Investigating Language-Related Episodes during Mechanical and Meaningful Output Activities

Shirin Abadikhah

Department of English Language and Literature, University of Mazandaran

Pasdaran Street, 47415, P.O. Box: 416, Babolsar, Iran

E-mail: sh_abadikhah@yahoo.com

Abstract

The present study examines how EFL learners consciously reflect on their language during a set of mechanical and meaningful output activities. Thirty-six Farsi learners of English negotiated on linguistic features and completed six activities over a period of six weeks. The transcripts from the learners' interaction were analyzed for instances of language-related episodes (LREs), their principal focus on meaning or grammar and their nature and outcome. The results showed that (1) the meaningful output activities elicited significantly more LREs than did the mechanical output activities, (2) while approximately half of the LREs in the meaningful activities focused on lexis and meaning, the majority of LREs in the mechanical activities were directed towards grammatical forms and a small portion was focused on meaning, (3) the two output groups differed significantly in the continuous and correctly solved episodes. The study provides support on the effectiveness of collaborative output activities in pushing learners to verbalize their internal linguistic processing and focusing their attention on a wide range of linguistic features.

Keywords: Language-related episode, Collaborative dialogue, Meaningful and mechanical output

1. Introduction

Recent studies have indicated a need to investigate a range of grammar-based output activities featuring collaborative dialogue. In particular, the role of meaningful output has been emphasized by a number of SLA researchers (e.g., Izumi, 2003, Izumi & Izumi, 2004; Swain, 1997, 2000). The emphasis arouse out of the observation that the output tasks employed in some of the studies did not result in the promotion of L2 learning (e.g. Izumi & Izumi, 2004). This was contrary to expectations and inconsistent with the main claim of the *output hypothesis* which posits that language production plays a significant role in the acquisition of a second language. Swain (1985, 1995) proposed that output provides a unique opportunity for the use of linguistic resources and allows learners to test hypotheses about the second language. Output may encourage learners to move from semantic processing prevalent in comprehension to syntactic processing necessary for production. In fact, by being 'pushed' to produce language, learners are required to pay attention to syntactic and morphological features of the language in order to produce accurate, precise and appropriate language.

The output tasks employed by several researchers attempt to fix attention on form by holding overall text meaning constant. Among the well researched collaborative output tasks are dictogloss and text reconstruction task (Izumi, 2002; Kowal & Swain, 1994; Leeser, 2004; Nabei, 1994; Swain, 1998; Swain & Lapkin 2001). The former involving learners in reconstructing a text after being read aloud by the teacher or researcher, and the latter requiring learners to reconstruct the text after reading it. Both tasks propose to provide a meaning-focused context to raise learners' awareness of the discoursal use of the target linguistic feature. However, unstructured production tasks such as text reconstruction and dictogloss may not direct learners' attention to the predetermined target linguistic forms. Since the learners are involved in conveying the content of the text, they tend to pay little attention to the form of their language (Storch 1998). Therefore, an additional attention drawing technique may be required to orient the learners' attention to the target linguistic forms and increase the likelihood of their discussion on those forms. Further research is required to explore the nature of other grammar-focused activities ranging from mechanical to meaningful output. It is not clear to what extent these activities would help learners negotiate over the target form and what features of language they would focus on during their collaboration. A further limitation of the previous studies is that the participants have been either from mixed L1 backgrounds or, in the case of the Canadian studies, grade 8 French immersion students. There is consequently a need for research into collaborative output in classes where the learners share an L1 background, as is the case in most teaching situations around the world. Thus, the purpose of the current study is to examine the interaction among Iranian EFL learners as they

carry out a range of output activities, providing a form-focused and a meaning-focused context in two intact classrooms.

2. Review of Literature

Many SLA researchers have recently applied sociocultural theory to L2 learning studies (Lantolf, 2000, 2007). According to sociocultural theory, originated in Vygotsky's (1978) work, humans organize and direct their mental activities, such as thinking and problem-solving by symbolic tools (e.g. language). It is argued that new knowledge is initially acquired in interpersonal or social level (Lantolf, 2006). The socially constructed knowledge is then internalized during interaction and collaboration between the learners. Adopting a sociocultural approach, Swain (2000, p.113) extended the scope of the output hypothesis and proposed that 'internal mental activity has its origins in external dialogic activity' and 'language learning occurs in collaborative dialogue'. Thus, the attention from individual one-way production of output was shifted to interactive dialogic production, referred to as 'collaborative dialogue'. Swain and her colleagues explained that during collaborative dialogue, 'learners work together to solve linguistic problems and co-construct language or knowledge about the language' (Swain, Brooks & Tocalli-Beller, 2002, p. 172). It is further argued that collaborative dialogue can objectify thought and make it available for further scrutiny (Swain, 2000). Donato (2004) holds the same view and believes that the input-output approach, which is emphasized in interactionist theory, disregards learners' goals and participation in their own learning process. Similarly, Lantolf and Thorne (2007, p.207) note the importance of dialogue and believe that learning in collaboration, especially in instructional setting results in development, in fact 'development arises in the dialogic interaction that transpires among individuals'.

In her recent explication of the concept of the output hypothesis, Swain (2000) stresses the role of meaningful output in L2 learning. During production, learners need to connect the linguistic form to its meaning. If their attention is focused on the intended structures in a meaningful way, then they may develop their interlanguage more deeply and with more mental effort than when simply comprehending those structures. As a result, learners might be able to internalize new linguistic forms and improve the grammatical accuracy of their production. Therefore, meaningful output may involve learners in a cognitive process required for consolidating form-meaning connections of L2 knowledge. Many other researchers have also speculated that for output to be effective for learning, it must trigger certain cognitive processes. For example, Izumi (2003) and Izumi and Izumi (2004) claimed that a mechanical production task does not engage learners in natural production mechanisms such as *conceptualizing* the message content and *formulating* the language (as discussed in Levelt, 1989). Referring to a study by Fotos and Ellis (1991), Kowal and Swain (1994) also argue that the quantitative analysis of the data does not always bear out the qualitative analysis. In Fotos and Ellis (1991), although the interactions generated a great deal of interactional modifications and exchanges, these were mechanical in nature, with little extension of the use of the target language (as cited in Kowal & Swain, 1994). Thus, as Swain (1997) proposes, in determining the task type, the characteristics of the task should be considered carefully because not all tasks elicit meaningful output. Despite the abundance of empirical studies and theoretical reasons on the use of the output tasks, there is relatively little research on the effect of mechanical and meaningful output on learners' focus of attention and the nature of their exchanges. With this in mind, there is an urgent need for a study exploring the context of production and how participants interpret and complete a task in intact classroom setting. The present study, thus, examines the influence of output task type on focusing learners' attention on their language and fostering the use of metatalk.

