Benefits of mobile instant messaging to develop ESL writing

Alberto Andujar

Department of Linguistics, University of Almería, Spain

Abstract

This study investigated the benefits of Mobile Instant Messaging (MIM) through an analysis of grammatical, lexical and mechanical accuracy as well as syntactic complexity in second-language learners’ writing. A WhatsApp group was created where 80 Spanish students taking a B1 English course participated in a daily interaction during six months. A quasi-experimental research design with an experimental and control group and a pre-post test was followed. Students were divided into two main groups according to treatment type with 40 students in each group. This research focused on the interaction in the application and attempted to measure, through a qualitative and quantitative analysis, the students’ degree of writing development. The ratios of lexical, grammatical and mechanical errors as well as error-free clauses per clause and error-free T-unit per T-unit indicated significant differences between the control and experimental group in terms of accuracy. Nevertheless, measures of syntactic complexity together with lexical diversity were not conclusive as the independent parameters for syntactic complexity showed no significant differences between the two groups. WhatsApp constitutes a powerful educational tool to encourage second language interaction among participants and its tremendous potential to activate students’ involvement remains one of the least exploited functionalities of mobile phones.

1. Introduction

Mobile learning breaks temporal and spatial lines, placing educative resources at the service of the students. Through mobile phones, Internet users now have the possibility of accessing the Internet from different points, having access to the net from almost any place. However, although its importance as a research field is increasing rapidly, as reflected in the numerous research articles, conferences, seminars and workshops taking place all over the world (see Gülbahar, Jacobs, & König, 2015), several reviews (see Burston, 2014; Cobcroft, Towers, Smith, & Bruns, 2006; Trifonova, 2003) indicate that MALL remains marginal, as experiments and class trials carried out lack follow-up reports showing curricular integration.

This study is based mainly on the need for extending research related to the use of Mobile Instant Messaging (MIM) in the English as a Second Language (ESL) classroom and its benefits. MIM is an asynchronous and synchronous communication tool that uses wireless networks and handheld devices, allowing students to hold conversations and exchanges (Dourando, Parker, & De la Harpe, 2007; Rambe & Bere, 2013). This type of messaging is characterized by instant delivery through a pop-up mechanism, a list of users, and a mechanism to indicate when they are available. At the same time, it allows access to diverse content, acquisition of specific data and information sharing among users. MIM constitutes one of the functionalities which holds great potential in a mobile phone; typed messages have evolved to become a new hybrid of spoken, written and

E-mail address: aav447@ual.es.

http://dx.doi.org/10.1016/j.system.2016.07.004
0346-251X/© 2016 Elsevier Ltd. All rights reserved.
electronic chat discourse (Blake, 2000; Muniandy, 2002; Warschauer, 1996; Werry, 1996). Through this technology, students find an environment where their interactions facilitate acquisition of second language communication skills (Blake, 2000; Bonk & Kim, 1998).

Despite its tremendous potential to activate students’ involvement, MIM remains one of the least exploited functionalities of mobile phones. This could be explained by several reasons such as (1) the difficulty and cumbersomeness of typing on a small mobile screen, as well as the screen resolution itself (Alexander, 2004; Shudong & Higgins, 2006), (2) the distractive nature of text messages, (3) limited academic conceptualization of how text messages can be integrated in ordinary instruction, and (4) lack of certainty that the comments generated via text message have real academic rigour (Rambe & Bere, 2013). Regardless of these hypotheses, this social practice allows contact with new information, builds social networks, supports the sharing of ideas and promotes mutual understanding and learning in informal contexts. Due to the inherent characteristics of mobile devices, this practice allows synchronous and asynchronous interaction virtually at any time and place.

Focusing on these interactions, this study attempted to measure, through a qualitative and quantitative analysis, the students’ degree of writing development. In particular, this research explored the pedagogical value of WhatsApp, its assets and its academic potential.

2. Literature review

In order to contextualize this research, it is necessary to take into consideration notions such as communicative competence, as well as which communication skills this study focuses on. Subsequently, we will explain the theoretical models that underlie the research.

From a functionalist perspective, communicative competence is defined as ‘an instrument of communication whose properties can be studied as long as they are instruments for communication’, as stated by Hymes (1972). Other authors like Canale (1983) and Canale and Swain (1980), point out that communicative competence is the combination of knowledge and skills (linguistic, sociocultural, discursive and strategic) which allow an effective performance of the speaker, suitable for a particular communicative situation and a specific speaking community.

