

Speech Errors in English as Foreign Language: A Case Study of Engineering Students in Croatia

Mirjana M. Kovač

FESB, University of Split

Ruđera Boškovića 32

21 000 Split, Croatia

E-mail:mirjana.kovac@fesb.hr

Received: March 2, 2011 Accepted: March 22, 2011 doi:10.5539/ells.v1n1p20

Abstract

The study reported in this paper investigates the frequency and distribution of speech errors, as well as the influence of the task type on their rate. The participants of the study were 101 engineering students in Croatia. A recorded speech sample in the English language (L2) lasting for approximately ten hours was transcribed, whereby more than three and a half thousand speech errors were recorded. Morphological errors were dominant due to a significantly frequent omission of articles. The distribution of different subcategories of lexical errors pointed to a relatively low frequency of unintended L1 switches, indicating that the participants were able to separate the two languages during lexical access.

Statistical testings of the influence of the task type on speech errors displayed that the retelling of a chronological order of events resulted in a significantly higher rate of syntactic errors if compared to other tasks. Due to limited attentional resources and insufficient knowledge, the speaker cannot process the message within the time constraints. The rate of lexical and phonological errors depended on the frequency of use, that is, less frequently used words were more susceptible to lexical errors than high-frequency words. The retelling of a chronological order of events is a demanding task, for this reason, this task type should be more practiced in foreign language teaching.

Keywords: Models of speech production, Monitoring, Disfluencies, Speech errors, Syntactic errors

1. Introduction

The creation of an utterance is a demanding, complex and continuous process, where different components in the mechanism of language production can fail causing speech errors (Kovač and Horga, 2010). The process of speech production comprises four main activities which proceed in successive order, as follows: a) conceptualization, that is, planning of the content of the utterance; b) formulation, which includes grammatical, lexical and phonological encoding; c) articulation, the phase that represents overt speech, and finally d) self-monitoring, that includes the verification of the correctness or appropriateness of the produced utterances (Kormos, 2006). While content planning in the native language requires increased conscious attention on the part of the speaker, the formulation and articulation are automated processes that can run in parallel without any conscious speaker's effort. However, despite the automated nature of the native language, the speakers do not produce perfect speech, on the contrary, their speech displays different forms of disfluencies, such as hesitations, false starts, repetitions, vowel prolongations, speech errors and the like. Speech errors are deviations from the speaker's communicative intention and are an important source of information for understanding the complex mechanisms of language production. The researches dealing with speech errors in the native language (e.g. Fromkin, 1973; Dell and Reich, 1981; Stemberger, 1985; Levelt, 1989; van Hest, 1996; Erdeljac, 2005) and in the foreign language (e.g. Poulishse, 1999; van Hest, 1996) indicated that lexical errors, which occur as a result of erroneous retrieval of lexemes in the mental lexicon, are very frequent. On the other hand, in the foreign language, errors are often results of insufficient knowledge of language or the semantic noise in the communication process (Jaeger, 2005). There are different theories regarding the reason for the occurrence of speech errors. Dell (1986) argued that the erroneous activation of certain nodes causes speech errors, explaining that the realization of a particular unit depends on the degree of its activation, but also on the degree of activation of other units that are 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.

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, that is, the interlocutor can correctly interpret the message without correction. However, most errors are corrected without the intervention of the interlocutor, which confirms the existence of a system for speech monitoring and the feedback loops. Their task is to control the correctness of speech utterances and consequently, if necessary, to execute the correction or enrichment of the propositional content. In recent decades a growing interest in language production has resulted in numerous psycholinguistic models which try to explain the mechanisms in the service of speech production, which is speech errors, self-corrections and various forms of disfluencies.

The aim of this study was to investigate the distribution of different categories of errors, and to determine the influence of the task type on the rate of certain categories of errors in the speech of technical studies students. The paper describes Kormos' bilingual model (2006) of speech production, which is based on Levelt's model (1989) as the empirically most accepted theory for monolingual speech processing. In addition, a brief overview of the findings dealing with speech errors and self-monitoring is presented, as well as the classification of errors based on the described model for the purpose of this study. Afterwards, the research methodology is presented, followed by the obtained results and corresponding conclusions.

2. Speech production

2.1 Speech production models

Kormos' bilingual speech production model (2006) is based on Levelt's model (Levelt 1989; 1993; 1995; 1999), as the most accepted and most widely used model in studies of speech production (Figure 1). She assumed that the bilingual speech production is modular and can be described as a series of relatively independently functioning processing components: the conceptualizer, the formulator, the articulator and the acoustic-phonetic processor, and finally, the speech comprehension system or parser. There are also three knowledge stores: the mental lexicon, the syllabary and the store for the knowledge of external and internal world. She proposed that the new model contains one large memory, called long-term memory, which is subdivided into several subcomponents: episodic memory, semantic memory including the mental lexicon, the syllabary, and a store for declarative knowledge of L2 rules. All knowledge stores are shared between L1 and L2, that is, there is a common episodic and semantic memory for L1 and L2, a shared store for L1 and L2 lemmas and lexemes, and for L1 and L2 articulatory scores. In L2 production, however, she postulated the existence of a fourth and L2 specific knowledge store: a declarative memory of syntactic and phonological rules in L2.

In Kormos' model the processing components are "specialists" in certain functions that must be executed, which means that they do not share processing functions. A component or module will start processing only if it receives a distinctive input. Kormos assumed that processing is incremental, that is, a fragment of a module's characteristic input can trigger encoding procedures in this module. Consequently, the articulation of an utterance can start even before the speaker has finished planning the entire message. This also means that in the case of more proficient speakers parallel processing is possible. Incremental and parallel processing as well as the automated nature in the native language may explain the high rate of speech production. In L1 production message conceptualization requires attention, whereas formulation and articulation are automatic processes which work in parallel, which makes L1 speech generally smoother and faster compared to L2.

Abutelabi et al. (2001; 2005) conducted neuroimaging studies and provided support for the similar nature of L1 and L2 speech processing, that is, neither the extent of brain activation nor the regions involved in the processing in L1 and L2 are different for bilinguals who learned the L2 early in their lives and for high proficient speakers with extensive L2 exposure. On the other hand, low proficient speakers and those who have had a low exposure to the target language, were found to activate larger and slightly different cerebral areas when speaking in L2 than in L1 (Kormos, 2006). Kormos' model accounts for this finding because proficient bilinguals do not rely on the separate knowledge store of declarative rules, whereas for learners at lower levels of proficiency grammatical and phonological rules are stored in a separate brain region.

The second theory explaining speech production is Dell's (1986) spreading activation theory. According to this theory, the lexicon, which contains non-productive knowledge, is seen as a network of interconnected units, the so-called nodes, which represent linguistic units such as concepts, words, morphemes, phonemes and so on. In the lexicon conceptual nodes are associated with nodes defining the words, which are in turn associated with morphemic nodes, which represent specific morphemes. Furthermore, there is a relationship between morphemic and phonemic nodes that determine the phoneme, and finally phonemic nodes that are connected with nodes for phonemic features. The mechanism responsible for the production of a sentence is the process of spreading activation. During speech processing the node of the required category (e.g. the node representing a word) is approached, in other words, it is the node that has the greatest degree of activation or arousal. Activation spreads from one level to another, from nodes representing semantic features to the corresponding word or lemma nodes,

which is then forwarded to phonemic nodes. This means that activation cascades from one level to another.

Priming produces increased subthreshold activity in a node. It spreads in parallel across nodes, with its strength being positively related to how well it matches the input. In the same fashion, priming spreads to the phonological nodes and, finally, to the lexical nodes. At each level, priming strength is a function of the match to the input from the preceding level. Although several nodes may be primed at once, only the node that is primed to a higher degree than all the others becomes activated (Shoaf i Pitt, 2002:1).

