International Review of Social Sciences and Humanities
Vol. 8, No. 1 (2014), pp. 113-137
www.irssh.com
ISSN 2248-9010 (Online), ISSN 2250-0715 (Print)

The Influence of Task Type on Speech Monitoring

Mirjana M. Kovač
Department of Languages, Faculty of Electrical Engineering
Mechanical Engineerig and Naval Architecture
Fakultet Elektrotehnike, Strojarstva i Brodogradnje
Ruđera Boškovića bb 21 000 Split, Croatia
E-mail: mirjana.kovac@fesb.hr

(Received: 2-7-14 / Accepted: 2-9-14)

Abstract

The main goal of this study was to find out how L2 learners allocate their attentional resources for speech monitoring in different task types, thus providing an insight into their self-repair behaviour. The retelling of a predetermined chronological order of events resulted in a significantly higher rate of syntactic repairs compared to all other tasks. Namely, the learners displayed a tendency to monitor erroneous syntactic structures, to interrupt them, and to "withdraw" from situations in which they became aware of the difficulties in preparing the message under time pressure. Also, in the predetermined retelling task the learners aimed to monitor their speech for morphological aspects in comparison to other tasks. For this reason, tasks based on retelling a story with a predetermined chronological order of events might enhance learners' conscious attention toward speech monitoring which is necessary for learning to take place.

Keywords: Speech monitoring, self-repairs, task complexity, speech production.

1. Introduction

The speech act production involves more than just information transfer, it also includes the communicative intention that needs to be properly recognized and interpreted by the listener. Levelt (1989) explained in detail the process of speech production starting from content planning to the production of the spoken word. The overall process comprises four main activities which proceed in successive order, as follows: a) conceptualization or the creation of the preverbal conceptual structure, which is a controlled process; b) formulation which includes grammatical, lexical, and phonological encoding, which is for L1 speakers an automatic process; c) articulation or the phase representing overt speech, and finally d) self-monitoring that includes the verification of the correctness or appropriatness of the produced utterances (Kormos, 2006). Goldrick and Rapp (2007) argued three different levels of word-form processing: the retrieval of sound information form long term memory, post-lexical processes which include the specification of more detailed aspects of sound structure, and motor programming as well as execution processes. The formulation of message is argued to be controlled in the case of L2 speakers, thus, considering this ability as a more complex skill (Fortkamp, 2000). In other words, speaking a foreign language involves processes which are partly automatized, therefore competing for the limited attentional resources of the cognitive system (Anderson, 1995; Kormos, 2006). Skehan (1989) claimed that L2 speakers rely on the memory-based system related to the retrieval of ready-made chunks, which requires less processing. On the other hand, during the conceptualization phase, the speakers rely on the rule-based system which requires more attention and more processing.

Nevertheless, the speakers rarely produce perfect speech either in the native or in the foreign language. On the contrary, the produced speech displays different forms of disfluencies, such as hesitations, silent and filled pauses, false starts, repetitions, vowel prolongations, speech errors, and self-repairs. Speech errors are deviations from the speaker's communicative intention, thus, they are an important information source for the understanding of the complex language production mechanisms. In contrast to errors, self-repairs are self-initiated corrections of one's speech which are a normal phenomenon in spontaneous speech. They occur as a response to some linguistic problem which may arise at any stage of language production. A simplified diagram of the process of self-monitoring is provided by Pillai (2006), Figure 1. In the error repair process Levelt (1989) distinguished three major stages: firstly, monitoring and speech interruption when trouble is detected; secondly, different forms of hesitations and pausing; and finally, repairing disfluent speech. Self-repairs can be regarded as a manifestation of „quality control“ (Hieke, 1981:148). Research in this field proved that monitoring is a conscious process which requires attentional control (Levelt, 1989).

Problem detected → Speech interrupted → Hesitation → Self-repair initiated

Figure 1: Process of Error-Detection, Hesitation, and Self-Repair (Pillai, 2006)

In a series of studies Schmidt (1990, 1993, 1994) implied that noticing is necessary for learning to take place. The rate of self-repairs has been one of the most extensively investigated appearances of the self-repair behaviour of L2 learners, even though the role of attention in speech monitoring has been a neglected area of research (Kormos, 2000). Van Hest (1996) assumed that with the development of L2 competence, the monitor becomes more sensitive to problems arising at the level of discourse.

There are different theories regarding the reason for the occurrence of speech errors in general. Current psychological models of word production (e.g. Dell,1986; Levelt, Roelofs, & Meyer, 1999) only delineate how we plan the phonological content of words, and not how we articulate them. Dell (1986) argued that the erroneous activation of certain nodes causes speech errors, explaining that the production of a particular unit depends on its activation degree, but also on the activation degree of other units organized in an associative network. This means that the unit which is in the process of realization has to be deactivated at some point in order to empty the place for another unit. However, Moat and Hartsuiker (2008) used the computational implementations to extend the cascading and non-cascading models, assuming that models based on the classic spreading activation account of word production may require modifications in order to accurately explain the key aspects of human error patterns. The findings suggested that cascading from unselected phonemes to articulation is not necessary to explain voicing traces of intended phonemes on erroneous productions. Each type of disfluency can be linked to a certain stage of the speech production process from conceptual planning through grammatical encoding to articulatory planning (Gósy 2005). One of the main theoretical models explaining speech monitoring and repairing was provided by the Perceptual Loop Theory (Levelt, 1989). According to this theory, audible self-produced speech goes through the Speech Comprehension System, where it is processed in the same way as we process other people’s speech that we hear (Pillai, 2006). When trouble is detected, the monitor makes the speaker aware of it, consequently, an alarm signal is sent to the working memory, and central controlled corrective action arises. Nooteboom (1980) concluded that 50% of all errors remain uncorrected for several reasons. Sometimes the monitoring mechanism does not register and respond to an error, or, in the speaker's point of view, the speech is sufficiently redundant, therefore, the listener can correctly interpret the message without any correction or adjustment.

Studies dealing with L2 speech production claim that fluent foreign language speech involves complexity, fluency, and accuracy (Fortkamp, 2000; D'Ely, 2006). There are probably trade-offs among these three goals of speech production. Complexity and accuracy are more controlled processes linked to message conceptualization, or the rule-based system (Skehan, 1998). Most of the studies on speech production concentrated on measuring the following three aspects, namely, fluency, complexity, and accuracy. Fluency is defined as the ability to maintain real-time communication with the focus on meaning, whereas complexity involves the willingness to choose more challenging language. Finally, accuracy is the learners' orientation towards the control over more stable elements in the interlanguage system (Skehan and Foster, 2001). The speakers are balancing among these three aspects, devoting more attention to some aspects at the expense of some others. Various tasks used in researches have been used as a context in which these three goals have been integrated in communication. According to Bygate (1999), tasks are a means of framing, reframing and unframing language so that the speakers' attention can be devoted to different aspects of oral performance in each encounter with the task. The author analyzed how task repetition affected these three goals respectively. The results of the research showed that improvements were noticed in terms of complexity, but at the expense of accuracy and fluency.

