Topics in Linguistics (2022), 23(1), pp. 83-96

10.2478/topling-2022-0007

Forms and functions of self-repetitions in spoken discourse: A corpus linguistics analysis of L1 and L2 English

Ilyas Yakut*
Fatma Yuvayapan
Kahramanmaras Istiklal University, Turkey

Abstract

Self-repetitions as a compensation strategy potentially have a powerful influence on maintaining communication in L1 and L2 contexts. By subjecting two comparable corpora to contrastive interlanguage analysis, in the present study we explored the use of self-repetitions by L1 speakers of English and Turkish speakers of English and their structural and functional distributions. The results indicated that while L1 English speakers tended to utilize repetitions as fillers, L2 English speakers employed considerably more repetitions as self-initiated self-repairs. The results of the study suggest that, despite being used for different purposes, self-repetition occurs in L1 and L2 English speech indicating that self-repetitions cannot be associated with disfluency as long as they do not hinder the flow of speech. For this reason, raising L2 learners' awareness might help them prevent breakdowns in L2 communication.

Key words

form, function, filler, self-initiated self-repair, self-repetition, speaking, corpus research

1. Introduction

Developing a native-like fluency and accuracy in oral communication has always been the ultimate aim of the language learning process. When learners face difficulties during the cognitive planning of communication, they utilize strategies to gain time and prevent misunderstandings and breakdowns (Yakut & Bada, 2021). For instance, they use their mimics and gestures, code-switch words, and explain the words they do not know. All these efforts are called compensation strategies (CSs). Kellerman (1998) explains that definitions of these strategies vary based on two perspectives. From a psycholinguistic perspective, they are defined as "the mental processes involved in creating solutions to problems arising from a relatively limited linguistic ability". When grounded in an interactional perspective, "they represent the means by which participants in conversations work together to solve the learner's linguistic problem" (p. 92). The best-known exponents of the first perspective are Faerch and Casper (1984) and the Nijmegen group. From an interactional standpoint, one can think of Tarone and Yule (1983). Despite their various definitions, CSs enable learners "to operate with a small vocabulary, and permit speech to remain fluent" (p. 93).

In 1981, Tarone labelled CSs under five main categories: a) paraphrase (approximation, word coinage, and circumlocution); b) borrowing (literal translation and language switch); c) appeal for

* Address for correspondence: Ilyas Yakut, Faculty of Humanities and Social Sciences Department of Translation and Interpreting, Kahramanmaras Istiklal University, Yamactepe 93009 Street, 5, Kahramanmaras, Turkey. E-mail: yakutilyas@gmail.com

assistance; d) mime; e) avoidance (topic avoidance and message abandonment). In their review article, Dörnyei and Scott (1997) summarized CSs thusly: a) avoidance and reduction strategies (message abandonment and topic avoidance); b) achievement or compensatory strategies (circumlocution, approximation, use of all-purpose words, word-coinage, use of nonlinguistic means, literal translation, foreignizing, code-switching and appeal for help); c) stalling or time-gaining strategies (use of fillers and hesitation devices).

Over four decades, scholars have suggested various taxonomies to refer to CSs depending on the two perspectives described above. Schegloff, Jefferson, and Sacks (1977) coined a new term called *repair* as a CS. This particular strategy is used to overcome some difficulties during the flow of speech. Following this study, Fox and Jesperson (1995) examined repair in L1 American English conversations from a syntactic view and observed changes in repair usage depending on syntactic categories. According to Rieger (2003), repair is "the search for a word, and the use of hesitation pauses, lexical, quasi-lexical, or non-lexical pause fillers, immediate lexical changes, false starts, and instantaneous repetitions" (p. 48). Among these devices, repetitions of lexical items fulfil prominent functions.

Tarone and Yule (1987) define repetition as the repetition of a word or a phrase in a speech. These devices give greater clarity to the meaning implied by the speaker in addition to a means of hearing the information again by the listener. For Crible and Pascual (2020), repetition appears to operate as reformulating speech. According to Crible and Pascual, self-repetitions may emerge in the form of repairs where an interlocutor interrupts the original utterance and repairs the erroneous form, while repetitions may also occur as fillers where the original utterance can be repeated identically without adding new information, or repetition with modification may emerge by inserting new items between the repeated words or phrases.

Repetitions as fillers are cognitive processes employed to gain time, keep one's turn in speech, and avoid non-filled pauses when speakers plan the next utterance or seek a specific word or phrase to transmit the intended message (Bada, 2010, Dörnyei & Scott, 1997). According to Dörnyei and Scott (1995a, 1995b), fillers are indirect CSs that are not "meaning related" since they do not offer alternative meaning forms. However, they are indirect problem-solving strategies. Fillers are used as a time-stalling device to gain time, which helps speakers avoid communication breakdowns and keep the communication channel open. Hence, repetitions as fillers do not seem to be different from other forms of fillers such as a*h*, *er*, *erm*, or the other vocalized fillers (Bada, 2010).

Repetitions as self-repairs indicate more complex cognitive processes since speakers need to monitor what is being uttered, detect the erroneous form or function and repair structurally or socially idiosyncratic utterances. According to Tarone (1980), self-repairs have interactional functions because the mere purpose of repairing utterances might not be correcting only linguistic forms but making sure that the intended message is sent to the other interlocutor. Self-repetitions as repairs have two essential components: initiation of repetition and fixing the erroneous form. Repetitions as repairs seem to allow speakers to avoid face-threatening acts that might arise due to the mistakes in the speech. Hence, once the speaker detects that an utterance is idiosyncratic, i.e., erroneous, they might resort to self-repetitions to repair their mistakes (Dörnyei & Scott, 1995a; 1995b; 1997; Tarone & Yule, 1987; Willems, 1987). Here, we need to note that repetitions with self-repair functions might co-occur with pragmatic markers or hesitation markers that Clark calls "editing terms" (Clark, 1994). Such co-occurrences might indicate that the speaker recognizes the mistake being made and corrects it themselves.