The model suggests the existence of inhibitory connections between nodes at the same level of processing and thus ensures the correct selection of nodes. Dell (1986) and his followers assumed that activation can spread in both directions. This means that a wrongly selected and activated node starts spreading activation in the opposite direction, that is "bottom up". In the case of wrongly chosen words, activation will spread from the level of nodes that represent words to conceptual nodes. It is assumed that this same feedback mechanism operates in the perception of speech and makes the monitor an inherent property of the process of perception and speech production.

2.2 Speech errors and self-monitoring

Errors can occur at each phase of speech production, that is, during lemma retrieval, grammatical and phonological encoding, as well as articulation. When the monitor notices an error, an alarm signal is sent to the conceptualizer, which then in turn issues the same preverbal plan, hoping that the renewed message will now be properly articulated (Levelt, 1989). Levelt (1983) distinguished lexical, syntactic and phonetic errors that correspond to the three basic levels of processing. Lexical errors are "any lexical item, color words, direction terms, prepositions, articles, etc." (Levelt, 1989:54). Levelt (1989) assumed that in the case of lexical errors an incorrectly activated lexical entry is retrieved and then articulated. The results of speech error studies indicated that errors as a result of faulty L2 lexical access are very common (Poulisse, 1993), but they also often occur in L1 (Dell and Reich, 1981; Fromkin, 1973; 1980; Stemberger, 1985). In Levelt's theory (1989) content and functional words as well as collocations and idioms are considered lexical entries. Errors of derivational morphology, such as *different* instead of *difference* also belong to the category of lexical errors, because derivations in Levelt's model of the lexicon represent different lexical entries. Syntactic errors include the syntactic structure that leads to a deadlock after which the speaker can not continue the utterance (Levelt, 1983:54). Unfortunately, Levelt did not give a precise definition of phonetic errors and did not provide a definition of morphological errors.

Nevertheless, speakers correct most errors without being interrupted by the interlocutor, which points to the existence of the speech monitoring system. Levelt's (1983; 1989) perceptual loop theory explains the detection of erroneous output data. The model includes three loops for checking the utterance. The first or conceptual loop (Levelt, 1989; Blackmer and Mitton, 1991) compares the preverbal plan with the original intention of the speaker before it is forwarded to the formulator for further processing. Its task is to check the appropriateness of the expression, that is, to detect conceptual errors. Blackmer and Mitton (1991) and Van Hest (1996) concluded that conceptual errors are significantly slower corrected than lexical and phonological errors. It is explained by the assumption that an erroneous concept is more difficult to reject and replace with a new one and therefore more time is required for self-correction of conceptual errors (Postma, 2000).

The second or inner loop is responsible for monitoring the phonetic plan or the inner speech before articulation, which is known as covert monitoring (Wheeldon and Levelt, 1995). It allows the speaker to reveal the error before articulation. Nootboom's (2010) analysis based on a collection of experimentally elicited spoonerisms and a collection of speech errors in spontaneous Dutch supported an important aspect of Levelt's perceptual loop theory of monitoring, that both inner and overt speech are being monitored for speech errors. There is at least a delay of 200 or 250 ms between the two (Hartsuiker and Kolk, 2001), and potentially much more, depending on how much material is buffered in inner speech. If an error is not detected in inner speech, detection of this error in overt speech is of course perfectly appropriate. It has been assumed that monitoring for speech errors is under attentional control (Hartsuiker, Kolk and Martensen, 2005) which means that the speaker can pay either more attention to inner speech or more to overt speech. Postma and Kolk (1993) suggested that the articulator has an autonomous ability to restart, and will implement the old program for the second time if it does not receive new input data on time. This usually happens at higher speech rates, which prevent the temporary storage and thus increase the possibility of time miscalculation, resulting in repetitions of shorter speech segments. Error detection of overt speech depends on the availability and capacity of the articulatory buffer. Higher speech rates reduce temporary storage of speech (Levelt, 1989; Blackmer and Mitton, 1991; Van Hest, 1996).

Finally, the produced utterance is checked after articulation, which constitutes the outer loop of monitoring or the

acoustic-phonetic processor. According to Levelt, the monitor is located in the conceptualizer but receives the information from a separate speech comprehension system, the parser, which is associated with the mental lexicon. The monitor will issue an alarm signal if an error or inappropriateness is detected in any of the stages of processing, whereby the same mechanisms of speech production will run for a second time. In order to avoid reduplication of knowledge, Levelt considers that the same lexicon is used for the production and perception of one's own speech, and the same system is used for understanding and monitoring of the interlocutor via the acoustic-phonetic module.

Studies of speech errors in the native language confirmed the superiority of this theory over others (e.g. Blackmer and Mitton, 1991; Nooteboom, 2005; 2010). Moreover, theories of foreign language speech production (e.g. Kormos, 2006) also accepted this theory and consider that there are no qualitatively different monitoring mechanisms in a foreign language.

3. The research project

The aim of this study was to determine the distribution and frequency of different categories of speech errors in English as foreign language, and to examine the influence of the task type on their occurrence.

3.1 The experimental procedure and tasks

The participants of this study were 101 native speakers of Croatian, 51 male and 50 female first-year students at the Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture in Split. The students 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. The participants were randomly chosen and no test was previously administered to measure the level of proficiency. The error data were collected by means of five different tasks which were performed in English (L2). All data were collected in an ordinary room 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 with the microphone were placed between the subjects and the researcher. Each student was individually audio-recorded and afterwards the speech samples were transcribed by two independent researchers. Time parameters were measured in *Audicity*, the programme for acoustic analysis of speech. The number of words was counted manually.

Errors, as variables of speech fluency, were collected by means of five tasks. Before carrying out the first 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 his/her own words in the English language. The participant had one minute to prepare and was not time limited.

The second task was preceded by the researcher's instructions which explicitly stated that the subject had to describe the picture of a room (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 second task (ROOM2) was almost identical to the first one, except for the furniture that was slightly differently arranged.

In the fourth task (UTTER) the subjects formed twenty utterances based on different, semantically unrelated drawings. The syntactic frame was not defined by the researcher, and the only requirement was that the drawing and the corresponding colour (e.g. the rose which was red or the blue shirt) appeared in the utterance. As in the case of previous tasks, each participant had one minute to prepare.

The fifth task (STORY) was a story narration. The participant had to make up a story based on five unrelated drawings and none of them was allowed to be omitted.

3.2 Classification of errors and analysis

A system of classification of errors was devised on the basis of previous taxonomies of errors and self-corrections (Levelt, 1983). The reason for choosing Levelt's model was because it was the most complete and the most widely applied system available. For the analysis of errors, Lennon's (1991) definition was used.

An error is a linguistic form or combination of forms, which in the same context and under similar conditions of production, would in all likelihood, not be produced by the speaker's native counterparts (Lennon, 1991:182).

On the basis of this definition, the researcher identified all the possible instances of errors in the transcripts of 101 students. 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. Only the cases which were considered unacceptable by the researcher and native speakers were counted as errors.

Errors were analyzed in accordance with Levelt's (1983:44) structure of error self-repair which includes three parts:

<u>'Go from left again to</u>	<u>uh..</u>	<u>from pink again to blue'</u>
Original utterance	editing	alteration
	phase	REPAIR

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*'). 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 overt articulation, but 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. After the identification of errors, all the instances of erroneous utterances were classified in accordance with the classification taxonomy.

