Discovering language variation in *WhatsApp* text interactions

Carmen Pérez-Sabater
Universitat Politècnica de València
España

Carmen Pérez-Sabater: Department of Applied Linguistics, Universitat Politècnica de València, España.
| Correo electrónico: cperezs@idm.upv.es

ONOMÁZEIN 31 (junio de 2015): 113 - 126
Carmen Pérez-Sabater
Discovering language variation in *WhatsApp* text interactions

114

Abstract

This paper reports the process and outcomes of an instructional project to familiarise postgraduate students with research. The study presents *Mobile-Var*, a project which complies with the philosophy of the Bologna Declaration regarding students' acquisition of research competences. By studying language variations in *WhatsApp* text interactions, participants were encouraged to develop innovative ideas to make a tangible contribution to the field of language and Computer-mediated Communication. Specifically, the research undertaken consisted in analysing new trends in digital communication by observing if the language used in *WhatsApp* text interactions constituted a new language variety. Authentic messages in Catalan, English and Spanish were studied from the perspective of the strategies of 'oralised' written texts addressed by Yus (2011) and other scholars. The debriefing session revealed that age is a determining factor in the pervasive use of non-standard language. From an educational perspective, students' involvement in a project for research purposes allowed them to develop the strategies needed to carry out research in tertiary education, such as critical thinking, data collection and interpretation, and collaboration with others. Consequently, the relevance of the study is twofold: on the one hand, it increased the competence of postgraduate students in discourse studies research; on the other hand, it connects with Baron's (2013) proposal regarding the fundamental role of mobile communication in the way written and spoken language is used today, given the current profusion of mobile communication technologies.

Keywords: development of research competence; language variation; new media research; mobile text-based interactions; *WhatsApp* text interactions.

ONOMÁZEIN 31 (junio de 2015): 113 - 126
Carmen Pérez-Sabater
Discovering language variation in *WhatsApp* text interactions

115

1. Theoretical justification for this project

1.1. Research Competences

Over the past millennia, human beings have developed strategies to better understand worldly phenomena (Leedy & Ormrod, 2010). However, the need to understand and contextualise new data accurately and quickly in the modern world seems to have created a need to further develop these strategies. The acquisition of research skills at tertiary level, especially in social sciences, is now a critical asset, as Payne and Israel (2010) suggest. As a result, organizations worldwide are trying to promote the incorporation of research competences in several domains, not only in postgraduate studies but also at other educational levels including graduate and secondary studies (see, for example, Payne & Israel, 2010). For instance, in April 2009 the Bologna Declaration stated that "the acquisition of research competences should be increased, research should be better integrated within doctoral programmes and the career development of early stage researchers should be made more attractive". This is the context in which a project called *MobileVar* was devised to enhance the research competences of students taking a master's degree subject entitled "Digital Communication". The project aimed to foster some of the abilities needed to undertake research as formulated by Beyer (1985), Leedy and Ormrod (2010), and Payne and Israel (2010), such as critical thinking, data collection and interpretation, and collaboration with others. These abilities involve the following:

a) Critical thinking, i.e., the ability to evaluate information or arguments in terms of their accuracy and worth (Beyer, 1985). This generally requires a dispositional component to prepare learners for demanding cognitive work. It may involve aspects such as verbal reasoning, argument analysis, decision making, and critical analysis of prior research (Halpern, 1998, 2002; Leedy & Ormrod, 2010).

These aspects should be addressed at the beginning of the research.

b) Data collection and interpretation. This is a critical issue, especially in Computer-mediated Communication (CMC) approaches, since online interactions may be private and so should be handled sensitively and responsibly. In this process, the involvement of many participants may be needed to provide enough data to yield statistically significant findings (Dürscheid & Stark, 2011).

The use of computer analysis tools to process these corpora is common practice in CMC discourse analysis. In the *MobileVar* project, some elements which played an important role in data collection need further explanation:

· Computer analysis tools. Corpus-based research uses computers to analyse naturally-occurring texts (corpus) to perform both quantitative analyses on structure and on information regarding language use. These tools have great potential for linguistic studies. In CMC, corpus-based research can not only answer research questions but can also find patterns not considered in previous studies (Dürscheid & Stark, 2011). In addition to pure linguistic studies, corpus analysis is now being used in a pedagogical context (Scott, 2009). The implementation of a corpus-based approach in education has gained prominence over the past few decades. In this regard, Conrad (1999) posited that the use of corpora and software programs, whose primary purpose is to display words or simple grammatical items with their surrounding context, "facilitates the use of authentic language, makes students more active and independent analyzers of language, and provides empirical evidence about language use" (Conrad, 1999: 2). Likewise, quantitative statistical studies drawn from corpus analysis can allow

ONOMÁZEIN 31 (junio de 2015): 113 - 126
Carmen Pérez-Sabater
Discovering language variation in *WhatsApp* text interactions

116

students to identify major patterns in language, such as to what extent a feature is used, or how common one feature is compared to another. A data-driven language method can empower learners to find out things for themselves (Scott, 2009).

