

Document downloaded from:

<http://hdl.handle.net/10251/62952>

This paper must be cited as:

Pérez-Sabater, C.; Montero Fleta, MB. (2015). A first glimpse at Mobile Instant Messaging: some sociolinguistic factors. Poznan Studies in Contemporary Linguistics. 51(3):411-431.
doi:10.1515/psycl-2015-0016.



The final publication is available at

<http://dx.doi.org/10.1515/psycl-2015-0016>

Copyright De Gruyter

Additional Information

Please cite as: Pérez-Sabater, C. & Montero-Fleta, B. (2015). A first Glimpse at Mobile Instant Messaging: Some Sociolinguistic Determining Factors. *Poznań Studies in Contemporary Linguistics*, 51(3), 411-431.

doi: 10.1515/pscl-2015-0017

A first Glimpse at Mobile Instant Messaging: Some Sociolinguistic Determining Factors

Abstract

Despite the vast body of research on the linguistic peculiarities of Instant Messaging and Short Messaging Service, little is known about the language used in Mobile Instant Messaging in a cross-generational and cross-cultural context. To fill in this gap, the cross-linguistic study addressed here is an attempt to approach age-specific variation from a blended ethnography approach. The current research is grounded on an analysis of a naturally-occurring dataset of *WhatsApp* messages from the point of view of oralisation and deviations from standard forms. Two distinct generations of English and Spanish texters provide empirical data on the parameters of the oralised written discourse suggested by Yus (2011): emoticons, orthographic mistakes, phonetic orthography, abbreviations, acronyms and clippings, and the use of words in other languages. Subsequently, an analysis of the interviews held with the writers approaches the factors that may determine language variation in the messages. The conclusions drawn highlight the persistent use of deviations from standard language of English and Spanish teenagers. The study confirms a higher frequency of a conversational style in Spanish than in English. Not only brevity and speed reasons but also familiarity, playfulness and intimacy with the addressee are behind the intentional variations used.

Keywords: Language variation; *WhatsApp* text messages; Oralised written discourse; Cross-cultural and cross-generational linguistic research.

1. Introduction

In the last few years, we have witnessed a communications revolution which has changed the way we interact with each other on a daily basis. A new wave of mobile communication services called Mobile Instant Messaging (MIM) has gained a relevant position in the market. At the moment, one of the most popular applications for mobile phones is *WhatsApp* Messenger, a proprietary, cross-platform Instant Messaging (IM) application for smartphones, which allows users to send and receive text messages as well as images, video, and audio media messages¹. The affordances of *WhatsApp* have played an important role in its popularity. These, according to our own research and that of Baron (2008), Yus (2011), and Jones and Hafner (2012), stem from the following characteristics: the low price of *WhatsApp*, the lack of a need for immediate feedback; it allows multitasking, status information and notifies when the message has arrived at its destination and when it has been read, among other affordances that were not possible in traditional text messaging. *WhatsApp* also permits community interaction with the creation of groups of friends or contacts in order to send and share messages among the group. All these features, which mainly imply the “convergence” of formerly distinct technologies and modes of communication in a single platform (Thurlow and Poff, 2013), have contributed to the current ubiquitous character of *WhatsApp*. Notwithstanding this, it is important to recognise “the interplay between what a

¹ *WhatsApp* is currently the fastest growing global messenger social networking tool in terms of users per month, with 500 million users in 2014, and an additional million a day according to Dredge's figures (2014).

technology itself allows (or affords) and what the communicator herself/himself brings to the technology” (Thurlow and Poff, 2013:164).

The linguistic characteristics of traditional instant messaging services such as Short Messaging Service (SMS) or Instant Messaging (IM) have been broadly analysed over the last decade (e.g., Baron 2008); however, little is known about how users communicate in MIM through the new applications developed for smartphones since these applications have other new affordances. Thurlow and Poff (2013) recommend the realisation of ethnographic or situated analyses to address a variety of issues since the technologies of texting are constantly varying, “so too are the practices and meaning of texting; any research on texting needs to be constantly updated” (*Ibid.*, 180). Linguistic studies of text-based interactions have addressed issues in sociolinguistics, for instance, the sociolinguistics of orthographic practices (Squires, 2012). Cross-cultural variation in text messaging has been also attested by several studies of German (e.g., Beiswanger, 2007), Greek (Spilioti, 2009) and Nigerian English (Chiluwa, 2008), among others; however, much research still covers English (Thurlow and Poff, 2013). It is this gap that the present study seeks to fill.

Focusing on the strategies of oralisation developed by Yus (2011), we will observe online writing practices cross-culturally, English-Spanish, and cross-generationally, teenagers versus adults. Our reason for analyzing teenagers’ texting follows Anderson (2001:8) who documented that teenagers are “often found at the forefront of linguistic deviation and innovativeness. Linguistic innovation is both inherent in the expression of social identity and crucial to the development of new linguistic forms and norms”. As Glaznieks and Stemle (2014) put it, we assume that users of different age groups can be

distinguished by their linguistic behaviour. The interrelation between linguistic behaviour and age is a topic in many sociolinguistic studies (Glaznieks and Stemle, 2014), but these variables have not been the focus of attention in the context of mobile interactions from an English-Spanish cross-cultural blended ethnography approach following Androutsopoulos (2008).

To situate the present study, we briefly review the main findings of research with regard to language variation and its implications in a Computer-mediated Communication (CMC)² context, and specifically, the parameters that affect language change which also include word, message and sentence length, significant elements that show conversational deviation of written registers.

2. Sociolinguistic variables in CMC: language and age

Language variation is a highly productive subfield of research in sociolinguistics. Scholars such as Labov (1994, 2001) have claimed its central status in linguistic theory.