The main criterion for distinguishing lexical from grammatical errors was the way of accessing lexical entries. Lexical entries can be accessed via syntactic building procedures or on the basis of their conceptual specifications, or lemma activation. Firstly, it is proposed that errors of derivational and inflectional morphology should be handled separately. As derivations (*rotate*, *rotating*, *rotation*) are assumed to be different lexical entries, and word-formation is supposed to take place as a part of lexical encoding (Levelt et al. 1999), errors of derivational morphology should be classified as lexical errors. On the other hand, inflectional morphemes are encoded and processed in the grammatical encoding phase and therefore are handled as morphological errors. Lexical, syntactic, morphological and phonological errors are errors arising at lower levels of processing, whereby conceptual errors (Van Hest, 1996) have their roots at higher levels and occur as a result of an erroneous conceptual plan. Conceptual errors are not found in Levelt's classification system, nor are they included in the classification of this study, because the author believed that only a retrospective analysis could reliably determine whether an error was a result of faulty lemma retrieval or occurred at the level of conceptualization.

a) Lexical errors

The first type of error to be discussed is lexical errors. Three subcategories of lexical errors have been established: a) idioms, collocations, functional and content words, errors of derivational morphology (1); b) unintentional use of L1 lexemes (2); c) non-existent words (3).

(1) *And then they hear some **beautifully** singing and er, er, knew it was Johnny's mother.*

(2) *Johnny Bravo went to the **šum-sorry**, to the wood.*

(3) *er, so, Johnny Bravo, his mother was making him a sweater of **/pineko/**.*

One of the problems in the classification of lexical errors are errors within idioms and collocations. In example (4) the subject chooses the preposition '*on*' and then replaces with '*in*'.

(4) ***On** the second -in the second picture I can see a lamp and a vase.*

It can be assumed that in the case of L2 speakers many idioms and collocations are not fully automated, therefore they are not stored as complete lexical entries. It is assumed that the mechanisms by which the L2 speakers produce these phrases are similar to those by which new words are created (Kormos, 1998), and for this reason these errors are classified as lexical errors, which do not occur at the stage of lemma activation, but in the process of lexical encoding.

b) Grammatical errors

i) Syntactic errors

Opposite to lexical errors that arise as a result of incorrect lexical approach, grammatical errors occur as a result of problems in grammatical encoding (Levelt, 1989). The differentiation between lexical and grammatical errors is not simple, since the processes of lexical access and grammatical encoding are closely related. For example, prepositions are in some cases classified as lexical, whereas in some other instances as grammatical errors. In order to overcome these difficulties of classification, it is important to distinguish function words which have conceptual specifications from those which do not have them. The preposition "under" (e.g. the ball is under the desk) has its conceptual specification, whereas the preposition "to" (e.g. She listens to music) has no conceptual specification and is addressed via the syntactic building procedures and therefore regarded as a syntactic error.

Syntactic errors were analyzed according to where they occurred in the the three stages of the grammatical encoding phase, that is, (i) when the various complements, specifiers and parameters are encoded and (ii) handled by the different subroutines, and (iii) when these processed materials are ordered (Kormos, 1999). The next phase involves the production of speech errors that occur at the stage of grammatical encoding.

The following subcategories of syntactic errors have been established: a) wrong word order (5); b) unfinished expressions or false starts (6); c) completely unacceptable morpho-syntactic and/or semantic structure (7) (Note 1); d) wrongly encoded complements and specifiers, which are accessed via syntactic building procedures (8).

(5) *Johnny met a bear who was, er, sleeping...er, he woke **up him** in the middle of, er, winter.*

(6) *And he came to er, er, cave...**er, er, that cave was er, er** -in that cave he saw er, a bear.*

(7) *They start to arguing and hunting into the woods **but er, it er, didn't er, was success.***

(8) *The bear and Johnny listened--- beautiful singing of J ohhny's mother.*

ii) Morphological errors

include a) inflexional errors, when the speaker chooses the wrong verb form (9); b) incorrect plural of nouns (10);

(9) *well Johnny Bravo's mum **was mading er**, was making a coat for J ohhny.*

(10) *Lots of **womans** wear black shoes.*

c) errors for "time and aspect" (11), when the speaker uses the wrong time or aspect. In Levelt's classification of errors (1989) this category does not exist, probably because these errors are very rare in L1 (Van Hest, 1996);

(11) *He er, er, was trying to find er, er, to find his mum and he **wants** to escape from er, forest.*

d) omitted article (12) or misused article (13).

(12) *er Johnny was in -- wood with his mum*

(13) *we are now er in front of **the** my room.*

c) Phonological errors

The next steps in speech processing are phonological encoding and articulation. Phonological errors mainly included articulatory clumsiness (14), whereas other forms of phonological errors 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.

(14) *his mother, his mother er, er, sang a **/ole/** lullaby and er, they fell asleep.*

Table 1 presents the classification of errors used in this study.

4. Results and discussions

4.1 Error distributions

The participants of the study produced a total sum of 56 862 words, Table 2, whereby almost ten hours of recorded speech was collected.

Table 3 displays the error distribution for all tasks. It can be observed that a total number of 3518 errors was recorded, which outnumbered the total number of errors in the native language (Croatian, L1-943) (Kovač and Horga, 2010) for the same task types. Morphological errors were the most frequent ones (over 60% of the total number of errors), followed by syntactic and lexical errors which occurred in approximately equal proportion (18%). Phonological errors constituted 2.5% in the error corpus. The main reason for such a large number of morphological errors can be found in the article system, which is, omitted or incorrectly used articles. If this subcategory had been left out, the total number of morphological errors would have been only slightly higher than the number of lexical and syntactic errors respectively.

Considering the proportions of certain subcategories of syntactic errors in their total number, Figure 2, it can be noted that the subcategories: wrong word order or unfinished utterances (false starts) and morphosyntactic and/or semantically completely unacceptable utterances occurred more frequently than the subcategory of wrongly encoded complements and specifiers, which are accessed via syntactic building procedures.

The distribution of the subcategories of lexical errors, Figure 3, pointed to a rare presence of L1 lexemes and L2 non-words in relation to lexical errors caused by incorrect selection of L2 lexical units. It follows that the speakers at this level of language competence generally manage to keep separate two languages while lexical selection. Separation of the two languages is hypothesized to be controlled by the language cue added to the concepts in the

conceptualization phase (Kormos, 2006).

The analysis of the distribution of morphological errors showed that the participants of this study displayed severe difficulties in English article use, Figure 4. They were found to employ bare nouns without the definite/indefinite article. The omission was the main problem and notorious source of difficulty, largely because it is based on a complex set of abstract distinctions, which are to a certain degree, arbitrarily mapped onto surface forms (*a, the* and *zero*) in article use. This difficulty only occurs when learner's first language does not share the same surface features and/or the same semantic conceptualizations with the second language (L2) (Ekiert, 2007). This statement can be confirmed by the results shown in Table 4. In 96% of all cases of difficulties in article usage, the definite/indefinite article was omitted, whereas in only 4% of all cases it was overused or incorrectly used. This study confirmed the results of a recent study (Avery and Radisic, 2007), emphasizing omission as the main problem in article use when the L1 semantic and conceptual system is different from L2, which points to lower levels of proficiency than expected.

Considering the distribution of errors for each task separately, Table 3, it may be noted that morphological errors occurred in the highest proportion in each task, whereas the difference between syntactic and lexical errors for each task was within the limits of $\pm 6\%$, that is, their proportions were very similar. It is noticeable that incorrect verb form as a subcategory of morphological errors outnumbered the total number of lexical errors in CARTOON and STORY. The explanation can be found in the nature or cognitive complexity of the retelling task, which imposes a frequent usage of the past tense and the participants very often faced the problem of verb formation, as they often did not know the inflexion of the verb forms, due to insufficient grammar knowledge. On the other hand, in designing simple utterances, as in tasks ROOM1, ROOM2 and UTTER, speakers deliberately used simple tenses and simple syntactic frames.

4.2 The influence of the task cognitive complexity on error categories in L2

The present paper examined how the task type influenced the specific error category, in other words, whether there was a statistically significant difference in the rate of certain error categories among different tasks.