Robinson (2001) distinguished among task complexity, difficulty and task condition. In the author's view, narratives are more complex than picture descriptions. In a picture description task speakers have visual support and the memory is less loaded than in a narrative task. A number of previous research studies which focused on speech errors and repairs (Levelt, 1983; Verhoeven, 1989) mainly included the description of static objects and constellations, in particular spatial relationships (e.g. up-down, left-right). Thus, special attention was paid to the linearization process of macroplanning, that is, to the organization of natural order. The content of three-dimensional spatial information must be lineary organized since there is no natural order to proceed, as it is in the case of retelling a chronological order of events. Whenever speakers want to express any communicative intention, regardless of its complexity, they must solve the linearization problem. Levelt (1989:138) explained it in the following way: "Deciding what to say first, what to say next, and so on ...". In order to determine an order, the speakers need to memorize what has already been said and what needs to be said. When describing spatial relationships, Levelt (1989) concluded that the participants follow three major principles which apply to spatial and nonspatial description areas, since they reflect the most general characteristics of perception and memory. *The Principle of connectivity* is a general ordering principle in perception and memory, and the dominant way of describing hierarchically organized structures. However, it is not always possible to introduce new items of information without repeating old items. The *Stack principle* is known in psychology of problem solving and is a dominant way of keeping track of hierarchically organized structures. The speaker's bookkeeping for return address is like putting them on a stack and always returning to the top item on the stack after reaching the end of a connected string (Levelt, 1989:143). Finally, *the Minimal-load principle* is applied when there are more alternatives, and the principle says “Do the simplest thing first” (Levelt, 1989:144), implying that the speakers describe less demanding constellations first. However, it is difficult to apply the principle of natural order to the description of dynamic constellations, and the reasons can be found in the dynamic aspects of the event, as opposed to static objects and constellations. If the information to be conveyed is complex involving several consecutive speech acts, the speakers must decide how to organize information. This is the speakers' linearization problem (Levelt, 1989), by which the principle of natural order requires that the speech acts must have a chronological order of appearance. Studies based on tasks involving the description of spatial constellations (Levelt, 1982) aimed to determine how the static spatial structure is used for the construction of a linear order. Tasks given to participants in earlier research included picture description (van Hest, 1996), spatial description (Levelt, 1983; Verhoeven, 1989), interview tasks (Fathman, 1980; van Hest, 1996), storytelling (Fathman, 1980; Lennon, 1990; van Hest, 1996; Verhoeven, 1989), and information gap activities (Kormos, 2000). Finardi (2008) investigated the effects of repeating a picture description task on learners' L2 oral performance. Four measures of speech performance were calculated following Fortkamp (2000): fluency, accuracy, complexity, and lexical density. The results indicated gains in complexity, thus confirming Bygate's (2001) findings for this task condition. Moreover, the results showed that the complexity and accuracy aspects of L2 speech production are based on the same cognitive system, that is, the rule-based system focused on form, whereas the dimension of fluency is based on the memory-based system (Skehan, 1998). The trade-offs were interpreted in terms of a focus on meaning, which explained why learners gained in complexity (focus on form) but not on fluency (focus on meaning).

The present paper examines the amount of attention toward monitoring of errors and inappropriaces in foreign language by means of analyzing the influence of a specific task type on respective self-repair categories, thus providing an insight into the self-repair behaviour of L2 learners. Firstly, Levelt's, Kormos' and Van Hest's classification systems will be discussed since they represent the starting point for the data analysis. This is followed by the data collection procedures and task selection. The final two sections will present the results and conclusions based on the speech sample analysis of engineering students.

2. The Classification of Self-Repairs

2.1. Levelt's Self-Repair Classification System

Levelt (1983) was the first psycholinguist who proposed a precise classification of self-repairs based on his speech production model, and which was accepted as the best empirically validated model for monolingual as well as for bilingual speech processing (Kormos, 2006).

The following categories are proposed:

1. Different information repair modifies the content of the original message. The reasons for repairing refer to the conceptualizer which has not properly ordered or has encoded a false information, resulting in both cases in an inadequate preverbal plan. The following subcategories are established within this category (Levelt, 1989):

a) **Inappropriate information repair** occurs when the speaker repairs because the informational content of the utterance is false;

b) **Ordering error repair** is issued when the speaker decides to encode the intended parts of the utterance in a different order. This repair category reveals the extent to which the speech production system is focused on the past, the present and the future, thus providing information on how the system solves the problem of component sequencing. The theory of ordering components must satisfy a number of functional requirements: the present system has to be activated, the past deactivated, and plans for the future activation should be prepared (Dell et al., 1997). The speaker must decide how to order the components in order to express a complex information. The speaker may also realize that a different arrangement of components could be acceptable, and can decide to encode parts of the intended message in a different order (Levelt, 1983). Levelt (1989) explains this matter in the following way: "Deciding what to say first, what to say next, and so on…" (Levelt, 1983:138).

c) **Message abandonment repair** is issued when the speaker rejects the intended message and exchanges it with a new one.

2. Appropriacy repair aims at a further specification of the informational content, in other words, there is no actual error that needs to be repaired. Levelt (1983) distinguished three subcategories:

a) **Ambiguity repair** occurs when the speaker repairs the utterance because the interpretation might be ambiguous for the listener;

b) **Appropriate level of information** is issued when the speaker wants to specify the utterance;

c) **Coherence repair** appears in case when the utterance is not coherent with the previously used terminology.

Brédart (1991) enriched Levelt's classification by introducing the *Pragmatic appropriacy repair* which aims at repairing pragmatic errors. In this instance the speaker repairs some parts of the utterance which are pragmatically inappropriate in a given situation. Furthermore, the speaker can assume that some parts of the utterance are pragmatically acceptable but insufficiently sophisticated, thus using *Repairs for good language*.

3. Error repairs occur as results of imperfect functioning at the level of the formulator, where wrongly activated words, false syntactic structures, false morphemes or phonemes are selected. Accordingly, Levelt (1983) differentiated lexical, syntactic, and phonological errors, which correspond to the three basic processing levels. Namely, the first phase in the processing of the preverbal plan involves the lemma retrieval belonging to the corresponding concept, by which the concept is specified by the preverbal plan. The repair of a wrongly activated lemma is known as a *lexical error repair* (Levelt, 1983, 1989). According to Levelt's theory (1989), content and functional words, collocations and idioms, are considered *lexical entries*. Therefore, lexical repairs include the repairs of wrongly activated content and functional words, as well as idioms and collocations. The repairs of derivational morphology, for instance *different* instead of *difference*, also belong to the category of lexical repairs, since in Levelt's (1989) lexicon model, derivations are different lexical entries.

Lexical errors are defined as “any lexical item, colour words, direction terms, prepositions, articles, etc” (Levelt, 1989:54). Levelt (1989) presumed that in the case of a lexical error, a wrongly lexical entry is activated, and finally, articulated.

A syntactic repair is defined as a repair of a syntactic structure which leads to a deadlock and cannot be continued by the speaker (Levelt, 1983:54). Unfortunately, Levelt did not provide a definition of phonetic repairs and did not explain morphlogical errors.

Table 1 represents Levelt's system of classification with corresponding examples, enriched by Brédart's (1991) classification system.

Table 1: Taxonomy of repairs in L1 (Levelt, 1983)

Repair type	Example
Different information repair	*We gaan rechtdoor offe...We komen binnen via rood, gaan dan.* *We go straight on or... We come in via red, go then straight to green.* (Levelt:1983:51)
Appropriacy repair	
Ambiguity repair	*We beginnen in het midden met... in het midden van het papier met een blauw rondje.* *We start in the middle with... in the middle of the paper with a blue disc.* (Levelt 1983: 52)
Appropriate level of information repair	*Met een blauw vlakje, een blauw aan de blovenkant* *With a blue spot, a blue disc at the upperend.* (Levelt 1983: 53)
Coherence repair	*Ga je een naar boven, is uh... kom je bij geel.* *Go you one up, is uh... come you to yellow.* (Levelt 1983: 53)
Repair for good language	*C'est qu'un con, un idiot pardon.* *He is nothing but a damn fool, an idiot sorry.* (Bredart 1991:127)
Error repair	
Lexical repair	*Rechtdoor rood, of sorry, rechtdoor zwart* *Straight on red, or sorry, straight on black.* (Levelt 1983:53)
Syntactic repair	*En zwart...van zwart naar rechts naar rood.* *And black... from black to right to red.* (Levelt 1983:54)
Phonological repair	*Een eenheed, eenheid vanuit de gele stip.* *A unut, unit from the yellow dot.* (Levelt 1983: 54)

2.2 Kormos' and Van Hest's Classification of Self-Repairs

Kormos (1998, 1999, 2006) extended Levelt's taxonomy of repairs by introducing rephrasing repairs. As opposed to appropriacy and different information repairs which occur at the conceptualization level, this repair category includes the revision of the preverbal plan without the modification of the propositional content. Thus, the speakers use a slightly modified version of the word or phrase, since they are not certain about its correctness.

Table 2 and Table 3 display Kormos' (2000) and van Hest's (1996) repair classification systems, respectively.