Bayley (2002) documents that linguistic factors (i.e., structural composition and linguistic environment) may combine with social factors (i.e., age, situational context) to explain variability. In this line of thought, distinct peculiarities of language may be associated with different communities of practice or even communication environments. In the context of the present study, these communities are teenagers and adult writers of mobile text interactions, as “age is a relevant category in conversations and thus

² According to Herring et al. (2013:5) “CMC is based on established tradition and remains the term preferred among communication scholars”.

influences communication in general and the communicative behaviour of the participants in particular” (Glaznieks and Stemle, 2014: 34).

In online language studies, variation has been a key concept understood in a number of ways (Androutsopoulos, 2011). Early studies such as Crystal’s (2001) focused on medium variables. He defended that the language of the Internet displays features that are unique to the Internet, which he calls “Netspeak”. Crystal’s postulations of the language used on the Internet as a distinctive *variety* have been overtly criticised; he has been accused of overgeneralisation (Bieswanger, 2013). Research findings have revealed that there is no universal language of texting or “textspeak” (Bieswanger, 2013); rather there is considerable variation in language use in texting depending on the language used, the intention of the writers, the purpose of the message, and factors such as gender and age. For example, variation in usage, particularly in typography and orthography, has been explained related to participants’ status, regional dialect, gender, and CMC mode (Herring et al., 2013). The effects of digital communication on written language, especially with young people’s writing practices, have also been broadly researched (e.g., Baron, 2004).

Specifically in studies about texting, much variation is shown in international research (e.g., Thurlow and Poff, 2013). These authors have observed that spelling variation is a resource in individual identity performance; variables such as gender, nation, mother tongue and age showed noticeable differences of texting style. Indeed, researchers have claimed that spelling variation is functional, principled and meaningful. Concerning this, Tagg et al. (2013:219) comment that spelling variation is functional because spelling variants are not prescribed or learnt, they emerge during interaction “as a response to

immediate functional demands". Along with functional, spelling variation is principled because it follows orthographic principles of the language, and meaningful because it contributes to the performance of social identities (Ibid: 2013). Yus (2011) proposes a global model that includes these spelling variations, or text deformation in his words, in what he calls 'oralised written text'. The main aim of these oralisation strategies is:

...to turn typed text into a more expressive and speech-connote kind of discourse that allows for the communication, to a certain extent, of the nonverbal behaviour (vocal and visual) that typically accompanies human interactions in situations of physical co-presence and makes it possible to convey not only thoughts, but also the feelings and emotions attached or associated with them in a more effective way. (Yus, 2011: XIII)

From a sociological perspective, researchers like Thurlow and Poff (2013) echoing Androutsopoulos (2000) justify the need of young texters to affirm their personality by the use of deviations from conventional forms. Other powerful means of differentiating themselves from adults, for example, are to personalise and informalise their messages. Researchers have identified some features that characterise the informalisation of discourse (Authors et al., 2009). Among these, this study will focus on word, message and sentence length, parameters that can show conversational deviation of written registers, since short words, messages or sentences are characteristic of spoken discourse (Chafe, 1982) and a relevant characteristic of informality (Baron, 2000).

Particularly regarding young texters' writing practices, oralisation strategies or anti-normative textese language are claimed to be destroying the language by the media (e.g., Thurlow, 2006). However, assumptions widely expressed in the media about the fact that texting is threatening people's ability to write in standard language have not been corroborated by serious studies on the topic. As Plester et al. (2009) posit, only a

minority of scientific studies affirm that texting has a negative influence on standard writing, grammar and spelling. On the contrary, most empirical studies suggest that texting “does not pose a threat to standard language teaching and learning” (Thurlow and Poff, 2013: 171). For example, the findings of Plester et al. (2009) prove that frequent teenage texters perform better in standardised measures of English proficiency.

To conclude this section, we would like to highlight that much past research has contrasted oralisation strategies and spelling variation across English and certain other western languages. These articles have examined the differences among texting practices in languages like English, German or Japanese (Beiswanger, 2013). Nevertheless, in high impact dissemination platforms, little has been said about the writing practices in Spanish and English contrastively of young and adult texters. This is what the research questions below address.

3. Research questions

Considering the theoretical perspectives laid out in the previous sections, the empirical research described here is based on the sociodemographic variable *age* and the language of communication used. The analysis of these variables may contribute to answering the specific research questions devised for this study:

- To what extent does age act as a variable in the use of deviations from conventional language in *WhatsApp* written messages?
- To what extent is this process attested cross-culturally in *WhatsApp* written messages?

To respond to the research questions posed, a cross-linguistic study was carried out on the traits of the language used in English and Spanish *WhatsApp* messages written by teenagers and adults. Following Cheshire (2006), we tried to observe generation-specific language and cross-cultural differences regarding the deviations from standard forms in mobile interactions in the languages under study and the contribution of oralisation strategies to language variation.