Speech duration per participant may be of varying size in each task. In order to facilitate the comparison of various tasks by applying appropriate statistical tests, it was necessary to perform the calculations by reducing the corresponding categories of errors per subject to a pre-specified number of words. In the available literature that number is in most of the cases 100 words, so the same approach was chosen in this paper. The analysis of the influence of the task type on the rate of examined error categories was performed on the same subjects (101 undergraduate student). Consequently, there were five dependent samples.

4.2.1 Syntactic errors

Table 5 presents descriptive statistics related to the rate of syntactic errors in L2. The main reason for such a detailed data presentation was a requirement for a good knowledge of samples, in order to decide which samples to compare with each other, if the statistical test results displayed that the population from which the samples come, were not identical.

Shapiro-Wilk's test of normality (Marques de Sá, 2007) was selected, Table 6, being highlighted in recent works as the most efficient test of normality (Thode, 2002; Zhang and Yuehua, 2005; Keskin, 2006; Henderson, 2006; Coin, 2008). According to this test, it may be noted that the distributions of all five populations significantly deviated from the normal distribution. Therefore, the non-parametric Friedman's test was selected (Field, 2005; Marques de Sá, 2007; Demšar, 2006), as probably the most commonly used (Al-Subaihi, 2000) and most popular (Van de Wiel, 2004) nonparametric repeated measures test in the case of three or more dependent samples, Table 7. Nonparametric tests are called distribution-free tests, which require neither normally distributed data, nor homogeneity of variance (Montgomery and Runger, 2003). However, one of the drawbacks of nonparametric tests are that they are less efficient than parametric ones, but only if the assumptions for using parametric tests are fulfilled, which was not the case with regard to the result of Shapiro-Wilk's test. If the assumptions for using parametric tests are not met, the nonparametric tests often provide considerable improvement (Montgomery and Runger, 2003). The value of Friedman's test was lower than 0.0001, Table 7, displaying that statistically significant differences in the rate of syntactic errors among individual tasks were obtained. Thus, all populations were not identical considering that the p-value was lower than or equal to the threshold value α ($p < 0.0001 \leq \alpha = 0.05$).

The results of Friedman's test revealed that the null hypothesis of identical populations should be rejected, yet, it did not answer the question which tasks, when compared, displayed statistically significant differences in the rate of syntactic errors. 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) answered this question.

Comparisons can be performed only for selected pairs of tasks or for all possible combinations of pairs. 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 being reduced. The decision which pairs to compare can be made based on a good knowledge of descriptive statistics related to each sample, Table 5. In conclusion, it is important to be restrictive in choosing comparison pairs (Marques de Sá, 2007).

The lower quartile was different from zero in CARTOON, while in other tasks the lower quartile and the median equaled zero, Table 5, and therefore Dunn's multiple comparison test for the rate of syntactic errors was implemented for combinations of task pairs that included CARTOON as a member of the pair. In other words, CARTOON was compared with each task, Table 8. The results of Dunn's test, Table 8, showed that the retelling of a chronological sequence of events resulted in a high rate of syntactic errors. An analog conclusion was presented in the case of the rate of syntactic errors in the native language (Kovač and Horga, 2010). In recounting the chronological sequence of actions, it was necessary to organize the speech acts under time constraints and therefore monitoring was very important. However, monitoring is a conscious process that requires special attention (Levelt, 1989) and it can be assumed that attentional resources are very limited in the case of less proficient speakers, consequently, the amount of attention available for monitoring will have an impact on the efficiency of the process of speech production. This is especially important for L2 speakers whose production processes are partly automated and more attention is required than the encoding process in the native language. Since L2 speakers at a lower level of language competence have a limited amount of declarative knowledge and the mechanisms of linguistic encoding are not automated, it can be concluded that the speaker's attention was oriented towards the processes of lexical, grammatical and phonological encoding, and that less attention was available for other stages of messages production. This referred to the conceptualization of the message, formulation and articulation, and finally monitoring. Since the speech acts followed each other in successive order, which happened in the case of the retelling of a chronological order of events, the speaker needed more time for message planning, and very often did not have sufficient time to properly form the utterance.

4.2.2 Lexical errors

Descriptive statistics for the rate of lexical errors per 100 words can be seen in Table 9. According to Shapiro-Wilk's test for the rate of lexical errors per 100 words, Table 10, all distributions significantly deviated from the normal distribution. The p-value of Friedman's test was 0.0108, Table 11, indicating that statistically significant differences in the rate of lexical errors were obtained. In CARTOON the lower quartile was different from zero, whereas in all other tasks (except UTTER) even the median equaled zero, Table 9, therefore, CARTOON was compared with the tasks ROOM1, ROOM2 and STORY. Dunn's test, Table 12, showed that in CARTOON speakers made significantly more lexical errors than in the tasks ROOM2 and STORY. The findings could be explained by the fact that the same lexemes occurred in ROOM1 and ROOM2, some of which were high frequency words and more familiar to the speakers (e.g. *chair, table, desk, bed*). A very similar conclusion could be made for STORY, due to the occurrence of high frequency words such as *church, forest, children* etc.

4.2.3 Morphological errors

Descriptive statistics for the rate of morphological errors are presented in Table 13. The results of the Shapiro-Wilk's test, Table 14, lead to the same conclusion as in the previous analysis of error categories, pointing that all distributions significantly deviated from the normal distribution. The p-value of Friedman's test was 0.0036, Table 15, which indicated that statistically significant differences in the rate of morphological errors among different tasks were obtained.

The lower quartile differed from zero in CARTOON, only. CARTOON was compared with all tasks since it is evident that the greatest difference in medians and means existed between ROOM2 and UTTER, Table 13. For that reason, the pair ROOM2-UTTER was added to the chosen ones. Dunn's test confirmed the statistically significant differences in the rate of morphological errors between CARTOON and ROOM2, as well as between ROOM2 and UTTER, Table 16. A significantly greater rate of morphological errors was obtained in the task ROOM2, compared to CARTOON and UTTER. As can be seen in Table 3, in CARTOON 1.683 omitted or incorrectly used articles were recorded per 100 words, whereas that number significantly

increased in ROOM2 and amounted to 5.034 per 100 words. This discrepancy could be clarified by the nature of task ROOM2 which required frequent article usage (*There is a bed, a table...on the right side, on the left side, near the table...*).

Kormos (2000) concluded that formal methods of language teaching influence the degree of importance that students pay to grammatical accuracy. In a research review on the impact of formal instruction on grammatical accuracy, Ellis (1994) argued that if formal language teaching is directed towards grammatical accuracy at the appropriate stage of speaker's L2 language development, it will result in improved accuracy. From the results of this study it can be concluded that speakers paid very little importance to the use of articles, which can be explained by: a) the influence of the native language (Croatian) which does not have the article system, b) the students' attitude that articles are not important for message understanding and c) formal language teaching in primary and secondary schools does not give adequate importance to article acquisition.

4.2.4 Phonological errors

Descriptive statistics for the rate of phonological errors per 100 words can be seen in Table 17. According to Shapiro-Wilk's test, the distribution of all five populations significantly deviated from the normal distribution, Table 18. Table 19 displays the results of Friedman's test. Statistically significant differences between tasks were obtained, since the p-value of Friedman's test was 0.0002.

Only two pairs were compared by implementing Dunn's test: ROOM2 – UTTER and UTTER – STORY, since the upper quartile was different from zero in UTTER, whereas in the tasks ROOM2 and STORY even the ninth decile equaled zero. The results are shown in Table 20. A statistically significant difference was obtained between tasks ROOM2 and UTTER. The nature of the task UTTER required the use of low frequency words. The speaker attempted to retrieve the lexeme for which only incomplete phonological information (e.g. usually the initial phonemes) was available and finally pronounced it incorrectly.