2.1 Language-related episodes in various tasks

A number of studies have examined the learners' discussion during the completion of different output tasks. Kowal and Swain (1994) argued that the choice of task and how participants interpret and complete it must be considered in using collaborative tasks in L2 classroom. This was the first study by Swain and colleagues investigating the contribution of collaborative output on L2 learning. They were particularly interested in knowing a) what happens if learners talk about form in relation to a meaning they wish to convey, b) the relationship between the learners during task completion and the effect of this relationship on the quality of the interactions and c) the nature of feedback (accurate or inaccurate) the peers provide for each other. They hypothesized that the activities which promoted learners' output and encouraged them to focus on their output may enhance the accuracy of their production. The task they employed was a dictogloss technique, introduced by Wajnryb (1990). In this task, learners listen to a short passage (either tape-recorded or read by the class teacher) and take separate notes. Next, they reconstruct their own version of the passage using their notes. The participants were 19 mixed ability students from an intact grade 8 French immersion classroom. They focused on the acquisition of French grammar with particular attention to present tense. The study involved four dictoglosses employed at fortnightly sessions over a two-month period. The learners' interaction during the completion of the third dictogloss was audio-recorded, transcribed and analyzed. Following Duff (1986) and Doughty and Pica (1986), who proposed dyadic interaction as the most appropriate grouping for the L2 classroom, they assigned learners to self-selected pairs. In the transcripts of the interactions, they identified Critical Language-Related Episodes (CLREs) and categorized them

into meaning-based, grammatical and orthographic episodes. According to Kowal and Swain (1994, p. 80), 'a CLRE began with the identification of a grammatical point to be discussed or a sentence or phrase which needed to be reconstructed and finished once the discussion was completed'. The total number of CLREs produced by the groups was 224, 42% of which focused on grammar, 31% on meaning and 28% on orthographic episodes. As mentioned above, the proficiency level of the participants differed; they were grouped as high, upper-middle, lower-middle and low proficiency levels in French. By extending Vygotsky's *Zone of Proximal Development* ─ which was originally defined as a situation in which an adult provides assistance for a child ─ Kowal and Swain (1994, p. 85) assumed that 'the more able peer will provide the same sort of assistance to the less able peer' in the L2 learning context, which encouraged them to assign learners to heterogeneous groupings. However, analysis of the transcripts revealed that in a pair of heterogeneous ability, the more proficient learner tended to do most of the hypothesizing. They reasoned that the less proficient learner may not have contributed because they were a) willing for the more proficient learner to do the task, b) too intimidated to say anything or c) not allowed to contribute to the discussion and task completion. Kowal and Swain further observed that in the more homogenous pairs, the contributions of the participants were more balanced, with both members contributing to the discussion and the role of 'expert' being fluid, alternating between the students. With respect to the grouping of the learners, they concluded that 'perhaps what needs to be avoided are extreme degrees of heterogeneity' (e.g., upper-middle and low), since there were some pairs with mixed abilities (upper-middle and lower-middle) who displayed successful collaboration (Kowal and Swain, 1994, p. 86). Considering the task type, Kowal and Swain concluded that the dictogloss was successful in encouraging learners to attend to the accuracy of their language and form-function relationship.

Storch (1998) argued that communicative tasks do not focus learners' attention on the grammatical forms of the target language. Thus, to promote negotiation over grammatical forms, she employed four grammar-focused tasks: multiple choice, text reconstruction, cloze and composition. The tasks focused on the choice of article, verb tense, word forms and singular and plural nouns. Eleven students from various linguistic backgrounds at two undergraduate and postgraduate levels completed the tasks. They completed the tasks in the same self-selected dyads and one triad in two sessions. In the first session, the learners completed a text reconstruction task and in the second session, they worked on the multiple choice, cloze and composition tasks. Data analysis took place in two stages: firstly, the learners' talk during the completion of the tasks was examined for the number of LREs, and secondly, the way that the learners made their decisions on grammatical features was considered in detail. The results indicated that almost all learners' talk in the more structured tasks, i.e. multiple choice and text reconstruction, focused on grammar (100% and 93%, respectively). Less structured tasks such as composition elicited less attention to language and grammar (53%). Most of their talk was devoted to planning, brainstorming, generating ideas and producing the content rather than providing correct grammatical forms. Storch also found that the text reconstruction task produced the highest amount of metatalk focusing on linguistic features. In the second stage of analysis, she identified the taxonomy of knowledge resources in the learners' transcripts, such as application of a grammatical rule, offering the meaning of the words or phrases, intuition and contextual clues in defending their grammatical decisions.

Fortune and Thorp (2001) also examined the metalinguistic function of output and amended LRE framework introduced in the previous study (i.e. Kowal and Swain, 1994). Their study involved two linguistically heterogeneous classes of EFL learners from 14 different L1 backgrounds. Each class consisted of five triadic groups and was divided into three proficiency bands based on a grammar test. They argued that 'analysis based on LRE counts, although valuable, fails to capture completely the complexity of the interaction' (Fortune & Thorp, 2001, p.143). Therefore, they introduced two further categorizations of *nature* and *value* to demonstrate major features of LREs. With respect to the type of LREs, they amended Kowal and Swain's (1994) taxonomy by subcategorizing grammatical episodes into inflectional and derivational morphology, verb tense, verb form, gerund and infinitive. They further divided discourse episodes into reference, linking text elements with an appropriate connector, and lexical cohesion.

Swain and Lapkin (2001) studied the nature of the two communicative tasks of dictogloss and jigsaw in further detail. They wanted to know which task type would generate more noticing the gap, hypothesis testing and metatalk. As mentioned earlier, the dictogloss task engaged learners in listening to a text read at normal speed. The students took separate notes on the content of the text and worked together to reconstruct the passage. The jigsaw task involved learners in constructing a short story based on a series of pictures. Both tasks depicted the same story, thus, they were 'similar in content but different in form' (Swain & Lapkin, 2001, p.100). The researchers anticipated that since the jigsaw is a 'meaning negotiation' task (Pica *et al.*, 1993), it would generate less focus on form than the dictogloss task would do. Two classes of grade 8 French immersion students with mixed abilities attended the study. The data were collected in five stages. LREs in the transcripts of the pair talk were coded as lexis-based or form-based. They hypothesized that the dictogloss would generate more attention to form and the jigsaw would produce more attention to meaning than their counterparts would do. The transcripts highlighted

three salient differences between the two tasks. Firstly, they differed in the type of stimulus ─ while the stimulus in the jigsaw task was visual, it was auditory for the dictogloss task. Secondly, the dictogloss offered a linguistic model on the basis of which learners could establish their own story, whereas the jigsaw did not. Finally, the two tasks differed in their cognitive demands on the learners' understanding. That is, while the pairs working on dictogloss produced their narratives in a paragraph form, those working on jigsaw produced their narratives in separate numbered sentences. Thus, the pairs in the former group had to deal with discourse requirements, such as linking their sentences together and giving coherence to them. Comparison of the LREs revealed that the two tasks did not differ significantly in a) the average number of LREs they produced, b) the average time spent on the completion of the task, and c) the average number of lexis-based and form-based LREs. Thus, contrary to the researchers' expectations, the learners in both tasks similarly focused on form. However, with respect to accuracy, the jigsaw learners produced fewer correct pronominal verbs compared to the dictogloss learners. As regards discourse structures, the dictogloss learners attended to logical and temporal sequencing of their sentences, which resulted in composing paragraphs, whereas such attention was not present in the numbered sentences produced by the jigsaw students. Furthermore, the linguistic nature of the stimulus in the dictogloss and the less open-endedness in linguistic focus, compared to jigsaw, constrained the range of vocabulary the pairs used and the time they spent on the task.