Following the integrated approach espoused by Oxford (2001), four primary skills used to reach communicative competence can be distinguished in the EFL/ESL classroom: reading, writing, listening and speaking. We will take into consideration those belonging to the written medium (writing and reading), paying special attention to the writing skill.

2.1. Framework and Vygotsky’s zone of proximal development

The Framework for the Rational Analysis of Mobile Education (FRAME) (Koole, 2009) describes ‘mobile learning’ as a process resulting from mobile technologies’ convergence with human capacities of learning and social interaction. This model takes into consideration the technical characteristics of mobile devices as well as social and personal aspects of learning (Koole, 2009). The theory underlying the Framework echoes Vygotsky’s (1978) theorization about mediation and zone of proximal development. Vygotsky’s social development theory stresses the fundamental role of social interaction in the development of cognition (Vygotsky, 1978), as he deeply believed that community plays a central role in the process of ‘making meaning.’ Therefore, in this sociocultural approach to cognitive development, individual development cannot be understood without reference to the social and cultural context within which it is embedded (Lantolf & Thorne, 2006). Hence, higher mental processes taking place in the individual have their origin in social processes. This theory directly affects our study, as the students are part of a social activity in which interaction plays a key role.

FRAME (Koole, 2009) underlines the role of technology and, at the same time, gives an important emphasis to constructivism. In this model of learning, students can use different physical and virtual situations, interacting with other people, information or systems at any time and place. Learning experiences take place in an information context where interaction is mediated through technology.

2.2. Mobile text chat and interaction

Swain (1985) regarded interaction as a key element in second language development since interaction constitutes a fundamental factor where students construct knowledge in an active manner. This constructivist approach is based on building active meaning, focusing the process on primary learning concepts and not on isolated facts.

Due to their real-time nature, text chats and consequently MIM resemble face-to-face interaction, and thus may carry language development benefits such as repair moves or negotiation for meaning (Smith, 2003; Sotillo, 2000; Warschauer, 1996). Repair moves are more frequent due to the fact that participants can send messages simultaneously regarding unrelated or abandoned topics, hence creating a discourse sequence different from face-to-face interaction where they speak in turn on a single topic (Doughty & Long, 2003; Negretti, 1999). These messages appear within seconds on their interlocutors’ screens and each textual turn contains the learner’s language and typographical errors (Jepson, 2005). The negotiation of meaning that is produced when students suffer or perceive problems with communication is defined as a ‘process in which a listener requires message clarification and confirmation and a speaker follows these requests, often through repeating,
elaborating or simplifying the original message’ (Pica, 1994, p.497). Their discourse is modified in order to reach an understanding, turning into comprehensible input that will rise to intake acquired by the students.

Research on task-based computer-mediated text chats has shown that about one third of the total turns taken by English learners are related to negotiation, and it engenders SLA among NNs (Non-native speakers) (Chen, Belkada, & Okamoto, 2004; Pellettieri, 1996, 2000; Smith, 2003, 2004). In studies such as Payne and Whitney (2002) about task-based chats, Spanish L2 participants of an experimental teacher-led group demonstrated greater improvement in their oral proficiency than those in the control group. Furthermore, as several studies show, text chats seem to allow participants from different levels to help one another in co-constructing social activity (Ortega, 1997; St. John & Cash, 1995; Warner, 2004) as well as knowledge, following the cooperative learning theory stated by Slavin (1983). Conversely, although these studies address task-based teacher-led chats which may seem similar to MIM, some distinctive features of MIM can be addressed: (1) synchronous and asynchronous interaction, as opposed to CMC where only asynchronous interaction is carried out, (2) easy access to content and possibility of seeing online users, (3) wider range of chat features such as negotiation of meaning, repair moves, comprehensions and confirmation checks, clarification requests and self-repetition, (4) greater ease in media sharing: voice, images and text, and (5) availability to use Internet resources.

Furthermore, students do not use technology they consider intrusive or if they need time to learn how to use new tools and software (Kukulksa-Hulme & Shield, 2007). In contrast, WhatsApp is an application that many students have already mastered, which is a key advantage to avoid these drawbacks previously mentioned. Some of the assets of this application are: promoting contact between students and teachers; fostering interactions among students and promoting academic cooperation; encouraging active learning; providing instant feedback; developing high communicative expectations (Desai & Graves, 2006; Dourando et al., 2007; Farmer, 2003; Rambe & Bere, 2013).