4. Methodology

4.1. Data collection

Gathering a significant amount of interactions in the new media is a challenging task which is even harder when these interactions have a private character such as those exchanged through *WhatsApp*. The collection of private exchanges needs the collaboration of multiple interactants to obtain a representative sample of written messages (Spooren and van Charlidorp, 2014). The corpus of *WhatsApp* text interactions in Spanish used in the present study was made up from a sample of naturally-occurring messages in Spanish collected with the help of teachers from our department³. As for the corpus in English, we contacted native teachers of English currently living in Spain. Regarding the corpus of messages written by teenagers in English, we counted on a group of native English teenagers participating in an exchange program in a high school in Spain, who voluntarily agreed to collaborate in our study. Each participant was asked to provide *WhatsApp* messages sent and received during the previous week. The Spanish students of the survey were also high school students studying in this Spanish high

³ All the participants gave their written consent for our use of their personal interactions for research purposes.

school. The participants were selected from similar social and cultural backgrounds, as these sociolinguistic variables may have an influence on the language used (see, for example Squires, 2012). As in the recent study by Lyddy et al. (2014), the participants were informed of the purpose of the research and were encouraged to provide *WhatsApp* messages that were representative of those generally sent. The participants recruited, called ‘primary subjects’ following Squires (2012), were 3 male and 3 female texters for each group that forms the corpus; the gender variable was not considered in this study, although this variable can provide interesting findings as in the article by Squires (2012), where differences of language output were found according to the gender of the author. The primary subjects contributed with MIM chat chains exchanged with interlocutors, i.e., ‘secondary subjects’ according to Squires (2012), of the same age and native language. The participants of the study recruited were the following:

- 6 English teenager chat writers from 12 to 17 years old⁴.
- 6 English adult mobile chat writers aged around 40.
- 6 Spanish teenager chat writers from 12 to 17 years old.
- 6 Spanish adult mobile chat writers aged around 40.

The written interactions of these 24 interlocutors with secondary subjects of the same characteristics form the unit of analysis, a corpus of original messages and of everyday conversation that, in our opinion, maybe representative, and is manageable as far as the linguistic analysis and interviews.

⁴ Following Livingstone and Helsper ‘s (2010) study of teenagers’ Internet literacy.

Once the corpus had been formed, a second step of this case study implied corpus preparation, a process which required the removal of names, times and dates from the text of the message for further processing. A number of messages of a non-naturally-occurring source, such as forwarded chat chains and texts in images, were removed. The participants' contributions were balanced so as to contain the same word count in English and in Spanish in each group. After this process, the total number of words processed amounted to 10,000 words. This sample facilitated the linguistic framework of *WhatsApp* chats from the two different age ranges, and in the two languages under study. We would like to clarify that for this study we only took into account the total number of words gathered. The number of the messages and threads varies so much that, in our view, these data may not provide relevant facts in the overall analysis.

4.2. Coding procedure

In a study of the analysis of politeness in text messages, Spilioti (2011:71) posits that: “Systematic observation and interviews, together with detailed linguistic analysis, can contribute to an emic understanding of texters’ practices and categories”. Thus, similarly to the study of Spilioti, the methodology undertaken followed a blended-ethnography approach (Androutsopoulos, 2008). This “discourse-centred online ethnography”, although not wholly ethnographic in the traditional sense, draws on ethnographic principles and methods; it combines two basic techniques: a linguistic analysis and systematic observation and interviews with the participants.

4.2.1. Linguistic analysis

The linguistic analysis consisted of a cross-generational linguistic study of the corpus. In some steps of the process, we incorporated *WordSmith* tools (Scott 2004). Since the analysis on language variation in speech and writing by Biber (1989), software packages have been used in the field of corpus linguistics to search patterns in a language. The parameters measured with *WordSmith* were word, message and sentence length, significant parameters to show conversational deviation of written registers.

Furthermore, in order to observe oralisation and the frequent use of non-standard language in online written conversations, the computerised analysis was completed by a manual study of other parameters chosen from the proposals put forward by Baron (2008) and Yus (2011). The following aspects of oralised written discourse and intentional variations were studied: emoticons⁵, orthographic mistakes, phonetic orthography, abbreviations, acronyms and clippings, and the use of words in other languages, that is, code-switching. Several academics have classified these strategies differently (see, for example, Yus, 2011; Tagg et al., 2012; Lyddy et al., 2014). For this study, we have followed the systematic classification of Yus (2011).

Generally speaking, emoticons, initially iconic compositions of characters are “nowadays many galleries of fully iconic faces … offered by the different interfaces for virtual interactions” (Yus, 2014:511). Rather than on the form of emoticons, the present study focuses on their function as non-verbal emphasisers of emotions, and their role in contributing to the eventual relevance of the text they accompany. In Orthographic mistakes we have distinguished between unintentional misspellings and intentional

⁵ In this article we include in the emoticons emoji that depict facial expressions. We have not considered holiday symbols, activities or animals since they change daily in the application.

misspelling. On the one hand, unintentional mistakes are usually words unintentionally misspelled in general writing (MacDonald et al., 2013). In turn, in CMC, economy reasons due to the pressure to type and send the messages as fast as possible, have made many intentional mistakes pervasive such as the absence of capitalisation, accents and apostrophes. As for phonetic orthography, it “is the strategy of reproducing textually the text as it would be pronounced orally” (Yus, 2011: 176). In this section we detailed the subcategories established by Yus (2011): phonetic spellings, i.e., reproducing the text as it would be pronounced; colloquial spellings, the transcription of the colloquial reduction of words; regiolectal spellings or eye dialect, the transcription of regional variations of a language; prosodic spellings, the textual transcription of prosodic contours of the voice; interlingual spellings, transferring the phonetic attributes of a word from a foreign language but making it fit the orthographic conventions of the importing language; and homophone spellings, which includes lexical and grapheme substitution.

Then, we observed abbreviations, acronyms and clippings, which turn “paragraphs into a kind of hieroglyphic that only those users who master the conventions of these textual strategies can decipher” (Yus, 2001: 177). Finally, the use of *code-switching* by using words from another language is a politeness strategy that reinforces community links according to Lan (2000), and it can also act as a marker of elite social identities in the opinion of Herring (2011).