- Teaching students to exploit corpus data. Fligelstone (1993: 98) groups corpus-related activities into three categories which involve a direct use of corpora in teaching:

  - "Teaching about (i.e., teaching about corpora/corpus linguistics)

  - Teaching to exploit (i.e., teaching students to exploit corpus data)

  - Exploiting to teach (i.e., exploiting corpus resources in order to teach)".

As will be detailed below, teaching to exploit is the approach followed in this project, which consisted in teaching students to examine a corpus made up of *WhatsApp* messages in a CMC environment.

c) Collaboration with others. Leedy and Ormrod (2010) established that bringing several professionals together in a collaboration can increase the cognitive resources available to address the research problem, through their different views, and to find greater meaning in the data obtained.

1.2. Mobile Communication

In addition to the research competences described, this project is deeply rooted in new trends in CMC discourse studies. The new technological devices which emerged at the dawn of this century have caused people to start using mobile phones differently (De Souza, 2013), specifically in terms of frequency, purpose and patterns of engagement. The incorporation of Internet services to smartphones has been crucial in this process. At present, as Horst (2013) indicates, mobile phones and mobile media are broadly recognised as one of the most global

communications platforms. The global use of these platforms is seen to facilitate emergent social/spatial interactions that "are either unprecedented or occurring on a scale not hitherto witnessed" (Lim, 2013: 97). This is the reason why the affordances and the language used in mobile interactions have attracted the attention of scholars such as Naomi Baron, who explains that "the field of mobility and language is highly fluid at this moment in our history and therefore especially promising for study" (2013: 139). However, the field is evolving so quickly that these applications are in some way provisional and, so, difficult to study consistently. Over the last few years many online interaction platforms have been developed while others have disappeared.

With regard to the current state of research on mobile written interactions, different approaches have addressed everyday communication (see, for example, Baron, 2008; Dürscheid & Stark, 2011; Ling & Baron, 2013). While the earliest studies focused mainly on the affordances and linguistic characteristics of traditional instant messaging services such as Short Messaging Service (SMS) or Instant Messaging (IM), the publication of the first volume of the *Journal of Mobile Media & Communication* in 2013 offered these studies a wider scope. For example, Linke (2013) proposed the integration of mobile, online media and communication in a common research perspective on the practices of everyday communication. On the whole, these articles highlight the importance of the field and defend academic research on mobile media and communication as a well-defined field in its own right (Campbell, 2013). The peculiarities of the language used in this new communication environment present opportunities for the development of a line of research in this field that may make a rich contribution to the analysis of language variation, as written mobile communication appears to have generated a new language of abbreviations, acronyms, word combination and punctuation (e.g., Dürscheid & Stark, 2011).

ONOMÁZEIN 31 (junio de 2015): 113 - 126
Carmen Pérez-Sabater
Discovering language variation in *WhatsApp* text interactions

117

In general, language variation can be observed in different domains of language and linguistics, as languages may vary from one communicative situation to another and from one social group to another. Online communication has increased the recent interest of linguists in the field. Androutsopoulos (2011), for example, in his overview of the state of the art, emphasises the important role of language variation in new media discourse studies. Indeed, the role of language variation in online communication is particularly relevant since contributors to a particular social network tend to use unique "social languages" so as to be accepted as respected members of the group (Jones & Hafner, 2012).

Another innovative aspect of this research involved the analysis of digital discourse from the perspective of young researchers while they were being introduced to the study of new media language, a new point of view which, in Herring's (2011) opinion, should be included in CMC research. Moreover, the methodological format of the project follows one of the greatest innovations put forward by the Bologna Declaration: that courses leading to a qualification should focus on the student's acquisition of competences, rather than on the simple transfer of knowledge. Through *MobileVar*, the students enrolled in the master's degree subject "Digital Communication" were encouraged to develop research competences while analysing the language variation involved in mobile communication.

2. Language variation and mobile text-based interactions

Language variation studies have viewed CMC as the variation between speech and writing, or between standard and vernacular forms. Concerns have been raised as to what constitutes an online standard with no institutional control (Androutsopoulos, 2011). In the literature, the independent variables studied have included gender, genre, culture and region (Herring & Paolillo, 2006; Baron, 2008). In mobile language studies, specific variables such as gender and age have been proved crucial in mapping language behaviour (Baron, 2008).