Table 2: Classification of repairs (Kormos, 2000)

Self-repair type	Definition	Example
Different information repair	Involves the content modification of the pre-verbal plan.	
Inappropriate information repair	The speaker decides to repair the message, since the informational content has proven to be inappropriate or incorrect (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 repair	Parts of the intended message need to be ordered differently (Levelt, 1983).	*'Well, we..it's it's about a thousand forints...'*
Message abandonment repair	The speaker abandons the originally intended message and replaces it with another one (Levelt, 1983).	*'We have some er er v-maybe you have vegetarians in your group.'*
Appropriacy repair	The speaker decides to encode the originally intended message in a modified way (Levelt, 1983).	
Ambiguity repair	The speaker modifies the message because of the ambiguous information that needs to be further specified (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.'*
Appropriate level of information repair	The speaker decides to provide further details because the information is not precise enough (Levelt, 1983).	*'There are very wide choice of er main courses er er steak er er several kind of steak...'*
Coherence terminology repair	The speaker corrects incoherent terminology (Levelt, 1983).	*'...er but this letter is er- the order-er your request...'*
Repair for good language	The speaker modifies a pragmatically acceptable, but not sophisticated enough utterance.	*'Can I what can I do for you? If you want the room, I mean if you decide on it...'*
Pragmatic appropriacy repair	The speaker repairs part of the message which he/she feels to be pragmatically unacceptable in a given situation (Bredart, 1991).	*'If you want the room, I mean if you decide on it...'*
Rephrasing repair	The speaker is uncertain about the correctness of the word or phrase and adds something and/or uses paraphrase (Kormos, 2000).	*'We will er reflect er to you in another letter we will answer you.'*
Error repair	Involves the re-issuing of the same pre-verbal plan due to accidental lapses which occur in the formulator (Levelt, 1983).	
Grammatical error-repair:		
a) Inaccurate use of inflectional morphology		*'It × have to be er uhm uhm thirty-five people.'*
b) Inappropriate choice of tense or aspect of the verb phrase		*'But I × don't mention er the room is er only on the er eighteen and on the nineteen of December free.'*
c) Faulty encoding of complements and specifiers		*'You have to pay er the uhm twenty-five er percent × the uhm the price.'*
d) Wrong word order		*'Minimum er thirty-five er people er have to be er then × can I er let it for you then × can I er let it for you.'*
e) Inappropriate choice of prepositions and auxiliaries accessed by syntactic building procedures		*'We may make a contract if you er if you × will pay more.'*
Lexical error-repair:		
a) Inappropriate choice of content words		*'My chef can make × cancer very good.'*
b) Inappropriate choice of prepositions and auxiliaries with independent conceptual specifications		*'If you need this room you need to tell me × before twenty hours.'*
c) Collocational error		*'We can cook er × to taste.'*
d) The erroneous production of a derivative form		*'You have to write a × confirmament.'*
Phonological error-repair	Involves the correction of a phoneme, an allophone, an allomorph, the metrical and intonation structure of a word or of a string of words (intonational phrases).	*'We could arrange er more smaller *[taibiz] tables if you would like that better.'*

Table 3: Classification of repairs (Van Hest, 1996)

Self-repair type	Definition	Example
Error repair		
Lexical error repair	The speaker has selected a wrong word and substitutes the correct one for it. This repair category includes: content words (nouns, adjectives, adverbs, verbs) or function words (demonstratives, relative pronouns, prepositions etc.).	*'then you see the man uhm which..who is entering the room'*
Syntactic error repair	The speaker produces an utterance which cannot be finished without violating the grammatical rules. The utterance is interrupted and a fresh start issued.	*'it's not you do..something you do every day'*
Repair for tense and aspect	The speaker has used a wrong tense or aspect and corrects it.	*'yesterday it has rained..uh rained'*
Phonological error repair	The speaker corrects a phonological error, which occurred as a result of mispronunciation or exchange of phonemes.	*'they have a /naif/ nice boat'*
Morphological error repair:		
a) Inflectional error repair	The speaker has selected a wrong ending of a verb or noun.	*'so the man have..has got his hats back'*
b) Derivational error repair	The speaker has selected a wrong derivational morpheme.	*'the car man drove away very quick..uh quickly'*
c) Morphological/phonological error repair	This category includes repairs which can be classified as either morphological or as a combination of the two.	*'we have /ri:/ .. read some books'*
Conceptual error repair	Errors as a result of a wrong conceptual plan, usually due to misinterpretations.	*'now wee see the two boy..one boy and one girl'*
Appropriacy repair		
Appropriateness repair: lexical	The speaker replaces one term with another, usually a more precise one.	*'a shore of a lake..or the sea'*
Appropriateness repair: insertion	The speaker repeats part of the original utterance and inserts one or more words to specify it.	*'you see a policeman..an English policeman'*
Appropriateness repair: syntactic	The speaker replaces the original syntactic construction with a more appropriate one.	*'you see a door with a card named closed uh that the shop is closed'*
Appropriateness repair for tense and aspect	The speaker repairs for tense or aspect, even though the original verb forms are not erroneous.	*'a girl that's that was addicted to drugs'*
Different information repair	The speaker interrupts the current message to introduce a totally different topic.	*'you see two strange men, they are, on the background you see the the car'*
Rephrazing repair	All overt self-repairs that: a) do not fit into the above categories, b) can be classified in more than one category, and c) repairs of correct items.	*'last year we had..we've had'*

3. The Research Project

3.1. Data Collection Procedures and Task Selection

The sample group consisted of 101 participants, first-year students of technical studies in Croatia. The learners received 8-9 years of formal English instruction and it was presumed that the participants' instructional background was very similar, since all of them claimed to have received a mixed form-focussed and communicative syllabus as a method of instruction in their primary and secondary school. Therefore, the speakers should have reached the B level of English proficiency according to the Common European Framework of Reference for Languages. The participants were randomly selected to participate in the study. The error and repair data were collected by means of five different tasks which were performed in English (L2). All data were collected in an ordinary office with no special facilities at the Faculty. It was preferred to an experimentation room because it would contribute to the informal atmosphere in which the tasks were performed. The subjects were seated opposite the researcher at the same table. The computer and the microphone were placed between the subjects and the researcher. Each student was individually audio-recorded and afterwards the speech samples were transcribed and categorized by two coders to avoid inconsistency and data bias. When there was a discrepancy in the data analysis between the two coders, the third coder was consulted in order to come to a final agreement. Time parameters were measured by means of a speech analysis programme.

Before carrying out the selected speech tasks, the researcher gave explicit instructions for each task. In the story retelling task (CARTOON) the students watched the cartoon Johnny Bravo in the English language. This cartoon was chosen under the assumption that it was relatively unknown to the subjects and that its content represented a significant cognitive effort in terms of discourse organization. After watching the cartoon which lasted for six minutes, each participant described the chronological order of events in their own words in the English language. The participant had one minute to prepare and was not time limited.
In the second task the students had to describe a room picture (ROOM1) that had six pieces of furniture in such a way that someone who could not see the picture could produce a global setting on the basis of the description. The third task (ROOM2) was almost identical to the second one, except for the furniture that was differently arranged.

In the fourth task (UTTER) the subjects had to form utterances based on different semantically unrelated drawings. The syntactic frame was not defined by the researcher, therefore, the only requirement was that the drawing and the corresponding colour appeared in the utterance (e.g. *The rose is red*; or *The blue shirt is very nice*).

The fifth task (STORY) was an invented and undetermined story narration. The subject had to make up a story based on five unrelated drawings and none of them were allowed to be omitted. As in the case of previous tasks, each participant had one minute to prepare.

In conclusion, the selected tasks consisted of descriptions of static objects and constellations with visual support (ROOM1, ROOM2, UTTER), story narration with visual support (STORY), as well as the narration of space-time entities (CARTOON).

3.2. The Classification of Self-Repairs

The repair classification system was devised on a slightly modified version of Levelt's classification system, since it has been the most complete and the most widely applied system available (Kovač, 2011; 2013). The classification was enriched by some subcategories adopted from Kormos (2000) and van Hest (1996), as well as by a current frequently occurring subcategory of syntactic errors introduced by the researcher. Levelt's overt repair structure consisting of three main parts was followed as the main criterion in the identification of repairs (Levelt, 1983:44):

'Go from left again to	*uh..*	*from pink again to blue'*
original utterance	editing phase	alteration REPAIR

Namely, the first part *'Go from left again to'* is the original utterance and consists of a problematic spot or reparandum which needs to be corrected ('*left*') or in some cases further specified, since an utterance does not necessarily have to be wrong in order to be repaired. Furthermore, shorter or longer speech segments can be erroneous, that is, ranging from an erroneous phoneme to a whole utterance. The speaker may interrupt the speech before or after the overt articulation, and an interruption delay can also occur. An interruption may be

followed by different kinds of disfluencies, such as editing terms, silent pauses, vowel prolongations which represent the editing phase, and finally, the third phase is the repair.
This study only focused on overt repairs in which the following categories were established: Error-, Appropriateness-, Different information-, and Rephrasing repairs. Some of them were further subdivided which will be described and illustrated by corresponding examples.

