, the more attention is required for it (for a review see Schmidt 1992). Thus it seems that intermediate learners need to pay so much attention to other aspects of encoding their message anew that they are often unable to keep the syntactic structure of their original message in their working memory. In turn, this frequently results in ill-formed self-corrections.

As regards the varying percentage of well-formed and ill-formed repairs in the case of different subtypes of self-repairs in L2, it is understandable that grammatical error-repairs do not always follow the well-formedness rule. When implementing changes in the syntactic structure of the message (e.g., in the word order), the coordination of the same phrasal category is often impossible. When producing message-replacement repairs, L2 speakers need to conceptualise a completely new message instead of the one to be replaced. The structure of this new utterance might not follow the previous one due to the lack of attentional resources, as learners might be so overloaded with the cognitive planning of the repair that they have no attention to spare to adjust the syntactic structure of the reparatum to that of the reparandum. Results concerning the time necessary for re-planning message-replacement repairs also support the increased cognitive load in this case (Kormos 2000b). In the case of L1 self-corrections, the corpus is too small to observe any meaningful relationship between the type of repair and well-formedness.

5.2. The use of editing terms

With respect to the editing terms used upon making the repair, the research project investigated two important issues: the relationship of the types of self-correction and the editing terms used by the participants as well as the effect of language proficiency on the frequency and type of editing terms.

As can be seen in Table 3 (overleaf), the results of the research project replicated previous findings concerning the default repair structure in L2, since 55% of the self-corrections contained no editing term. Similarly to earlier studies in this field, it was found that in the majority of the cases when editing expressions were used, the repair was merely signalled by the presence of filled non-lexicalized pauses, such as *er* or *uhm*. Among lexicalized editing terms *or* was most frequently applied, and interestingly, Hungarian expressions were very rarely used for this purpose. Certain editing terms tended to cooccur with certain types of repairs. *Sorry* was employed exclusively with error-repairs, whereas *I mean* always indicated appropriacy or different information repairs.

The relationship of the main types of self-corrections and editing terms was also studied by means of Chi-square analysis. The results of this investigation indicated that certain types of repairs cooccurred with specific editing terms ($\chi^2 = 19.47$, $p = 0.0001$). Error-repairs contained unfilled pauses with higher frequency than lexicalized editing terms, while different information-repairs tended to be signalled by lexicalized expressions more often than by unfilled pauses.

Editing terms were used differently in Hungarian than in English. Interestingly the default structure of repairs in Hungarian was not the lack of editing term, but the use of the term *tehát* 'so'. Due to the small number of L1 self-repairs, no meaningful Chi-square statistics could be computed concerning the relationship of types of repairs and editing terms in Hungarian.

The influence of the level of proficiency on the frequency of the use of editing terms was also analysed by means of one-way analysis of variance. The results of the study indicate limited effect of proficiency in this respect. The frequency of the editing terms was not significantly affected by the participants' level of competence in English. It could be observed, however, that certain editing terms such as *I mean* or *well* were merely applied by highly proficient speakers, while Hungarian editing terms occurred only in the speech of participants with low level of competence.

Table 3

The distribution of editing terms across different types of self-repairs in L2

Type of repairs (count, row %)	Unfilled pause	er, uhm	so	or	sorry	I mean	and	well	other	ja (ugh)	vagy (or)
EL-repairs	22 53.7	10 24.4	0	2 4.9	4 9.8	0	0	0	1 2.4	0	2 4.9
EG-repairs	33 67.3	15 30.6	0	0	0	0	1 2.0	0	0	0	0
EF-repairs	19 11.9	3 3.8	0	0	0	0	0	0	0	0	0
AL-repairs	21 50.0	8 19.0	4 9.5	7 16.7	0	2 4.8	0	0	0	0	0
AA-repairs	2 66.7	1 33.3	0	0	0	0	0	0	0	0	0
ALC-repairs	1 50.0	1 50.0	0	0	0	0	0	0	0	0	0
AP-repairs	8 72.7	2 18.2	0	0	0	0	1 9.1	0	0	0	0
AG-repairs	6 75.0	1 12.5	0	0	0	1 12.5	0	0	0	0	0
DM-repairs	18 41.9	18 41.9	2 4.7	1 2.3	0	2 4.7	0	1 2.3	1 2.3	0	0
DI-repairs	6 35.3	1 5.9	0	5 29.4	0	0	0	0	4 23.5	1 5.9	0
DO-repairs	1 33.3	1 33.3	0	0	0	0	0	1 33.3	0	0	0
R-repairs	22 50.0	16 36.4	1 2.3	2 4.5	0	0	1 2.3	1 2.3	1 2.3	0	0
Total	159 55.0	80 27.7	7 2.4	18 6.2	4 1.4	5 1.7	3 1.0	3 1.0	7 2.4	1 0.3	2 0.7