5. Conclusion

The distribution of more than three and a half thousand elicited speech errors in the English language (L2) displayed a distinct dominance of morphological errors in relation to other categories of errors. The significantly frequent omission of definite and indefinite articles contributed to the dominance of morphological errors. The students have been found to use bare nouns without the definite/indefinite article very frequently, that is, in most of the cases the article was omitted. This study supports the results of previous findings, indicating that omission of articles is the main problem when the L1 semantic and conceptual system differs from L2. Also, wrong verb formation, as a subcategory of morphological errors, outnumbered the total number of lexical errors in tasks that required retelling. The participants had to use the past tenses and very often did not know the inflexion of the verb forms. On the other hand, in the design of simpler expressions, the speakers deliberately avoided complex tenses. Arguably, the most numerous subclass of syntactic errors which appeared in the retelling task was the completely unacceptable morphosyntactic and/or semantic expression. The retelling of the chronological order of events has proven to be very demanding for the learners, consequently, due to insufficient language competence, they tended to create expressions which were completely unacceptable as far as form and meaning are concerned. The distribution of different subcategories of lexical errors pointed to a relatively low occurrence of unintended L1 switches and non-existing L2 lexemes (non-words) if compared to errors as a result of lexical access failures. Unintended L1 use was proven to be the consequence of an error in accessing the intended L2 lemma, therefore it can be concluded that the participants of the study can separate the two languages during lexical access.

The results of statistical testings of the influence of the task type on the occurrence of speech errors have displayed that the retelling of the chronological sequence of actions of longer duration resulted in a significantly higher frequency of syntactic errors, in relation to the tasks aimed at describing the static, spatial constellations. Since L2 speakers at a lower level of language competence have a limited amount of declarative knowledge and since the mechanisms of linguistic encoding are not automated, the speaker's attention is directed towards the processes of lexical, grammatical and phonological encoding, and less attention is available for other phases of message production. Story retelling is proven to be a complex cognitive process, whereby the speaker's attention is oriented towards planning and organization of speech acts under time pressure, for this reason, the speaker's system very often does not manage to coordinate production processes on different levels in real time.

The rate of lexical and phonological errors depended on the frequency of use, that is, low-frequency words were more susceptible to lexical errors than high-frequency words. Certain task types, such as precise descriptions, required more often the definite and indefinite article, resulting in a significantly higher rate of morphological errors. The speakers paid very little importance to the use of articles, which can be explained by: a) the influence of the native language (Croatian) which does not have the article system, b) the students' attitude that articles are not important for message understanding and c) formal language teaching in primary and secondary schools does not

give adequate importance to article acquisition.

The conclusions of this study might find a practical application in foreign language teaching, that is, formal language teaching in primary and secondary schools should be more directed towards grammatical accuracy and the students should be reminded that the correct usage of the definite and indefinite article is an important prerequisite for the development of communicative competence in general. Finally, it may be concluded that the retelling of a chronological order of events is a demanding task, for this reason, this task type should be more practiced in foreign language teaching.

The presented findings of this research are related to the population of engineering students. Future speech error research might answer the question if the corresponding conclusions are also valid for the general population.

References

- Abutalebi, J., Cappa, S. F., & Perani, D. (2001). The bilingual brain as a function of functional neuroimaging. *Bilingualism: Language and Cognition*, 4, 179-190, doi:10.1017/S136672890100027X, <http://dx.doi.org/10.1017/S136672890100027X>
- Abutalebi, J., Cappa, S. F., & Perani, D. (2005). What can functional neuroimaging tell us about the bilingual brain. In J. Kroll, & A. M. B. De Groot (Eds.). *Handbook of bilingualism: Psycholinguistic perspectives* (pp. 497-515). New York: Oxford University Press.
- Al-Subaihi, A. A. (2000). A Monte Carlo Study of the Friedman and Conover Tests in the Single-Factor Repeated Measures Design. *J. Statist. Comput. Simul.* 65, 203-223, doi:10.1080/00949650008811999, <http://dx.doi.org/10.1080/00949650008811999>
- Avery, P., & Radišić, M. (2007). Accounting for Variability in the Acquisition of English Articles. *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*. MA: Cascadia Proceedings. Project.
- Blackmer, E. R., & Mitton, J. L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition*, 39, 173-194, doi:10.1016/0010-0277(91)90052-6, [http://dx.doi.org/10.1016/0010-0277\(91\)90052-6](http://dx.doi.org/10.1016/0010-0277(91)90052-6)
- Coin, D. (2007). A goodness-of-fit test for normality based on polynomial regression. *Computational Statistics and Data Analysis*, 52, 2185-2198, doi:10.1016/j.csda.2007.07.012, <http://dx.doi.org/doi:10.1016/j.csda.2007.07.012>
- Daniel, W. W. (1990). *Applied Nonparametric Statistics*. PWS-Kent Publishing Company.
- Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, 93, 283-321, doi:10.1037/0033-295X.93.3.283, <http://dx.doi.org/10.1037/0033-295X.93.3.283>
- Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: an analysis of speech error data. *Journal of verbal learning and verbal behaviour*, 20, 611-629.
- Demšar, J. (2006). Statistical Comparisons of Classifiers over Multiple Data Sets. *Journal of Machine Learning Research*, 7, 1-30.
- De Smedt, K., & Kempen, G. (1987). Incremental sentence production, self-correction and coordination. In G. Kempen (Ed.). *Natural language generation: Recent advances in artificial intelligence, psychology and linguistics* (pp. 365-376). Dordrecht, Netherlands: Kluwer.
- Ekiert, M. (2007). The Acquisition of Grammatical Marking of Indefiniteness with the Indefinite Article a in L2 English. *Teachers College, Colombia University Working Papers in TESOL & Applied Linguistics*, 7 (1), 1-43.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Erdeljac, V. (2005). Govorne greške-pogled u strukturu mentalnoga leksikona. In J. Granić (Ed.). *Semantika prirodnog jezika i metajezik semantike* (pp 213-222). Zagreb-Split, Hrvatsko društvo za primijenjenu lingvistiku.
- Erdeljac, V. (2009). *Mentalni leksikon: modeli i činjenice*. Ibis grafika: Zagreb.
- Field, A. (2005). *Discovering Statistics using SPSS*. SAGE Publications: UK.
- Fromkin, V. A. (Ed.). (1973). *Speech errors as linguistic evidence*. The Hague: Mouton Publishers.
- Fromkin, V. A. (Ed.). (1980). *Errors in Linguistic Performance*. New York: Academic Press.