3.2.1 Error Repairs

a) Lexical Repairs: In this case the speaker has selected a wrong word and substitutes it for a correct one. Three subcategories of lexical repairs have been established: Repairs of a) idioms, collocations, functional and content words, errors of derivational morphology (1); b) unintentional use of L1 lexemes (2); c) non-existent words (3). Numbers in brackets refer to examples collected from the repair corpus.

(1) There is a wooden table in the lower left, right part of the room.
(2) Johnny Bravo went to the šum- sorry to the wood.
(3) His mother was making him a sweater of /pineko/ well I mean something like pines.

b) Syntactic Errors were analyzed according to where they occurred in the three stages of the grammatical encoding phase, that is, (i) when the various complements, specifiers and parameters are encoded, (ii) handled by the different subroutines, and finally (iii) when these processed materials are ordered (Kormos, 1999). Thus, following sub-categories of syntactic repairs have been identified: a) repairs of wrong word order (4); b) repairs of unfinished expressions or false starts (5); c) repairs of completely unacceptable morpho-syntactic and/or semantic structures (6); d) repairs of wrongly encoded complements and specifiers, which are accessed via syntactic building procedures (7).

(4) Johnny met a bear who was, er, sleeping, he woke up him- he woke him up in the middle of, er, winter.
(5) And he came to er, er, cave, er, er, that cave was er ,er in that cave he saw er, a bear.
(6) They start to arguing and hunting into the woods but er, it er, didn't er, was success.
(7) The bear and Johnny listened**---** beautiful singing, er, er, listened to beautiful song of Johhny's mother.

c) Morphological Repairs include: a) repairs of inflexional errors, when the speaker chooses the wrong verb form (8); b) repairs of incorrect plural of nouns (9); c) repairs of errors for "time and aspect" (10), when the speaker uses the wrong time or aspect (Van Hest, 1996); d) repairs of omitted article (11) or misused article (12).

(8) And when she find out found out that she, er, hadn't enough material she went, er, deeper in the woods
(9) It is a place with beautiful woman, er, with beautiful women.
(10) Johnny Bravo was waiting in the woods his mum wants wanted to make a suit.
(11) On**---**left side is a bed and a table
(12) His mother was singing song which which helped the Johnny Bravo, er, Johnny Bravo to sleep

d) Phonological Repairs: Phonological repairs mainly included repairs of articulatory clumsiness (13), whereas other forms of phonological errors and repairs were not analyzed, since a systematic differentiation of inaccuracies due to the speaker's accent and lack of knowledge of phonological forms of lexical entries could potentially cause serious problems.

(13) His mother, his mother er, er, sang a /lole/ lullaby and er, they fell asleep.

3.2.2 Appropriacy Repairs

The speaker decides to encode the originally intended message in a modified way (Levelt, 1983). This category includes instances when the speaker modifies the message because of the ambiguous information that needs to be further specified (Levelt, 1983), and when the speaker decides to provide further details since the information is not precise enough (Levelt, 1983) (14).

(14) A whale attacked him with his flipper er- I mean black flipper.

3.2.3 Different Information Repairs

Within this category two subcategories have been established: a) message abandonment, when the speaker rejects the original message because of limited L2 competence, or because a more relevant information is considered adequate (15); b) ordering error repair, when parts of the intended message need to be ordered differently (Levelt, 1983) (16).

(15) On the left we have, er, er, clo-closer to the wall we see a table and a desk.
(16) I saw a snowman with a big nose and two, er, and a black hat and two arms made of sticks with a big nose.

3.2.4 Rephrasing Repairs

The speaker is uncertain about the correctness of the word or phrase and adds something and/or uses paraphrase (Kormos, 2000) (17).

(17) Johnny Bravo and, er, his mother were in a wood- in a forest.

On the basis of the threefold repair structure all possible instances of overt repairs in the transcripts of 101 students were identified. All the cases where no unambiguous judgements could be made, data were collected and were shown to two educated native speakers of English. The native speakers were informed about the nature of the tasks the subjects had to perform and were shown the errors together with their contexts. All other cases which lacked this structure (utterances in which a reparandum was lacking but where hesitations, repetitions of lexical units, vowel prolongations, filled and silent pauses pointed to some kind of difficulties in speech) were excluded from the analysis. Consequently, covert repairs were not included in the analysis since it is rarely possible to determine the source of trouble the speaker is facing without a thorough retrospective analysis.

4. Results and Discussion

Since monitoring involves the noticing of errors and inappropriaces, the following research question will be investigated: How do L2 learners allocate their attention for speech monitoring in different task types?

Speech duration per participant may be of varying size in each task. In order to provide the comparison of various tasks by applying appropriate statistical tests, it was necessary to perform the calculations by reducing the corresponding categories of repairs per subject to a prespecified number of words (Bortfeld and al., 2001). In the available literature that number is in most cases 100 words, therefore, the same approach was chosen in this research. The analysis of the influence of the task type on the rate of examined repair categories was performed on the same subjects (101 undergraduate student). Consequently, there were five dependent samples.

4.1 Error Repairs in L2

Table 4a shows descriptive statistics for the rate of error repairs in L2. The main reason for such a detailed data presentation is a requirement for a good sample knowledge, in order to decide which samples to compare with each other, if the statistical test results display that the populations from which the samples come, are not identical.

The Shapiro-Wilk normality test (Marques de Sá, 2007) is selected, Table 4b, highlighted in recent works as the most efficient normality test (Thode, 2002; Zhang and Yuehua, 2005; Keskin, 2006; Henderson, 2006; Coin, 2008). According to this test, it may be noticed that the distributions of all five populations significantly deviate from the normal distribution. Therefore, the Friedman test is selected (Field, 2005; Marques de Sá, 2007; Demšar, 2006), as probably the most commonly used (Al-Subaihi, 2000) and the most popular (Van de Wiel, 2004) nonparametric repeated measures test in the case of three or more dependent samples, Table 4c. Nonparametric tests are known as distribution-free tests, which require neither normally distributed data, nor homogeneity of variance (Montgomery and Runger, 2003). However, the drawback of the nonparametric tests is that they are less efficient than the parametric ones, but only if the assumptions for using parametric tests are fulfilled, which is not the case with regard to the results of the Shapiro-Wilk test. If the assumptions for using parametric tests are not met, the nonparametric tests often provide considerable improvement (Montgomery and Runger, 2003). The p-value of the Friedman test is lower than 0.0001, Table 4c, displaying that significant differences in the rate of syntactic repairs among individual tasks have been obtained. Thus, all populations are not identical considering that the p-value is lower than or equal to the threshold value α ($p < 0.0001 \leq \alpha = 0.05$).

Table 4a: Descriptive statistics for error repairs in L2

ER	*CARTOON*	*ROOM1*	*ROOM2*	*UTTER*	*STORY*
n	101	101	101	101	101
x_{min}	0	0	0	0	0
D_1	0	0	0	0	0
Q_1	0.721	0	0	0.651	0
Me	1.429	0.714	0	1.042	0.645
Q_3	1.990	2.083	1.307	1.919	2.198
D_9	3.227	3.709	2.488	2.817	3.524
x_{max}	5.294	6.977	5.263	6.250	6.452
d	5.294	6.977	5.263	6.250	6.452
d_Q	1.269	2.083	1.307	1.268	2.198
$\bar{x}$	1.522	1.260	0.697	1.359	1.279
σ	1.130	1.643	1.170	1.187	1.602
σ^2	1.277	2.699	1.369	1.409	2.566
$S_{\bar{x}}$	0.112	0.164	0.116	0.118	0.159
V	74.24 %	130.39 %	168.02 %	87.40 %	125.28 %
α_3	0.916	1.456	1.805	1.281	1.245
α_4	0.676	1.893	2.861	2.269	1.044

n – sample size, x_{min} – sample minimum, D_1 – first decile, Q_1 – lower quartile, Me – median, Q_3 – upper quartile, D_9 – ninth decile, x_{max} – sample maximum, d – sample range, d_Q – interquartile range, $\bar{x}$ – sample mean, s – sample standard deviation, s^2 – sample variance, $S_{\bar{x}}$ – standard error of the mean, V – coefficient of the variation, α_3 – skewness, α_4 – kurtosis