The interest in studying repetitions from the perspective of functions is two-fold. Firstly, a large body of literature has shed light on the multifaceted occurrence of repetitions in classroom discourse. In a qualitative analysis of self-repetitions, Lyster (1998) concentrated on the repetitions and recasts made by teachers and found that they fulfil identical functions. Recasts were the most frequently utilized form of corrective feedback and suggest using a broader range of feedback techniques leading to student-generated repair by teachers. Examining repetitions in the discourse of a Spanish-speaking learner of Portuguese in three different settings (language lessons, conversation sessions, and interviews), Silva and Santos (2006) observed different functions of the repetitions in these three learners' discourse. Since they enhance interaction and conceptual understanding of the discourse, repetitions have significant roles in the learning process. They thus propose explicit teaching of different types and functions of repetitions through various critical awareness activities. In a specific context of spoken interaction in ELF, Kaur (2012) studied self-repetitions in English and stated that self-repetitions were the means of making meaning clear and preventing misunderstandings for ELF speakers. In another L2 context,

Kovac (2016) demonstrated that self-repetitions were used in the service of gaining time for cognitive planning of conversation.

Secondly, comparing repetition usage in different cultures and languages has also received attention. Fujimura-Wilson (2007) compared the frequency and functions of exact repetitions in Japanese and English conversations and revealed the culture-specific employment of these two nations' repetitions. Due to social characteristics, the Japanese and the English tended to apply different exact repetitions. While the English prefer more self-repetitions than the Japanese to pause and repair overlapping speech, Japanese people tend to use other-repetitions more than English people to emphasize collaboration and empathy with interlocutors. Focusing mainly on self-repetitions as repair strategies employed by English-German bilinguals, Rieger (2003) found different language-specific preferences, which these languages' structures might explain. Crible and Pascual (2020) concentrated on repetitions and self-repairs in native English, French and Spanish spoken discourse in a recent study. They reported that combinations with self-repetitions depend on languages and repair types.

Some research has also been conducted in Turkish contexts. Bada (2010) investigated the prevalence, type and systematicity of repetition in spoken discourse by non-native speakers of English and non-native speakers of French and observed two basic functions of repetitions as vocalized fillers and self-repairs. Genç, Mavaşoğlu, and Bada (2010) studied the repetitions in French learners' discourse and revealed similar functions of self-repetitions: a) maintaining the flow of interaction; b) self-repair. Kellerman (1998) emphasizes that CSs are concerned with both L1 and L2 discourse. Based on this view, Tekdemir Yurtdaş (2012) calculated that launching repair and showing listenership and agreement/disagreement were the most frequent functions of self-repetitions in L1 Turkish interaction.

As can be concluded from the studies above, scholars examined functions of self-repetitions mainly in two different contexts. All these studies, concerning classroom discourse (Bada, 2010; Genç, Mavasoglu & Bada, 2010; Kaur, 2012; Kovac, 2016; Lyster, 1998; Silva & Santos, 2006) and natural spoken discourse (Crible & Pascual, 2020; Fujimura-Wilson, 2007; Rieger, 2003), showed that repetitions play an important role in fostering cognitive planning of discourse and self-repair. Also, they characterize both L1 and L2 speech. Hence, the lack of repetitions in L2 speech hinders language production leading to misunderstandings.

The present study adopts a corpus-based approach to examine the frequency, structures, and functions of self-repetitions in the speech of Turkish learners of English and L1 speakers of English using comparable corpora. Of particular importance in this study is the potential role of repetitions in influencing L1 and L2 speech production. It is hoped that the comparison of the functions of repetitions in the two corpora will extend our understanding of these devices' role in L1 and L2 oral performance, which might bring about solid suggestions on the teachability of repetitions. According to Bialystok and Fröhlich (1980), L2-based CSs strategies are different from L1-based CSs in that they require the use of more information. Additionally, Dörnyei (1995) claims real-life L2 communication is challenging for L2 learners; thus, instruction involving CSs helps them deal with performance problems. The following research questions will constitute the essence of this study:

1. To what extent do L1 speakers of English and Turkish speakers of English employ self-repetitions in spoken discourse?
1.1. To what extent are word-level and group-level repetitions employed in the two datasets?
1.2. What is the ratio of self-repetitions as self-initiated self-repairs and fillers in the two datasets?
2. Do L1 speakers of English and Turkish learners of English significantly differ with regard to the structural and functional distribution of self-repetitions?

2. Methods

According to Granger (2008), corpus data in second-language research represents authenticity. For this reason, we extracted data from corpora compiled from L1 and L2 English spoken interactions: the Louvain Corpus of Native English Conversation (LOCNEC) (De Cock, 2004), and the Turkish component of the Louvain International Database of Spoken English Interlanguage (LINDSEI-TR) (Kilimci, 2014) compiled as sub-corpora of the Louvain International Database of Spoken English Interlanguage (LINDSEI) (Gilquin, 2012). All the corpus compilation processes of LOCNEC and LINDSEI followed the same structure. LOCNEC was constructed as a comparable corpus of interviews

with L1 speakers of English allowing researchers to investigate and compare L1 and L2 English usage in spoken interactions. LOCNEC contains 50 informal interviews with British university students, all L1 speakers of English. LINDSEI-TR consists of 58 informal interviews with L1 speakers of Turkish. As revealed by Kilimci (2014), the participants of LINDSEI-TR were advanced speakers of English determined through "the external criteria that they are third and fourth-year students at Faculty of Education, the Department of English Language Teaching, Çukurova University" (p. 402).

Interviewee turns in LOCNEC total approximately 119,000 words, while the learner turns in LINDSEI-TR amounts to some 64,000 words. The interviews in both corpora consisted of three sections: a warm-up activity, a free informal conversation, and a picture description.

This study is a corpus-based study aiming to explore structures and functions of self-repetitions in L1 and L2 English spoken interactions. As the two corpora were compiled according to the same criteria and transcribed with the same conventions set by the Centre for English Corpus Linguistics, we conducted a contrastive interlanguage analysis (Granger, 1996). In this study, we annotated the data considering two main criteria: structures and functions of self-repetitions. The purpose of the structural analysis was to explore the syntactic elements being repeated. Hence, structural analyses of repetitions dealt with whether repetitions were actualized as a one-word linguistic item or an at least two-word linguistic item. With the functional analysis, we aimed to discover pragmatic aspects of self-repetitions as a type of CS.