- Hartsuiker, R., Kolk, H. H. J., & Martensen, H. (2005). The division of labour between internal and external speech monitoring. In R. Hartsuiker, Y. Bastiaanse, A. Postma, F. Wijnen (Eds.). *Phonological Encoding and Monitoring in Normal and Pathological Speech* (pp. 187-205). Psychology Press Hove.
- Henderson, A. R. (2006). Testing experimental data for univariate normality. *Clinica Chimica Acta*. 366, 112-129, doi:10.1016/j.cca.2005.11.007
<http://dx.doi.org/10.1016/j.cca.2005.11.007>
- Jaeger, J. J. (2005). *Kids' slips. What young children's slips of the tongue reveal about language development*. Lawrence Erlbaum Associates Publishers, Mahwah, New Jersey and London.
- Keskin, S. (2006). 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), 296-300.
- Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning*. 49, 303-342, doi:10.1111/0023-8333.00090,
<http://dx.doi.org/10.1111/0023-8333.00090>
- Kormos, J. (2000). The role of attention in monitoring second language speech production. *Language Learning*. 50, 343-384, doi:10.1111/0023-8333.00120,
<http://dx.doi.org/10.1111/0023-8333.00120>
- Kormos, J. (2006). *Speech Production and Second Language Acquisition*. Lawrence Erlbaum Associates, New York London.
- Kovač, M. M., & Horga, D. (2010). Govorne pogreške studenata tehničkih studija. *Slavistična revija*. 58 (4), 419-443.
- Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*. 33, 41-103, doi:10.1016/0010-0277(83)90026-4,
[http://dx.doi.org/10.1016/0010-0277\(83\)90026-4](http://dx.doi.org/10.1016/0010-0277(83)90026-4)
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. (1993). Language use in normal speakers and its disorders. In G. Blanken, J. Dittmann, H. Grimm, J. C. Marshall, & C-W. Wallesch (Eds.). *Linguistic disorders and pathologies* (pp. 1-15). Berlin: deGruyter.
- Levelt, W. J. M. (1995). The ability to speak: From intentions to spoken words. *European Review*. 3, 13-23, doi:10.1017/S1062798700001290,
<http://dx.doi.org/10.1017/S1062798700001290>
- Levelt, W. J. M. (1999). Language production: A blueprint of the speaker. In C. Brown & P. Hagoort (Ed.). *Neurocognition of language* (pp. 83-122). Oxford, England: Oxford University Press.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioural and Brain Science*. 22, 1-38, doi:10.1017/S0140525X99001776,
<http://dx.doi.org/10.1017/S0140525X99001776>
- Marques de Sá, J. P. (2007). *Applied Statistics Using SPSS, STATISTICA, MATLAB and R*. Springer Verlag: Berlin Heidelberg.
- Mizera, G. J. (2006). *Working memory and L2 oral fluency*. University of Pittsburgh.
- Montgomery, D. C., & Runger, G. C. (2003). *Applied Statistics and Probability for Engineers*. John Wiley and Sons: USA.
- Nakatani, C. H., & Hirschberg, J. (1994). A corpus-based study of repair cues in spontaneous speech. *The Journal of the Acoustical Society of America*. 95 (3), 1603-1616, doi:10.1121/1.408547,
<http://dx.doi.org/10.1121/1.408547>
- Nooteboom, S. (1980). 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* (pp. 87-96). Academic Press: New York.
- Nooteboom, S. G. (2005). Lexical bias revisited: Detecting, rejecting and repairing speech errors in inner speech. *Speech Communication*. 47, 43-58, doi:10.1016/j.specom.2005.02.003,

<http://dx.doi.org/10.1016/j.specom.2005.02.003>

Nooteboom, S. (2010). Monitoring for speech errors has different functions in inner and overt speech. In M. Everaert, T. Lentz, H. De Mulder, O. Nilsen (Eds.). *The Linguistic Enterprise. From Knowledge of Language to Knowledge in Linguistics* (pp. 231-234). Amsterdam: John Benjamins Publishing Company.

Pett, M. A. (1997). *Nonparametric Statistics for Health Care Research*. (2nd ed.). Sage.

Poulisse, N. (1999). *Slips of the tongue: Speech errors in first and second language production*. Amsterdam: Benjamins.

Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition*, 77, 97-131, doi:10.1016/S0010-0277(00)00090-1,

[http://dx.doi.org/10.1016/S0010-0277\(00\)00090-1](http://dx.doi.org/10.1016/S0010-0277(00)00090-1)

Postma, A., & Kolk, H. (1993). The covert repair hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. *Journal of Speech and Hearing Research*, 36, 472-487.

Poulisse, N. (1993). A theoretical account of lexical communication strategies. In R. Schreuder & B. Weltens (Eds.). *The bilingual lexicon* (pp. 157-189). Amsterdam: Benjamins.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64, 249-284, doi:10.1016/S0010-0277(97)00027-9,

[http://dx.doi.org/10.1016/S0010-0277\(97\)00027-9](http://dx.doi.org/10.1016/S0010-0277(97)00027-9)

Shoaf, L. C., & Pitt, M. A. (2002). Does node stability underlie the verbal transformation effect? A test of node structure theory. *Perception and Psychophysics*, 64 (5), 795-803, doi:10.3758/BF03194746,

<http://dx.doi.org/10.3758/BF03194746>

Shriberg, E. E. (1994). *Preliminaries to a Theory of Speech Disfluencies*. Doctoral thesis. University of California, Berkley.

Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.). *Progress in the psychology of language* (pp. 143-186). Hillsdale NJ: Lawrence Erlbaum Associates.

Thode, H. C. (2002). *Testing for normality*. Marcel Dekker: New York.

Van de Wiel, M. A. (2004). Exact null distributions of quadratic distribution-free statistics for two-way classification. *Journal of Statistical Planning and Inference*, 120, 29-40, doi:10.1016/S0378-3758(02)00498-6,

[http://dx.doi.org/10.1016/S0378-3758\(02\)00498-6](http://dx.doi.org/10.1016/S0378-3758(02)00498-6)

Van Hest, G. W. C. M. (1996). *Self-repair in L1 and L2 production*. Tilburg University Press.

Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring in the time course of phonological encoding. *Journal of Memory and Language*, 34, 311-334, doi:10.1006/jmla.1995.1014,

<http://dx.doi.org/10.1006/jmla.1995.1014>

Zhang, J., & Yuehua, W. (2005). Likelihood-ratio tests for normality. *Computational Statistics and Data Analysis*, 49, 709-721, doi:10.1016/j.csda.2004.05.034,

<http://dx.doi.org/10.1016/j.csda.2004.05.034>

Note:

Note 1. An L2 learner whose speech contains so many lexical and/or morphosyntactic errors that it is incomprehensible cannot reasonably be considered "fluent" (Mizera, 2006:3).

This subcategory was suggested by Kormos (personal communication, London, 2008). If the utterance had a completely incorrect syntactic structure or incomprehensible propositional content, after being interpreted by two native speakers, the same was categorized as completely unacceptable.

Table 1. Classification of errors

Errors	Definition	Example
Lexical errors	Lexical errors constitute: a) collocations, idioms, content and functional words and errors of derivational morphology; b) unintentional use of L1 lexemes; c) non-words.	<i>And then they hear some beautifully singing and er er knew it was Johnny's mother.</i> <i>Johnny Bravo went to the şum-sorry to the wood.</i> <i>er so Johnny Bravo, his mother was making him a sweater of /pineko/.</i>
Grammatical errors <i>a) Syntactic errors</i>	Syntactic errors include: a) wrong word order; b) unfinished expression or false start; c) completely unacceptable morpho-syntactic and / or semantic structure; d) wrongly encoded complements and specifiers, which are accessed by syntactic building procedures	<i>Johnny met a bear who was er sleeping...er he woke up him in the middle of er winter.</i> <i>And he came to er er cave...er er that cave was er er -in that cave he saw er a bear</i> <i>They start to arguing and hunting into the woods but er it er didn't er was success.</i> <i>The bear and Johnny listened beautiful singing of Johhny's mother.</i>
<i>b) Morphological errors</i>	Depending on the type of errors, morphological errors are divided into a) inflexional errors, when the speaker chooses the incorrect structure (wrong verb form); b) incorrect plural of nouns; c) errors of "time and aspect," when the speaker uses the wrong time, or aspect; d) omitted or misused article	<i>well Johnny Bravo's mum was mading er was making a coat for Johhny.</i> <i>Lots of womans wear black shoes.</i> <i>He er er was trying to find er er to find his mum and he wants to escape from er forest.</i> <i>we are now er in front of the my room.</i>
Phonological errors	Error in the articulation stage	<i>his mother his mother er er sang a /lole/ lullaby and er they fell asleep.</i>