Table 4b: Shapiro-Wilk normality test for error repairs in L2

Shapiro-Wilk	*CARTOON*	*ROOM1*	*ROOM2*	*UTTER*	*STORY*
W	0.9343	0.7808	0.6621	0.8899	0.7953
p	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001

W – Shapiro Wilk test statistic, p – p value

Table 4c: Friedman test for error repairs in L2

Friedman	*n*	*df*	F_r	*p*
	101	4	30.967	< 0.0001

n – sample size, df – degrees of freedom, F_r *– Friedman test statistic, p – p value*

The results of the Friedman test reveal that the null hypothesis of identical populations should be rejected, yet, it does not answer the question which tasks, when compared, display significant differences in the rate of syntactic repairs. Dunn's multiple comparison test (Daniel, 1990; Pett, 1997), an effective test with a careful assessment of statistically significant differences between compared pairs (Pett, 1997), answers this question. Comparisons can be performed only for selected pairs of tasks or for all possible pair combinations. In the case of implementing multiple comparisons with a threshold value $\alpha = 0.05$ for each pair, it is necessary to take into account that the greater number of comparisons considerably increases the overall probability of error type I, that is, the probability of rejecting the null hypothesis when it is true in at least one case of comparison (e.g. it is 0.226 for 5 compared pairs, whereas in the case of 10 comparisons, which is the maximum number in the case of 5 samples, it exceeds 0.4). The solution may be to lower the threshold of significance for each comparison to a value which ensures that the overall probability of error type I does not exceed 0.05. On the other hand, by reducing the threshold value for each comparison, the risk of error type II also increases (null-hypothesis is not rejected, even though the alternative hypothesis is correct), which leads to the conclusion that it is useful to compare only the chosen pairs. Thus, the risk of error type II is reduced. The decision which pairs to compare can be made based on the good knowledge of descriptive statistics related to each sample, Table 4a. In conclusion, it is important to be restrictive in choosing comparison pairs (Marques de Sá, 2007).

The median equals zero in ROOM2, Table 4a, as opposed to the medians referring to all other tasks. Consequently, all tasks are compared to ROOM2. The multiple comparison test revealed significant differences in the total number of error repairs in L2 between ROOM2 and all other tasks, except STORY, Table 4d. A detailed analysis of syntactic, lexical, morphological, and phonological repairs will be carried out in the continuation of the paper in order to provide the explanation of obtained differences.

Table 4d: Dunn's multiple comparison test for error repairs in L2

Dunn	*Difference in rank sum*	*Significant difference*
CARTOON – ROOM2	109	Yes
ROOM1 – ROOM2	60.5	Yes
ROOM2 – UTTER	–88	Yes
ROOM2 – STORY	–47.5	No

4.1.1 Syntactic Repairs

Descriptive statistics for the rate of syntactic repairs in L2 can be seen in Table 5a. The results of the normality test for syntactic repairs indicate that all distributions significantly deviate from the normal distribution, Table 5b. The results of the Friedman test for syntactic repairs are presented in Table 5c. The p-value is lower than 0.0001, showing that significant differences among certain tasks have been obtained. In CARTOON the median differs from zero, compared to other tasks, Table 5a. Therefore, Dunn's test has been implemented for all combinations containing CARTOON.

The results presented in Table 5d suggest that the retelling of a predetermined chronological sequence of events resulted in a significantly higher rate of syntactic repairs compared to all other tasks. Overall, in retelling a story it is necessary to organize the speech acts under time constraints and therefore special attention for monitoring is very important. Self monitoring of speech is meant to limit damage in communication that has already been made public (Nooteboom, 2004).

In CARTOON the learners display a tendency to monitor erroneous syntactic structures, to interrupt them, and to “withdraw“ from a situation in which they become aware of the difficulties in formulating the message under time pressure, if compared to other tasks.

Table 5a: Descriptive statistics for syntactic repairs in L2

ER_S	*CARTOON*	*ROOM1*	*ROOM2*	*UTTER*	*STORY*
n	101	101	101	101	101
x_{min}	0	0	0	0	0
D_1	0	0	0	0	0
Q_1	0	0	0	0	0
Me	0.380	0	0	0	0
Q_3	0.850	0.943	0	0.551	0
D_9	1.373	1.909	1.316	1.109	1.582
x_{max}	2.703	5.128	3.846	2.679	3.509
d	2.703	5.128	3.846	2.679	3.509
d_Q	0.850	0.943	0	0.551	0
$\bar{x}$	0.522	0.534	0.294	0.310	0.332
σ	0.622	1.024	0.737	0.559	0.731
σ^2	0.387	1.049	0.543	0.312	0.534
$S_{\bar{x}}$	0.062	0.102	0.073	0.056	0.073
V	119.19 %	191.64 %	250.58 %	180.52 %	220.05 %
α_3	1.481	2.349	2.740	2.089	2.233
α_4	2.186	5.920	7.442	4.547	4.527

n – sample size, x_{min} – sample minimum, D_1 – first decile, Q_1 – lower quartile, Me – median, Q_3 – upper quartile, D_9 – ninth decile, x_{max} – sample maximum, d – sample range, d_Q – interquartile range, $\bar{x}$ – sample mean, s – sample standard deviation, s^2 – sample variance, $S_{\bar{x}}$ – standard error of the mean, V – coefficient of the variation, α_3 – skewness, α_4 – kurtosis

Table 5b: Shapiro-Wilk test of normality for syntactic repairs in L2

Shapiro-Wilk	*CARTOON*	*ROOM1*	*ROOM2*	*UTTER*	*STORY*
W	0.8122	0.6000	0.4639	0.6269	0.5225
p	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001

W – Shapiro Wilk test statistic, p – p value

Table 5c: Friedman test for syntactic repairs in L2

Friedman	*n*	*df*	F_r	*p*
	101	4	28.969	< 0.0001

n – sample size, df – degrees of freedom, F_r *– Friedman test statistic, p – p value*

Table 5d: Dunn's multiple comparison test for syntactic repairs in L2

Dunn	*Difference in rank sum*	*Significant difference*
CARTOON – ROOM1	59	Yes
CARTOON – ROOM2	82.5	Yes
CARTOON – UTTER	60	Yes
CARTOON– STORY	76	Yes

4.1.2 Lexical Repairs

Table 6a presents descriptive statistics for the rate of lexical repairs in L2. According to the normality test, Table 6b, all obtained distributions significantly deviate from the normal distribution. Table 6c presents the Friedman test results. The p-value is lower than *0.0001*, that is, significant differences have been obtained among specific tasks.

Table 6a: Descriptive statistics for lexical repairs in L2

ER_L	*CARTOON*	*ROOM1*	*ROOM2*	*UTTER*	*STORY*
n	101	101	101	101	101
x_{min}	0	0	0	0	0
D_1	0	0	0	0	0
Q_1	0	0	0	0	0
Me	0.446	0	0	0.629	0
Q_3	0.943	1.042	0	0.948	0
D_9	1.495	2.165	1.385	1.522	2.632
x_{max}	2.941	4.348	3.226	2.817	5.000
d	2.941	4.348	3.226	2.817	5.000
d_Q	0.943	1.042	0	0.948	0
$\bar{x}$	0.578	0.527	0.287	0.610	0.539
σ	0.626	0.968	0.699	0.688	1.126
σ^2	0.392	0.937	0.489	0.473	1.268
$S_{\bar{x}}$	0.062	0.096	0.070	0.068	0.112
V	108.21 %	183.80 %	243.28 %	112.90 %	208.76 %
α_3	1.158	1.880	2.501	1.146	2.133
α_4	1.259	3.055	5.648	1.157	3.668

n – sample size, x_{min} *– sample minimum,* D_1 *– first decile,* Q_1 *– lower quartile, Me – median,* Q_3 *– upper quartile,* D_9 *– ninth decile,* x_{max} *– sample maximum, d – sample range,* d_Q *– interquartile range,* $\bar{x}$ *– sample mean, s – sample standard deviation,* s^2 *– sample variance,* $S_{\bar{x}}$ *– standard error of the mean, V – coefficient of the variation,* α_3 *– skewness,* α_4 *– kurtosis*

Table 6b: Shapiro-Wilk normality test for lexical repairs in L2

Shapiro-Wilk	*CARTOON*	*ROOM1*	*ROOM2*	*UTTER*	*STORY*
W	0.8561	0.6151	0.4745	0.8246	0.5514
p	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001

W – Shapiro Wilk test statistic, p – p value

Table 6c: Friedman test for lexical repairs in L2

Friedman	*n*	*df*	F_r	*p*
	101	4	29.911	< 0.0001

n – sample size, df – degrees of freedom, F_r – Friedman test statistic, p – p value

From Table 6a it can be observed that in CARTOON and UTTER the medians differ from zero, whereas in ROOM2 and STORY even the upper quartiles equal zero. Therefore, Dunn's test is employed to compare the following pairs: CARTOON – ROOM2, CARTOON–UTTER, ROOM2 – UTTER, and UTTER – STORY.