This research on CMC has highlighted that the interactions through media like chats, Short Message Service (SMS) messaging, Instant Messaging (IM) and mobile-text chats contain clear oral traits along with features characteristic of written language (Baron, 2008; Yus, 2011; Jones & Hafner, 2012; Pérez-Sabater, 2012). These characteristic features include: the use of emoticons, lexicalisation of vocal sounds, orthographic and punctuation mistakes, phonetic orthography and eye dialect, abbreviations, acronyms, clippings, ellipsis, use of contractions, one-word transmissions, the use of words in other languages, absence of openings and closings, short messages, short words and short sentences.

Focusing on non-standard distinctive linguistic characteristics, such as phonetic orthography, some past research on CMC has tried to "exoticise" digital discourse, according to Herring (2011). However, as Androutsopoulos (2011) contends, the results of the studies on language variation and CMC have contributed to a "deexoticised" understanding of new media discourse. In this vein, Baron (2008) comments on the popular belief that English is in decline because of the literary practices promoted by emergent technologies. In her view, the persistent belief that "text messaging corrupts all languages" (*The Economist* 2008, in Baron, 2011: xii) is not backed up by research; her analysis of IM reaches the conclusion that "electronic language is at most a very minor dialectal variation" (2008: 163). Despite these findings, Baron (2008) regretfully states that there is an international perception that computers and mobile phones are adversely affecting everyday language. In general, this perception has been that of non-digital natives, or what Herring (2011) calls "adult lenses". However, in *MobileVar* young adults observe language variation in CMC through their own eyes during

ONOMÁZEIN 31 (junio de 2015): 113 - 126
Carmen Pérez-Sabater
Discovering language variation in *WhatsApp* text interactions

118

their initiation into digital discourse analysis through a current controversial issue.

3. Aims and scope of the instructional research project, *MobileVar*

The project's main aim was to promote the acquisition of research competences such as critical thinking, data collection and interpretation, and collaboration with others. The promotion of research competences was carried out through a cross-linguistic discourse analysis of mobile interactions and the linguistic variation implied by this new form of communication. The main purpose of the project was to provide young researchers with appropriate tools to study new trends in digital communication; participants examined the linguistic and typographic features of *WhatsApp* written messages, authentic online discourse that is familiar to them.

This platform was chosen because *WhatsApp* is currently the most popular and 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). *WhatsApp* Messenger is a proprietary, IM application for smartphones which allows users to send and receive text messages as well as images, video, and audio media. These are exchanged with individuals or groups of contacts at practically no cost, provided the users have an adequate Internet data plan for their smartphone. Despite *WhatsApp*'s current global popularity, Church and De Oliveira (2013) believe that these interactions are yet to be sufficiently addressed by academics.

4. Methodology

The basic methodological orientation of the CMC discourse used in this study is what Herring terms Computer-Mediated Discourse Analysis (CMDA), in other words "language-focused content analysis" (Herring, 2004: 343). CMDA

analysis is applied to four domains or levels of language: structure, meaning, interaction and social behaviour. In Herring's approach, some requirements must be fulfilled in order to obtain successful results; for instance, researchers should take into consideration logs of verbal interaction that must represent a sufficient corpus of data, and a research question must be posed that is, in principle, answerable.

Relying on CMDA, we focused on what Herring (2004) calls structure to study typography, orthography and morphology. The other issues addressed in this project, among those proposed by Herring, were orality and expressivity. The methods applied were structural/descriptive linguistics and text analysis, using a computer corpus analyser. *WordSmith* (Scott, 2004), a commercially available software package, was the corpus analyser that facilitated students' initial research on data for *MobileVar*. A final Likert scale questionnaire examined social behaviour and discourse styles.

Participants: As discussed above, the project is contextualised within a university subject as part of a master's degree program. 20 students participated, their ages ranged from 22 to 27 years old. An equal number of male and female students took part.

Project *MobileVar* consisted of the following steps:

1. Corpus formation:

   a. Participants downloaded *WhatsApp* chat chains following these instructions:

   i. Group A *WhatsApp* texts: formed by messages written by interlocutors who were 12-17 years old[1]. Students asked relatives or friends for authentic interactions authored by writers of these ages.

   ii. Group B *WhatsApp* texts: formed by chat messages by writers aged around 25

---

1   Following Livingstone and Helsper's (2010) study of teenagers' Internet literacy.

ONOMÁZEIN 31 (junio de 2015): 113 - 126
Carmen Pérez-Sabater
Discovering language variation in *WhatsApp* text interactions

119

(the average age of the students taking the subject "Digital Communication").

    iii. Participants had to choose, if possible, more than one language of interaction, preferably Catalan and Spanish, the two official languages in the Valencian Region.

  b. *WhatsApp* messages in English. Since most studies on CMC have been about English, this project invited native speakers of English to share their authentic interactions for analysis. In this way, the English language used in these interactions by writers belonging to the same age ranges as those used in Groups A and B could be compared with the results in Catalan and Spanish. In the discussion session, the findings in Catalan and Spanish can be related to other studies published on CMC, most of which have been produced by members of Anglophone fora about English.