Following Bada (2010), we divided self-repetitions into two groups as word-level repetitions (WLR), which were further divided into their word classes, and group-level repetitions (GLR), which consist of repetitions encompassing two or more linguistic items being repeated. GLR were categorized according to word classes initiating self-repetitions. Functional analyses of repetitions, however, were conducted to figure out why they were employed in L1 and L2 spoken interactions. Depending on Bada's (2010) suggestions, we focused on two main reasons for self-repetitions: fillers (repetitions used to fill pauses and plan the next utterance) and self-initiated self-repairs (repetitions employed to fix language mistakes).

We performed horizontal reading in the data annotation, i.e., tagging data by going into the context of the emergence of the item in question and deciding the function of a linguistic element manually, and each self-repetition was tagged according to the criteria set for this study. After completing the data annotation process, the corpus searches were conducted using AntConc 3.5.8. (Anthony, 2019). The search queries in AntConc were defined as combinations of specific search conventions used to tag self-repetitions. Using Antconc, we extracted instances of self-repetitions through vertical reading, i.e. listing key words vertically in a corpus, which enabled us to figure out and share the results with respect to their overall distribution and their structural and functional distributions in tabular forms.

Depending on the results we obtained from the data, we normalized occurrences in every 1000 words. Then, we compared self-repetition usages between the two corpora, as well as comparisons of repetitions within each corpus. We administered Log-likelihood (LL) tests for all stages, except the distribution of word classes, which we explained via descriptive statistics, to test differences between the two sub-corpora in the use of self-repetitions. Where appropriate in the analysis, differences in repetitions within each corpus were tested for statistical significance using the Chi-square test. All self-repetitions of utterance elements given as excerpts are in italics for illustrative purposes.

3. Results

Initial analysis was made through frequency calculation of all self-repetitions in LOCNEC and LINDSEI-TR. Table 1 shows overall and normalized frequencies of repetitions per 1000 words in each.

Table 1. The overall distribution of self-repetitions

	Corpus size	**Repetitions**	
		n	n/1000
LOCNEC	118.553	2669	22.51
LINDSEI-TR	63.924	1728	27.03

As shown in Table 1, the corpus searches in LOCNEC yielded 2669 emergences of repetitions accounting for 22.51 usages per 1000 words, while 1728 repetitions were found in LINDSEI-TR corresponding to 27.03 occurrences in every 1000 words in the corpus. Since the normalized frequency

counts revealed overuse of repetitions by Turkish speakers of English relative to L1 speakers of English, we ran a Log-Likelihood test and found a statistically significant difference between the two corpora at $p<0.05$ level with an LL value of -34.60. Such a result might indicate overuse of self-repetitions by L2 speakers of English in their spoken interactions, agreeing with Stuart and Lynn (1995), as repetitions in L2 were more common than L1 in their study as well. According to the overall results, both L1 and L2 speakers of English resorted to self-repetitions frequently in their impromptu speech despite observing their occurrences with different ratios. Hence, we can state that self-repetitions are an indispensable part of spoken discourse in both the L1 and the L2 contexts.

From the two sub-corpora, we identified two structural types of self-repetitions. While the repetition of one specific word class, such as *he*, *last*, *pretty*, *talent*, *talk*, *on*, etc., was considered a word-level repetition (WLR), the repetition of at least two linguistic items, namely *about the same*, *it seemed*, *bought a*, *just to waste*, etc., was categorized under group-level repetitions (GLR). Since group-level repetitions did not always constitute a specific phrase, we preferred not to annotate GLR as phrase-level repetitions. Table 2 exhibits the overall structure-based distributions of self-repetitions in each corpus.

As can be seen, a total of 1448 WLR tokens (n/1000=12.21) were found in LOCNEC, while 1221 GLR tokens (n/1000=10.30) emerged in the corpus. A statistically significant difference between the two types of repetitions in LOCNEC was found at $p<0.05$ level (X^2=19,306; p=,000), which revealed that L1 speakers of English feel disposed to utilize more WLR compared to GLR. Contrarily, GLR (n=928; n/1000=14.52) tokens were more common compared to WLR (n=800; n/1000=12.51) tokens in LINDSEI-TR. The differences in the proportion of WLR and GLR tokens in the corpus proved to be statistically significant $p<0.05$ (X^2=9.481; p=.002). The results observed in LINDSEI-TR contradict Genc's (2007) and Yakut's (2013) findings because Turkish learners of English employed WLR more frequently than GLR in those studies. Such a contradiction might be partly related to the differences in participants' proficiency levels between the present study and the abovementioned studies. Contrary to Genc (2007) and Yakut (2013), Bada's (2010) results revealed that Turkish speakers of English opted for GLR more frequently compared to WLR.

Table 2. The overall distribution of structures of self-repetitions

	Corpus size	**Repetition Type**			
		WLR		GLR	
		n	n/1000	n	n/1000
LOCNEC	118,553	1448	12.21	1221	10.30
LINDSEI-TR	63.924	800	12.51	928	14.52

Since the current study mainly aimed at comparing self-repetitions in L1 and L2 English, structure-based comparisons of self-repetitions between the two corpora were made. The results revealed that WLR were barely more common among Turkish speakers of English than L1 speakers of English because LL statistical test results did not reveal a statistically significant difference between the two groups (LL=-0.30). Despite observing roughly similar usage rates in the WLR tokens between the two corpora, GLR tokens in LINDSEI-TR were more common than those in LOCNEC. We observed a statistically significant difference between the two corpora concerning the frequency of GLR (LL=-60.89). Such comparisons between L1 and L2 speakers English regarding repetitions at WL and GL might explain how cognitive processes function in L1 and L2 speech. In addition, the results suggest that by repeating more than one item, L2 speakers tried to gain more time or they might have thought that the message would be more clear for the addressee if a unit instead of a single linguistic element was repeated. Examples of both WLR and GLR in each corpus are presented below:

(1) It means that *you: you* can't afford to stay up late and work. [WLR] (LOCNEC)
(2) No *there weren't there weren't* a very lot of people actually no. [GLR] (LOCNEC)
(3) So (eh) she (eh) go to a painter and *want* (eh) *want* him to make her picture. [WLR] (LINDSEI-TR)
(4) (eh) *in your in my* free time *I like doing I like* (eh) *dancing.* [GLR] (LINDSEI-TR)

Examples [1] and [3] represent WLR, while [2] and [4] are GLR. In examples [1] and [3], the pronoun *you* and the verb *want* were repeated successively. In these excerpts, repetitions emerged with pauses or

hesitation markers. In [2], we observed a repetition in the initial position in which the pronoun *there* + negated form of *were* were the repeated forms. In [4], we see two different GLR. The first repeated part was initiated with the preposition of *in* and followed by a determiner, whereas the second repetition consisted of pronoun+verb+gerund form. In [4], the first self-repetition starts with a hesitation marker while the second one includes a hesitation in medial position just before the linguistic element fixed by the interlocutor. To shed light on what word classes were repeated at word level and what word classes initiated repetitions at group level, we tagged each token in the data with respect to their syntactic classes. The results were presented in Table 3 and Table 4. Table 3 illustrates the frequency and percentage of repeated linguistic items at WLR.