This manual analysis was intended to provide a tentative approximation to the most striking features in naturally-occurring text messages in *WhatsApp* discourse. As a point of departure, we analysed a random sample of 500 words from the messages of each

group in order to observe if this variation could be scholarly interesting, following Biber (1989). These preliminary results confirmed our initial intuitions about the relevance of the task. A manual codification followed, which entailed the identification, categorisation and analysis of the oralisation strategies mentioned above in the whole corpus. Some problems arose in the coding phase since some of the deviations from standard writing could be coded in several categories. For example, in this sentence taken from adult writers in English: '*fieessstaaaa at mine on fri!!!!*', the word '*fieessstaaaa*' could be included in orthographic spelling, missed capitalisation and in code-switching. We then decided to count this word in the three parameters. After the coding, the corpus was again independently analysed by a second rater. The interrater agreement was 92%.

4.2.2. Systematic observation and interviews

Interviews with participants, a technique necessary to carry out this discourse-centred online study (Androutsopoulos, 2008), involved the direct contact with social actors. In our case, this was done by means of personal interviews to get feedback from the participants' writing practices on smartphones and their perception of deviation from standard language. Two participants from each category were selected randomly for the systematic observation and interviews. The face-to-face conversations followed the same format: took 20 minutes, were recorded and transcribed. The topics of these interviews were addressed through open questions regarding the senders' awareness of the difference between formal and informal registers and also of their writing habits in *WhatsApp* production, as for example:

| Topics dealt with in interviews |
|---|
| <ul style="list-style-type: none"> - teenagers'/adults' revision of the messages before sending; - teenagers'/adults' use of predictive texting; - teenagers'/adults' style used when the messages were sent to one recipient or to more than one; - teenagers'/adults' different style used in the written interactions depending on the age of the prospective recipient/s; - teenagers'/adults' feedback on the oralised strategies and creative language used; - teenagers'/adults' feedback on the purpose of the strategies used. |

The processing and interpretation of these interviews added information about the writing habits of these groups.

5. Results

5.1. Linguistic analysis

From the language analysis undertaken, the data shown below summarise the results obtained on the 10,000 words processed from *WhatsApp* messages written by participants of different age ranges and in the two languages studied. All the examples we show and discuss are of naturally-occurring mobile interaction. Table 1 shows the results obtained after processing the corpora using *WordSmith*.

| | English | | Spanish | |
|--------------------------------|----------------|--------|----------------|--------|
| | Teenagers | Adults | Teenagers | Adults |
| Mean message length in words | 4.71 | 10.02 | 4.64 | 6.21 |
| Mean word length in characters | 3.82 | 3.71 | 3.95 | 3.66 |

| | | | | |
|-------------------------|------|------|------|------|
| Mean words per sentence | 4.59 | 4.87 | 5.78 | 5.61 |
|-------------------------|------|------|------|------|

Table 1. WordSmith data in English and Spanish.

These data show that, both in English and Spanish, teenagers favoured shorter messages. The difference is more outstanding in English where adults' interactions double teenagers' in the number of words used. However, teenagers used slightly longer words in both languages and their sentences were a bit shorter than those of adults in English and in Spanish.

Table 2 details the average number of occurrences of the parameters analysed in each corpus. The figures are given per 500 words to facilitate their explanation.

| Parameter | English | | Spanish | |
|---------------------------------------|-----------|--------|-----------|--------|
| | Teenagers | Adults | Teenagers | Adults |
| Emoticons | 10 | 14 | 16 | 14 |
| Orthographic mistakes | 76 | 24 | 76 | 16 |
| Phonetic orthography | 68 | 14 | 66 | 15 |
| Abbreviations, acronyms and clippings | 64 | 19 | 62 | 16 |
| Words in other languages | 2 | 20 | 4 | 10 |
| Total number of occurrences | 220 | 91 | 224 | 71 |

Table 2. Occurrences of each parameter in the corpora per 500 words.

Globally speaking, the figures above attest that 44% of the content of the messages of English teenagers was non-standard vs. 44.4% of the content of Spanish teenagers. As for adults, 18.2% of the content of English messages and 14.2% of Spanish messages was non-standard.

Table 3 shows the rate that each parameter exhibited in the global number of oralised strategies in the variables age range and language.

| Parameter | English | | Spanish | |
|---------------------------------------|-----------|--------|-----------|--------|
| | Teenagers | Adults | Teenagers | Adults |
| Emoticons | 4.54% | 15.38% | 7.14% | 19.71% |
| Orthographic mistakes | 34.54% | 26.37% | 33.92% | 22.53% |
| Phonetic orthography | 30.90% | 15.38% | 29.46% | 21.12% |
| Abbreviations, acronyms and clippings | 29.09% | 20.87% | 27.67% | 22.53% |
| Words in other languages | 0.90% | 21.97% | 1.78% | 14.08% |

Table 3. Frequency of each parameter in the total amount of oralisation strategies.

To illustrate the use of the deviations from standard forms, some representative examples of each parameter are shown below.

a. Emoticons

The figures in Table 3 demonstrate that the participants displayed a similar use of these strategies. In general, emoticons were mainly used to express laughter, surprise and approval.

b. Orthographic mistakes.

In this parameter we include both types of mistakes: those unintentionally misspelled and those intentionally created.

- i. Unintentional misspellings. In English, teenagers confounded *your* and *you are* or *there house* instead of *their house*; adults did not usually make this type of

mistakes. In Spanish, teenagers made unintentional mistakes of the type *aber* instead of *haber* [have], since the letter *h* is silent in Spanish. Some mistakes may have been due to the contiguity of the letters *b* and *v* in the keyboard: *buelo* instead of *vuelo* [flight].