2. Students formatted messages in plain text so they could be processed by the corpus linguistic analyser tool, *WordSmith*. According to the ages of the authors of the messages, two groups of corpora were formed for each language, teenagers formed Group A and adult writers formed Group B. In each corpus, names, times and dates were removed from the text. Chunks of 500 words were prepared following the strategy proposed by Biber (1998), so as to balance the corpus with the same number of words.

3. In class, these corpora were processed through *WordSmith* to measure message, word and sentence length. A discussion followed.

4. A research question was proposed to observe the corpora: To what extent does language used in *WhatsApp* text interactions constitute a new language variety? The two variables, language and age, were the focus of research.

5. The hypothesis to be confirmed was that teenagers would make a higher contribution to language variation as a result of their persistent use of non-standard language in mobile interactions.

6. Students applied the research question to confirm this hypothesis. Following the strategies proposed by Yus (2011) for 'oralised' written CMC texts, and Jones and Hafner (2012), students circled the following aspects of orality and expressivity in their corpora:

  a. emoticons,

  b. lexicalisation of vocal sounds,

  c. orthographic mistakes,

  d. phonetic orthography and eye dialect,

  e. abbreviations, acronyms and clippings,

  f. words in other languages.

7. A debriefing session focused on oralised communication as a language variety in the languages studied. Variation across age was also discussed in this debriefing session.

8. Finally, the students collaborated in class to devise an example of a Likert scale questionnaire as a tool for measuring media and social effects of mobile text interactions.

With regards to privacy, it is important to note that the corpora were formed by real *WhatsApp* message interactions. *WhatsApp* chat conversations are stored in users' mobile phones and can be forwarded to an email account as an attachment, thus preventing transcription errors. The conversations held were private; therefore, to prevent participants being reluctant to share them in class for analysis, names and sensitive information were removed from the chat conversations before they were examined.

5. Debriefing: processing and interpreting data

The data shown below summarise the results obtained by the students from their research on the *WhatsApp* messages written by

ONOMÁZEIN 31 (junio de 2015): 113 - 126
Carmen Pérez-Sabater
Discovering language variation in *WhatsApp* text interactions

120

participants in different age ranges and in the three languages studied. Each student incorporated their corpus of 500 words. The total amount of words processed was 41,000. Table 1 shows the results obtained after processing the corpora using *WordSmith*.

From these data, the most significant result is that teenagers (Group A) favoured shorter messages but longer words, mainly because they frequently resorted to reduplication as we will discuss below. Their sentences were also much shorter than those of adults.

To respond to the research question posed, let us refer to some examples of the most relevant results circled by the students in their corpora with regard to emoticons, lexicalisation of vocal sounds, orthographic mistakes, phonetic orthography and eye dialect, abbreviations, acronyms, clippings, and words in other languages.

a. Emoticons

Emoticons, i.e., iconic compositions of characters (Yus, 2011), showed the following occurrences per 500 words (cf. table 2).

In Catalan, the younger participants' corpus contained an abundant use of emoticons. Meanwhile in English, participants often employed emoticons but sometimes also showed a preference for expressions that lexicalise vocal sounds such as *hahaha*, *woops*, or *cool* instead of an emoticon. *X*, a traditional substitute for a kiss in written correspondence, was the most frequently used emoticon. Texts in Spanish, on the contrary, displayed a lower use of emoticons in both groups to express laughter, as authors preferred lexicalisation of vocal sounds to the use of these iconic compositions.

b. Lexicalisation of vocal sounds

Teenagers writing in Catalan (Group A) as well as adult writers (Group B) made noteworthy use of lexicalisation of vocal sounds to express laughter but without using a consistent spelling pattern: *jaajajajaja*, *jaja*, or *jeje*. These lexical patterns for laughter expressions were also frequently noticed in *WhatsApp* messages in English, although also without consistent spelling. These spellings were, for example, *haha*, *ha-*

TABLE 1

Language variation across Catalan, English and Spanish

	Catalan		English		Spanish	
	Group A	Group B	Group A	Group B	Group A	Group B
Mean message length in words	4.91	8.32	4.81	10.23	4.72	6.23
Mean word length in characters	4.27	3.68	3.80	3.67	3.97	3.65
Mean words per sentence	4.14	5.25	4.62	4.97	5.80	5.52
Sentences per 500 words	117	96	108	97	87	89