E-repairs = error repairs; EL-repairs = lexical error-repairs, EG-repairs = grammatical error-repairs; EF-repairs = phonological error-repairs; A-repairs = appropriacy-repairs; AL-repairs = appropriate level of information-repairs; AA-repairs = ambiguous reference-repairs; ALC-repairs = coherent terminology repairs; AP-repairs = pragmatic appropriacy repairs; AG-repairs = repairs for good language; D-repairs = different information repairs; DM-repairs = message replacement-repairs; DI-repairs = inappropriate information repairs; DO-repairs = ordering error-repairs; R-repairs = rephrasing-repairs

The findings of the present study lend strong support to previous assumptions concerning Dutch and English repairs, namely, that in most of the cases speakers do not use any signals except for an unfilled pause for indicating that a repair will be made. One of the reasons for this might be

that making repairs hinders fluent expression, and the use of long editing terms would make the flow of speech even more disrupted. Another possible explanation is that in most of the cases when a correction is implemented, it is obvious that the speaker intends the listener to ignore the previously uttered message (e.g., in the case of non-existing words and grammatical errors); thus, there is no need to signal it. This seems to be especially true in the case of L2 error-repairs, which are most frequently accompanied by unfilled pauses in the present corpus.

With respect to different-information repairs in L2, however, listeners frequently need to be warned that the message to be conveyed next might not logically follow the previously uttered one. For example, in cases when part of a message is totally replaced by a new one (message-replacement repairs), or when the previously provided information has been false (inappropriate information repair), the listeners' activated schemata of what to expect as a continuation can be disturbed by the repair. Thus, speakers might perceive that a warning signal is more essential upon producing these types of corrections than in cases when it is the reparandum (e.g., the slip of the tongue) and not the reparatum which does not fit the expectations of the listener. The use of editing terms in these cases is very similar to that of **verbal strategy markers**, which can stand "before or after a (communication) strategy to signal that the word or structure does not carry the intended meaning perfectly in the L2 code" (Dörnyei-Scott 1997, 191). The similarity of these two markers lies in the fact that they both aim to elicit cooperation from the interlocutor and to achieve that the listener and the speaker share the same meaning.

The results concerning the cooccurrence of different types of editing terms and repairs in L2 indicate that the participants of the study use these expressions in a similar way as native speakers of English do. They are aware of the fact that *I mean* is applied for further specifying the informational content of the message, and apologising (*sorry*) is only necessary after having made an error (DuBois 1974; cf. Levelt 1989). Moreover, these results also indicate that the use of editing terms is not random, but systematic both in L1 and L2 speech. Most editing terms including filled pauses tend to cooccur with specific types of repairs, that is, they signal a given reason for communication breakdown. This supports Clark's and Fox Tree's (2002) proposal concerning the process of self-corrections, in which they claim that speakers consciously and systematically select the signal for making a repair.

The infrequent use of L1 editing terms by the participants of this project seems to contradict Poulisse and Bongaerts' (1994) and van Hest's (1996) findings. In both of these studies beginning and intermediate learners transferred editing expressions from L1 to L2 considerably often, whereas in the present corpus L1 editing terms occurred only in the speech of intermediate learners and with very low frequency. One of the possible explanations for this difference might be that due to the fact that the L1 of the participants of both Poulisse and Bongaert's (1994) and van Hest's (1996) study was Dutch. Dutch learners of English might be more willing to transfer the editing expressions from their L1 to L2 than Hungarian speakers, whose mother tongue has few characteristic features in common with English (for the effect of the proximity of languages on transfer see Kellerman 1979).

The effect of language proficiency on the use of editing terms seems to be limited, the reason for which can be that it is rather the nature of the repair that determines what type of editing expression will be used than the level of L2 competence. Nevertheless, it can be observed that only advanced learners apply *well* and *I mean* for signalling the correction, which is probably the consequence of the lack of instruction of discourse markers in most of the language courses in Hungary.