Table 3. Word classes constituting WLR

Part of Speech	**LOCNEC**		**LINDSEI-TR**	
	n	%	n	%
Adjective	13	0.90	21	2.63
Adverb	67	4.63	31	3.88
Auxiliary	11	0.76	8	1.00
Conjunction	151	10.43	26	3.25
Determiner	347	23.96	133	16.63
Negation	13	0.90	6	0.75
Noun	24	1.66	45	5.63
Preposition	161	11.12	86	10.75
Pronoun	603	41.64	356	44.50
Verb	58	4.01	88	11.00
Total	1448	100.00	800	100.00

As can be seen in Table 3, the most common repetitions at WL were *pronouns* (41.64%) followed by *determiners* (23.96%), *prepositions* (11.12%), and *conjunctions* (10.43%) in LOCNEC. We observed a similar pattern in the LINDSEI-TR corpus, except the ratio of *verbs* (11.00%), as *pronouns* (44.50%), *determiners* (16.63%), and *prepositions* (10.75%) constituted the most frequent word classes being repeated. The general picture in the table shows that function words seem to constitute the majority of WLR in LOCNEC (88.81%) and LINDSEI-TR (76.88%). However, this does not mean that both L1 and L2 speakers of English repeated exactly the same linguistic items with the same amount in their spoken interactions. Even though function words constitute the majority of repeated items in both corpora, we observed a 12% difference between the rates of function-word repetitions. Hence, we can conclude that the sequence of function words in the two corpora seemed to be similar, yet the density of function word repetitions was different supported by inferential statistics (LL=+6,07). In addition, the density of repeated content words showed differences between LOCNEC (11.19%) and LINDSEI-TR (23.12%) and the difference between the two corpora was statistically significant at $p<0.05$ level (LL=-48.31). These results suggest that the sequences of repeated items at WLR in both corpora indicated similar patterns, yet the rates of function and content words in the two corpora were different. Hence, we can deduce that being an L1 or L2 speaker does not affect the sequence of items being repeated in spoken discourse, yet speaking English as L1 and L2 does have a significant effect on the rates of items being repeated in impromptu speech.

Examples [5] – [11] show instances of the most common word classes repeated at word level in LOCNEC and LINDSEI-TR. At the end of each example, the items in square brackets represent the word classes being repeated at WLR. According to the excerpts given below, it can be stated that when an interlocutor needed to repeat an item, hesitations such as *(eh)* or pauses generally accompanied repetitions in the utterance.

(5) *they they* really are respectful. [pronoun] (LOCNEC)
(6) in the second year I moved out *onto into* town .. and now. [preposition] (LOCNEC)
(7) then eh she sort of eh contests *his* (eh) *his* drawing *and and* eh *and* eh she probably says no I don't look like that I'm far more attractive. [determiner, conjunction] (LOCNEC)
(8) when woman looks at the picture *he she* doesn't like his drawing. [pronoun] (LINDSEI-TR)

(9) the man also put something from his soul and made *the that* woman (eh) like happy. [determiner] (LINDSEI-TR)
(10) (eh) she (er) *shows shows* this picture to her friends. [verb] (LINDSEI-TR)
(11) (eh) first of all I stayed *with with* one of my friends. [preposition] (LINDSEI-TR)

Considering the results in Table 3, we can state that function words were more frequently repeated compared to content words in both corpora. Our results correlate with Lickley (1994), Maclay and Osgood (1959), and Rieger (2003), who highlighted the common utilization of function words in self-repetitions. Despite working with participants possessing a different interlanguage, Genc et al. (2010) suggested that pronouns and determiners were among the most prominent repeated items by Turkish learners of French. Although the most frequently repeated items were function words in the current study, the results also suggested that verbs were amongst the most frequently repeated items in LINDSEI-TR, which supports Bada's (2010) results indicating a common utilization of verb repetitions by Turkish speakers of English. Interestingly, Bada's result also suggested that the repetition of a determiner in French was among the most common word classes in the data collected from Turkish L2 users of French.

While we focused on word classes being repeated at WLR, we identified word classes initiating GLR in LOCNEC and LINDSEI-TR. The frequency and percentage of each part of speech initiating GLR in each corpus are presented in Table 4.

Table 4. Word classes initiating GLR

Part of Speech	**LOCNEC**		**LINDSEI-TR**	
	n	%	n	%
Adjective	4	0.33	19	2.05
Adverb	17	1.39	29	3.13
Auxiliary	14	1.15	29	3.13
Conjunction	58	4.75	49	5.28
Determiner	70	5.73	112	12.07
Negation	6	0.49	3	0.32
Noun	12	0.98	34	3.66
Preposition	63	5.16	80	8.62
Pronoun	940	76.99	496	53.45
Verb	37	3.03	77	8.30
Total	1221	100	928	100

Looking at Table 4, we can see the dominance of function words initiating GLR in LOCNEC. *Pronouns* constituted 76.99% of the GLR and *determiners* (5.73%), *prepositions* (5.16%), and *conjunctions* (4.75%) followed pronoun-initiated repetitions. According to these results, we can assume that L1 English speakers followed similar patterns in both GLR and WLR. Distribution of GLR concerning word classes initiating these repetitions in LINDSEI-TR revealed that the most prominent word class was *pronouns* (53.45%), and it was followed by *determiners* (12.07%), *prepositions* (8.62%), and *verbs* (8.30%), revealing similar results we observed in WLR in LINDSEI-TR. Despite the prominence of the four main word classes initiating GLR in LINDSEI-TR, we need to highlight that the distribution of word classes in LINDSEI-TR seems to be more balanced than LOCNEC. Having said that, function words in each corpus had ascendancy over content words.