Significant differences in the rate of lexical repairs have been obtained in CARTOON, if compared to ROOM2 and STORY, Table 6d. ROOM2 is interesting for two reasons. Firstly, most lexemes are high frequency words (e.g *chair, table, desk, bed*), and secondly, they have already been retrieved from the mental lexicon since it is a repeated task. On the other hand, in CARTOON the speakers retrieve the lexemes for the first time and pay attention to monitoring of content words since they provide understanding of the intended message. Moreover, the learners exhibit less lexical repairs in STORY compared to CARTOON, which might be explained by the fact that in STORY the learners have to invent a narrative under time pressure and less attention is available for monitoring.

Table 6d: Dunn's multiple comparison test for lexical repairs in L2

Dunn	*Difference in rank sum*	*Significant difference*
CARTOON – ROOM2	82.5	YES
CARTOON – STORY	61	YES
ROOM2 – UTTER	−77.5	YES
UTTER – STORY	56	NO

Furthermore, a significant difference between ROOM2 and UTTER has been obtained for the rate of lexical repairs. High frequency words belonging to the same semantic field appear in ROOM2, therefore, different substitutes might have passed unnoticed by the monitor (for instance, *chair* instead of *sofa*), and lexical units corresponding to the concept have already been selected previously. It has commonly been agreed that the three levels of processing are involved in the production of the spoken word. The first phase known as the conceptual planning refers to the propositional content of the intended message; lexical encoding includes the selection of the lexical units corresponding to the concept; and finally, the phonological encoding. In the opinion of many researches all three mechanisms are based on the principle of "competition". Accordingly, concepts, words, and phonemes compete for selection. The choice is the result of the activation level, namely, the selection depends on the activation degree and accessibility. The unit having the greatest activation level is the "Competition winner", and is further processed. The greater the difference in the activation level between the intended and the related unit, the easier and faster the selection becomes. If the difference is lower, the selection among different “competitors“ becomes more difficult. During message planning not only the intended but also related concepts become activated. In ROOM2, when the speaker wants to say “desk“, concepts representing other pieces of furniture also become activated to a certain degree. The activation is further forwarded to the lexical level, as a result of which lexical units such as “bed“ and “chair“ will also compete for selection (Levelt, 1983, 1989, Levelt et al. 1999). In contrast to ROOM2, UTTER requires a large number of words that can be classified as low frequency words belonging to different semantic fields (e.g. *cherry, glove, high heels*), but their usage cannot be avoided, therefore, the speakers make errors while trying to find the corresponding lexical unit in the mental lexicon. Moreover, the monitor responds because these words are crucial for the understanding and correct interpretation of the utterance.

4.1.3 Morphological Repairs

Table 7a displays descriptive statistics for the rate of morphological repairs in L2. According to the results of the Shapiro-Wilk test, Table 7b, four distributions significantly deviate from the normal distribution. The test has not been conducted for ROOM1, since no morphological repairs have been recorded. Table 7c represents the Friedman test results where the p-value is lower than 0.0001, indicating significant differences among individual tasks.

Table 7a: Descriptive statistics for morphological repairs in L2

ER_M	*CARTOON*	*ROOM1*	*ROOM2*	*UTTER*	*STORY*
n	101	101	101	101	101
x_{min}	0	0	0	0	0
D_1	0	0	0	0	0
Q_1	0	0	0	0	0
Me	0.232	0	0	0	0
Q_3	0.542	0	0	0	0
D_9	0.897	0	0	0.671	1.303
x_{max}	2.222	0	1.587	2.273	3.030
d	2.222	0	1.587	2.273	3.030
d_Q	0.542	0	0	0	0
$\bar{x}$	0.326	0	0.016	0.103	0.236
σ	0.412	0	0.158	0.311	0.648
σ^2	0.170	0	0.025	0.097	0.420
$S_{\bar{x}}$	0.041	0	0.016	0.031	0.064
V	126.36 %	-----	1004.99 %	302.22 %	274.55 %
α_3	1.583	0	10.050	4.241	2.902
α_4	3.670	0	101.000	23.648	8.000

n – sample size, x_{min} – sample minimum, D_1 – first decile, Q_1 – lower quartile, Me – median, Q_3 – upper quartile, D_9 – ninth decile, x_{max} – sample maximum, d – sample range, d_Q – interquartile range, $\bar{x}$ – sample mean, s – sample standard deviation, s^2 – sample variance, $S_{\bar{x}}$ – standard error of the mean, V – coefficient of the variation, α_3 – skewness, α_4 – kurtosis

Table 7b: Shapiro-Wilk test of normality for morphological repairs in L2

Shapiro-Wilk	*CARTOON*	*ROOM1*	*ROOM2*	*UTTER*	*STORY*
W	0.7838	-----	0.075	0.3680	0.4193
p	< 0.0001	-----	< 0.0001	< 0.0001	< 0.0001

W – Shapiro Wilk test statistic, p – p value

Table 7c: Friedman test for morphological error repairs in L2

Friedman	*n*	*df*	F_r	*p*
	101	4	112.128	< 0.0001

n – sample size, df – degrees of freedom, F_r *– Friedman test statistic, p – p value*

In CARTOON the median differs from zero, whereas even the upper quartiles of the other tasks equal zero, Table 7a, therefore, CARTOON has been compared to other tasks, Table 7d. The significant differences for the rate of morphological repairs have been obtained between CARTOON and all other tasks. The nature of the retelling task imposes a frequent requirement for the past tense in contrast to other tasks. Consequently, the participants oftentimes faced the problem of verb formation in CARTOON, yet, they paid attention to monitoring of grammatical accuracy in comparison to other tasks.

Table 7d: Dunn's multiple comparison test for morphological repairs in L2

Dunn	*Difference in rank sum*	*Significant difference*
CARTOON– ROOM1	124	YES
CARTOON – ROOM2	120.5	YES
CARTOON – UTTER	93.5	YES
CARTOON – STORY	84.5	YES

4.1.4 Phonological Repairs

Descriptive statistics for the rate of phonological errors per 100 words can be seen in Table 8a. According to the Shapiro-Wilk test, the distribution of all five populations significantly deviate from the normal distribution, Table 8b. Table 8c displays the Friedman test results. Significant differences among different tasks have been obtained, since the p-value of the Friedman test is 0.0002.

Only two pairs have been compared by implementing Dunn's test: ROOM2 – UTTER and UTTER – STORY, since the upper quartile is different from zero in UTTER, whereas in ROOM2 and STORY even the ninth decile equals zero. The results are shown in Table 8d. A significant difference has been obtained between ROOM2 and UTTER.

The nature of UTTER requires the use of some infrequently used lexemes. The speakers repair simple phoneme replacements since these words carrying meaning are of vital importance for the listener. On the other hand, ROOM2 is a repeated task and the speakers' attention shifts from repairing simple structural errors arising at lower levels to problems occurring at the level of discourse (Kormos, 2006).

Taking into account the obtained analysis of error repair categories (syntactic, morphological, lexical, and phonological), the significant difference in the frequency of error repairs between CARTOON and ROOM2, Table 4d, can be better understood. This difference is influenced by the differences in syntactic, lexical, and morphological repairs between the same tasks, Tables 5d, 6d, and 7d, respectively. Significant differences in the repair frequencies between ROOM2 and UTTER, Table 4d, are certainly partly due to the differences in the number of lexical and phonological repairs between these tasks, Tables 6d and 8d. Besides these pairs, a significant difference has been obtained between ROOM1 and ROOM2. The existing difference cannot be explained by the significant difference in the frequency of particular repair categories, since the pair ROOM1-ROOM2 has not been chosen for comparison in any case. Therefore, the findings remain inconclusive, but it can be assumed that the attention available for monitoring in the repeated task is shifted towards other aspects of oral performance.