It is interesting to observe that despite the fact that there were very few error repairs in the Hungarian corpus, the speakers used a lexical editing term in 69.4% of the cases and that Hungarian speakers used the editing term *tehát* 'so' the most frequently. Due to the small number of participants and self-corrections, one can only speculate that perhaps Hungarian speakers prefer to signal corrections with the term *tehát* 'so' rather than simply to use a non-filled pause.

5.3. The point of interruption

The point of interruption was also investigated both in L2 and in L1 of the participants. As can be seen in Table 4, interruptions within the troubleword occurred infrequently in the L2 corpus of self-repairs of the present study. The highest frequency of within-word cut-off points could be found in the case of phonological error-repairs, and words containing lexical or grammatical errors were also frequently interrupted. The results suggest that the flow of speech is mainly halted after the reparandum has been uttered. Late interruptions were rare in the corpus.

Table 4

The placement of cut-off points in the case of various sub-types of self-repairs in L2

Type of repairs (count, row %)	Within the trouble word	Immediately following the trouble word	Within the word fol- lowing the trouble word	1 word after the trouble word	2 words after the trouble word	Later than 2 words after the trouble word
EL-repairs	7 17.5	23 57.5	2 5.0	5 12.5	2 5.0	1 2.5
EG-repairs	6 12.0	32 64.0	1 2.0	3 6.0	6 12.0	2 4.0
EF-repairs	13 59.1	8 36.4	0 0.0	1 4.5	0 0.0	0 0.0
ALC-repairs	2 4.8	32 76.2	1 2.4	3 7.1	2 4.8	2 4.8
AA-repairs	0 0.0	3 100.0	0 0.0	0 0.0	0 0.0	0 0.0
AC-repairs	0 0.0	1 50.0	0 0.0	1 50.0	0 0.0	0 0.0
AP-repairs	1 9.1	8 72.7	1 9.1	1 9.1	0 0.0	0 0.0
AG-repairs	1 12.5	5 62.5	0 0.0	1 12.5	1 12.5	0 0.0
DM-repairs	2 4.7	38 88.4	1 2.3	2 3.0	0 0.0	0 0.0
DI-repairs	0 0.0	13 76.5	0 0.0	1 5.9	2 11.8	1 5.9
DO-repairs	0 0.0	3 100.0	0 0.0	0 0.0	0 0.0	0 0.0
R-repairs	1 2.3	34 77.3	1 2.3	2 4.5	3 6.8	3 6.8
Total	33 11.6	200 70.2	7 2.5	20 7.0	16 5.6	9 3.1

EL-repairs = lexical error-repairs; EG-repairs = grammatical error-repairs; EF-repairs = phonological error-repairs; AL-repairs = appropriate level of information-repairs; AA-repairs = ambiguous reference-repairs; ALC-repairs = coherent terminology repairs; AP-repairs = pragmatic appropriacy repairs; AG-repairs = repairs for good language; D-repairs = different information repairs; DM-repairs = message replacement-repairs; DI-repairs = inappropriate information repairs; DO-repairs = ordering error-repairs; R-repairs = rephrasing-repairs

As the number of cells with values lower than 5 was high when all the types and possible cut-off points were computed, meaningful Chi-square statistical analyses concerning the interruption pattern of the different

main groups of self-repairs in L2 could only be performed when the cut-off points were divided into three categories: (1) interruptions within the trouble word, (2) interruptions immediately following the trouble word, and (3) interruptions later than following the trouble word. The main groups of self-repairs varied as regards the placement of the cut-off points to a significant extent ($\chi^2 = 29.97$, $p = 0.0001$). The results indicate that erroneous words were interrupted with a higher frequency than expected, whereas inappropriacies, words containing false information, and lexical items in the correctness of which speakers were uncertain tended to be completed.

As error-repairs were found to behave in a different manner concerning the placement of cut-off points, the interruption pattern of each subtype of this group of corrections was also analysed by means of Chi-square statistics. The results indicate that it is mainly the class of phonological error-repairs that contributed to the above observed differences, as words containing phonological errors were more frequently interrupted than the expected value ($\chi^2 = 21.68$, $p = 0.0001$). As regards the other two types of error-repairs, no considerable deviations from the usual distribution of cut-off points could be observed. The level of proficiency was not found to affect the placement of cut-off points ($\chi^2 = 1.07$; $p = 0.89$).