- ii. Intentional misspelling. The absence of apostrophe, lack of accentuation and missed capitalisation were common features of *WhatsApp* messages. In the case of English, intentionally created mistakes such as a generalised absence of apostrophes was observed in both age groups, as in this example by adult users: *if i cant find it ill get it somewhere else*. English teenagers' use of apostrophes was even more seldom. Double letter reduction was common practice in *sory* for *sorry* and *wory* for *worry*. Comparatively speaking, Spanish adults made a low number of mistakes, while teenagers' writing in Spanish demonstrated a systematic lack of accentuation; errors like *practica* instead of *práctica* [practical] were commonly detected. As for the absence of capitalisation, it was frequent in all the groups.

- c. Phonetic orthography

Intentional variations in orthography were persistently observed in the corpus but were outstandingly preferred by teenagers who scored three times higher than adults in the use of this strategy of reproducing prosodic contours of the voice: in 500 words, teenagers used 68 in English and 66 in Spanish while adults only had 14 occurrences in English and 15 in Spanish.

- i. Phonetic spellings. In English, both groups used this strategy a lot: *Drawing somewhere wud b better* [would be]. Spanish teenagers used *uapaah* (*guapa*) [pretty] or *weno* (*bueno*) [ok, I see].
- ii. Colloquial spellings. We found examples such as *hafta* [have to], *wanna* [want to] or *gonna* [going to] in English teenagers. In Spanish, teenagers wrote *He acabao ya* for *he acabado ya* [I have already finished].
- iii. Regiolectal spellings and eye dialect. Teenagers writing in English frequently used dialectal forms such as *da* [the], *heyas* [hello]. In Spanish, teenagers wrote *comio* instead of *comido* [eaten], or *qué animalà* instead of *qué barbaridad* [that's incredible].
- iv. Prosodic spellings. This strategy was used in both languages, especially by teenagers. These are some of the most common examples found: *helloooooo*, *yessssss*, *okkkkk*, *fuckkk*, *noooo*, *soooo*, *maaaaybe*. *Ok* was sometimes spelt *okey* or *okipokiii* (for *okey-dokey*). Interestingly, prosodic spellings were very frequently used in laughter expressions in *WhatsApp* messages in English and Spanish, although without consistent spelling. English teenagers often applied the mechanism of explicit letter-sound skills, e.g., *haha*, *hahaa* or *hehe*. Some examples used by Spanish texters were, e.g., *ajajajajajaj*, *uajjjajajajaja*, *ajajajja*, *jaja*, *jeje*, *jajajaja* or *jajajaj*; prosodic spelling was preferred to the use of emoticons to express the same feeling. Not only letter repetition was observed, exclamation mark repetition was particularly present in all the groups: *vaaaale!!!!!!!* [Ok!!!!].

- v. Interlingual spellings. We only found examples in Spanish of this strategy. Participants wrote in their WhatsApp texts the English word *busy* as *bisi* reflecting its pronunciation, or *people* as *pipol*.
- vi. Homophone spellings. Some instances of English homophones used by adult writers were *sink* for *synch* or *plain* for *plane*. Grapheme substitution, as in *k* or *q* for *que* [that] and lexical substitution such as *xq* for *porque* [because], was a pattern of use in Spanish where the mathematical symbol *x* substitutes the preposition for.

- d. Abbreviations, acronyms and clippings.

As shown in Table 3, these shortenings were constant (29.09% of the non-standard forms by English teenagers and 27.67% in Spanish), although teenagers used them far more than adults: 64 occurrences and 62 in English and Spanish teenagers respectively versus 19 and 16 by adults per 500 words.

We found examples of abbreviations, acronyms and clippings in English of the type *fri* [Friday], *gf* [girlfriend], *prob* [problem], *u* [you] or *2moro* or *2moz* [tomorrow] in both age groups. Spanish teenagers made an abundant use of these strategies through the omission of vowels and some consonants in certain words or expressions in common use: *cn* instead of *con* [with]; *cd* for *cuando* [when], *qtal* for *qué tal* [how are you]; *bs* instead of *besos* [kisses], *spo* instead of *espero* [I hope].

English and Spanish adult writers also employed these strategies exemplified in messages of the kind: *toi* [estoy] *en la pelu* [peluquería] [I am at the hairdresser's].

e. Words in other languages

The *WhatsApp* sample authored by native adult speakers of English regularly included words in Spanish (20.87% of the deviations) because they live in Spain. The use of *fiesta* or *calle* [instead of *street*] or *cena* [instead of *dinner*] was normal practice, e.g.: *What about cena at mine?* However, English teenagers' code-switching was very rare (0.90%).

Spanish adult texters frequently resorted to words in English. The most popular words in English were *please* and *thanks*, whose spellings were often adapted to Spanish as we have already mentioned: the phonetic spellings *plis* or *pls* for *please* or *zanks* for *thanks*. Other words in English detected were *of course*, *break*, *sorry*.

5.2. Systematic observation and interviews

The second stage of our research consisted in personal interviews with the participants aiming at discovering some determining factors and attitudes that may have an influence on the production of *WhatsApp* messages. To find the answer to the frequent orthographic mistakes in MIM, the usual revision of the messages before being sent was questioned. There were more adults than teenagers who habitually re-read their written production before sending. Adults manifested a higher tendency to proofread their messages and revised an average of 1 out of 2 messages before sending. Teenagers revealed that their misspellings were due to speed in texting; they were able to write messages much faster than when using Standard English. The informants were asked about their use of predictive texting in relation to orthographic mistakes. Most adults, 70%, had this facility incorporated while most teenagers did not. Finally, whereas adult

texters took into account the age of the recipients as a determining factor for language correctness, teenagers used similar language patterns disregarding the age of the addressee. All the writers manifested different writing habits depending on the online platform used. In their opinion, their messages on *WhatsApp* accounted for the same degree of informality and language variation that they applied to *Twitter*, for example, whereas their messages on *Facebook* were far more formal in their view. The results of this survey confirmed a general awareness of the writers of the difference between formal and informal registers.