Research in the field of MIM is very scarce. Nevertheless, some studies, such as Levy and Kennedy (2005), have explored the possibilities of SMS messages in mobile devices. In their study the instructor sent messages at specific times and days, not taking into consideration that students could answer anytime and anywhere. Furthermore, Vázquez-Canó et al. (2015) found that in MIM the amount of shortened words was reduced, as opposed to SMS where students did not have a word limit. Their research investigated the linguistic and paralinguistic characteristics of MIM in Spanish teenagers and emphasized a suprasegmental enrichment due to the use of emoticons and stickers. Other studies such as Diaz (2014), Morató Payá (2014) or Padron (2013) focused on the potential that WhatsApp offers to promote collaborative learning but failed to demonstrate linguistic gain.

2.3. Noticing and collaborative learning

Regarding the kind of learning that is produced through MIM, DeKeyser and Larson-Hall (2005) describe implicit learning as that which is produced without any awareness that such learning is taking place. This type of learning is characterized as being unintentional, unconscious and, on some occasions, involving consciousness, but only at the level of perception. This perception is considered by Schmidt (1995) to be necessary for the transformation of input into intake. Noticing is a mechanism that mediates between communication and acquisition (Gass, 2003), where students notice the gap between their current language and their target language, appreciating the feedback given by the teacher. In an output-input cycle, learners are pushed in their production and can compare it with the input subsequently received from other students as well as the teacher in the chat group. This comparison aids learners in seeking relevant linguistic forms in order to convey meaning more accurately (Swain & Lapkin, 1995; Swain, 1998; Thornbury, 1997). According to Hanaoka and Izumi (2012), noticing engages learners’ cognitive processing and they become aware of the possibilities as well as problems in their IL (Interlanguage) capabilities and development, which will alter the subsequent input processing. Throughout the interaction, students explicitly focused their attention on language, questioning their language use or correcting themselves or others, which Swain and Lapkin (1998) called language-related episodes (LREs). In LREs attention is drawn towards formal features of language, which leads to L2 acquisition (Izumi & Bigelow, 2000; Izumi, 2002; Izumi, Bigelow, Fujiwara, & Fearnow, 1999). Hypotheses are created by the learners who experience a cognitive process of making meaning, and shaping knowledge and experience through reflecting on language, a process called ‘linguaging’ (Swain, 2006), which becomes a source of L2 learning. According to Dobao (2012), novice learners construct knowledge in collaboration with more capable individuals, in this case, the teacher. Other studies have pointed out the benefits of learner-learner interaction where participants act as both novice and expert (Ohta, 2000, 2001; Storch, 2002; Swain & Lapkin, 1998).

2.4. Research questions

This study seeks to answer the following questions: (1) has the programme implementation in the experimental group involved important effects in the development of accuracy as well as lexical and syntactic complexity of the subjects? (2) Is the use of Mobile Instant Messaging profitable for this development? Regarding this last question, the study tries to answer aspects in relation to the potential value of the application in ESL writing as well as how this MIM provides opportunities for interaction and, consequently, second language development.
3. Methodology

3.1. Research design

The purpose of this exploratory research was to investigate writing development in ESL learners, through analysing the interactions in a mobile text-based teacher-led chat. Regarding the research design, a quasi-experimental design with an experimental and control group and a pre-post test was followed (Cambell & Stanley, 1963). The experimental group underwent the prescribed treatment while the control group received no treatment at all and served as the benchmarking point of comparison. Naturalistic observation was used as the researcher looked at the behaviour and interaction of students in a natural setting with no attempts at intervention (Nunan, 1992).

Null and alternative hypothesis were posed and tested in order to reveal significant differences between two pre-existing groups (control and experimental group).

- Alternative hypotheses: The level of accuracy as well as lexical diversity together with syntactic complexity of the students in the treatment group will be higher than those in the control group.
- Null hypothesis: No significant differences will be found between in terms of accuracy, lexical diversity and syntactic complexity in the two groups.