In the Hungarian corpus 71.4% of the repairs were interrupted immediately after the trouble word, 16.3% of the repairs were within-word interruptions, in 10.2% of the repairs the interruption took place one word after the trouble word and in 2% of the cases two words after the trouble word. Due to the small number of repairs, no meaningful Chi-square statistics could be computed concerning the relationship of the type of repair and the point of interruption.

The placement of cut-off points in relation to word boundaries is an unreliable indicator of detection processes, therefore, this aspect will not be discussed here. Thus, the conclusions one can draw from the results concerning the interruption pattern of different types of repairs are limited to the structural and pragmatic aspects of this issue.

Similarly to previous studies in this field (e.g., Brédart 1991; van Hest 1996), the findings of the present project suggest that Hungarian learners of English also interrupt erroneous words, while inappropriate lexical items tend to be completed. This can be explained with reference to the fact that erroneous words need to be ignored by the listener when decoding the interlocutor's message, consequently, they tend to be interrupted. Inadequate information, however, only needs further speci-

fication, and as words containing inappropriacies often carry meaning for the listener (Brédart 1991), they will be completed. No such tendency could be observed in the case of Hungarian native speakers, which might be due to the small number of repairs in the corpus. The findings concerning L2 self-repairs indicate that not only the choice of editing terms and lexical fillers is conscious and systematic but also decisions concerning the interruption of the flow of speech when a correction is needed. This shows that the process of monitoring and self-repair is a conscious process, which is governed by universal pragmatic constraints that require communication to be clear and unambiguous.

6. Conclusion

The investigation of the structure of self-repairs in the speech of Hungarian learners of English yielded a number of results which confirm earlier studies on the monitoring behaviour of L2 speakers from a Dutch native speaking background. From the analysis of the well-formedness of self-repairs, it was concluded that both self-corrections in Hungarian and in the speech of Hungarian learners adhere to the rules of syntactic coordination. This implies that in most cases L2 speakers also store the syntactic structure of the reparandum in their working memory and adjust the reparatum to it. In addition, these findings also support earlier assumptions that the majority of self-repairs do not involve the mere substitution of one lexical entry with another one, but the encoding of the relevant part of the speech plan completely anew. Therefore, the study lends additional support for modular models of speech production (e.g., Levelt 1989; Levelt et al. 1999). A new finding of the research was that the level of proficiency has a significant effect on the rate of well-formed repairs because intermediate learners produced fewer well-formed repairs than advanced students. This was explained with reference to the limited attentional resources of low level learners.

The findings concerning the use of editing terms in L2 also confirmed the assumptions of previous studies, namely, that the default repair structure contains no editing term, and if editing terms are used, they are most frequently unfilled pauses. L2 speakers only signal the repair with a lexical editing term if they assume that the reparatum might not fit the listener's schemata activated by the reparandum, for example in the case of the abandonment of the original message and when the original information provided was inappropriate. In this case editing terms

can also be interpreted as verbal strategy markers, which indicate that the speaker needs the interlocutor's attentive cooperation (Dörnyei–Scott 1997). When L2 speakers presume that their listeners will easily perceive that the previously uttered message is erroneous (e.g., because it contains a non-existing word or a grammatical error), and it needs to be ignored, they do not signal that a repair will be made, as the use of an editing term would make the flow of speech even more disruptive. The fact that the level of L2 competence was not found to affect the frequency of the use of various editing terms also supports the assumption that the major factor that determines the use of editing terms is the type of repair, that is, the nature of error. The interruption pattern of errors and inappropriacies upon making the repair in L2 was found to show analogous tendencies. Erroneous words that speakers intended to be ignored tended to be interrupted more frequently than parts of utterances conveying inadequate information as the latter types of reparandum could carry meaning for the listener and needed only further specification.

The results of the study indicate that it is not only the process of monitoring that is conscious and controlled in nature, but speakers also make conscious and systematic decisions concerning the interruption of the flow of speech and the selection of editing terms. These decisions are primarily governed by the Gricean (1975) maxim of clarity, which requires that utterances in a conversation should be clear and unambiguous. This constraint involves considering the communicative situation and context and the interlocutor's perspective, which shows a strong relationship between pragmatic and discourse knowledge and psycholinguistic processing.