Although the descriptive results indicated that the overall sequences of GLR were similar between the two corpora, we can also observe that the rate of function word repetitions was more prominent in LOCNEC (94.27%) compared to LINDSEI-TR (82.87%) which revealed an approximately 12% difference between the groups. In addition, we observed a relationship between L1 background and the density of content word repetitions at GLR as L1 speakers' GLR constituted 5.73% of their GLR while content words at GLR level among L2 speakers constituted 17.14% of their GLR repetitions and LL ratio did also reveal a statistical difference between the groups (LL=-111.99).

Similar to WLR, our findings at GLR in both corpora partly correlate with Bada's (2010) findings because his findings suggested that *pronouns* and *prepositions* were among the most prominent word classes initiating self-repetitions in the data collected from both Turkish speakers of English. Examples

of the most common GLR initiators in both corpora are given in [12] – [19]. The items in square brackets at the end of each example represent the word classes initiating GLR. The qualitative analyses of GLR repetitions indicated that L2 speakers seemed to struggle more when they needed to resort to self-repetitions either to reformulate the utterance by fixing erroneous forms or to plan upcoming utterance(s) as they used hesitation markers and pauses.

(12) yeah we all lost weight you know *I'm I was* quite thin to start with so I didn't lose too much but . some people lost two or three stone. [pronoun] (LOCNEC)
(13) *our heating our heating* is on all night. [determiner] (LOCNEC)
(14) like on *the way* yeah *on the way* there. [preposition] (LOCNEC)
(15) *when I .. when I went . when I went* through school I was the last year where wasn't compulsory to take a second language. [conjunction] (LOCNEC)
(16) *I don't* (eh) *I will not* get married this plan is here and (eh) this plan is (eh) *completely* (eh) *completely* different. [pronoun, adverb] (LINDSEI-TR)
(17) (eh) ... *my husband* (eh) *my husband* (eh) have. [determiner] (LINDSEI-TR)
(18) because (eh) . (eh) *in picture in picture* (eh) the woman (eh) .. (eh) isn't *the same the same* one to (eh) real woman. [preposition, determiner] (LINDSEI-TR)
(19) (eh) one day (eh) . my friend and I (em) *decided to do decided to go* on a picnic. [verb] (LINDSEI-TR)

This study consists of both structural and functional analyses of repetitions in L1 and L2 English data. While structural analysis deals with syntactic deciphering of the repetitions at WLR and GLR, the functional analysis explains the possible reasons for employing these routines. Hence, we focused on the functions of self-repetitions in the second part of the study. Depending on the suggestions made by Bada (2010) and the integrated reading of the data at hand, we divided self-repetitions into two prime functions: fillers and self-initiated self-repairs. While filler refers to "an attempt to plan for a new utterance", self-repair describes "an attempt to monitor and modify an utterance" (Bada, 2010, p. 1685). In other words, fillers are intralingual CSs used by interlocutors as a time-stalling strategy to avoid silence and to hold the floor while processing the next utterance in their cognition.

On the other hand, self-repairs are intralingual CSs employed by speakers indicating that the trouble source in the language is identified and fixed via repetition. The modification of the utterance might occur due to several reasons, including grammatical mistakes, ambiguities, or pragmatically inappropriate usages. No matter the specific reason behind a self-repair, self-repair allows interlocutors to negotiate and fix the problem in the language. Table 5 provides the distribution of self-repetitions concerning their functions in both corpora.

Table 5. Functions of self-repetitions in LOCNEC and LINDSEI-TR

	Corpus size	**Functions of Self-repetitions**			
		Filler		Self-repair	
		n	n/1000	n	n/1000
LOCNEC	118.553	2218	18.71	451	3.80
LINDSEI-TR	63.924	965	15.10	763	11.94

As can be seen in Table 5, self-repetitions were recruited as fillers 2218 times whereas self-repetitions with a repair function emerged 451 times in LOCNEC, indicating that fillers were employed five times more than self-repairs in the L1 English. Such a large discrepancy between the functions of self-repetitions in LOCNEC yielded a statistically significant difference (X^2=1169.835; p=.000). In LINDSEI-TR, fillers (n=965) were more common compared to self-repairs (n=763). A statistical difference between the two types of functions was found at $p<0.05$ level (X^2=23.613; p=.000) in LINDSEI-TR, yet the difference was not as substantial as the one in LOCNEC. These findings corroborate previous research on the more frequent employment ratio of self-repetitions as fillers in L2 English (Bada, 2010; Genc, 2007; Rabab'ah, 2013; Yakut, 2013), while few studies indicate more frequent employment of self-repairs compared to fillers in L2 English (Badem, 2014).

As for the distribution of functions of self-repetitions between the two corpora, it was observed that self-repetitions with the filler function in LOCNEC (n/1000=18.71) was more common than the ones in

LINDSEI-TR (n/1000=15.10), and the difference between the two corpora was statistically significant at p<0.05 level (LL=+31.80). Contrarily, self-repetition as a self-repair strategy was overused in LINDSEI-TR (n/1000=11.94) compared to the usage rates in LOCNEC (n/1000=3.80) since a statistically significant difference between LOCNEC and LINDSEI-TR was found with respect to the self-repair usage rates (LL=-387.80).

Table 6 and Table 7 presents the distributions of functions of self-repetitions at WLR and GLR in the two data Table 6 lists functions of self-repetitions at WLR in LOCNEC and LINDSEI-TR.

Table 6. Functions of WLR in LOCNEC and LINDSEI-TR

	Corpus size	**Functions of Self-repetitions**			
		Filler		Self-repair	
		n	n/1000	n	n/1000
LOCNEC	118,553	1342	11.32	106	0.89
LINDSEI-TR	63,924	586	9.17	214	3.35

Out of 1448 WLR, 1342 were fillers, while only 106 were self-repairs in LOCNEC. Fillers emerged 13 times more frequently than self-repairs in L1 English discourse, which yielded a statistically significant difference between the two functions of self-repetitions (X^2=1055,039; p=,000). Out of 800 WLR in LINDSEI-TR, 586 emerged as fillers, whereas 214 repetitions functioned as self-repairs which resulted in a statistically significant difference between the two functions at WLR in L2 English discourse (X^2=172,980; p=,000) which supports previous claims that using repetitions as time stalling devices is more common than repetitions as self-repairs in the English interlanguages of Turkish individuals (Bada, 2010; Yakut, 2013). Looking at the chi-square test results within each corpus, we might assume a more balanced distribution of WLR in LINDSEI-TR compared to LOCNEC. Examples of these types at WLR in both corpora are given in [20] – [27]. The items in square brackets at the end of each example represent the function of the self-repetitions.