Table 8a: Descriptive statistics for phonological repairs in L2

E_P	*CARTOON*	*ROOM1*	*ROOM2*	*UTTER*	*STORY*
n	101	101	101	101	101
x_{min}	0	0	0	0	0
D_1	0	0	0	0	0
Q_1	0	0	0	0	0
Me	0	0	0	0	0
Q_3	0	0	0	0.539	0
D_9	0.483	0.808	0	1.508	0
x_{max}	1.176	4.651	2.632	3.571	5.556
d	1.176	4.651	2.632	3.571	5.556
d_Q	0	0	0	0.539	0
$\bar{x}$	0.096	0.199	0.100	0.336	0.171
s	0.231	0.698	0.419	0.669	0.171
s^2	0.053	0.487	0.176	0.448	0.029
$S_{\bar{x}}$	0.023	0.069	0.042	0.067	0.071
V	241.01 %	350.45 %	421.23 %	199.00 %	418.21 %
α_3	2.601	4.319	4.469	2.343	5.500
α_4	6.653	21.015	20.355	6.047	34.899

n – sample size, x_{min} – sample minimum, D_1 – the first decile, Q_1 – lower quartile, Me – median, Q_3 – upper quartile, D_9 – the ninth decile, x_{max} – sample maximum, d – sample range, d_Q – interquartile range, $\bar{x}$ – sample mean, s – sample standard deviation, s^2 – sample variance, $S_{\bar{x}}$ – standard error of the mean, V – coefficient of variation, α_3 – skewness, α_4 – kurtosis

Table 8b: Shapiro-Wilk test of normality for phonological repairs in L2

Shapiro-Wilk	*CARTOON*	*ROOM1*	*ROOM2*	*UTTER*	*STORY*
W	0.4836	0.3232	0.2540	0.5792	0.2616
p	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001

W - Shapiro-Wilk's test statistic, p – p value

Table 8c: Friedman test for phonological error repairs in L2

Friedman	*n*	*df*	F_r	*p*
	101	4	21.971	0.0002

n - sample size, df - degrees of freedom, F_r – Friedman's test statistic, p - p value

Table 8d: Dunn's multiple comparison test for phonological repairs in L2

Dunn	*Difference in rank sum*	*Significant difference*
ROOM2 – UTTER	–51	Yes
UTTER – STORY	46	No

4.2 Appropriacy Repairs In L2

Table 9a presents descriptive statistics for the rate of appropriacy repairs in L2, whereas Table 9b displays the corresponding results of the Shapiro-Wilk test related to particular tasks. It can be noted that all distributions significantly deviate from the normal distribution.

Table 9a: Descriptive statistics for appropriacy repairs in L2

AR	*CARTOON*	*ROOM1*	*ROOM2*	*UTTER*	*STORY*
n	101	101	101	101	101
$\mathbf{x_{min}}$	0	0	0	0	0
$\mathbf{D_1}$	0	0	0	0	0
$\mathbf{Q_1}$	0	0	0	0	0
Me	0	0	0	0	0
$\mathbf{Q_3}$	0.320	0	0	0	0
$\mathbf{D_9}$	0.575	0.842	1.045	0.552	0
$\mathbf{x_{max}}$	0.971	2.222	2.439	1.242	2.941
d	0.971	2.222	2.439	1.242	2.941
$\mathbf{d_Q}$	0.320	0	0	0	0
$\bar{x}$	0.152	0.143	0.172	0.102	0.075
σ	0.244	0.449	0.520	0.280	0.391
σ^2	0.060	0.202	0.270	0.078	0.153
$S_{\bar{x}}$	0.024	0.045	0.052	0.028	0.039
V	160.41 %	313.18 %	302.07 %	276.00 %	521.05 %
α_3	1.319	3.069	3.006	2.703	5.668
α_4	0.534	8.430	8.109	6.226	34.160

n – sample size, x_{min} – sample minimum, D_1 – first decile, Q_1 – lower quartile, Me – median, Q_3 – upper quartile, D_9 – ninth decile, x_{max} – sample maximum, d – sample range, d_Q – interquartile range, $\bar{x}$ – sample mean, s – sample standard deviation, s^2 – sample variance, $S_{\bar{x}}$ – standard error of the mean, V – coefficient of the variation, α_3 – skewness, α_4 – kurtosis

Table 9b: Shapiro-Wilk test of normality for appropriacy repairs in L2

Shapiro-Wilk	*CARTOON*	*ROOM1*	*ROOM2*	*UTTER*	*STORY*
W	0.6670	0.3560	0.3730	0.4111	0.1930
p	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001

W – Shapiro Wilk test statistic, p – p value

Table 9c presents the results of the Friedman test for the rate of appropriacy repairs, showing that the p-value is lower than 0.0001, pointing to the conclusion that significant differences have been obtained among individual tasks.

Table 9c: Friedman test for appropriacy repairs in L2

Friedman	*n*	*df*	F_r	*p*
	101	4	28.000	< 0.0001

n – sample size, df – degrees of freedom, F_r – Friedman test statistic, p – p value

Taking into account that the upper quartile in CARTOON differs from zero if compared to other tasks, Table 9a, CARTOON is examined in contrast to other tasks. The results of the multiple comparison test for the rate of appropriacy repairs can be observed in Table 9d. The significant difference has only been confirmed between CARTOON and STORY. The nature of STORY does not require a more precise description since it has an undetermined chronological order, in other words, it is made up by the speakers, whereas in CARTOON the speakers aim to retell the story more accurately, since their listener has not seen the cartoon and does not know its content.

Table 9d: Dunn multiple comparison test for appropriacy repairs in L2

Dunn	*Difference in rank sum*	*Significant difference*
CARTOON – ROOM1	44.5	NO
CARTOON – ROOM2	42	NO
CARTOON – UTTER	43.5	NO
CARTOON – STORY	62.5	YES

4.3 Different Information Repairs and Rephrasing Repairs in L2

Even though significant differences have been confirmed by means of the Friedman test, Dunn's test has not confirmed which pair combinations exhibit these significant differences. Therefore, no relevant conclusions can be drawn for these two repair categories.

5. Conclusion

Since monitoring involves noticing of errors and inappropriaces, the main goal of this study was to find out how L2 learners allocate their attention for speech monitoring in different task types, thus providing an insight into their self-repair behaviour. By means of statistical testings of all repair categories, the following can be concluded:

a) The retelling of a predetermined chronological sequence of events resulted in a significantly higher rate of syntactic repairs compared to all other tasks. Obviously, in that task type the learners displayed a tendency to interrupt erroneous syntactic structures and to “withdraw“ from situations in which they become aware of the difficulties in formulating the message under time pressure;

b) Significant differences in the rate of lexical repairs have been obtained in the predetermined story retelling, in contrast to the repeated room description and the invented story. In the repeated room description most lexemes are high frequency words (e.g *chair*, *table*, *desk*, *bed*) belonging to the same semantic field and different substitutes might have passed unregistered by the monitor. Furthermore, those lexical units have a higher activation level since they have already been selected previously. On the other hand, in the story narration the speakers retrieve the lexemes for the first time and pay attention to monitoring of content words because they provide understanding of the intended message.

The learners exhibit less lexical repairs in the invented story compared to the predetermined chronological order of the cartoon, since the learners had to make up a story under time pressure and less attention was available for monitoring.

Moreover, a significant difference in the rate of lexical repairs has been obtained between the repeated room description and unrelated utterances. The latter task required a large number of words that can be classified as low frequency words belonging to different semantic fields, but their usage cannot be avoided, therefore, the speakers make errors while trying to find the corresponding lexical unit in the mental lexicon. Obviously, the monitor responds because these words are crucial for the understanding and correct interpretation of the utterance;

c) Significant differences in the rate of morphological repairs have been confirmed between the predetermined retelling of the cartoon and all other tasks. The nature of the cartoon retelling task required the usage of the past tense in contrast to other tasks. Consequently, the participants oftentimes faced the problem of verb formation, yet, they paid attention to monitoring of morphological aspects in comparison to other tasks;

d) A significant difference in the rate of phonological repairs has been obtained between the repeated task and unrelated utterances. In the latter task the learners repaired simple phoneme replacements since these words carrying meaning are of vital importance for the listener, whereas in the repeated task the learners' attention shifted from repairing simple structural errors arising at lower levels to other aspects of oral performance;

e) The pair which compares two almost identical room descriptions has not been chosen for comparison in any repair category. Therefore, a significant difference in the rate of error repairs remains inconclusive, yet, it can be assumed that the attention available for monitoring in the repeated task is shifted towards other aspects of oral performance;

The findings suggest that the participants tend to self-monitor their speech for grammatical aspects in the predetermined retelling task. Therefore, it can be assumed that the retelling tasks might be considered as quality control tasks.