Participant selection and initial differences were determined by the tests given at the beginning of the course to both groups. Internal threats to validity (Campbell & Stanley, 1963) and effects such as the Hawthorne effect (Mayo, 1933), the halo effect (Thorndike, 1911) and participant expectancy were minimized as the students were not conscious that they were being observed and the teacher did not participate in the correction of the tests given. An external rater was used in this process. The two groups did not have the same timetable, the experimental one met in the mornings and the control in the evenings, thus students did not know the different activities carried out in both groups as they did not know each other.

3.2. Setting and participants

This study involved four groups of third year students from the University of Almeria that were taking a B1 English course. The materials and contents provided in the course met the parameters established by the Common European Framework of Reference for Languages (CEFR) for a B1 level. The expected entry level was lower than B1 as students who enrolled on the course aimed to achieve such level by the end of the course. Enrolment was determined by a A2+ placement test, which was administered to the students at the beginning of the course in order to ensure they were able to take it. In case participants did not pass the test, they were not allowed to enrol on the course. The length of this course was six months, from November to May, and met for four and a half hours per week. The number of students was 80, 61 female and 19 male students, all of them with the same nationality and ages ranging from 20 to 26. Students were divided into two main groups according to treatment type with 40 students each. Intact classes were arbitrarily assigned as experimental and control groups. The researcher participated actively in the process, as he was the teacher of the groups studied. Although the control group did not receive any treatment, it was used for studying the differences between groups after the implementation. Pre-testing was undertaken to establish any differences between the control and experimental group, discussed below. In this manner, we were able to measure the effect of the treatment on experimental group subjects. After applying this treatment or implementation in the programme of the experimental group, we verified its effect through the comparison of the results obtained. The identities of the participants have been changed so as to maintain confidentiality in this research, meeting ethical guidelines set by the American Psychological Association (2002).

In both groups, during formal learning in class, the amount of opportunities to speak, as well as to be in contact with other students was very limited due to 1) lack of time, 2) psychological distance between teacher and student, and 3) students’ lack of confidence to ask questions in public.

3.3. Study

This study focuses on determining the potential of MIM services like WhatsApp in order to develop ESL writing. The WhatsApp application is presented as a tool to encourage participation and interaction during the course. WhatsApp is an instant messaging platform that uses the Internet to connect different mobile phone users, allowing the creation of groups in which different users can interact. Exchanges between users can be produced in a synchronous or asynchronous manner, and are received on phones as alerts. In this research, WhatsApp did not substitute for explanations in class, but was used as a tool to maintain communication out of class on a daily basis, becoming a constant support for language use. All the students in the experimental group participated in the activity, as they all had mobile phones with a 3G Internet connection and an available wi-fi at the University as well as the WhatsApp application previously downloaded. The expenses of 3G Internet connection were covered by the participants which may influence the accessibility of such affordances. Notwithstanding that it could be a major drawback, students made use of their own data plans when necessary as they were already using them before the beginning of the activity and in any case did not consider participation in the activity to be costly. After creating the WhatsApp
group, the activity was carried out taking into account a series of conditions established by the teacher: participation in the activity and English use was compulsory; a different question had to be formulated by a different student each day, 7 days per week; the topic of the question was free; each student had to give at least one answer per question and there was no minimum or maximum text length; sharing images and voice messages in English was allowed; the teacher participated as a student, responding to the questions. Availability of the teacher in the application was complete, and errors were not corrected in an explicit manner during interaction.

Regarding participation, a daily tracking of the minimum answer per student required was carried out in order to ensure student’s involvement in the activity. Participants sent 13,937 messages during the activity, which shows a remarkable degree of engagement. Grading was not determined by participation, hence no sanctions were imposed if they did not participate as the teacher tried to keep the activity as natural and authentic as possible in order to ensure reliability. Students thought they were just chatting but not being observed or evaluated.

3.4. Experimental methods

Due to the short-length of the samples in the application, a chronological axis was used to measure the difference between both groups. Written English tests at the beginning and end of the course were used to measure writing skill in a quantitative manner, by analysing grammatical, lexical and mechanical accuracy as well as syntactic complexity together with lexical diversity. The model used attempted to measure B1 writing in which students can write simple texts such as letters, events, experiences or impressions on topics which are familiar or of personal interest by linking a series of shorter discrete elements into a linear sequence as stated by the CEFR. This model consisted of: two texts to write which were the same for both groups with a certain number of words per text (100 words). Compositions were done in a classroom environment with a time limit of 30 min. Students of both the control and experimental groups were given these tests whose content dealt with a previous topic explained during the course. The purpose of these tests was to become a tool to measure the students’ writing level. Such tests proved to be very helpful, not only by allowing a comparison between the experimental and control groups, but also by measuring the development of experimental group students in relation to their tests at the beginning and end of the course.