Appendix

The classification of self-repairs (quoted from Kormos 2000a, 380–3)

Name of repair	Definition	Example
Different information (D-) repair	The speaker decides to encode different information from the one he/she is currently formulating (Levelt 1983)	
Inappropriate information (DI-) repair	The speaker repairs the message because its information content is faulty (Levelt 1983)	The room is er uhm eer thirty thirty thousand er too much er ten thousand er forint er forints per day

Ordering error (DO-) repair	The speaker decides to encode parts of the intended message in different order (Levelt 1983)	Well, we it's it's about a thousand Forints
		Retrospection: First I wanted to answer the second question, but then I realized that I should answer the first question first.
Message abandonment (DM-) repair	The speaker abandons the originally intended message and replaces it with a different one	we have some er er v. . . maybe you have vegetarians in your group
		Retrospection: Here the idea of vegetarians suddenly popped up, and I abandoned what I was going to say because I would not have been able to list any more types of food anyway.
Appropriacy (A-) repair	The speaker decides to encode the original information in a modified way (Levelt 1983)	
Appropriate level of information (AL-) repair	The speaker decides to further specify the original message (Levelt 1983).	There are very wide choice of er main courses er er steak er er several kind of steak
		Retrospection: I wanted to say it more precisely that we do not only have one kind of steak but several kinds of steak.
Ambiguous reference (AA-) repair	The speaker repairs the referring expression because of ambiguity (Levelt 1983).	And you have to pay extra for the drinks. Then you have to negotiate that and talk about the drinks with the barman.
		Retrospection: I corrected what I said because it was not clear whether you have to talk about the drinks or the price with the barman.
Coherent terminology (AC-) repair	The speaker repairs incoherent terminology (Levelt 1983).	in this case er if it is so urgent and important for you, we would like er you to:: to write us an order —er in er 24 hours that you make sure that you will er come and book this eel room.
		...
		R: I see, all right and then I can only pay the deposit next week when I er find out how many people come and when I have talked to all of the people.
		S: Er but this letter is er— the order —er your request is er anyway—needed and we::

		Retrospection: I remembered that I had used the word 'order' earlier, and I wanted to stick to the same terms, so I replaced 'letter' with 'order'.
Pragmatic appropriacy (AP-) repair	The speaker repairs part of the message which is pragmatically inappropriate in the given situation (based on Brédart 1991)	Can I what can I do for you? Retrospection: First I wanted to say 'can I help you', but I thought this is said in shops only, and I decided to say 'what can I do for you' because it was more appropriate in this situation.
Repair for good language (AG-)	The speaker repairs part of the message which he/she judges to be not sophisticated enough concerning the manner of expression	If you want the room, I mean if you decide on it Retrospection: I was not satisfied with this sentence with the word 'want' in it, I did not like it stylistically.

References

- Bot, Kees de 1992. A bilingual production model: Levelt's 'speaking' model adapted. In: *Applied Linguistics* 13: 1–24.
- Brédart, Simon 1991. Word interruption in self-repairing. In: *Journal of Psycholinguistic Research* 20: 123–37.
- Chomsky, Noam 1965. *Aspects of the theory of syntax*. MIT Press, Cambridge MA.
- Clark, Herbert H. – Jean E. Fox Tree 2002. Using *uh* and *um* in spontaneous speaking. In: *Cognition* 84: 73–111.
- Dell, Gary S. 1986. A spreading activation theory of retrieval in sentence production. In: *Psychological Review* 93: 283–321.
- Dell, Gary S. – Peter O'Seaghda 1991. Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). In: *Psychological Review* 98: 604–14.
- Dörnyei, Zoltán – Lucy Katona 1992. Validation of the C-test amongst Hungarian EFL learners. In: *Language Testing* 9: 187–206.
- Dörnyei, Zoltán – Mary Lee Scott 1997. Communication strategies in a second language: Definitions and taxonomies. In: *Language Learning* 47: 173–210.
- DuBois, John 1974. Syntax in mid-sentence. In: Charles Fillmore – George Lakoff – Robert Lakoff (eds): *Berkeley studies in syntax and semantics III.1–III.25*. University of California, Institute of Human Learning and Department of Linguistics, Berkeley CA.
- Ericsson, K. Anders – Herbert A. Simon 1980. Verbal reports as data. In: *Psychological Review* 87: 215–251.
- Ericsson, K. Anders – Herbert A. Simon 1993. *Protocol Analysis: Verbal reports as data*. MIT Press, Cambridge MA.