When informants were interviewed about their opinion on the oralised strategies employed, creative language and its purpose, both English and Spanish age groups were aware of the strategies utilised, which, in their opinion, conveyed solidarity with their addressees, playfulness, and innovativeness.

In brief, this survey clearly confirms the results of the linguistic analysis undertaken: the preference of teenagers for the inclusion of oralisation strategies in their messages.

6. Discussion and conclusion

This study tried to shed light upon some sociolinguistic determining factors of naturally-occurring text messages exchanged by interactants of different age spans and different languages. MIM has provided an important site in which to examine sociolinguistic variation and its adherence to standard language. The corpora of *WhatsApp* messages collected provided a concrete support and examples that reliably represent authentic language use in current electronic communication, in Herring's (2011) words, one of the

most difficult tasks in Computer-mediated discourse analysis. Some interesting facts could be outlined from the results drawn from the study.

Regarding our research questions, if age is a determining factor attested cross-generationally and cross-culturally in *WhatsApp* messages, the comparison between English and Spanish produced some unforeseen results. On the one hand, adults writing in English wrote longer messages than the other interactants, contrarily to results from previous studies on online communication (e.g., Authors et al., 2009). In this former cross-cultural examination of online written messages, English writers consistently displayed a preference for much shorter sentences than writers of Romance languages such as Spanish. Another interesting result has been that, in the present research, teenagers have employed slightly longer words in both languages, this can be accounted for the fact that they frequently resorted to letter repetition.

In general, our results have confirmed that the conversational style conveyed by means of oralisation strategies in MIM is a tendency more frequently present in Spanish than in English, unlike some previous studies carried out by the Authors (2008) that showed a higher preference for a less conversational style in online written practices in Spanish.

On the other hand, from the two variables studied, the variable language manifested fewer dramatic differences than age. Our cross-cultural analysis showed the higher contribution of English and Spanish teenage texters to language variation by incorporating a noticeable number of non-standard forms into their messages in contrast with adults' messages. Cross-generational differences were observed in the strategies of oralisation by means of the parameters analysed that, in consistence with

Androutsopoulus (2013), compensated for the lack of phonological feedback, posture and other cues used in ordinary spoken conversation.

Let us now see an overview of the examination of these oralisation strategies in detail. Firstly, the study of emoticons has provided revealing findings. Unlike the infrequent use of emoticons found by Baron (2008) as well as by the Authors (2008) in previous analyses of CMC, the corpora of these texters attested a high number of emoticons. This may be due to the fact that *WhatsApp* has recently incorporated a great variety of them as textual portrayals of the writer's mood or facial expressions; emoticons have taken the place of the former emoticons made with keyboard strokes, combinations of different typographic symbols, such as colon, semicolon or brackets to express, for example, smile, surprise or horror faces. In this case, the affordances of the medium have clearly shaped the message (Condon and Cech, 2001). In our study, participants mainly expressed their emotions online through facial expressions of smile or frown to express laughter, surprise or approval.

Interestingly, our study found out that *X*, a traditional symbol to signal emotion in written correspondence, was a frequent substitution for an emoticon in English when signing off a message to display emotion, as in the study carried out by Lyddy et al. (2014).

Secondly, our results have confirmed that, at the moment, there is a clear tendency to the use of non-standard language, especially in young texters' messages with nearly half of the word content with some oralisation strategy, unlike the results of previous empirical studies which consistently supported the view that the majority of text messaging is

standard form (see the recent overview carried out by Lyddy et al., 2014). The generalised intentional variation found in spelling and non-standard language, for example, was largely attested by teenagers' writings, which showed a higher contribution than adults to language variation. However, differences found in the writing practices of both age groups, such as revision of messages and predictive texting, may have had an influence on the use of fewer oralisation strategies in adult texters.

Mainly, these oralisation strategies stem from the brevity-speed imperative as the interviews suggested. Similarly, economy reasons may be behind the absence of capitalisation, whose inclusion needs a second keystroke to write uppercase letters, as happens with accentuation in Spanish and the use of apostrophes in English, all of them intentionally produced as "part of the language games that this medium favours" Yus (2011:176).

Past research on variation in texting (e.g., Thurlow and Brown, 2003) consistently revealed its motivation by three 'social maxims': brevity and speed, paralinguistic resolution and phonological approximation. In chatrooms and instant messaging, these maxims encourage variation in spelling which includes, for instance, lexical abbreviation and the use of capitals to indicate emphasis or attempts to capture informal speech. In this concern, the results drawn from this study have also highlighted the presence of the three 'social maxims' in Mobile Instant Messaging. However, what holds for texting does not always hold for MIM: it is worth observing the absence of capitalisation in our findings in sharp contrast to its frequent use in texting according to Thurlow and Brown (2003).

Thirdly, the use of code-switching by English adults was unsurprising, as the authors live in Spain and have a daily contact with Spanish. This incorporation of words in other languages is used as a marker of social identities according to Herring (2011), and can reinforce community links in Lan's (2000) view. The widespread use of words like "ok" by Spanish writers is worth mentioning, as this expression is currently integrated into everyday Spanish; it is a commonly accepted Anglicism (Otheguy, 2001).

Notwithstanding all the differences mentioned above, we want to highlight that individual style is an important component in variation which may show notable differences among writers. Furthermore, in consistence with Squires (2012), the writers' individual stylistic choice considered non-standard in the mainstream may be considered standard in the particular community under study. This implies that the deviations present in the sample analysed may be standard in this context but non-standard in other forms of written language. Indeed, the language used presented a less traditional writing style that may require the gradual emergence of norms, which seems to concord with the conclusions reached by Ferrara et al. (1991) on interactive written discourse in the early stages of CMC research. This may need some time to acquire a standard status or not. Nevertheless, this study has demonstrated that, despite personal styles, there is a general tendency to implement oralisation strategies in written registers in which time constraints are largely involved.