Following previous research of written texts (Dobao, 2012; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009), accuracy, lexical diversity and syntactic complexity were analysed. Fluency was not taken into consideration, as the number of words in each test was set in advance (100 words). Different measures of syntactic complexity were used: number of words per clause and per T-unit; number of clauses per T-unit produced by the students. These units of measurement analyze how students increase the syntactic complexity of the texts written, turning short discrete elements into a linear sequence as required by the CEFR for a B1 level. Furthermore, thanks to the increase of such complexity, students are able to meet B1 ‘can-do’ descriptors such as the possibility of writing essays, opinions, brief reports or stories as well as detailed descriptions, feelings and experiences.

Regarding lexical diversity together with syntactic complexity (see Skehan, 2009), a mean segmental type-token ratio (MSTTR) was calculated. The written texts were divided into segments of 50 words each, which were used to calculate the MSTTR of each student (see Dobao, 2012).

With regard to linguistic accuracy, three different types of errors were calculated: grammatical errors; lexical errors; and mechanical errors.

**Grammatical errors.** Participants’ mistakes regarding grammar such as choice of verb tense, use of infinitives, subject-verb agreement or use of the auxiliary verb are considered. Words in bold indicate errors made by students.

**Example 1**

Maria: I will **to** travel to Madrid next weekend.
Ana: Madrid is my favourite place. I **usually go** when I was younger.
Maria: It’s fantastic! Best place ever!

**Lexical errors.** Mistakes made due to the lexical choice of the participants are taken into consideration.

**Example 2**

Nuria: Are you similar to your parents?
Noelia: Yes! To my father ... he is very active and easy-going but my mother is very **bored**.
Nuria: Really? Why is she boring?
Noelia: Because she is all the time sad and crying. She is very **sensible**.

**Mechanical errors.** Learners’ spelling and punctuation problems are analysed throughout the interaction.

**Example 3**

Sara: Is it **difficult** to learn English?
David: English is not difficult for me Sara. However (,) pronunciation sometimes is very different.
Sara: Yes for me to. I think Spanish is easier.

The ratios of error-free clauses to total clauses, error-free T-units to total T-units and errors to words were also calculated to measure accuracy.

3.5. Qualitative methods

Through the language samples obtained in WhatsApp use, we can observe the individual development of each student, thanks to the analysis and tracking of the LRESes taking place in the application. These LRESes arose as a result of the interaction what encouraged learners to collaborate in order to reach correct solutions (Swain & Watanabe, 2012). Conversely, this process may have ended up generating non-target like forms, hence, feedback was required in order to guide the learning process (Hanaoka & Izumi, 2012). Students participated by helping partners with difficulties when using the language, reflecting on the language use and constructing knowledge. Mistakes were corrected by students throughout the interaction as well as by the teacher who monitored the activity in the application. This correction was implicit as both students and teacher rephrased the mistakes made by other classmates. Students’ performance was not marked in the application due the short length of text messages, which made it impossible to evaluate the conversations consistently. Students’ answers, if necessary, were reformulated, and following Hanaoka and Izumi’s (2012) recommendation for lower proficiency learners, reformulated parts were emphasized helping participants’ noticing. Reformulation by the teacher was used throughout the interaction providing instant feedback to the participants and promoting learners’ spontaneous noticing (Cohen, 1983). Such reformulation addressed different texts — short mobile text messages - from those reformulations carried out in other studies (e.g., Adams, 2003; Lapkin & Swain, 2004; Lapkin, Swain, & Smith, 2002; Yang & Zhang, 2010) where the aim was proficiency. Nevertheless, such reformulation of text messages may lead to changes in learners at all levels: lexical, syntactic, grammatical, clarity of thought, etc.