However, it must be pointed out that the presented findings of this research are related to the population of engineering students. Future research will answer the question if the corresponding conclusions are also valid for the general population.

References

[1] A.A. Al-Subaihi, A Monte Carlo study of the Friedman and Conover tests in the single-factor repeated measures design, *J. Statist. Comput. Simul.*, 65(2000), 203-223.

[2] J.R. Anderson, *Learning and Memory*, (1995), Singapore: Wiley.
[3] H. Bortfeld, S.D. Leon, J.E. Bloom, M.F. Schober and S.E. Brennan, Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender, *Language and Speech*, 44(2) (2001), 123-147.
[4] S. Brèdart, Word interruption in self-repairing, *Journal of Psycholinguistic Research*, 20(1991), 123-137.
[5] M. Bygate, Task as context for the framing, reframing and unframing of language, *System*, 27(1999), 33-48.
[6] M. Bygate, Effects of task repetition on the structure and control of oral language, In: M. Bygate, P. Skehan and M. Swain (Eds), *Researching Pedagogic Tasks Second Language Learning and Testing*, (2001), Longman.
[7] D. Coin, A goodness-of-fit test for normality based on polynomial regression, *Computational Statistics and Data Analysis*, 52(2007), 2185-2198.
[8] W.W. Daniel, *Applied Nonparametric Statistics*, (1990), PWS-Kent Publishing Company.
[9] G.S. Dell, A spreading activation theory of retrieval in language production, *Psychological Review*, 93(1986), 283-321.
[10] G.S. Dell, M.F. Schwartz, N. Martin, E.M. Saffran and D.A. Gagnon, Lexical access in aphasic and non aphasic speakers, *Psychological Review*, 104(1997), 801-838.
[11] R. D'Ely, A focus on learners' metacognitive processes: The impact of strategic planning, repetition, strategic planning plus repetition and strategic planning for repetition on L2 oral performance, *Doctoral Thesis: Florianopolis*, (2006), Universidade Federal de Santa Catarina.
[12] J. Demšar, Statistical comparisons of classifiers over multiple data sets, *Journal of Machine Learning Research*, 7(2006), 1-30.
[13] A.K. Fathman, Repetition and correction as an indication of speech planning and execution processes among second language learners, In: H.W. Dechert and M. Raupach (eds.), *Towards a Crosslinguistic Assessment of Speech Production*, (1980), 77-85, Frankfurt, Germany: Peter D. Lang.
[14] A. Field, *Discovering Statistics using SPSS*, (2005), Sage Publications: UK.
[15] K.R. Finardi, Effects of task repetition on L2 oral performance, *Trabalhos em Linguistica Aplicada*, 47(1) (2008), 1-9.
[16] M.B.M. Fortkamp, Working memory capacity and speech production: An exploratory study, *Doctoral Thesis Florianópolis*, (2000), Pós-Graduação em Inglês e Literatura Correspondente, UFSC.
[17] M. Goldrick and B. Rapp, Lexical and post-lexical phonological representations in spoken production, *Cognition*, 102(2007), 219-260.
[18] M. Gósy, *Pszicholoingvisztika*, (2005), Budapest: Osiris Kiadó.
[19] A.R. Henderson, Testing experimental data for univariate normality, *Clinica Chimica Acta*, 366(2006), 112-129.
[20] A.E. Hieke, A content-processing view of hesitation phenomena, *Language and Speech*, 24(1981), 147-160.
[21] S. Keskin, Comparison of several univariate normality tests regarding type I error rate and power of the test in simulation based small samples, *Journal of Applied Science Research*, 2(5) (2006), 296-300.
[22] J. Kormos, A new psycholinguistic taxonomy of self-repairs in L2: A qualitative analysis with retrospection, Even Yearbook, *ELTE SEAS Working Papers in Linguistics*, 3(1998), 43-68.
[23] J. Kormos, Monitoring and self-repair in L2, *Language Learning*, 49(1999), 303-342.
[24] J. Kormos, The role of attention in monitoring second language speech production, *Language Learning*, 50(2000), 343-384.
[25] J. Kormos, *Speech Production and Second Language Acquisition*, (2006), Lawrence Erlbaum Associates, New York London.

[26] M.M. Kovač, Speech errors in English as foreign language: A case study of engineering students in Croatia, *English Language and Literature Studies*, 1(1) (2011), 20-39.
[27] M.M. Kovač and B. Milatović, Analysis of repair distribution, error correction rates and repair successfulness in L2, *Studia Linguistica: Journal of General Linguistics*, 67(2) (2013), 225-256.
[28] P. Lennon, Investigating fluency in EFL: A quantitative approach, *Language Learning*, 40(1990), 387-417.
[29] W.J.M. Levelt, Linearization in describing spatial networks, In: S. Peters and E. Saarinen (eds.), *Processes, Beliefs and Questions*, (1982), Dordrecht: Reidel.
[30] W.J.M. Levelt, Monitoring and self-repair in speech, *Cognition*, 33(1983), 41-103.
[31] W.J.M. Levelt, *Speaking: From Intention to Articulation*, (1989), Cambridge, MA: MIT Press.
[32] W.J.M. Levelt, A. Roelofs and A.S. Meyer, A theory of lexical access in speech production, *Behavioral and Brain Sciences*, 22(1999), 1-75.
[33] J.P. Marques de Sá, *Applied Statistics using SPSS, STATISTICA, MATLAB and R*, (2007), Springer Verlag: Berlin Heidelberg.
[34] H.S. Moat, M. Corley and R.J. Hartsuiker, Connecting phonological encoding to articulation: Is cascading required? A computational investigation, In: B.C. Love and K. McRae and V.M. Sloutsky (eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, (2008), 1320-1325, Cognitive Science Society: Austin, TX.
[35] D.C. Montgomery and G.C. Runger, *Applied Statistics and Probability for Engineers*, (2003), John Wiley and Sons: USA.
[36] S.G. Nooteboom, Speaking and unspeaking: Detection and correction of phonological and lexical errors in spontaneous speech, In: V.A. Fromkin (Ed.), *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen and Hand*, (1980), 87-96, New York & London: Academic Press.
[37] S. Nooteboom, Self-monitoring of inner speech and overt speech, *From Sound to Sense at MIT*, (2004), 187-190.
[38] M.A. Pett, *Nonparametric Statistics for Health Care Research (Edition 2)*, (1997), Sage.
[39] S. Pillai, Self-monitoring and self-repair in spontaneous speech, *A Biannual Publication on the Study of Language and Literature*, 8(2) (2006), 114-126.
[40] P. Robinson, Task complexity, cognitive resources and syllabus design: A triadic framework for examining task influences on SLA, In: P. Robinson (Ed.), *Cognition and Second Language Instruction*, (2001), 287-318, Cambridge: Cambridge University Press.
[41] R. Schmidt, The role of consciousness in second language learning, *Applied Linguistics*, 11(1990), 129-158.
[42] R. Schmidt, Awareness and second language acquisition, *Annual Review of Applied Linguistics*, 13(1993), 206-226.
[43] R. Schmidt, Deconstructing consciousness in search of useful definitions for applied Linguistics, In: J.H. Hulstijn and R. Schmidt (Eds.), *Consciousness in Second Language Learning (Special Issue)*, AILA Review, 11(1994), 11-26.
[44] P. Skehan, *A Cognitive Approach to Language Learning*, (1998), Oxford: Oxford University Press.
[45] P. Skehan and P. Foster, Cognition and tasks, In: P. Robinson (ed.), *Cognition and Second Language Acquisition*, (2001), Cambridge: C.U.P.
[46] H.C. Thode, *Testing for Normality*, (2002), Marcel Dekker: New York.
[47] G.W.C.M.V. Hest, *Self-Repair in L1 and L2 Production*, (1996), Tilburg University Press.
[48] M.A.V. de Wiel, Exact null distributions of quadratic distribution-free statistics for two-way classification, *Journal of Statistical Planning and Inference*, 120(2004), 29-40.

[49] L.T. Verhoeven, Monitoring in children's second language speech, *Second Language Research*, 5(1989), 141-155.

[50] J. Zhang and W. Yuehua, Likelihood-ratio tests for normality, *Computational Statistics and Data Analysis*, 49(2005), 709-721.