TABLE 2

Emoticons per 500 words

	Catalan		English		Spanish	
	Group A	Group B	Group A	Group B	Group A	Group B
Emoticons in 500 words	11	6	11	10	5	5

ONOMÁZEIN 31 (junio de 2015): 113 - 126
Carmen Pérez-Sabater
Discovering language variation in *WhatsApp* text interactions

121

haa or *hehe*. As for teenagers in Spanish, they made an abundant use of them: *ajajajjajajaj, uajjjajajjajaja, ajajajja, jaja, jeje, jajajaja* or *jajajaj*; this lexicalisation was preferred to the use of emoticons to express the same feeling. Adult writers in Spanish (Group B), on the other hand, used very few examples of lexicalisation to express laughter.

c. Orthographic mistakes

In Catalan, teenagers' orthographic mistakes were cases of a systematic lack of accents, hyphens and apostrophes: *es* instead of *és* [is], *diro* instead of *dir-ho* [say it]; *dirli* instead of *dir-li* [tell him/her]. A total absence of apostrophes was also the case in English in both groups. Teenagers writing in Spanish (Group A) demonstrated a systematic lack of accentuation and made orthographic mistakes of the type *francess* instead of *francés* [French].

d. Phonetic orthography and eye dialect

Intentional variations in orthography were persistently observed in young writers in Catalan (Group A): *Carmeeen* [the female name Carmen], *siiiiii* for *si* [yes], *oleee* for *ole* [hurrah] or *guachap* and *wasap* for *WhatsApp*. In both groups in English, phonetic orthography and eye dialect were used in *hellooooooo, yessssss, okkkkk, fuckkkk, noooo. Ok* was sometimes spelt *okey* or *okipokiii* (for *okey-dokey*). Teenagers writing in English frequently used dialectal forms such as *hafta* [have to], *wanna* [want to] or *gonna* [going to].

Teenagers used phonetic orthography in *aaaah* [ah], *uapaah* (guapa) [pretty] in Spanish. Eye dialect was sometimes employed such as *depilah'* for *depilar* in this example: *despues de depilah'* [after waxing]. Conversely, adult writers in Spanish (Group B) rarely used them.

e. Abbreviations, acronyms and clippings

With regard to Catalan, teenagers made constant use of abbreviations such as *bs* for *besets* [kisses]. In English we found, for example, *fab* [fabulous], *prob* [problem], *u* [you] or *2moro* [tomorrow] in both groups. Adult writers in Spanish made use of abbreviations, acronyms and clippings profusely, but the use of abbreviations was more abundant in teenagers' interactions in Spanish through the omission of vowels 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]. Grapheme substitution, as in *k* or *q* for *que* [that], was a pattern of use both in Catalan and Spanish.

f. Words in other languages

In Catalan, the messages in Group A and Group B contained many words in Spanish because of the widespread bilingualism among speakers in the Valencian Region. Examples of words regularly used in Spanish and with a Spanish spelling were *chata* [babe] or *cariño* [sweetie]. Similarly, many words in Catalan were present in interactions in Spanish such as *adéu* [bye] or *fins demà* [see you tomorrow]; this linguistic practice may correspond to a more intimate relationship with the interlocutor, such as in the case of the closing *bona nit* [good night]. Moreover, as Lan (2000) suggests, the use of *code switching* by using words from another language is a politeness strategy that reinforces community links.

On the other hand, words in English were commonly used by all the groups. The most popular words in English were *please* and *thanks*, whose spelling was often adapted to Spanish or Catalan: the phonetic spellings *plis* for *please* or *zanks* for *thanks*. The widespread use of "ok" is also worth mentioning, as its use is integrated into everyday Catalan and Spanish; it is a commonly accepted Anglicism (Otheguy, 2001). Other words in English detected such as *of course* may act as a marker of elite social identities, as Herring (2011) suggests.

ONOMÁZEIN 31 (junio de 2015): 113 - 126
Carmen Pérez-Sabater
Discovering language variation in *WhatsApp* text interactions

122

Conversely, the *WhatsApp* corpus authored by native speakers of English regularly included words in Spanish. This fact was unsurprising as the authors live in Spain; consequently the use of *tapas*, *fiesta* or *calle* [instead of *street*] was normal practice.

In summary, some significant results drawn from the linguistic study should be highlighted. Firstly, the data circled and processed by the participants on the oralised characteristics of interactions confirmed our initial hypothesis that teenagers contributed more than adults to language variation as a result of their generalised use of non-standard language in mobile messages. Also, it is interesting to note that, although the analysis proved that these interactions contained a high number of non-standard language uses, the conclusion reached in the debriefing was that the participants would not call it a language variety, despite the use of creative language. In fact, the young researchers suggested that the language used simply exhibited a less traditional writing style that required the gradual emergence of norms. These conclusions agree with those reached in the early stages of CMC research by Ferrara et al. (1991) on interactive written discourse and with Baron's study of Instant Messaging in 2008.