4. Results and analysis

4.1. Results

The ANOVA related samples repeated measures test was conducted to confirm that the differences between pre and post-test scores were statistically significant. Alpha level of significance was set at 0.05 under a nondirectional (two-tailed) hypothesis. As shown in the pre-test results, no significant initial differences were found between groups. The improvement in accuracy scores by experimental and control group students are presented in Table 1. Significant differences were found between the experimental and control group in pre and post-test regarding error-free clauses per clause (F (1,78) = 8.766; p < 0.05) and error-free T-units per T-unit (F (1,78) = 6.485; p < 0.05). Sidak adjustment was used to counteract multiple comparisons, showing evidence of a greater improvement in experimental group students’ accuracy in the post-test: (0.19; p < 0.05) error-free clauses per clause and = + 0.17; p < 0.05) error-free T-units per T-unit.

As Table 2 demonstrates, statistically significant differences were found between experimental and control group in the independent analysis of grammatical, lexical and mechanical errors. The ratios of grammatical errors per word (F (1,78) = 41.194; p < 0.05), lexical errors per word (F (1,78) = 155.508; p < 0.05), and mechanical errors per word (F (1,78) = 115.50; p < 0.05) are significant between groups in pre and post-test. Sidak was used to compare pre-post measures in both groups, showing a significant decrease in the number of grammatical errors per word (−0.021; p < 0.05), lexical errors per word (−0.032; p < 0.05), and mechanical errors per word (−0.032; p < 0.05) of the experimental group.

The results of syntactic and lexical complexity analysis are presented in Table 3. No statistically significant differences were found in the total number of T-units between groups. Nevertheless significant differences were found in the number of T-units between the first measure (pre-test) and the second (post-test) which show that participants’ number of T-units decreased during the course (F (1,78) = 13.670; p < 0.05). An explanation for this drop in the number of T-units could be the high number of single clauses that participants used in the first measure, which was diminished in the post-test. As previously

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-post test measures of accuracy for texts written in control and experimental group.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Error-free clauses</td>
</tr>
<tr>
<td>Error-free clauses per clause</td>
</tr>
<tr>
<td>Error-free T-units</td>
</tr>
<tr>
<td>Error-free T-units per T-unit</td>
</tr>
<tr>
<td>Errors</td>
</tr>
<tr>
<td>Errors per word</td>
</tr>
</tbody>
</table>

<sup>a</sup> Post-test results in experimental and control group.

<sup>b</sup> Statistically significant difference between experimental and control group (p < .05).
mentioned, participants reduced the number of T-Units in both groups, however the quantity of words per T-unit in both experimental (1.21) and control (0.55) groups increased, although none of these differences were statistically significant.

Concerning the total number of clauses, significant differences were found between groups in the pre and post-test (F (1.78) = 30.147; p < 0.05). Sidak test for multiple comparisons showed a significant decrease in the number of clauses in the experimental group (≠ 2.1; p < 0.5). However, the number of words per clause increased more in the experimental group (= 1.16) than in the control (≠ 0.24), suggesting greater complexification, but yielded no statistically significant difference.

Lastly, a mean segmental type token ratio (MSTTR) was calculated and significant differences were found between groups in pre and post-test (F (1.78) = 10.038; p < 0.05). Although both groups showed greater ratios in the MSTTR, Sidak showed a significant increase in the experimental group (≠ 0.64; p < 0.05), therefore more lexical diversity together with syntactic complexity in the treatment group is regarded.

These analytic procedures and the results obtained from them illustrate major improvements in the number of errors at all levels (grammatical, lexical and mechanical), with the experimental group being significantly more accurate than the control. Furthermore, syntactic complexity and lexical diversity are observed in the experimental group, as the figures in the MSTTR show. Nevertheless, other measures of complexity such as words per clause and words per T-unit, which showed better scores in the experimental group, were not statistically significant.

4.2. Qualitative analysis

In order to show the evolution of the experimental group students, we made use of the conversations kept through WhatsApp, particularly several samples of LREs taking place in the application. These LREs were tracked in order to see if students reflected on the focus of each LRE in later messages, correcting the errors previously made. Three random samples of LREs were chosen and later messages of the students involved in such LREs were observed.

**Fig. 1** shows a LRE in which the student ‘Marta Alo B1 M’ misuses the word ‘used to’ and student ‘Ana Ferrer B1 M’ rephrases the sentence solving the mistake. The students ‘Marta Alo B1 M’ reflects on that mistake and agrees with her partner. In **Fig. 2** we can appreciate how the student ‘Marta Alo B1 M’ in a subsequent conversation, uses the adverb ‘usually’ appropriately.