Leeser's (2004) study concerned the effect of proficiency of dyads of learners on the number, type and outcome of LREs. Twenty-one pairs of adult L2 Spanish learners completed a text-reconstruction task, *i.e.* dictogloss. Based on the instructors' overall ability ratings, he classified the learners into higher-higher proficiency (H-H=8 dyads), lower- higher proficiency (L-H=9 dyads) and lower-lower proficiency (L-L=4 dyads) levels. The transcribed pair talk was analyzed for types of LREs (lexical and grammatical) and outcome of LREs (correct resolution of the problem, unresolved or abandoned problem and incorrect resolution of the problem). A total of 138 LREs were identified in the transcripts of 21 dyads. The learners solved their linguistic problems correctly on most occasions (77%) and the rest of their problems were approximately divided into either unresolved (11%) or resolved incorrectly (12%). With regard to the focus of LREs, 40% of the LREs addressed lexical features and 60% grammatical features. The comparison of the number and types of LREs produced by three proficiency groupings (H-H, H-L and L-L) showed a positive relationship between total number of LREs and the proficiency level of dyads. In other words, as the proficiency of the dyads increased, so did the mean number of total LREs. Moreover, the percentage of lexical and grammatical LREs varied according to the type of dyad: the higher proficiency learners (H-H) focused more on grammatical (67%) than on lexical (33%) items, whereas the lower proficiency learners (L-L) focused more on lexical (58%) than grammatical (42%) items. The comparison of the outcome of episodes across the three dyadic groupings showed that although all three groups solved most of their problems correctly, the higher proficiency learners solved more LREs correctly than did the other two dyadic types. Leeser concluded that the proficiency level of the dyads influenced the amount, type and outcome of LREs they produced during their discussions.

Malmqvist (2005) investigated the effects of small-group interaction on written German output. She employed three dictogloss tasks: the first and third were completed individually and the second collaboratively. Twelve students with mixed abilities (*i.e.* high and low proficiency levels) participated in this study. Malmqvist formed triadic groups based on such variables as gender and proficiency level. Her assumption was that the learners would benefit most in heterogeneous groups. To determine the focus of their attention in LREs, she analyzed audio-taped interactions of the learners during the reconstruction of the task. The LREs were divided into meaning-based, grammatical and orthographic episodes. The results of LRE comparison demonstrated that the less proficient learners attended primarily to meaning and lexical items rather than grammatical items, giving support to Leeser's (2004) finding. With respect to the outcome of episodes, the learners solved the problems correctly on most occasions. She further noted that, in addition to proficiency level, personality traits can also influence the outcome of LREs and collaborative tasks. She observed that sometimes a less proficient but more confident member of a group can convince the more proficient learner to accept a wrong decision, thus, 'it is not always the people with the best talent for convincing who are right' (Malmqvist, 2005, p. 138).

The initial coding system introduced by Kowal and Swain (1994) was further extended by Benson, Pavitt and Jenkins (2005), who conducted a small-scale study to examine the focus and nature of the discussion occurring among ESL learners. They employed dictogloss since it is assumed to be a 'planned, closed, convergent and two-way task' (p.1). The study involved three classes of learners of English from three proficiency levels ─ intermediate, upper-intermediate and advanced. They were adult learners coming from multilingual backgrounds, who were more exposed to a form-focused approach of learning than the Canadian immersion students participating in Swain and colleagues' studies. Thus, the researchers expected that the learners would have access to a wide range of meta-linguistic knowledge allowing them to discuss a variety of formal features of language during their interaction. The same text was used for all three levels of participants in order to increase the chance of discussion on similar linguistic items and allow for comparison of their performance. The text was read twice, with

the learners listening to the first reading and taking notes during the second. After the learners were assigned to dyads and triads, a third reading was conducted to stimulate further discussion. While the learners were reconstructing the text in separate rooms, they were tape-recorded. The researchers adapted Kowal and Swain's (1994) coding system, which included meaning-based, grammatical and orthographic episodes. They further divided the meaning-based episodes into meaning-definition, meaning explanation, sentence-level meaning and text-level meaning. They also added three categories including discourse (i.e. discussion on how to connect text parts), identification (i.e. discussion on what was said by the teacher) with sub-divisions of text, sentence and word and reading aloud, representing those segments of speech where learners seemed to vocalize what they were writing. Analysis of one group from each level showed that a) the grammatical episodes were less frequent than other episodes in all three groups, b) the intermediate group were more concerned about content and identification than grammar, and c) the highest number of identification and reading aloud episodes occurred in the upper-intermediate group, who seemed to be less successful in completing the task than other groups. Furthermore, the learners made very little use of meta-language in their discussions, spending most of the time on reading aloud and identification of text parts.

Storch (2007) repeated her earlier study with four intact ESL classes at university level. She wanted to examine the nature of the learners' talk during completion of the task. The participants were university students from different L1 backgrounds and in high intermediate level. The participants in class A completed the task in pairs (9 pairs and one group of 3) and in class B individually (16 students). In the other two classes (C and D), the participants were free to complete the task in pairs or individually. The interaction between the students in class A was audio-recorded. Storch employed a text editing task, which required them to make changes to the text in order to improve its accuracy. The text was deliberately seeded with errors in verb tense/aspect, use of articles and word forms. The learners' decisions on the grammatical accuracy and lexical appropriateness of the text elements were scored as either correct/acceptable or incorrect/unacceptable. The transcripts of the learners' interaction were analyzed for LREs, which were categorized as form-focused (dealing with morphology or syntax), lexis-based (dealing with word meaning and word choices) and mechanics (dealing with punctuation, spelling and pronunciation). The analysis of the edited text scores indicated that the learners in the two individual and collaborative conditions did not differ significantly in the mean accuracy score. The participants in pair group (class A) focused more on grammar (67% of all episodes) than lexis (31%). Similar to Leeser's (2004) study, most of the LREs were resolved correctly (80%). The results also indicated that pairs of learners spent more time on the task than did the individual learners.

Findings from previous studies suggest that learners' collaboration might result in modification or consolidation of their current linguistic knowledge. However, apart from the study by Storch (1998), the majority of the studies have employed dictogloss or text-editing tasks. The findings of some of the studies have also suggested that 'extreme heterogeneity' in students' proficiency level may hinder collaborative learning. To control for the variation in the performance of the pairs of learners, all participants should be selected from the same proficiency level, as is commonly found in real EFL classrooms.