All in all, the considerations set out on the oralisation strategies used in MIM confirm the sense of intimacy, expressiveness and connection, cross-generationally and cross-culturally speaking, among the interlocutors.

Finally, some considerations need clarification. The data obtained from this research cannot be linguistically conclusive due to the fact that the messages analysed could be biased because they were not randomly selected. We agree with Baron (2008) that they simply constitute a convenience sample available for study. With a broader sample, individual styles, for example, will likely show different patterns of variation, as suggested by Squires (2012). We would also like to say that the fact that the group of English adult participants were living in Spain may have influenced the results, at least in some parameters such as the use of code-switching.

This research provides a starting point for further investigations in this promising field of Computer-mediated Communication where cutting edge technological developments change rapidly. *WhatsApp* is constantly growing and adopting new channels and new facilities that may make our results obsolete almost overnight, as Spooren and van Charlidorp (2014) explained. We are aware that the factors that influence the adoption and usage of new technologies are, at most, complex and sometimes unpredictable. Notwithstanding this, since the primary objective of *WhatsApp* is to communicate events, the need to interact will provoke the adoption of another popular service with similar affordances, allowing the research published in this article to be scholarly relevant again. Nevertheless, and apart from this continuous change, what remains unchanged, in Thurlow and Poff's (2013: 179) words, "is people's determination and capacity to rework technologies (both mechanical and linguistic) for maximising sociality — in other words, for communication".

Further cross-linguistic studies will be needed to reach to more conclusive results in the field. Future research may observe gender differences in the use of *WhatsApp* or

communities of practice and their particular stylistic features of use by collecting online messages sent to groups of texters in *WhatsApp* since friends and peer groups usually establish their own local stylistic norms in order to maintain and build relationships (Thurlow and Poff, 2013).

References

- Andersen, G. 2001. *Pragmatic markers and sociolinguistic variation: A relevance-theoretic approach to the language of adolescents*. Pragmatics and Beyond New Series, 84. Amsterdam and Philadelphia: John Benjamins Publishing.
- Androutsopoulos, J. 2000. “Non-standard spellings in media texts: The case of German fanzines”. *Journal of Sociolinguistics*, 4(4). 514-533.
<http://onlinelibrary.wiley.com/doi/10.1111/1467-9481.00128/abstract;jsessionid=7DC7B606BCAAEDB3618057A1D616F863.f04t04?userIsAuthenticated=false&deniedAccessCustomisedMessage=> doi: 10.1111/1467-9481.00128
- Androutsopoulos, J. 2008. “Potentials and limitations of discourse-centred online ethnography”. *Language @ Internet*, 5. Available at <<http://www.languageatinternet.org/articles/2008/1610>>. Last accessed 8 July 2015.
- Androutsopoulos, J. 2011. “From language variation to heteroglossia in the study of Computer-mediated Discourse.” In: Thurlow, C. and K. Mroczek (eds.), *Digital Discourse: Language in the New Media*. Oxford and New York: Oxford University Press. 277-298.
- Baron, N. S. 2000. *Alphabet to Email*. New York and London: Routledge.

- . 2004. “‘See you online’: Gender issues in college student use of Instant Messaging”. *Journal of Language and Social Psychology*, 23(4). 397-423.
<http://jls.sagepub.com/content/23/4/397.short>. doi: 10.1177/0261927X04269585
- . 2008. *Always On. Language in an Online and Mobile World*. Oxford: Oxford University Press.
- Bayley, R. 2002. “The quantitative paradigm.” In: Chambers, J. K., P. Trudgill, and N. Schilling-Estes (eds.), *The Handbook of Language Variation and Change*. Oxford: Blackwell. 117-141.
- Beiswanger, M. 2013. “Micro-linguistic structural features of Computer-mediated Communication.” In: Herring, S.C., D. Stein and T. Virtanen (eds.), *Pragmatics of Computer-mediated Communication*. Berlin: Walter de Gruyter. 463-488.
- Biber, D. 1988. *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- Cheshire, J. 2006. “Age- and generation-specific use of language”. In: Ammon, U., N. Dittmar and K. Mattheier (eds.), *Sociolinguistics: an international handbook of the science of language and society*. Berlin, Germany: Walter de Gruyter. 1552–1563.
- Chafe, W. 1982. “Integration and involvement in speaking, writing, and oral literature”. In: Tannen, D. (ed.), *Spoken and written language. Exploring orality and literacy*. Norwood, NJ: Ablex. 35-53.
- Chiluwa, I. 2010. “Nigerian English in informal email messages”. *English World-Wide*, 31(1). 40-61. <https://benjamins.com/#catalog/journals/eww.31.1.02chi/details>. doi: 10.1075/eww.31.1.02chi
- Condon, S. L. and C. G. Cech. 2001. “Profiling turns in interaction: Discourse structure and function”. *Proceedings of the 34th Annual Hawaii International Conference on*