**Fig. 3** shows a LRE in which the student ‘Belen B1 M’ makes a lexical mistake with the word ‘cooker’. In this case the student ‘Ana B B1 M’, who knows the meaning of the word her partner is trying to convey, tries to explain the meaning of the word ‘cook’. In **Fig. 4**, the student ‘Belen B1 M’ days later in a conversation about dream jobs uses the correct form of the word.

**Fig. 5** shows a LRE in which the student ‘Laura Arts B1 M’ lacks awareness of the negation in English. In this case the student ‘Ana Ferrer B1 M’ points out the fact that the double negation in English is not possible. The teacher supports the argument of the student, confirming that this claim is true. In **Fig. 6**, we can observe how the student ‘Laura Arts B1 M’ in a different conversation months later, uses the negation appropriately.
**Fig. 1.** Screenshot of a LRE elicited from the interaction in WhatsApp.

**Fig. 2.** Screenshot of a subsequent conversation in which the appropriate form of the word in Fig. 1 is observed.
Fig. 3. Screenshot of a LRE elicited from the interaction in WhatsApp.

Fig. 4. Screenshot of a subsequent conversation in which the appropriate form of the word in Fig. 3 is observed.
Fig. 5. Screenshot of a LRE elicited from the interaction in WhatsApp.

Fig. 6. Screenshot of a subsequent conversation in which the appropriate use of negation is observed.
5. Discussion

The present study was designed to determine if the programme implementation in the experimental group involved greater improvements in writing skills. Returning to the research questions previously posed, as shown in the group comparison carried out in Tables 1 and 2, the ratios of lexical, grammatical and mechanical errors as well as error-free clauses per clause and error-free T-unit per T-unit indicated significant differences between the control and experimental group in terms of accuracy. Notwithstanding that MSTTR measuring syntactic complexity together with lexical diversity showed significant differences between the two groups, no significant differences were found in terms of syntactic complexity as the ratios of words per clause, words per T-unit and clauses per T-unit yielded no statistically significant difference. The control group, as shown in Table 1, experienced improvements in terms of accuracy too, which may be due to practice within the group during the course. Nevertheless, as it was previously mentioned, Sidak results showed significant differences in the development of accuracy between the two groups. These findings highlight the potential of text chat in MIM to improve second language accuracy, in line with the latest studies regarding linguistic accuracy via text chat in CMC (Adams, Alwi, & Newton, 2015; Alwi, Adams, & Newton, 2012; Jepson, 2005). Figures regarding MSTTR concerning syntactic complexity together with lexical diversity emphasize the idea that synchronous and particularly asynchronous interaction gives learners more opportunities to produce syntactically complex language in line with Sotillo’s (2000) research in CMC, as students are able to reflect on their language productions. Nevertheless, the independent parameters of complexity showed no significant difference between the groups studied, hence further investigation is required.

As shown in previous studies of CMC such as Bueno-Alastuey (2013), Jepson (2005) or Rouhshad, Wigglesworth, and Storch (2016), chat-based conversation is a fertile ground for LREs. In a similar vein, MIM, due to the similarities to CMC, contains many of the LREs which contribute to construct knowledge as well as to second language development. The differences in the LREs between MIM and CMC remain as an object of study as the characteristics of mobile devices offer an environment full of opportunities for interaction and feedback.

Regarding the theoretical implications of these findings, the combination of error perception and negotiation of meaning skills has a beneficial effect on writing development. Knowledge is built in an active manner by students, who perform a key role in the process paying attention not only to the feedback provided by the teacher but also to their language productions. A rich input environment is created where instead of only one source of input, all participants take part in its creation, satisfying real communicative needs (Rixon, 2001). Cooperative learning taking place in the activity becomes an essential source of knowledge, and is an indispensable factor to encourage students. As Slavin (1983) claims, this kind of learning is conceived of as a team process in which the members help each other and trust in themselves to reach a goal. Some of these goals are: developing and sharing a common project; contributing to problem-solving, through questions, reflections and solutions; encouraging participation among students.

Long’s (1983) interaction theory — although it addresses communicative oral skill — is applicable in this study as the similarities between MIM and oral language correlate highly: spontaneity; high degree of interaction; naturalness; informality. Interaction, as opposed to studies such as Levy and Kennedy (2005) - where messages were sent at set times, on set days - takes place in a synchronous and asynchronous manner, exploiting the functionalist advantages of MALL.