3. Research questions

The present study seeks answers to the following research questions:

1). Do learners working collaboratively on the meaningful output activities produce more language-related episodes (LRE) than learners working collaboratively on the mechanical output activities?

2). Do learners in the two output groups (Mechanical and Meaningful) predominantly focus on grammatical or meaning-based features of the language?

3). Do output activity types (Mechanical and Meaningful) affect the nature and outcome of LREs?

4. The study

4.1 Participants

The study was carried out in two intact classes in a private language school in Tehran, Iran. The participants were 36 Farsi learners of English, all females and within the age range of 15 to 28. They were attending an intensive English programme, which had several levels of instruction, ranging from beginning to advanced level of proficiency. The data for the present study were collected from two low intermediate intact classes. The two classes were randomly assigned to either the Mechanical or the Meaningful group and the participants in each class (18 learners) were asked to choose their partners as they were going to work in pairs.

4.2 Instructional tools

Following Loschky and Bley-Vroman (1993), two sets of material with differing degrees of control and meaningfulness were developed across a continuum. Each set consisted of three activities, which were designed to elicit the use of subject, direct object and object of preposition relative clauses. The activities at the more

controlled end of the continuum represented the Mechanical output and consisted of *substitution*, *transformation* and *fill in the blank*. They involved learners in producing target linguistic forms without necessarily knowing the meaning or function of the words. The responses were highly controlled and only one correct answer was possible for each item. The activities at the less controlled end of the continuum represented the Meaningful output and consisted of *picture description*, '*let's complain*' and *dictogloss*. In these activities, the linguistic forms were not isolated from their meaningful context. In fact, they were designed to promote constant attention to the form-meaning relationship in production. The learners in this group had more freedom in choosing the linguistic forms than the former group.

The first Mechanical activity employed in this study was a substitution drill, which is, according to Dakin (1973), the simplest way of requiring learners to produce an utterance without paying attention to meaning. In this drill, the participants were provided with the main clause of the sentence as a prompt. They had to produce the relative clause using the words and phrases given above the picture and with a slight change in the model sentence. The second drill for the Mechanical group was a transformational drill designed to give practice in the structure of relative clauses by varying the original sentence in a predetermined way. Dakin (1973) defines transformational drill as one of the meaningless drills which require changes in the word order of the sentence involving the addition or deletion of grammatical constituents, changes in voice from active to passive, or changes in sentence type from simple to complex or compound. Following this definition, a transformational drill was developed to give practice in changing simple sentences into complex sentences containing relative clauses (Hutchinson, 1992). In the third activity, fill in the blank, a passage, adapted from low intermediate level EFL textbooks, was presented to the participants with its relative clauses missing. The learners were required to complete the passage using the information provided in a box. Working together in pairs, they discussed choosing the appropriate sentences and attaching them to the text.

In the picture description task, a series of pictures were presented to the participants in the Meaningful group. Following Chalker (1987) and Seidl (1992), they were asked to look at each picture and produce a relative clause to describe the person or the object in the picture. The second activity for the Meaningful group, 'let's complain', was intended to give practice in the structure and function of the relative clauses. Following Ur (1988), the participants were told that they were going to have a complaining session and they had to complain about the things that bother them. Since brainstorming and finding a topic to write about would take some time, they were suggested topics such as people, problems, surroundings, course books and homework. The third activity was dictogloss which encourages learners to reflect on their own output (Swain, 1998; Wajnryb, 1990). The activity involved the participants in listening to a text read at normal speed and to reconstruct it through collaboration. The passage used for this activity was a narrative text with a clear structure and sequence of events. Following Swain (1998), the learners were instructed that they were going to reconstruct the text, played in the tape, matching the content and grammar of the original text as closely as possible. After listening to the text for the second time and taking notes, the learners pooled their resources to reconstruct the passage. Finally, they were supplied with the original text in written form and compared their written output with it.

In addition to these activities, the learners in both groups were supplied with an input sheet to draw their attention to the target linguistic form and increase the likelihood of their discussion on it. The input page served as a warm-up activity to initiate the discussion and contained a brief description of English relative clause structure in Farsi, accompanied by relevant examples in English.

4.3 The procedure

The participants in each group completed three activities in fortnightly sessions over a period of 6 weeks. The Mechanical group worked on substitution, transformation and fill in the blank on the first, third and fifth week; the Meaningful group completed the picture-description, 'let's complain' and dictogloss on the second, fourth and sixth weeks. They were tape-recorded while interacting with each other in dyads. To ensure the clarity of recording and to prevent distraction from other pairs, each pair was tape-recorded separately at a specified time of the week.

As Swain and Lapkin (2000) put it, learners may benefit from using L1 during collaborative dialogue. Thus, the participants in the current study were advised to use either Farsi or English while completing the activities. In order to encourage joint production, each pair was given only one copy of the task and the input sheet. They had no access to dictionary or any other aid during the sessions and were requested not to refer to any textbook about relative clause structure for the duration of the study. Prior to the study, they signed consent forms to agree to participate in the study and allow their recorded voice to be used.

4.4 Categorization of language-related episodes

One source of data to explore L2 learning process is the language-related episodes (LREs) produced during pair-talk. According to Swain (1998, p. 69), LREs induce a kind of focus on form which 'may serve the function of helping students to understand the relationship between meaning, forms, and function in a highly context-sensitive

situation'. An LRE refers to 'any part of the dialogue where learners talk about the language they are producing, question their language use, or other- or self-correct' (Swain & Lapkin, 1998, p. 326). Previous studies have analyzed transcripts of the tape-recorded interaction for the occurrence of LREs in each pair's discussion (Nabei, 1996). LREs are divided into several major categories including type, nature and outcome (*e.g.* Benson, Pavitt & Jenkins, 2005; Fortune & Thorp, 2001; Kowal & Swain, 1994; Leeser, 2004). These categorizations of LREs guided the analysis of learner interaction in the present study. Once the major categories were established, the subcategories emerged from the data.

The types of episode included grammatical, meaning-based, orthographic, discourse and identification. The grammatical episodes constituted those parts of the interaction in which the participants discussed syntactic and morphological features of the language. These episodes were subdivided into categories involving verb form (verb tense/aspect, auxiliary verb, verb form: passive or active verb), relative clause structure (choice of the relative pronoun/clause, omission or retention of the pronoun, choice of the defining/non-defining clause, clause position: right-embedded or centre-embedded clauses, finding the referent of the relative pronoun), word order, subject-verb agreement, choice of preposition, conjunction, definite/indefinite article, gerund/infinitive, genitive S, pronoun and adverb of time. The meaning-based episodes constituted those parts of the interaction where learners' attention was drawn into semantic components of the language such as negotiating the meaning of the words or clauses and the content of the sentences to be constructed. This category was subdivided into considering lexical or clause choices, word/phrase meaning and vocabulary search. The orthographic episodes were subdivided into spelling, punctuation and pronunciation. Since the pronunciation of the words was sometimes followed by a request for their spellings, this category was subsumed under orthographic episodes. Following Benson, Pavitt and Jenkins (2005), the identification category was introduced to code those segments of speech in which the learners identified words, phrases and sentences mentioned in the tape (in the dictogloss task). Finally, discourse episodes constituted those parts of interaction where learners discussed the order of their sentences or sentence parts and identified the preceding or following parts of the sentences in question. Although most of the activities did not involve learners in connecting text elements and discussion at discourse level, a small number of discourse episodes were observed in the dictogloss task.