Table 2. Word count and speech duration

	ALL TASKS	CARTOON	ROOM1	ROOM2	UTTER	STORY
<i>r</i>	56 862	24 536	6 656	6 118	13 068	6 484
<i>t_m [min]</i>	598.058	233.803	61.865	47.472	186.775	68.143

r – word count, *t_m* – speech duration in minutes

Table 3. Error distributions

Errors	ALL TASKS		CARTOON		ROOM1		ROOM2		UTTER		STORY	
	N _E	p _E (%)	N _E	p _E (%)	N _E	p _E (%)	N _E	p _E (%)	N _E	p _E (%)	N _E	p _E (%)
<i>E_{SI}</i>	256	7.277	98	6.617	29	6.729	27	5.921	81	11.035	21	5.048
<i>E_{S2}</i>	282	8.016	169	11.411	15	3.480	13	2.851	38	5.177	47	11.298
<i>E_{S3}</i>	123	3.496	79	5.334	17	3.944	13	2.851	6	0.817	8	1.923
<i>E_S</i>	661	18.789	346	23.363	61	14.153	53	11.623	125	17.030	76	18.269
<i>E_{L1}</i>	546	15.520	221	14.922	80	18.561	67	14.693	129	17.575	49	11.779
<i>E_{L2}</i>	50	1.421	29	1.958	0	0	1	0.219	17	2.316	3	0.721
<i>E_{L3}</i>	50	1.421	17	1.148	5	1.160	2	0.439	20	2.725	6	1.442
<i>E_L</i>	646	18.363	267	18.028	85	19.722	70	15.351	166	22.616	58	13.942
<i>E_{M1}</i>	441	12.536	310	20.932	2	0.464	1	0.219	67	9.128	61	14.663
<i>E_{M2}</i>	215	6.111	103	6.955	16	3.712	17	3.728	4	0.545	75	18.029
<i>E_{M3}</i>	1362	38.715	413	27.887	255	59.165	308	67.544	252	34.332	134	32.211
<i>E_{M4}</i>	103	2.928	21	1.418	0	0	1	0.219	77	10.490	4	0.962
<i>E_M</i>	2121	60.290	847	57.191	273	63.341	327	71.711	400	54.496	274	65.865
<i>E_P</i>	90	2.558	21	1.418	12	2.784	6	1.316	43	5.858	8	1.923
<i>E</i>	3518	100	1 481	100	431	100	456	100	734	100	416	100

E_{SI} – wrong word order or unfinished utterance (false start), *E_{S2}* – morphosyntactically and/or semantically unacceptable utterance, *E_{S3}* – wrongly encoded complements and/or specifiers accessed via syntactic building procedures, *E_S* – syntactic error, *E_{L1}* – L2 lexeme, *E_{L2}* – nonexisting L2 word (non-word), *E_{L3}* – L1 lexeme, *E_L* – lexical error, *E_{M1}* – wrong verb form, *E_{M2}* – errors of “time and aspect”, *E_{M3}* – omitted or misused article, *E_{M4}* – wrong singular or plural ending, *E_M* – morphological error, *E_P* – phonological error, *E* – error, *N_E* – total number of errors, *p_E* – percentage of a particular error category

Table 4. Percentages of omitted and incorrectly used articles

<i>E_{M3}</i>	<i>N_{ar}</i>	<i>p_{ar}</i> (%)
Omitted article	1 310	96.182
Incorrectly used article	52	3.818
Total	1 362	100

N_{ar} – number of articles, *p_{ar}* – percentages

Table 5. Descriptive statistics for syntactic errors in L2

E_s	<i>CARTOON</i>	<i>ROOM1</i>	<i>ROOM2</i>	<i>UTTER</i>	<i>STORY</i>
n	101	101	101	101	101
x_{\min}	0	0	0	0	0
D_1	0	0	0	0	0
Q_1	0.471	0	0	0	0
Me	1.042	0	0	0	0
Q_3	2.165	1.878	1.464	1.351	2.120
D_9	3.278	2.564	2.688	2.343	3.371
x_{\max}	4.918	5.085	5.769	6.349	6.452
d	4.918	5.085	5.769	6.349	6.452
d_Q	1.694	1.878	1.464	1.351	2.120
\bar{x}	1.460	0.924	0.806	0.887	1.128
s	1.250	1.265	1.324	1.327	1.621
s^2	1.563	1.600	1.753	1.761	2.628
$S_{\bar{x}}$	0.124	0.126	0.132	0.132	0.161
V	85.58 %	136.95 %	164.26 %	149.62 %	143.76 %
α_3	0.976	1.327	1.767	2.055	1.453
α_4	0.235	1.142	2.880	4.535	1.460

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 6. Shapiro-Wilk's test of normality for syntactic errors in L2

Shapiro-Wilk	<i>CARTOON</i>	<i>ROOM1</i>	<i>ROOM2</i>	<i>UTTER</i>	<i>STORY</i>
W	0.9036	0.7576	0.6725	0.7110	0.7364
p	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001

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

Table 7. Friedman's test for syntactic errors in L2

Friedman	n	df	F_r	p
	101	4	42.893	< 0.0001

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

Table 8. Dunn's multiple comparison test for syntactic errors in L2

Dunn	Difference in rank sum	Statistically significant difference
CARTOON – ROOM1	99	Yes
CARTOON – ROOM2	113	Yes
CARTOON– UTTER	107.5	Yes
CARTOON– STORY	90.5	Yes

Table 9. Descriptive statistics for lexical errors in L2

E_L	CARTOON	ROOM1	ROOM2	UTTER	STORY
n	101	101	101	101	101
x_{min}	0	0	0	0	0
D₁	0	0	0	0	0
Q₁	0.494	0	0	0	0
Me	0.985	0	0	0.901	0
Q₃	1.618	2.003	1.942	1.924	1.575
D₉	2.635	3.955	4.249	2.849	3.169
x_{max}	3.774	8.974	7.576	5.854	6.452
d	3.774	8.974	7.576	5.854	6.452
d_Q	1.124	2.003	1.942	1.924	1.575
\bar{x}	1.160	1.263	1.157	1.200	3.169
s	0.907	1.733	1.745	1.315	1.401
s²	0.823	3.003	3.045	1.729	1.963
S_{\bar{x}}	0.090	0.172	0.174	0.131	0.139
V	78.15 %	137.22 %	150.87 %	109.59 %	152.39 %
α_3	0.750	1.768	1.562	1.381	1.620
α_4	0.006	3.798	1.787	1.916	2.306

n – sample size, *x_{min}* – sample minimum, *D₁* – the first decile, *Q₁* – lower quartile, *Me* – median, *Q₃* – upper quartile, *D₉* – the ninth decile, *x_{max}* – sample maximum, *d* – sample range, *d_Q* – interquartile range, \bar{x} – sample mean, *s* – sample standard deviation, *s²* – sample variance, *S _{\bar{x}}* – standard error of the mean, *V* – coefficient of variation, α_3 – skewness, α_4 – kurtosis

Table 10. Shapiro-Wilk's test of normality for lexical errors in L2

Shapiro-Wilk	CARTOON	ROOM1	ROOM2	UTTER	STORY
W	0.9359	0.7565	0.7141	0.8368	0.7130
p	0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001

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

Table 11. Friedman's test for lexical errors in L2

Friedman	<i>n</i>	<i>df</i>	<i>F_r</i>	<i>p</i>
	101	4	13.095	0.0108

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

Table 12. Dunn's multiple comparison test for lexical errors in L2

Dunn	Difference in rank sum	Statistically significant difference
CARTOON-ROOM1	33.5	No
CARTOON-ROOM2	59	Yes
CARTOON-STORY	60	Yes

Table 13. Descriptive statistics for morphological errors in L2

E_M	<i>CARTOON</i>	<i>ROOM1</i>	<i>ROOM2</i>	<i>UTTER</i>	<i>STORY</i>
n	101	101	101	101	101
x_{\min}	0	0	0	0	0
D_1	0.953	0	0	0	0
Q_1	1.903	1.258	1.227	1.598	1.424
Me	3.361	3.704	5.128	2.857	4.211
Q_3	5.401	7.692	10.396	4.578	6.351
D_9	7.449	12.279	14.857	6.357	10.130
x_{\max}	11.765	17.949	23.881	16.667	22.222
d	11.765	17.949	23.881	16.667	22.222
d_Q	3.498	6.434	9.169	2.980	4.927
\bar{x}	3.752	4.829	6.153	3.322	4.641
s	2.406	4.489	5.658	2.796	4.232
s^2	5.789	20.151	32.013	7.818	17.910
$S_{\bar{x}}$	0.239	0.447	0.563	0.278	0.421
V	64.12 %	92.95 %	91.95 %	84.17 %	91.17 %
α_3	0.703	0.902	0.718	1.640	1.415
α_4	0.099	0.015	-0.258	4.868	3.080