Secondly, it is worth mentioning that, although the *WhatsApp* application does not limit the number of words sent, unlike SMS and other platforms such as *Twitter*, the number of words per message was remarkably short in all the groups analysed. Moreover, word and sentence length were also shorter than the normal average word and sentence length found in written discourse (on average, a written sentence in English was 20 words long in the 1980s, as indicated by Baron, 2000).

Thirdly, the comparison between the three languages produced some unforeseen results. On the one hand, young writers using Catalan resorted regularly to very short sentences, the shortest of the three languages, unlike the results from previous studies carried out by Montero-Fleta et al. (2009), which found that writers using Catalan preferred to use long sentences. On the other, adults writing in English used longer messages than the other interactants, contrary to results from previous studies on online communication (e.g., Montero-Fleta et al., 2009). In this previous cross-cultural examination of online written messages, writers using English consistently displayed a preference for much shorter sentences than writers using Romance languages such as Catalan and Spanish.

Finally, some considerations need clarification. The data obtained from the participants' research cannot be linguistically conclusive due to the small amount of data processed and the fact that the messages analysed could be biased because they were not selected randomly. I agree with Baron (2008) that they simply constitute a convenience sample available for study. Nevertheless, I believe that the study has been successful since the methodology followed has given the participants a first glimpse into the study of new written interactions in mobile applications for smartphones.

6. Developing a Likert scale questionnaire

Once the debriefing process was completed, a class activity was designed to focus on developing the students' ability to collaborate with others (Leedy & Ormrod, 2010). To this end, the students devised a group research proposal, a questionnaire, which included a battery of items with 5-point Likert-type response options on *WhatsApp* interactions. The objective of the proposal was to raise students' awareness of the specific research purposes and structure of a collaborative scientific project where critical thinking is in play: argument analysis, decision making and critical analysis of prior research. They were instructed on the characteristics of Likert scales, the number of categories, and the clear definitions of constructs, among other items. The research proposals developed were

ONOMÁZEIN 31 (junio de 2015): 113 - 126
Carmen Pérez-Sabater
Discovering language variation in *WhatsApp* text interactions

123

reviewed by the teacher. A student-centred methodology was then put into practice that engaged the students in an inquiry-focused approach to learning.

To conduct this Likert scale questionnaire on *WhatsApp* interactions, the students created their own surveys to collect the information they required. The planning of the research strategy started with a brainstorming phase where the research question was the centre of a debate. Then they discussed some hot topics in CMC, such as the panic expressed in the popular press that new media is compromising and corrupting language standards (Baron, 2011). Moreover, echoing Daft and Lengel's (1983) classical distinction between "rich" and "lean" media, the class discussed the concept of media richness and transaction costs, which implies that the richer the medium, the higher the transaction costs. The participants agreed that *WhatsApp* interactions were an example of "lean" media with lower transaction costs that provide advantages such as optional salutation formulae, the optional nature of an immediate answer or the possibility of multitasking. All these social affordances were considered by Jones and Hafner (2012) as one of the main reasons for the great success of mobile interactions.

These discussions provided the students with the information they needed to frame the inquiry. They worked in groups of four and suggested items for the survey based on a Likert scale, which asked respondents to report on these issues and their own attitudes, opinions or beliefs about media effects and user effects. A Likert scale questionnaire was proposed by each group with different options indicating their level of agreement or disagreement with the statements proposed. The point system agreed was: strongly disagree, disagree, neither agree nor disagree, agree, and strongly agree. The five different surveys produced, one per group, were examined and a consensus was reached among the groups on the categories and descriptions.

Finally, the teacher synthesised the discussion, helping the students identify criteria. A refinement of the statements was proposed to adjust them to the newest trends on digital-based written interaction put forward by CMC researchers (see, for example, Jones & Hafner, 2012). Appendix A contains the agreed version of the survey.

7. Final remarks: Implications of *MobileVar*

As Baron (2011: XIV) posits: "in thinking about language used with new technologies, the relationship between surface phenomena and root causes may be less than obvious. As with any scientific venture, the study of new media language demands both creative sleuthing and hard work". Teaching students to exploit mobile interactions corpus data involved hard work, which brought about fruitful discussions on the orality and expressivity of mobile interactions from the perspective of young researchers.

Through the implementation of *MobileVar*, the students were personally involved in hands-on research that facilitated specific linguistic analysis on cross-cultural variation in *WhatsApp* interactions and the particular features that characterise them.