Fotune and Thorp (2001) proposed a fourfold classification for the nature of LREs: continuous, discontinuous, embedded and overlapping. They defined continuous episode as an episode in which 'learners discuss a language form and conclude the discussion without returning to the form later'; in fact, the episode 'remains on the same language point without any other obvious focus' (p.155)'. On the other hand, in discontinuous episodes, the learners 'leave the point and return to it later, sometimes more than once' (ibid: 155). Some of the episodes are also embedded within other episodes. The embedded episode 'is necessarily preceded and followed by a discontinuous one' (ibid: 156). The final category, within the nature of LREs, is overlapping episode in which two or more episodes overlap, that is, within one exchange two points are discussed (ibid: 157). The preliminary analysis of the data showed frequent examples of these sub-categories; therefore, this categorization was also incorporated into the framework.

The final feature of the LREs is the outcome, which has been considered in a number of previous studies (Leeser, 2004; Malmqvist, 2005; Storch, 1998, 2007; Swain, 1998). Based on this feature, LREs were categorized into three types of correctly solved, incorrectly solved and unresolved episodes. As the terms suggest, in correctly solved episodes, the problem is solved correctly by the two learners, but in incorrectly solved episodes, the problem is solved incorrectly. Unresolved episodes constitute those LREs where the problem is left unresolved 'either because the topic of their discussion is dropped or because the pair could not reach a joint decision' (Leeser, 2004).

After coding the transcribed data based on the established framework, a sample of LREs was submitted to an inter-rater reliability test. The sample consisted of two continuous extracts from two different pairs' interaction. These two extracts together with the complete LRE framework accompanied by examples, were given to two raters, who were already trained in this regard. Since the measurement was categorical and the raters checked which category each LRE falls in, instead of calculating correlation, the percent of agreement between the raters was obtained. The nature of episodes showed the highest agreement percentage (87.5%) followed by the types (80%), and outcome (70.58%) of episodes.

5. Results

5.1 Number of language-related episodes

The first research question addressed the number of LREs in the two sets of output tasks. A total of 1348 LREs were identified in the transcripts of the learners obtained from 28 hours of tape-recorded conversation. Table 1 shows the total number of LREs, their percentage, means and standard deviations in each activity.

Findings from the quantification of LREs indicate that the total number of LREs produced by the Meaningful group was higher than that produced by the Mechanical group. Furthermore, the mean number of total LREs for

the former group (86.2) is more than that for the latter group (63.5). To determine whether the Meaningful activities prompted more LRE production, a t-test analysis was carried out. The result of this analysis showed a significant difference between the two groups (p<.05) indicating that the Meaningful output activities produced more instances of LREs compared to the Mechanical output activities (t=1.95, *df* =16, p=.03).We can further examine the distribution of LREs produced within the two sets of output activities. As the pie chart in Figure 1 shows, the picture description task and the let's complain in the Meaningful output seem to be more successful than the dictogloss in focusing the learners' attention on linguistic features. The two activities had approximately produced similar number of episodes, slightly more than twice the number produced in the dictogloss.

The low number of LREs associated with the dictogloss may be attributed to the nature of this activity. It seems that the learners' discussion of linguistic features in this activity is affected by the degree of their access to input. Unlike the other activities, which were abundant in input and were delivered in written form, the dictogloss was delivered orally in one minute, whereby the learners had no control over the speed at which the information was presented (Lynch, 1996). Thus, the low number of LREs in this activity can be accounted for by the learners' limited access to input, which may have inhibited them in extending their discussions of the linguistic features. While the distribution of episodes in the Meaningful activities differed, all the Mechanical activities stimulated similar number of LREs. As figure 2 shows, the three activities in the Mechanical group had produced almost similar proportions of LREs.

5.2 Types of language-related episode

The second research question addressed the linguistic focus of LREs in the two output groups: 'Do learners in the two output groups predominantly focus on grammatical or meaning-based features of the language?'. In order to answer this question, a comparison should be made between the groups in terms of the amount of attention they generated to linguistic features. Based on the framework, the major categories included in the types of LRE were *grammatical*, *meaning-based*, *orthographic*, *identification* and *discourse*. Tables 2 and 3 and Figures 3 and 4 show the types of LREs produced in each activity by both the Mechanical and Meaningful groups.

As can be seen, grammatical episodes were produced more frequently than other episodes by the Mechanical group. While the meaning-based episodes constituted a small portion in this group's episodes, their proportion was practically the same as grammatical episodes in the three activities of the Meaningful group. In other words, learners in the Meaningful group tended to pay attention to both meaning-based and grammatical features of their language. Orthographic features attracted the learners' attention in both groups, whereas identification episodes were produced only by the Meaningful group in the dictogloss task. This is not surprising, since the latter category was introduced solely to encode the discussion about the identification of words on the tape. Finally, discourse episodes constituted a marginally smaller proportion compared to other episodes in the Meaningful group. This suggests that the learners had dealt with discourse requirements such as linking their sentences together in the dictogloss.

Overall, a relatively good spread of attention on various linguistic features can be observed in the Meaningful group, whereas in the Mechanical group, it was the grammatical features that captured the learners' attention much more frequently than any other episode. Tables 2 and 3 further demonstrate that the Meaningful activities (with the exception of the dictogloss) were stronger than the Mechanical activities in generating LREs. By comparing the number of the grammatical LREs in the five activities, it is revealed that the picture description and let's complain not only produced almost equal number of grammatical LREs as the Mechanical activities (see the range of grammatical LREs in these activities), but also almost equally generated meaning-based LREs within the same range (141-150). The trend, however, was different for the dictogloss task. Not only the number of LREs was less than that in other activities but also the meaning-based episodes were produced much more frequently than the grammatical episodes (three times more).

To find out whether the two output groups differed in the mean number of grammatical and meaning-based LREs, two independent samples t-test analyses were carried out. Since no instances of identification and discourse LREs were identified in the Mechanical group's discussion and the orthographic episodes seemed to have similar means between the two groups, these three categories were excluded from the analysis. The result of this analysis, presented in Table 4, shows that the two groups differed in the mean number of grammatical and meaning-based LREs (p<.05). This means that the grammatical LREs were produced more frequently in the Mechanical activities and the meaning-based LREs were produced more frequently in the Meaningful activities. The result confirms the speculation that the Meaningful output activities invoke more discussions on the form-meaning relationship of the target language.