Students undergo improvements in terms of syntactic complexity together with lexical diversity and accuracy, which correlate with can-do descriptors established by CEFR such as the possibility of writing about experiences, feelings, reactions and a detailed description on a range of familiar subjects as students are able to elicit a wider range of words than in previous levels. Regarding grammatical and mechanical skills, apart from the ones previously mentioned, we can observe several correlations with other descriptors set by CEFR such as the possibility of correcting mix-ups with tenses or expressions that lead to misunderstandings, or the ability to start again if communication breaks down. Furthermore, the possibility of narrating a story, writing a short or simple essay or summarising an opinion, would not be possible without such skills.

Despite the fact that the researcher in this inquiry tried to avoid all the possible limitations, some were found throughout the study. Participation was a major concern and although the acceptance of the activity was remarkable with a total number of 13,937 messages sent and a participation of 97.5% of the students, the amount of time students spent chatting through the application was not balanced as some limited themselves to the daily answers whereas others constantly followed the thread of the conversation. As such, the effects were more salient in certain students who participated to a higher degree.

6. Conclusion

Overall, the study strengthens the idea that the use of MIM contributes to the development of accuracy in second language writing. The significant decrease in the number of errors in the experimental group in relation to the control, whether grammatical, lexical or mechanical, illustrates the positive effects of the treatment in terms of linguistic accuracy. Regarding complexity and lexical diversity, the MSTTR emphasizes greater syntactic complexity together with lexical diversity. Nevertheless, the independent parameters for syntactic complexity yielded no statistically significant difference, hence there is no strong evidence for the MIM exchanges themselves to be significant. This may be due to its simplified register and short length, therefore further studies are required. In addition, the LREs taking place in the application constitute a fundamental source of knowledge for second language development, as students reflect and construct knowledge in an active manner, providing feedback and negotiating meaning.
Leaving aside those questions related to learning that were previously detailed in the discussion and throughout our research, the acceptance earned by the use of WhatsApp among students needs to be emphasized. According to Coll (2004), referring to didactics, the interest in studying the impact of new information and communication technologies (TIC) in educative processes has grown progressively during the last years, in parallel with an increasing inclusion of technologies at all levels of teaching. This makes us wonder why the tremendous potential that the WhatsApp application has proved to have as an educational tool during this research study has gone unnoticed in other studies. MALL encourages collaboration and co-construction of knowledge where learners had to find information and share it with their peers (Kukulska-Hulme & Shield, 2007). In this manner, the high involvement of the students in the activity and throughout the process was remarkable, even when the teacher was not participating actively. Moreover, the activity helped reduce the teacher-student distance, making the students feel more comfortable and increasing their participation and interest in this field of knowledge, thus facilitating learning. The number of studies related to MIM in education is extremely scarce and the tremendous potential of this tool has not been taken into consideration, thus a new field of research is open, where education and technology must progress in unison.

**Funding**

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Acknowledgements**

This research was carried out at the University of Almeria Language School, Spain, as part of a B1 English language course in 2013–2014. I wish to thank professor Phil Benson who provided insight and expertise that greatly assisted the investigation.

**References**


Leaving aside those questions related to learning that were previously detailed in the discussion and throughout our research, the acceptance earned by the use of WhatsApp among students needs to be emphasized. According to Coll (2004), referring to didactics, the interest in studying the impact of new information and communication technologies (TIC) in educative processes has grown progressively during the last years, in parallel with an increasing inclusion of technologies at all levels of teaching. This makes us wonder why the tremendous potential that the WhatsApp application has proved to have as an educational tool during this research study has gone unnoticed in other studies. MALL encourages collaboration and co-construction of knowledge where learners had to find information and share it with their peers (Kukulska-Hulme & Shield, 2007). In this manner, the high involvement of the students in the activity and throughout the process was remarkable, even when the teacher was not participating actively. Moreover, the activity helped reduce the teacher-student distance, making the students feel more comfortable and increasing their participation and interest in this field of knowledge, thus facilitating learning. The number of studies related to MIM in education is extremely scarce and the tremendous potential of this tool has not been taken into consideration, thus a new field of research is open, where education and technology must progress in unison.

**Funding**

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Acknowledgements**

This research was carried out at the University of Almeria Language School, Spain, as part of a B1 English language course in 2013–2014. I wish to thank professor Phil Benson who provided insight and expertise that greatly assisted the investigation.