(20) it must be what erm Lady Thatcher's experience <X> *quite quite quite* a few times or at least the queen. [filler] (LOCNEC)
(21) we got a little bit of money sort of thing but again *not not* very much. [filler] (LOCNEC)
(22) erm I think it's about two or three years ago erm *I w*e went work experience for a week there. [self-repair] (LOCNEC)
(23) there was two people *my our* corridor who actually went mad like. [self-repair] (LOCNEC)
(24) I want to be a teacher *in .. in* state school. [filler] (LINDSEI-TR)
(25) *if* (eh) *if* they are her friends (eh) they know that she's not happy. [filler] (LINDSEI-TR)
(26) and then he *decides decide* to *decides* to (eh) . give his son. [self-repair, self-repair] (LINDSEI-TR)
(27) and material evaluation *course courses* they affected me very much. [self-repair] (LINDSEI-TR)

As for the comparison of functions of self-repetitions between the two data, we found that repetitions with the filler function were more common in LOCNEC (n/1000=11.32) than the occurrences in LINDSEI-TR (n/1000=9.37), and a statistically significant difference was found between the two data (LL=+18.63). Contrary to the results found in filler function between the two corpora, repetitions as self-repairs occurred more frequently in LINDSEI-TR (n/1000=3.35) compared to the emergences in LOCNEC (n/1000=0.89), yielding a statistically significant difference at p<0.05 level (LL=-133.93). Hence, L1 speakers of English tended to utilize self-repetitions as fillers at WLR more frequently than L2 speakers of English, whereas a reverse result was observed in self-repairs.

Table 7. Functions of GLR in LOCNEC and LINDSEI-TR

	Corpus size	**Functions of Self-repetitions**			
		Filler		Self-repair	
		n	n/1000	n	n/1000
LOCNEC	118.553	876	7.39	345	2.91
LINDSEI-TR	63.924	379	5.93	549	8.59

Table 7 illustrates the distribution of functions of self-repetitions at GLR in the two data. In LOCNEC, 876 repetitions were fillers, and 345 emerged as self-repairs. Such a large discrepancy yielded a statistically significant difference between the two functions of repetitions at GLR in LOCNEC (X^2=230.926; p=.000). In LINDSEI-TR, however, there were 379 fillers and 549 self-repairs, which resulted in a statistically significant difference between the functions of repetitions at GLR (X^2=31.142; p=.000). Contrary to observing more common facilitation of fillers than repairs in the L1 English data, we found that L2 English speakers utilized self-repetitions as repairs more frequently than repetitions as fillers, which correlates with the previous research (Genc, 2007; Bada, 2010; Yakut, 2013). Repetitions with each of the functions at GLR are exemplified in [28] – [35]. The items in square brackets at the end of each example represent the function of the self-repetitions.

(28) I'm. erm an *independent student independent student.* [filler] (LOCNEC)
(29) we= *went back went back* to York . and did some arc= archaeology. [filler] (LOCNEC)
(30) it ha= is having an effect *on the on my* teaching. [self-repair] (LOCNEC)
(31) because *I'm my mother's* half Scottish [self-repair]. (LOCNEC)
(32) *Alex thought Alex thought* that I think he (erm) *hurt her friend's* brea= (eh) hu= hurt (eh) *break .. her friend's* heart. [filler, self-repair] (LINDSEI-TR)
(33) they a little bit *make fun of her make fun of her.* [filler] (LINDSEI-TR)
(34) *he didn't. (eh) he: wasn't* pleased (eh) of the outcome. [self-repair] (LINDSEI-TR)
(35) she likes it (eh) . and ... show it *for your friends for her friends.* [self-repair] (LINDSEI-TR)

A comparison made between the two corpora concerning functional distributions of self-repetitions revealed that the frequency of repetitions employed to gain time to plan following utterances was more common in LOCNEC (n/1000=7.39) than the occurrences in LINDSEI-TR (n/1000=5.93). The difference in the proportion of fillers in the two corpora proved to be statistically significant when examined with the LL test (LL=+13.18). Usage rates of self-repairs in LOCNEC (n/1000=2.91) and LINDSEI-TR (n/1000=8.59) proved that Turkish learners of English resorted to repetitions to fix their utterances much more frequently when compared to L1 speakers of English. The difference in the use of repetitions as repairs proved to be statistically significant at $p<0.05$ level (LL=-256.91).

When the functions of WLR and GLR are compared, we found that filler function was more prominent in LOCNEC suggesting that the level of repetition does not have a direct impact on the function of repetitions. However, the comparison of WLR and GLR with respect to distribution of filler and self-repair in LINDSEI-TR showed that repetitions with filler function were more prominent at WLR while repetitions with repair function were more common at GLR among L2 speakers of English. These results suggest that speakers' L1 background directly affects the function of self-repetition considering the level of repetition as a variable. Depending on the results we can postulate that L2 speakers of English tended to employ GLR when they needed to repair their utterance which might be due to a longer cognitive processing demand. However, L2 speakers of English seemed to use WLR when they need to stall time for processing the following utterances. Although we can observe an increase in the function of self-repairs in LOCNEC at GLR compared to the ones at WLR, the change does not seem to be as drastic as the one we observed the change in LINDSEI-TR. According to the results, we can deduce that the rates of self-repair function at GLR are higher than the ones at WLR in both groups, yet the change seems to stand out in LINDSEI-TR, which might indicate the importance of speaking English as L1 or L2 in spoken discourse.

4. Discussion

This study confirms the existence of self-repetitions as a type of CS among L1 speakers of English and Turkish speakers of English, albeit with some differences between the two groups. This indicates that self-repetition as a CS is a natural reaction of L1 and L2 speakers which they resort to in order to avoid communication breakdowns in their spontaneous spoken interactions. Hence, we can say that repetition usage does not totally disappear in spontaneous interaction in the L2 context because repetitions as a CS emerge in L1, as well. For this reason, to change negative perceptions of language learners towards self-repetition usage in spoken interactions, language instructors might need to point out that self-repetition is a *sine qua non* feature of spontaneous speech in both L1 and L2 interactions. Language instructors also need to make sure learners comprehend that excessive use of self-repetitions in spoken interactions might cause communication breakdowns, as well. For this reason, meaningful learning of the functions of self-repetitions might be more beneficial for learners than rote learning. Although there is not a specific threshold for the appropriate amount of repetition usage in an interaction, learners could balance repetition usage as long as they are guided by language instructors.