With regards to the acquisition of research competences, critical thinking was enhanced, as the students carried out decision-making strategies by identifying and evaluating alternatives to a successful outcome. Moreover, the critical analysis of prior research favoured their evaluation of the research methods involved in *MobileVar* and the relevance of the conclusions drawn. The participants had to respond to a research question about which they had limited knowledge; they had to find information on the topic and report on their findings. Furthermore, the students' involvement in the process of designing a survey for research purposes involved their having an understanding of what must be done to fulfil the expectations and aims of the instructional research project. With the survey as a guide, the students learned to focus research on

ONOMÁZEIN 31 (junio de 2015): 113 - 126
Carmen Pérez-Sabater
Discovering language variation in *WhatsApp* text interactions

124

the sociolinguistics of *WhatsApp* interactions in progress.

As for collecting and interpreting data, the students assembled corpus data in order to answer the research question on the specific variety of the language under study. The corpora of *WhatsApp* interactions collected provided concrete evidence and examples of authentic online language use in current electronic communication. Finally, collaboration to carry out research was also fostered since the students worked in groups to decide upon the available alternatives to carry out the different stages of this project. In class, they were confronted with a learning situation that they may encounter in their future careers as researchers or in their working lives.

Overall, this paper has attempted to show the efforts being made in higher education to teach research competences in a specific subject of the students' curriculum in a process of discovery and synthesis. The approach used in *MobileVar* has tried to comply with the philosophy behind the Bologna Declaration regarding the student's acquisition of research competences. Future research may address practising other research competences such as deductive logic or inductive reasoning, as suggested by Leedy and Ormrod (2010).

8. References

Androutsopoulos, Jannis, 2011: "From language variation to heteroglossia in the study of computer-mediated discourse" in Crispin Thurlow & Kristine Mroczek (eds.): *Digital Discourse: Language in the New Media*, Oxford and New York: Oxford University Press, 277- 298.

Baron, Naomi S., 2000: *Alphabet to email: How written English evolved and where it's heading*, London and New York: Routledge.
Baron, Naomi S., 2008: *Always On. Language in an Online and Mobile World*, Oxford: Oxford University Press.

Baron, Naomi S., 2011: "Foreword" in Crispin Thurlow & Kristine Mroczek (eds.): *Digital Discourse: Language in the New Media*, Oxford and New York: Oxford University Press, XI-XV.
Baron, Naomi S., 2013: "Do mobile technologies reshape speaking, writing, or reading?", *Mobile Media & Communication* 1, 1, 134-140.

Beyer, Barry K., 1985: "Critical thinking: What is it", *Social Education* 49 (4), 270-276.

Biber, Douglas, 1988: *Variation across speech and writing*, Cambridge: Cambridge University Press.

Bologna Declaration. Leuven/Louvain-la-Neuve Communiqué of 28-29 April 2009 [http://europa.eu/rapid/press-release_IP-09-675_en.htm, date of retrieval: 27th May, 2013].

Campbell, Scott, W., 2013: "Mobile media and communication: A new field, or just a new journal?", *Mobile Media & Communication* 1, 8-13.

Church, Karen & Rodrigo de Oliveira, 2013: "What's up with Whatsapp?: Comparing Mobile Instant Messaging Behaviors with Traditional SMS" in *Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services*, ACM, 352-361.

Conrad, Susan M., 1999: "The importance of corpus-based research for language teachers", *System* 27, 1-18.

Daft, Richard L. & Robert H. Lengel, 1983: *Information richness. A new approach to managerial behavior and organization design* (No. TR-ONR-DG-02) [http://www.dtic.mil/dtic/tr/fulltext/u2/a128980.pdf, date of retrieval: July 5th, 2013].

De Souza e Silva, Adriana, 2013: "Location-aware mobile technologies: Historical, social and spatial approaches", *Mobile Media & Communication* 1 (1), 116-12.

ONOMÁZEIN 31 (junio de 2015): 113 - 126
Carmen Pérez-Sabater
Discovering language variation in *WhatsApp* text interactions

125

DREDGE, Stuart, 2014: "Why is WhatsApp worth up to $19bn to Facebook?", *The Guardian*, 20 February.

DÜRSCHEID, Christa & Elisabeth STARK, 2011: "SMS4science: An international corpus-based texting project and the specific challenges for multilingual Switzerland" in Crispin THURLOW & Kristine MROCZEK (eds.): *Digital Discourse: Language in the New Media*, Oxford and New York: Oxford University Press, 299- 320.

FERRARA, Kathleen, Hans BRUNNER & Greg WHITTEMORE, 1991: "Interactive written discourse as an emergent register", *Written Communication* 8 (1), 8-34.