5.3 Nature of language-related episodes

Another feature of LREs involved the nature of episodes. As discussed earlier, this category was divided into continuous, discontinuous, embedded and overlapping episodes. Table 5 presents the nature of episodes in the two sets of activities. What is apparent from this table is that the continuous episodes were produced far more

frequently than any other type of episode ─ discontinuous, embedded and overlapping episodes. By comparing the two groups across the four categories, it seems that their mean scores are different in the continuous and discontinuous episodes.

To determine whether these differences are significant, four independent samples t-tests were carried out. The summary of these analyses, presented in Table 6, shows a significant difference in the continuous LREs between the two groups, but the mean number of other subcategories does not seem to be different (p<.05). The higher occurrence of the continuous episodes in the Meaningful output activities might be due to the challenging context of the production in which the learners might have been encouraged to use focused attention and constantly engage with the task to solve the problem all at once.

5.4 Outcome of language-related episodes

The third research question also addressed the outcome of problems the learners encountered during the completion of the activities. As mentioned before, the outcome of LREs was categorized as correct, incorrect and unresolved episodes. Table 7 shows that the majority of the problems were solved correctly by the learners in both groups.

The comparison of the number and percentage of these subcategories between the two output groups shows that the two groups are very similar in these features. To determine whether there is a significant difference in the mean number of outcome episodes between the two output groups, three independent samples t-tests were carried out. The results, summarized in Table 8, reveal that the two groups significantly differed in the correctly solved episodes, with the Meaningful group solving more episodes correctly than the Mechanical group. No significant difference was found in the mean number of incorrectly solved and unresolved episodes between the two groups.

6. Discussion

The purpose of this study was to explore the effects of Mechanical and Meaningful output on learner interaction through the examination of language related-episodes. The first research question addressed how different output activities affect the occurrence of language-related episodes (LREs). Data from the learners' transcripts revealed that the Meaningful output activities stimulated more discussions than the Mechanical output activities. More briefly, the results indicated that the Meaningful output generated more LREs (58% of total LREs) compared to the Mechanical output (42%). The second research question addressed the focus of LREs in the output activities. The comparison of the grammatical and meaning-based episodes produced in the two sets of activities showed that the Meaningful activities not only generated a large number of grammatical episodes (n=325), but also frequently produced meaning-based episodes (n=358), which were strikingly higher than those produced in the Mechanical activities (n=81). While the Mechanical group talked about grammar in 77% and meaning (and lexis) in 14% of their episodes, the Meaningful group discussed grammar in 42% and meaning in 46% of their interactions. These findings suggest that, unlike the Meaningful activities, which involved a balanced focus of attention on grammar and meaning (except in the dictogloss), the Mechanical activities were predominantly focused on grammar. Therefore, the former activities promoted attention to the form-meaning relationship by engaging learners in processing the meaning of sentences more frequently than did the latter activities. Detailed analysis of the types of LRE in each activity showed that while in the substitution drill, the grammatical episodes (n=152) took place more frequently than the meaning-based episodes (n=24), in the dictogloss, the meaning-based episodes (n= 67) were produced more frequently than the grammatical episodes (n=22). It should be noted that the design of the research material anticipated that the activities at the less controlled end of the continuum would allow for more meaningful processing of the language. However, this meaningful processing rarely involved the target linguistic forms (i.e. relative clauses). We observed that out of 149 LREs in the dictogloss, only three LREs were focused on the relative clause structure. This apparently gives support to many SLA researchers' argument (e.g., Slimani, 1992; Swain, 2000) that 'learners appear to have their own agendas for which aspects of the language they decide to focus on at any given time. The agenda does not necessarily coincide with the intent of the instructor's' (Lantolf & Thorne, 2007, p. 206).

Analysis of the nature of episodes revealed that the Meaningful group produced more continuous episodes than did the Mechanical group. This means that, in most of the times, when learners encountered a linguistic problem and started discussion on a point, they did not give up the discussion until they solved the problem. One explanation could be that the Meaningful output activities involved focused attention due to their challenging nature and required learners to solve the problems all at once, otherwise their scattered attention, which might be associated with the discontinuity of the episodes, would not allow them to solve the problem correctly. As regards the outcome of LREs, the majority of episodes in both groups were resolved correctly on most occasions. That is, when learners encountered a linguistic problem during their discussions, they solved it correctly. The incorrectly solved and unresolved episodes less frequently occurred in both groups' interactions. This result is consistent with the findings of previous studies (Leeser, 2004; Malmqvist, 2005; Storch, 2007). The comparison of the outcome of episodes across the two output groups showed that although both groups solved most of their problems correctly,

the Meaningful group solved more LREs correctly than did the Mechanical group. One explanation could be that the learners in this group provided incorrect grammatical explanations during their discussions. Since they discussed grammatical features more than meaning-based features, they extended their incorrect metalinguistic knowledge and produced more incorrectly solved episodes than the Meaningful group.

Finally, it is worth mentioning that although the results of t-test analyses showed significant differences in some LRE features between the two groups, such results need to be interpreted cautiously. To answer the research questions, ten comparisons were carried out on the same set of data. It should be noted that there are possible dangers (type I error or false positive) in running multiple t-tests on the same set of data. That is, we may have observed a statistical difference between the two groups, when in truth, there is no difference. This suggests that the findings reported so far may not be definitive. One way of avoiding this problem is to adjust the significance level, for example, through Bonferroni approximation. However, adjusting the level of significance through this method would be too conservative and it may lead to underclaiming the number of significances that may truly exist between the two groups. Therefore, instead of applying this method for the present data, the frequencies of LRE features were compared through four separate Chi-square statistical tests, the results of which were consistent with the t-test analyses carried out so far. To confirm these findings, it is recommended that future studies address all these features across various output activities.

7. Conclusion

The activities examined in the present study provided various opportunities for discussion on a wide range of linguistic features. We observed that the Meaningful activities generated more episodes than the Mechanical activities. While the Mechanical pairs predominantly focused on grammar, the Meaningful pairs discussed grammar *and* meaning. Furthermore, in addition to the structure of English relative clauses, they generated discussions on various linguistic features. Therefore, these activities can be seen as more economical than the Mechanical exercises since various linguistic areas can be targeted by using a single activity. The analysis also revealed that the majority of the problems encountered during the pair interaction were resolved correctly. This may alleviate the concern of some SLA professionals about the incorrect provision of feedback during collaborative interaction. However, it is not clear whether individuals stick to the knowledge collaboratively constructed during pair work activities and transfer it to subsequent learning situation. Therefore, there is an urgent need for further research investigating individual learner's performance on linguistic features discussed during collaborative dialogue. In particular, it needs to be empirically supported whether or not engagement in a continuous episode which is correctly solved involves a deeper and sustained learning. Since the present study did not establish a strict laboratory setting such as control over exposure time to input, it may realistically reflect how interaction and focus on form occur in a classroom setting.