Even though it is now clear that self-repetitions emerge in L1 and L2 spoken interactions, it is also known that beginner-level L2 learners of a specific language might have to struggle more to maintain communication in their interlanguages. Thus, novice language learners' awareness might be raised with respect to self-repetitions to refrain from non-filled pauses or the use of erroneous language yielding communication breakdowns in their L2 communication. The results suggest that self-repetitions should not be regarded as disfluency in the L2 context as long as they do not hinder communication. Despite heavy reliance on repetitions in both corpora, we also need to highlight that repetitions were more common among Turkish speakers of English than L1 speakers of English.

According to Bada (2010), the two main reasons for utilizing self-repetitions as a CS in verbal interactions are fillers and self-repairs. Categorizing functions of repetitions based on Bada's suggestions, we found greater reliance on repetitions as self-repairs in L2 English compared to L1 English. This might indicate that even advanced level L2 speakers of English make more frequent mistakes in their impromptu speech compared to L1 speakers of English. However, this does not mean that L2 speakers of English could not communicate in the English language, yet it might indicate that they could recognize and correct many mistakes thanks to their active monitoring processes by employing self-initiated self-repairs as a CS. Usage rates of fillers in the two corpora were relatively close to each other, albeit observing statistically significant differences between the two data. Besides, rates of functions of self-repetitions within L1 data revealed a heterogeneous distribution since the majority of repetitions functioned as fillers. The distribution of functions of self-repetitions in L2 English, however, was relatively balanced. Comparison of functions of self-repetitions within each corpus responded to why L1 and L2 speakers of English resort to self-repetitions in their spoken interactions. Hence, we can extrapolate that being an L1 and L2 speaker of English impacts the employment of self-repetitions with respect to their structures and functions.

The corpus study results also suggest that structures of self-repetitions, i.e., repetitions at WL and GL, between the two data revealed both similarities and differences. While the overall WLR usage rates did not indicate statistically significant differences between the two groups, overall GLR usage rates between the two data revealed statistically significant differences. Such a discrepancy in GLR usage might indicate that L2 speakers needed longer time while constructing their following utterances or fixing erroneous forms. GLR might have emerged more frequently in LINDSEI-TR as repeating phrases can help speakers gain more time than repeating single words.

In addition to the overall comparisons of WLR and GLR, the results of the study indicated that repetitions at WL in L2 mainly functioned as fillers while self-repairs at GL were more common compared to fillers in L2 speech. According to the results, L2 speakers of English employed WLR mainly as a time-stalling device whereas GLR were predominantly used as a repair device, which might indicate that L2 speakers preferred to repeat phrases instead of words when they recognize a mistake in their speech. This might be due to the fact that they had to linger long in fixing their speech. In addition, GLR with repair function might emerge more frequently in L2 speech as they might think that repairing via GLR might help them transmit the intended message more clearly. Despite L2 speakers' preferences, L1 speakers of English resorted to both WLR and GLR to gain time while constructing their speech, as the results indicated predominant filler function of repetitions in L1 English speech. The comparison of L1 and L2 speakers of English with respect to the use of repetitions at WL and GL with the functions of

fillers and self-repairs indicate that self-repetitions are used with different purposes between the groups. Hence, such results suggest that L1 background, language repertoire of the speakers and individual interlanguages might render such fluctuation in the use of repetitions.

Even though we attempted to explore forms and functions of self-repetitions in L1 and L2 English in this study, further research might investigate and compare self-repetition usage among L2 English speakers coming from different L1 backgrounds. In addition, a longitudinal study aiming at revealing the relationship between L2 proficiency level and self-repetition usage might give paramount insight into whether repetition usage changes over time.

5. Conclusion

This study aimed at exploring forms and functions of self-repetitions in L1 and L2 English spoken conversations. Using corpora, we compared whether and how L1 and L2 users of English differed with respect to repetition usage. The results suggested that repetitions are recruited to gain time to prepare the next utterance cognitively or correct mistakes made by the speakers themselves. Since repetitions are not used only by L2 speakers, we concluded that such usages should not be regarded as disfluency in speech but rather a kind of CS that interlocutors resort to in times of difficulties in spoken interactions.

This study suggests that repetitions as either fillers or self-repairs are actively used in L1 and L2 English discourse. However, repetitions as fillers are more favoured in L1 English, while self-repairs are more common in L2 English. Since self-repetitions fulfil interpersonal functions in spoken discourse, we believe that they do not impair the fluency of communication as long as they are not used excessively affecting the natural flow of speech. As self-repetitions do naturally emerge in spoken discourse when interlocutors need to stall for time or correct erroneous forms or functions in their speech, language teachers should not explicitly teach novice learners to consciously use self-repetitions in times of difficulty, but raise their awareness regarding the functions of self-repetitions in spoken interaction. Hence, learners could avoid non-filled long pauses by employing various strategies, including self-repetitions as fillers. In addition, if they are familiarized with error-correction strategies, one of which is self-repetitions as repairs, learners will not hesitate to correct their linguistic and pragmatic mistakes when their language monitor warns them about erroneous forms.

References

Anthony, L., 2019. *AntConc (Version 3.5.8) [Computer Software]*. Tokyo, Japan: Waseda University. Available at: https://www.laurenceanthony.net/software

Bada, E., 2010. Repetitions as vocalized fillers and self-repairs in English and French interlanguages. *Journal of Pragmatics*, vol. 42, no. 6, pp. 1680-1688.

Badem, N., 2014. *Repetitions as communication strategies used by Turkish adult EFL learners: A cross-sectional study* [Unpublished master's thesis]. Adana: Çukurova University.

Bialystok, E. and Fröhlich, M., 1980. Oral communication strategies for lexical difficulties. *Interlanguage Studies Bulletin*, vol. 5, no. 1, pp. 3-30.

Clark, H. H., 1994. Managing problems in speaking. *Speech Communications*, vol. 15, no. 3-4, pp. 243-250.