FLIGELSTONE, Steve, 1993: "Some reflections on the question of teaching, from a corpus linguistics perspective", *ICAME Journal* 17, 97-109.

HALPERN, Diane F., 1998: "Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring", *American Psychologist* 53 (4), 449-455.
HALPERN, Diane F., 2002: *Thought and knowledge: An introduction to critical thinking*, London and New York: Routledge.

HERRING, Susan C., 2004: "Slouching toward the ordinary: Current trends in computer-mediated communication", *New Media and Society* 6, 26-36.
HERRING, Susan C., 2011: "Commentary" in C. THURLOW & K. MROCZEK (eds.): *Digital Discourse: Language in the New Media*, Oxford and New York: Oxford University Press, 340-347.

HERRING, Susan C. & John C. PAOLILLO, 2006: "Gender and genre variation in weblogs", *Journal of Sociolinguistics* 10 (4), 439-459.

HORST, Heather A., 2013: "The infrastructures of mobile media: Towards a future research agenda", *Mobile Media & Communication* 1, 147-152.

JONES, Rodney H. & Christoph A. HAFNER, 2012: *Understanding Digital Literacies*, London and New York: Routledge.

LAN, Li, 2000: "Email: a Challenge to Standard English?", *English Today* 16 (4), 23-29.

LEEDY, Paul D. & Jeanne E. ORMROD, 2010: *Practical research*, ninth edition, Boston: Pearson education international.

LIM, Sun S., 2013: "On mobile communication and youth 'deviance': Beyond moral, media and mobile panics", *Mobile Media & Communication* 1 (1), 96-101.

LING, Rich & Naomi S. BARON, 2013: "Mobile phone communication" in Susan C. HERRING, Dieter STEIN & Tuija VIRTANEN (eds.): *Pragmatics of Computer-Mediated Communication*, Berlin: Mouton de Gruyter, 199-216.

LINKE, Christine, 2013: "Mobile media and communication in everyday life: Milestones and challenges", *Mobile Media & Communication* 1 (1), 32-37.

LIVINGSTONE, Sonia & Ellen 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.

LORENZO-DUS, Nuria & Patricia BOU-FRANCH, 2013: "A cross-cultural investigation of email communication in Peninsular Spanish and British English emails", *Pragmatics and Society* 41 (1), 1-25.

MONTERO-FLETA, Begoña, Anna I. MONTESINOS LÓPEZ, Carmen PÉREZ-SABATER & Ed TURNEY, 2009: "Computer mediated communication and informalization of discourse: The influence of culture and subject matter", *Journal of Pragmatics* 41 (4), 770-779.

OTHEGUY, Ricardo, 2001: "Simplificación y adaptación en las lenguas en contacto" in II Congreso

ONOMÁZEIN 31 (junio de 2015): 113 - 126
Carmen Pérez-Sabater
Discovering language variation in *WhatsApp* text interactions

126

Internacional de la Lengua Española. [http://cvc.cervantes.es/obref/congresos/valladolid/ponencias/unidaddiversidaddelespanol/3elespanolenlosEEUU/otheguyr.htm, date of retrieval: July 17th, 2014].

PAYNE, Jarrod & Nicky ISRAEL, 2010: "Beyond teaching practice: Exploring individual determinants of student performance on a research skills module", *Learning and Individual Differences* 20 (3), 260-264.

PÉREZ-SABATER, Carmen, 2012: "The linguistics of social networking: A study of writing conventions on Facebook", *Linguistik Online* 56 (6), 81-96.

PLESTER, Beverly & Claire WOOD, 2009: "Exploring relationships between traditional and new media literacies: British preteen texters at school", *Journal of Computer–Mediated Communication* 14 (4), 1108-1129.

SCOTT, Mike, 2004: *WordSmith Tools Version 4*, Oxford: Oxford University Press.
SCOTT, Mike, 2009: "Corpora and language teaching" in Anke LÜDELINK & Merja KYTÖ (eds.): *Corpus Linguistics*, Berlin: Mouton de Gruyter, 112-131.

YUS, Francisco, 2011: *Cyberpragmatics: Internet-mediated communication in context*, Amsterdam and Philadelphia: John Benjamins.

9. Appendixes

Appendix A. Survey about *WhatsApp* use

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1. The language in mobile chat interaction represents a different language variety					
2. In these interactions, time and space affect written language					
3. I think that mobile chat messaging is corrupting language					
4. The age of the interlocutor affects the use of standard language					
5. Every time I send my message I review its content					
6. If the message is sent to a whole group I review my writing more carefully					
7. I use predictive text software to input my messages					
8. I use words in other languages					
9. I frequently multitask while I am sending a mobile chat message					
10. In most cases, I prefer sending a *WhatsApp* message instead of calling because it involves less social effort					