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 14. Shapiro-Wilk's test of normality for morphological errors in L2

Shapiro-Wilk	<i>CARTOON</i>	<i>ROOM1</i>	<i>ROOM2</i>	<i>UTTER</i>	<i>STORY</i>
W	0.9535	0.8967	0.9065	0.9777	0.8799
p	0.0013	< 0.0001	< 0.0001	< 0.0001	< 0.0001

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

Table 15. Friedman's test for morphological errors in L2

Friedman	n	df	F_r	p
	101	4	15.607	0.0036

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

Table 16. Dunn's multiple comparison test for morphological errors in L2

Dunn	Difference in rank sum	Statistically significant difference
CARTOON – ROOM1	-5.5	No
CARTOON – ROOM2	-59.5	Yes
CARTOON – UTTER	22.5	No
CARTOON – STORY	-25	No
ROOM2 – UTTER	82	Yes

Table 17. Descriptive statistics for phonological errors 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, *D₁* – the first decile, *Q₁* – lower quartile, *Me* – median, *Q₃* – upper quartile, *D₉* – the ninth decile, *x_{max}* – sample maximum, *d* – sample range, *d_Q* – interquartile range, \bar{x} – sample mean, *s* – sample standard deviation, *s²* – sample variance, *S _{\bar{x}}* – standard error of the mean, *V* – coefficient of variation, α_3 – skewness, α_4 – kurtosis

Table 18. Shapiro-Wilk's test of normality for phonological errors 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* – the *p* value

Table 19. Friedman's test for phonological errors in L2

Friedman	<i>n</i>	<i>df</i>	<i>F_r</i>	<i>p</i>
	101	4	21.971	0.0002

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

Table 20. Dunn's multiple comparison test for phonological errors in L2

Dunn	Difference in rank sum	Statistically significant difference
ROOM2 – UTTER	–51	Yes
UTTER – STORY	46	No

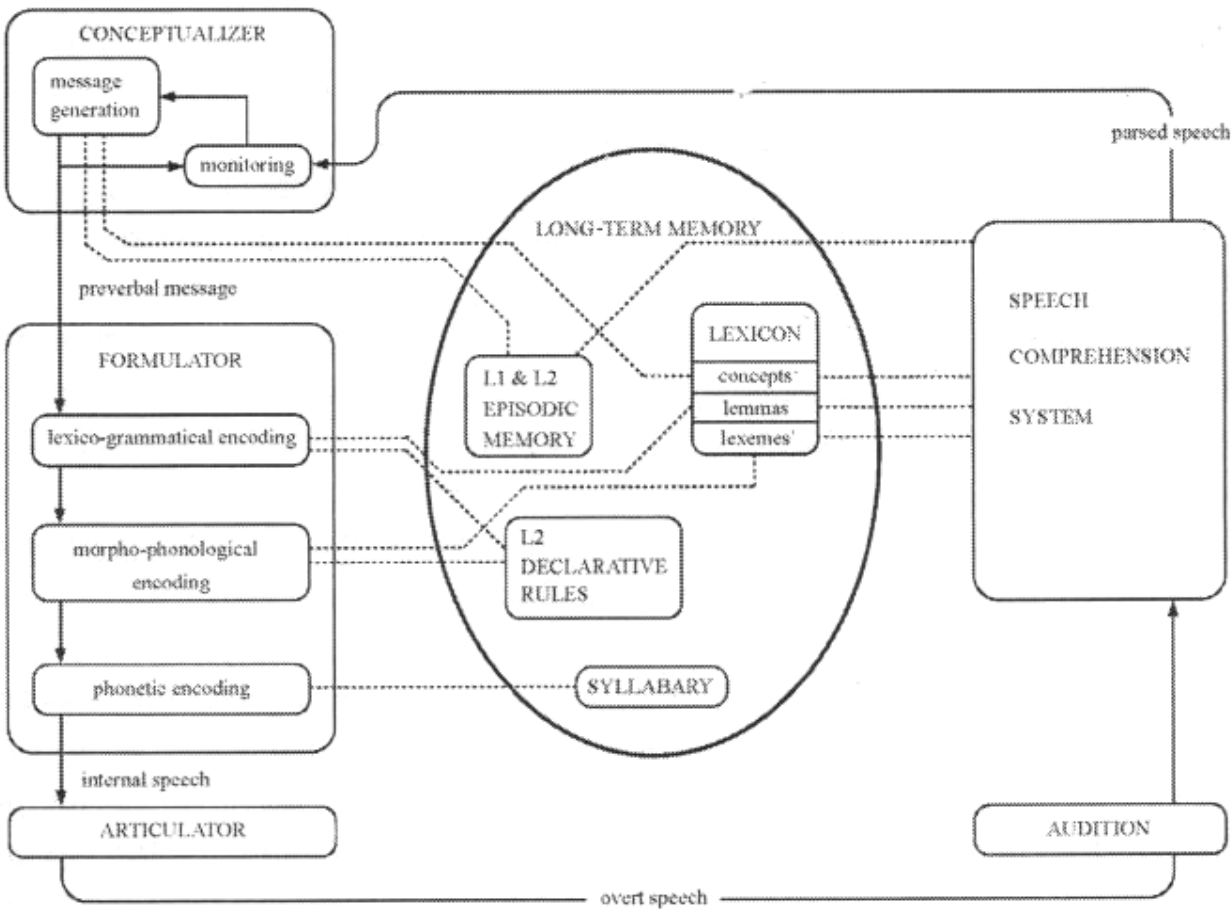


Figure 1. The model of bilingual speech production (Kormos, 2006: 168)

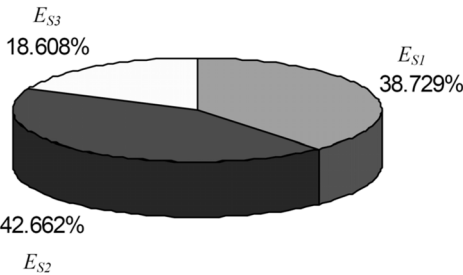


Figure 2. Percentages of subcategories of syntactic errors in their total number in all tasks

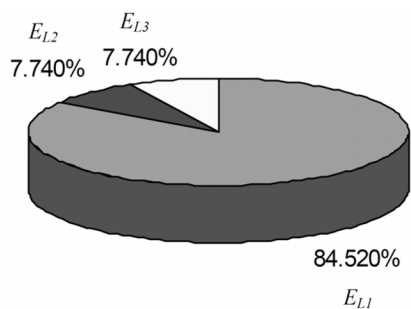


Figure 3. Percentages of subcategories of lexical errors in their total number in all tasks

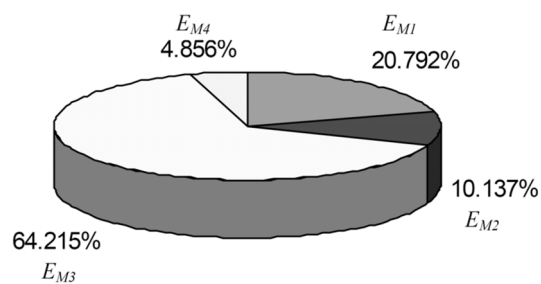


Figure 4. Percentages of subcategories of morphological errors in their total number in all tasks