Crible, L. and Pascual, E., 2020. Combinations of discourse markers with repairs and repetitions in English, French, and Spanish. *Journal of Pragmatics*, vol. 156, pp. 54-67.

De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literatures (BELL), New Series 2*, 225-246. http://hdl.handle.net/2078.1/75157

Dörnyei, Z., 1995. On the teachability of communication strategies. *TESOL Quarterly*, vol. 29, no. 1, pp. 55-85.

Dörnyei, Z. and Scott, M. L., 1995a. Communication strategies: What are they and what are they not? Paper presented at the *Annual Conference of the American Association for Applied Linguistics (AAAL)*, Long Beach,CA.

Dörnyei, Z. and Scott, M. L., 1995b. Communication strategies: An empirical analysis with retrospection. In: J. S. Turley and K. Lusby, eds. *Selected papers from the proceedings of the 21st Annual Symposium of the Deseret Language and Linguistics Society*. Provo, UT: Brigham Young University, pp.155-168.

Dörnyei, Z. and Scott, M. L., 1997. Communication strategies in a second language: Definitions and taxonomies. *Language Learning*, vol. 47, no. 1, pp. 173-210.

Faerch, C. and Kasper, G., 1984. Two ways of defining communication strategies. *Language Learning*, vol. 34, no. 1, pp. 45-63.

Fox, B. and Jasperson, R., 1995. A syntactic exploration of repair in English conversation. In: P. Davis, ed. *Descriptive and Theoretical Modes. The Alternative Linguistics*. Amsterdam: John Benjamins, pp. 77-134

Fujimura-Wilson, K., 2007. Japanese exact repetitions involving talk among friends. *Discourse Studies*, vol. 9, no. 3, pp. 319-339.

Genç, B., 2007. *An analysis of communication strategies employed by Turkish-speakers of English* (Publication no. 220634) [Doctoral dissertation] Adana: Çukurova University. Council of Higher Education Thesis Centre.

Genç, B., Mavasoglu, M. and Bada, E., 2010. Types and functions of repetitions in the narrations of Turkish speakers of French. *Novitas Royal*, vol. 4, no. 2, pp. 216-224.

Gilquin, G., De Cock, S. and Granger, S., eds. 2010. *The Louvain international database of spoken English interlanguage.* Louvain-la-Neuve: Presses Universitaires de Louvain.

Granger, S., 1996. From CA to CIA and back: An integrated contrastive approach to computerized bilingual and learner corpora. In: K. Aijmer, B. Altenberg and M. Johansson, eds. *Languages in contrast. Text-based cross-linguistic studies*. Lund: Lund University Press, pp. 37-51.

Granger, S., 2008. Learner corpora. In: A. Lüdeling and M. Kytö, eds. *Corpus linguistics: An international handbook*. Berlin: Walter de Gruyter, pp. 259–275.

Kaur, J., 2012. Saying it again: Enhancing clarity in English as a lingua franca (ELF) talk through self-repetition. *Text & Talk*, vol. 32, no. 5, pp. 593-613.

Kellerman, E., 1998. When words fail: from communication strategies to strategies for communication. In: K. Malmkjaer and J. Williams, eds. *Context in language learning & Language understanding*. Cambridge: Cambridge University Press, pp. 91-112.

Kilimci, A., 2014. LINDSEI-TR: A new spoken corpus of advanced learners of English. *International J. Soc. Sci. & Education*, vol. 4, no. 2, pp. 401-410.

Kovac, M. M., 2016. Repetitions as a communication strategy: A case study. *Studies in English Language Teaching*, vol. 4, no. 1, pp. 87-103.

Lickley, R., 1994. *Detecting disfluency in spontaneous speech* [Unpublished doctoral dissertation]. Edinburgh: University of Edinburgh.

Lyster, R., 1998. Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, vol. 20, no. 1, pp. 51-81.

Maclay, H. and Osgood, C. E., 1959. Hesitation phenomena in spontaneous English speech. *Word*, vol. 15, no. 1, pp. 19–44.

Rabab'ah, G., 2013. Strategies of repair in EFL learners' oral discourse. *English Language Teaching*, vol. 6, no. 6, pp. 123-131.

Rieger, C. L., 2003. Repetitions as self-repair strategies in English and German conversations. *Journal of Pragmatics*, vol. 35, no. 1, pp. 47-69.

Schegloff, E. A., Jefferson, G. and Sacks, H., 1977. The preference for self-correction in the organization of repair in conversation. *Language*, vol. 53, no. 2, pp. 361-382.

Silva, G. and Santos, D., 2006. Framing participation through repetition: the case of a Portuguese learner in different settings. *Portuguese Language Journal*, vol. 1, pp. 1-23.

Stuart, S. and Lynn, P., 1995. Development of communication strategies among foreign language learners. *Mountain Interstate Foreign Language Review*, vol. 5, pp. 122-127.

Tarone, E., 1980. Communication strategies, foreigner talk, and repair in interlanguage. *Language Learning*, vol. 30, no. 2, pp. 417–431.

Tarone, E., 1981. Some thoughts on the notion of communication strategy. *TESOL Quarterly*, vol. 15, no. 3, pp. 285-295.

Tarone, E., 1983. Some thoughts on the notion of "communication strategy." In: C. Faerch and G. Kasper eds. *Strategies in interlanguage communication*. London: Longman, pp. 61-74.

Tarone, E. and Yule, G., 1987. Communication strategies in east-west interactions. In: L. E. Smith, ed. *Discourse across cultures: Strategies in world Englishes*. New York: Prentice-Hall, pp. 49-65.

Tekdemir Yurttas, G., 2012. Repetitions in Turkish: Talk among friends. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, vol. 13, no. 2.

Willems, G. M., 1987. Communication strategies and their significance in foreign language teaching. *System*, vol. 15, no. 3, pp. 351–364.

Yakut, I., 2013. *Interlingual development of Turkish adult learners of English: L1 vs. L2-based communication strategy use* (Publication no. 327705) [Master's thesis] Adana: Çukurova University. Council of Higher Education Thesis Centre.

Yakut, I. and Bada, E., 2021. Interlanguage development of Turkish speakers of English: Exploring oral and written communication strategies. *The Journal of Teaching English for Specific and Academic Purposes*, vol. 9, no. 4, pp. 611-626.