- Fromkin, Victoria (ed.) 1980. Errors in linguistic performance: Slips of the tongue, ear, pen and hand. Academic Press, New York.
- Fry, Dennis 1969. The linguistic evidence of speech errors. In: Brno Studies of English 8: 69–74.
- Garrett, Merrill F. 1976. Syntactic processes in sentence production. In: Roger J. Wales – Edward Walker (eds): New approaches to language mechanisms, 274–93. North Holland, Amsterdam.
- Gathercole, Susan E. – Alan Baddeley 1994. Working memory and language. Hillsdale NJ, Lawrence Erlbaum.
- Gósy, Mária (ed.) 2004. Beszédkutatás 2004. „Nyelvbotlás”-korpusz tanulmányok [Speech research 2004. Studies of the “slip of the tongue” corpus]. MTA Nyelvtudományi Intézet, Budapest.
- Grice, H. Paul 1975. Logic and conversation. In: Peter Cole – Jerry L. Morgan (eds): Syntax and semantics, vol. 3: Speech acts, 41–58. Academic Press, New York.
- Hebb, Donald O. 1949. The organization of behavior. New York, Wiley.
- Hest, Erna van 1996. Self-repair in L1 and L2 production. Tilburg University Press, Tilburg.
- Huszár, Ágnes 2005. A gondolattól a szóig. A beszéd folyamata a nyelvbotlások tükrében [From thoughts to words. The process of speech production in the light of slips of the tongue]. Tinta Könyvkiadó, Budapest.
- Kellerman, Eric 1979. Transfer or non-transfer: Where are we now? In: Studies in Second Language Acquisition 2: 37–57.
- Kormos, Judit 1999. Monitoring and self-repair in L2. In: Language Learning 49: 303–42.
- Kormos, Judit 2000a. The role of attention in monitoring second language speech production. In: Language Learning 50: 343–84.
- Kormos, Judit 2000b. The timing of self-repairs in second language speech production. In: Studies in Second Language Acquisition 22: 145–69.
- Laver, John 1980. Monitoring systems in the neurolinguistic control of speech production. In: Fromkin (1980, 287–305).
- Levelt, Willem. J. M. 1983. Monitoring and self-repair in speech. In: Cognition 33: 41–103.
- Levelt, Willem. J. M. 1989. Speaking: From intention to articulation. MIT Press, Cambridge MA.
- Levelt, Willem J. M. – Ardi Roelofs – Antje S. Meyer 1999. A theory of lexical access in speech production. In: Behavioural and Brain Sciences 22: 1–75.
- Nooteboom, Sieb G. 1980. Speaking and unspeaking: Detection and correction of phonological and lexical errors in spontaneous speech. In: Fromkin (1980, 87–95).
- Postma, Albert – Herman Kolk 1992. The effects of noise masking and required accuracy on speech errors disfluencies and self-repairs. In: Journal of Speech and Hearing Research 35: 537–544.
- Postma, Albert – Herman Kolk 1993. The covert repair hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. In: Journal of Speech and Hearing Research 36: 472–487.

- Poullisse, Nanda 1999. Slips of the tongue. Speech errors in first and second language production. John Benjamins, Amsterdam & Philadelphia.
- Poullisse, Nanda – Theo Bongaerts 1994. First language use in second language production. In: *Applied Linguistics* 15: 36–57.
- Schmidt, Richard 1992. Psychological mechanisms underlying second language fluency. In: *Studies in Second Language Acquisition* 14: 357–85.
- Smedt, Konrad de – Gerard Kempen 1987. Incremental sentence production, self-correction, and coordination. In: Gerard Kempen (ed.): *Natural language generation: Recent advances in artificial intelligence, psychology and linguistics*, 365–76. Kluwer, Dordrecht.
- Stemberger, Joseph P. 1985. An interactive activation model of language production. In: Andrew W. Ellis (ed.): *Progress in the psychology of language*. Vol. 1, Hillsdale NJ. Lawrence Erlbaum, 143–86.
- Wheeldon, Linda R. – Willem J. M. Levelt 1995. Monitoring the time course of phonological encoding. In: *Journal of Memory and Language* 34: 311–34.