- System Sciences.* IEEE. 1-10. Available <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.105.6433&rep=rep1&type=pdf>>. Last accessed 8 July 2015.
- Crystal, D. 2001. *Language and the Internet*. Cambridge: Cambridge University Press.
- Dredge, S. 2014. "Why is WhatsApp worth up to \$19bn to Facebook?" *The Guardian*, 20 February. <http://www.theguardian.com/technology/2014/feb/20/facebook-whatsapp-acquisition-apps-messaging>
- Ferrara, K., H. Brunner and G. Whittemore. 1991. "Interactive written discourse as an emergent register". *Written Communication*, 8(1). 8-34. <http://wrx.sagepub.com/content/8/1/8.short>. doi: 10.1177/0741088391008001002
- Glaznieks, A. and E.W. Stemle. 2014. "Challenges of building a CMC corpus for analyzing writer's style by age: The DiDi project". *Journal for Language Technology and Computational Linguistics*, 29(2). 31-58. http://www.researchgate.net/publication/267448275_Challenges_of_building_a_CMC_corpus_for_analyzing_writers_style_by_age_The_DiDi_project
- Herring, S. C., D. Stein and T. Virtanen (eds.). 2013. "Introduction to the Pragmatics of Computer-mediated Communication". *Pragmatics of Computer-mediated Communication*. Berlin: Walter de Gruyter. 3-34. http://www.beck-shop.de/fachbuch/inhaltsverzeichnis/9783111738635_TOC_001.pdf
- Jones, R. H., and C. A. Hafner. 2012. *Understanding Digital Literacies*. London and New York: Routledge.
- Labov, W. 1994. *Principles of Linguistic Change, 1: Internal factors*. Malden and Oxford: Blackwell.

- . 2001. *Principles of Linguistic Change, 2: Social factors*. Malden and Oxford: Blackwell.
- Lan, L. 2000. "Email: a challenge to Standard English?" *English Today*, 16(4). 23-29.
<http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=1280724&fileId=S0266078400000511>. doi: <http://dx.doi.org/10.1017/S0266078400000511>
- Livingstone, S. and E. Helsper. 2010. "Balancing opportunities and risks in teenagers' use of the internet: The role of online skills and internet self-efficacy". *New Media & Society*, 12(2). 309-329.
<http://nms.sagepub.com/content/early/2009/11/24/1461444809342697.abstract>. doi: 10.1177/1461444809342697
- Lyddy, F., F. Farina, J. Hanney, L. Farrell and K.N O'Neill. 2014. "An analysis of language in university students' text messages". *Journal of Computer-mediated Communication*, 19(3). 546-561.
<http://onlinelibrary.wiley.com/doi/10.1111/jcc4.12045/full>. doi: 10.1111/jcc4.12045
- Otheguy, R. 2001. "Simplificación y adaptación en las lenguas en contacto". II Congreso Internacional de la Lengua Española. Available at <http://cvc.cervantes.es/obref/congresos/valladolid/ponencias/unidad_diversidad_del_esp/panol/3_el_espanol_en_los_EEUU/otheguy_r.htm>. Last accessed 8 July 2015.
- Plester, B., C. Wood and P. Joshi. 2009. "Exploring the relationship between children's knowledge of text message abbreviations and school literacy outcomes". *British Journal of Developmental Psychology*, 27(1). 145-161.
<http://onlinelibrary.wiley.com/doi/10.1348/026151008X320507/abstract;jsessionid=CCB847E77E0E870BA44EE4EDB740AAC4.f04t02?userIsAuthenticated=false&deniedAccessCustomisedMessage=>. doi: 10.1348/026151008X320507

- Scott, M. 2004. *WordSmith Tools Version 6.0*. Oxford: Oxford University Press.
- Spilioti, T. 2011. "Beyond genre: Closings and relational work in text-messaging". In: Thurlow, C. and K. Mrocze (eds.), *Digital Discourse: Language in the New Media*. Oxford and New York: Oxford University Press. 67-85.
- Spooren, W. and T. van Charldorp. 2014. "Challenges and experiences in collecting a chat corpus". *Journal for Language Technology and Computational Linguistics*, 29(2). 83-96.
- Squires, L. 2012. "'Whos punctuating what?' Sociolinguistic variation in instant messaging". *Orthography as Social Action: Scripts, Spelling, Identity and Power*, 3. 289-323.
- Tagg, C., A. Baron and P. Rayson, 2012. "'I didn't spel that wrong did i. Oops': Analysis and standardisation of SMS spelling variation". In: Cougnon, L.A. and F. Cédrick (eds.), *SMS Communication: A linguistic approach. Special issue of Lingvisticae Investigationes*, 35(2). 367–388.
- Thurlow, C. 2006. "From statistical panic to moral panic. The metadiscursive construction and popular exaggeration of new media language in the print media". *Journal of Computer-mediated Communication*, 11(3), 667-701.
<http://onlinelibrary.wiley.com/enhanced/doi/10.1111/j.1083-6101.2006.00031.x/>. doi: 10.1111/j.1083-6101.2006.00031.x
- Thurlow, C. and A. Brown. 2003. "Generation Txt? The sociolinguistics of young people's text-messaging". *Discourse Analysis Online*, 1(1). 30-57.
<http://extra.shu.ac.uk/daol/articles/v1/n1/a3/thurlow2002003-paper.html>
- Thurlow, C. and K. Mrocze (eds.). 2011. *Digital Discourse: Language in the New Media*. Oxford: Oxford University Press.

Thurlow, C. and M. Poff. 2013. “Text Messaging.” In: Herring, S.C., D. Stein and T. Virtanen (eds.), *Pragmatics of Computer-mediated Communication*. Berlin: Walter de Gruyter. 162-189.

Yus, F. 2011. *Cyberpragmatics: Internet-mediated Communication in Context*. Amsterdam and Philadelphia: John Benjamins

Yus, F. 2014. “Not all emoticons are created equal”. *Linguagem em (Dis) curso*, 14(3), 511-529. http://www.scielo.br/scielo.php?pid=S1518-76322014000300511&script=sci_arttext. <http://dx.doi.org/10.1590/1982-4017-140304-0414>