References

Benson, C., Pavitt, J. & Jenkins, M. (2005). The use of dictogloss to encourage discussion of language use. *Edinburgh Working Papers*, 14, 1–17.

Chalker, S. (1987). *Intermediate grammar workbook 2*. London: Longman.

Dakin, J. (1973). *The language laboratory and language learning*. London: Longman.

Donato, R. (2004). Aspects of collaboration in pedagogical discourse. *Annual Review of Applied Linguistics*, 24, 284–302. doi:10.1017/S026719050400011X, http://dx.doi.org/10.1017/S026719050400011X

Doughty, C., & Pica, T. (1986). Information gap tasks: Do they facilitate second language acquisition. *TESOL Quarterly*, 20(2), 305–325. doi:10.2307/3586546, http://dx.doi.org/10.2307/3586546

Duff, P. (1986). Another look at interlanguage talk: Taking task to task. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 147–181). Rowley, MA: Newbury House.

Fortune, A., & Thorp, D. (2001). Knotted and entangled: New light on the identification, classification and value of language related episodes in collaborative output tasks. *Language Awareness*, 10(2&3), 143–60. doi:10.1080/09658410108667031, http://dx.doi.org/10.1080/09658410108667031

Hutchinson, T. (1992). *Hotline: Pre-intermediate*. Oxford: Oxford University Press.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis. *Studies in Second Language Acquisition*, 24, 542-577. doi:10.1017/S0272263102004023, http://dx.doi.org/10.1017/S0272263102004023

Izumi, S. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, 24(2), 168–196. doi:10.1093/applin/24.2.168, http://dx.doi.org/10.1093/applin/24.2.168

Izumi, Y., & Izumi, S. (2004). Investigating the effects of oral output on the learning of relative clauses in English: Issues in the psycholinguistic requirements for effective output tasks. *Canadian Modern Language Review*, 60(5), 587–609. doi:10.3138/cmlr.60.5.587, http://dx.doi.org/10.3138/cmlr.60.5.587

Kowal, M., & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness*, 3(2), 73–93. doi:10.1080/09658416.1994.9959845, http://dx.doi.org/10.1080/09658416.1994.9959845

Lantolf, J. P. (2000). Introducing sociocultural theory. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 1–26). Oxford: Oxford University Press.

Lantolf, J. P. (2006). Sociocultural theory and L2: State of the art. *Studies in Second Language Acquisition*, 28, 67–109. doi:10.1017/S0272263106060037, http://dx.doi.org/10.1017/S0272263106060037

Lantolf, J.P. (2007). Sociocultural source of thinking and its relevance for second language acquisition. *Bilingualism: Language and Cognition*, 10(1), 31–33. doi:10.1017/S1366728906002768, http://dx.doi.org/10.1017/S1366728906002768

Lantolf, J. P., & Thorne, S. L. (2007). Sociocultural theory and second language learning. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 197–221). Mahwah, NJ.: Lawrence Erlbaum.

Lapkin, S., Swain, M. & Smith, M. (2002). Reformulation and the learning of French pronominal verbs in a Canadian French immersion context. *The Modern Language Journal*, 86 (4), 485–507. doi:10.1111/1540-4781.00157, http://dx.doi.org/10.1111/1540-4781.00157

Leeser, M. J. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research,* 8, 55–81. doi:10.1191/1362168804lr134oa, http://dx.doi.org/10.1191/1362168804lr134oa

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, Mass: MIT Press.

Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crooks & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 123–167). Clevedon: Multilingual Matters.

Lynch, T. (1996). *Communication in the language classroom*. Oxford: Oxford University Press.

Malmqvist, A. (2005). How does group discussion in reconstruction tasks affect written language output? *Language Awareness,* 14(2&3), 128–141. doi:10.1080/09658410508668829, http://dx.doi.org/10.1080/09658410508668829

Nabei, T. (1996). Dictogloss: Is it an effective language learning task? *Working Papers in Educational Linguistics, 12*(1), 59-74.

Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language instruction. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 9–34). Clevedon: Multilingual Matters.

Seidl, J. (1993). *Grammar three*. Oxford: Oxford University Press.

Slimani, A. (1992). Evaluation of classroom interaction. In J.C. Alderson & A. Beretta (Eds.), *Evaluating second language education* (pp. 197–220). Cambridge: Cambridge University Press.

Storch, N. (1998). Comparing second language learners' attention to form across tasks. *Language Awareness*, 7 (4), 176–191. doi:10.1080/09658419808667108, http://dx.doi.org/10.1080/09658419808667108

Storch, N. (2007). Investigating the merits of pair work on a text editing task in ESL classes. *Language Teaching Research*, 11(2), 143–159. doi:10.1177/1362168807074600, http://dx.doi.org/10.1177/1362168807074600

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, Mass.: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics; Studies in honour of William E. Rutherford* (pp. 125–144). Oxford: Oxford University Press.

Swain, M. (1997). The output hypothesis, focus on form and second language learning. In V. Berry, B. Adamson & W. Littlewood. *Applying linguistics: Insights into language in education* (pp. 1–21). Hong Kong: The University of Hong Kong, The English Centre.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition (pp.* 64–81). Cambridge: Cambridge University Press.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 99–116). Oxford: Oxford University Press.

Swain, M. & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, 82, 320–337. doi:10.2307/329959, http://dx.doi.org/10.2307/329959

Swain, M. & Lapkin, S. (2000). Task-based second language learning: The uses of the first language. *Language Teaching Research*, 4, 251-274. http://dx.doi.org/10.1177%2F136216880000400304, http://dx.doi.org/10.1177%2F136216880000400304

Swain, M. & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 99-118). London: Longman.

Swain, M., Brooks, L., & Tocalli-Beller, A. (2002). Peer-peer dialogue as a means of second language learning. *Annual Review of Applied Linguistics*, 22, 171–185. doi:10.1017/S0267190502000090, http://dx.doi.org/10.1017/S0267190502000090

Ur, P. (1988). *Grammar practice activities: A practical guide for teachers*. Cambridge: Cambridge University Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wajnryb, R. (1990). *Grammar Dictation*. Oxford: Oxford University Press.

Table 1. LREs produced in the six activities by the two groups

Mechanical activities (9 pairs of learners)	Total No. of LREs	% LRE	Mean	SD
Substitution	197	34.5 %	21.88	10.39
Transformation	166	29.0 %	18.44	6.57
Fill in the blank	209	36.5 %	23.22	10.56
Total	572 (42%)	100 %	63.5	27.52
Meaningful activities (9 pairs of learners)	Total No. of LREs	% LRE	Mean	SD
Picture description	322	41.5 %	35.77	12.33
'Let's complain'	305	39.3 %	33.88	11.35
Dictogloss	149	19.2 %	16.55	6.48
Total	776 (58%)	100 %	86.2	30.16

Table 2. Types of LRE in each activity: Mechanical group

Mechanical Group (9 pairs)	Substitution No. of LREs	Transformation No. of LREs	Fill in the blank No. of LREs	Total No. of LREs	% LRE	Mean	SD
Grammatical	152	141	149	442	77.3 %	49.11	16.88
Meaning-based	24	9	48	81	14.2 %	9.00	6.28
Orthographic	21	16	12	49	8.5 %	5.44	4.12
Total	197	166	209	572	100%	63.5	27.2

Table 3. Types of LRE in each activity: Meaningful group

Meaningful group: (9 pairs)	Picture-description No. of LREs	Let's complain No. of LREs	dictogloss No. of LREs	Total No. of LREs	%LRE	Mean	SD
Grammatical	155	148	22	325	41.9%	36.11	11.8
Meaning-based	150	141	67	358	46.1%	39.77	14.2
Orthographic	17	16	4	37	4.8%	4.11	2.52
Identification	0	0	48	48	6.2%	5.33	2.8
Discourse	0	0	8	8	1.0%	0.88	1.69
Total	322	305	149	776	100%	86.2	33.0

Table 4. Summary of the results of between group comparisons on the LRE types

LREs	*df*	Mean (Mechanical)	Mean (Meaningful)	t-value	Sig.
Grammatical	16	49.11	36.11	1.893	.038*
Meaning-based	16	9.00	39.77	5.945	.000*

Note: The mean difference is significant at the .05 level.

Table 5. Nature of LREs in the two output groups

Continuous	Mechanical group	Meaningful group
No. of LRES	499	670
% LRE	87.2 %	86.4 %
Mean	55.44	74.44
SD	17.19	19.90
Discontinuous		
No. of LRES	33	65
% LRE	5.8%	8.4%
Mean	3.66	7.22
SD	3.67	5.71
Embedded		
No. of LRES	24	35
% LRE	4.2%	4.5%
Mean	2.66	3.88
SD	3.00	4.59
Overlapping		
No. of LRES	16	6
% LRE	2.8%	0.7%
Mean	1.77	0.66
SD	2.53	1.4

Table 6. Summary of the results of between group comparisons on the nature of LREs

LREs	df	Mean (Mechanical)	Mean (Meaningful)	t-value	Sig.
Continuous	16	55.44	74.44	2.167	.023*
Discontinuous	16	3.66	7.22	1.569	.068
Embedded	16	2.66	3.88	.668	.257
Overlapping	16	1.77	0.66	1.147	.134

Note: The mean difference is significant at the .05 level.

Table 7. outcome of LREs in the two output groups

Correctly solved	Mechanical group	Meaningful group
No. of LRES	445	592
% LRE	77.8 %	76.3%
Mean	49.44	65.77
SD	16.18	19.49
Incorrectly solved		
No. of LRES	92	118
% LRE	16.1%	15.2%
Mean	10.22	13.11
SD	8.10	7.80
Unresolved		
No. of LRES	35	66
% LRE	6.1%	8.5%
Mean	3.88	7.33
SD	3.33	6.59

Table 8. Summary of the results of between group comparisons on the LRE outcome

LREs	df	Mean Mechanical	Mean Meaningful	t-value	Sig. (1-tailed)
Correctly solved	16	49.44	65.77	1.921	.036*
Incorrectly solved	16	10.22	13.11	.770	.226
Unresolved	16	3.88	7.33	1.398	.090

*. The mean difference is significant at the .05 level.
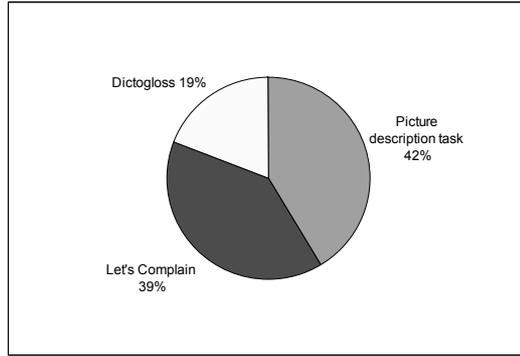
Figure 1. Distribution of LREs in the Meaningful activities
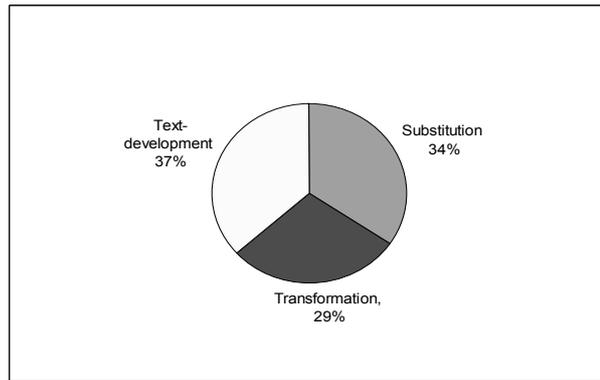

Figure 2. Distribution of LREs in the Mechanical activities
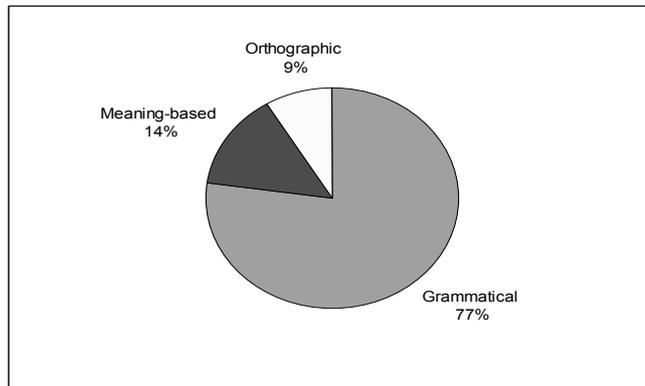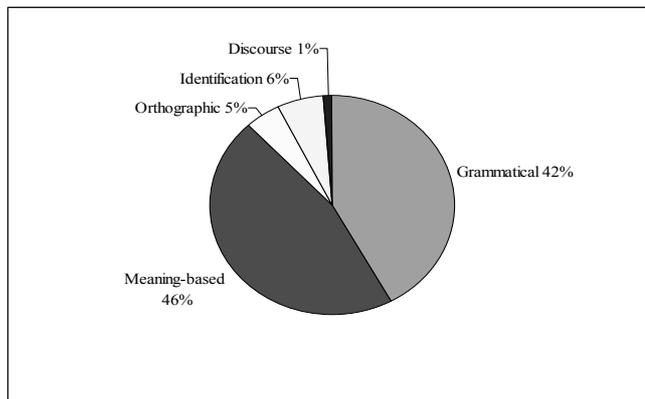

Figure 3. Types of LRE in the Mechanical activities

Figure 4. Types of LRE in the Meaningful activities