The case of Digital Writing in Instant Messaging:
When cyber written productions are closer to the oral code than the written code

Tonia Lanchantin[*1], Aurélie Simoës-Perlant[1] and Pierre Largy[1]

[1]Université de Toulouse 2 –
Le Mirail, Toulouse (France)

ABSTRACT

The use of New Information and Communication Technologies, or NICTs, has deeply changed the traditional reading and writing practices. It thus seems necessary to provide a definition of Digital Writing in Instant Messaging (DWIM) to better understand its grammatical, lexical and syntactic characteristics (these two last components define the traditional characteristics of both oral and written codes). Thirty-two French-speaking students around the age of 13 who were enrolled in 8th grade produced one hour of DWIM productions on an instant messaging website in groups of two. They were able to use as many cyber languages as they wanted (we preferred the expression *digital writing*). This corpus helped to understand that this written structure is closer to the oral code than the written code (the studied population developed their language skills in constant contact with the written in its dual form). Indeed, we showed for instance that users of DWIM sometimes produced repetitions (whereas it is forbidden in traditional writing), never use subject-verb inversions in interrogative sentences, can replace punctuation with emoticons, or used undefined deixises in their sentences. We have also been able to show that having traditional reading and writing habits is not sufficient to create a predisposition towards the use of the DWIM code.

---

1. Introduction

This work aims at identifying what a user needs to produce digital writing (which is the written form that appears on screens). Most studies on this topic have been carried out on the impact of its use on the quality of spelling and more recently on grammar

---

[*]Corresponding Author:
Tonia Lanchantin,
Maison de la recherche,
Laboratoire de Psychologie du Développement et Processus de Socialisation, E.A. 1687,
Université de Toulouse 2 – Le Mirail, Toulouse,
E-mail: tlanchan@univ-tlse2.fr

production. It seems therefore necessary to identify what is involved in this kind of production to then specify the structure of DWIM in order to promote a better understanding of the phenomenon. Some tools, with which we are all familiar, would help to define DWIM: we are talking about grammatical, lexical and syntactic characteristics. Traditional writing is indeed defined through all these characteristics; that is why we chose to use them to define DWIM also. Doing such work will contribute toward understanding of the impact of the use of DWIM on the quality of spelling. If we are able to show that DWIM is somehow closer to the oral code than the written code, we would be able to explain why DWIM would contain a lot of phonetic forms.

Firstly, we introduced (a) the description of literacy and semiotic tools that help written production to then provide (b) a definition to reading and writing in NICTs through a description of the different media. We also showed (c) how users adapt linguistic rules to digital written production thanks to the analysis of corpuses and typologies, and (d) how the oral code and written code are built.

Secondly, we chose to provide a definition to the DWIM code thanks to the descriptions that have been introduced in the first part, by analyzing a corpus resulting from the digital written productions of French adolescents enrolled in 8th grade. We then wanted to show if traditional reading and writing habits had a link with the use of the DWIM code.

1.1 Between literacy and semiotic tools

Memory, language, phonological awareness, and other abilities allow nowadays the transmission of information via semiotics, which is a result of the enormous progress that has been made throughout human evolution. Men and women developed sufficient memory capacities to then be able to create a structured language during prehistoric times. This memory was firstly episodic since they only remembered the immediacy of the events. It later became semantic, which enabled them (a) to create a language made of abstract signs, and (b) to keep everything in mind. It helped them to identify and to name referents, which were no longer required to be present when they actually had to use language (Defays & Englebert, 2009).

They have successfully developed language skills thanks to the use of the other semiotic tools and to the knowledge of their community partners (De Saussure, 1979). All these abilities have given rise to literacy, which includes anything that allows their effective integration into a social group by means of reading, writing (Burns, Espinosa & Snow, 2003; Tran, Trancart & Servent, 2008), spelling and grammar.

It is just the way things are for children who open doors to reading and writing today; they build progressively their literacy skills (Bazerman, 2006; Privat, 2006) by relying on their memory, as Prehistoric men and women used to do at the very beginning. This internal tool allows children to learn a code so that they become able to communicate with others. Memory gives all children the opportunity to learn and hold linguistic information emanating from their social group. In this sense, the members of this social group will help them to strengthen his phonic – or phonological – awareness in the very first years of his life (Demougin, 2003; Plester & Wood, 2009). They become able to break words down into phonemes, which are the smallest sound unit of language, and then to play with it.

They start to translate phonemes into graphemes when it is time to learn how to read and write (Cellier, 2003b). They have to understand that the oral encode the written to do so, thanks to their phonological awareness. Only then children can open doors to reading and writing and integrate a wide range of literacy skills. Children will therefore understand that it is capital to speak via a comprehensible code that includes spelling and grammatical rules to ensure comprehension.

Children will always develop all their literacy skills at the same time: they will never learn how to read and write separately but all together. This postulate rely on the interactive model of Bouillaud, Chanquoy and Gombert (2007), which includes that once children have acquired reading abilities, they draw on a variety of phonological, grammatical, spelling, semantic, etc., skills to understand and/or interpret one message. These processes perform automatically in a synchronized way. Acquiring literacy skills is resolutely implicit (see infra to read the definition) since children unintentionally learn these abilities. This is how they begin to produce correct sentences while being unable to explain how they proceed.

Some authors devoted part of their work to studying artificial grammars (e.g., Fischer, 1997; Reber, 1967; Witt, 2010), and showed it is related to implicit learning. They mentioned that their participants were able to say if a sentence was grammatically correct or not without knowing the artificial language they were using (the authors get the same results for both children and adults). Then, when children produce utterances that do respect spelling and grammatical rules, it is the result of implicit learning.

The adolescents born in the 1980s and 1990s used NICTs after having acquired these traditional skills, contrary to those of the present generation who only knew writing in its dual form. We have to ask if written production might have changed with the advent and the massive use of NICTs.

1.2 Reading, writing and NICTs

It seems impossible nowadays to consider traditional reading and writing apart from its digital equivalents. Children, adolescents and adults use digital writing and NICTs almost every day. Some authors like Mairesse, Cette and Kocoglu (2000) or Bekhuis (2007) define NICTs by making a distinction between material and virtual tools such as computers and mobile phones in the first case or software programs or the internet in the second. Blanc (2012) mentions that adolescents send about 83 SMS per day (or "Short Message System", Panckhurst, 2009).

Little studies related to instant messaging are available since the most part was focused on SMS. The authors who worked on the subject were furthermore interested in its social aspects rather than its linguistic production (e.g., Bonetti, Campbell & Gilmore, 2010; Fukking & Hermanns, 2009; Sjöberg, 2003). This linguistic aspect only gets the interest of psychology in the past decade. If we consider instant messaging, works are much rarer, probably because these two systems have a lot in common. The user needs a keyboard to send written messages, and can express his emotions through smileys in both cases. However, mobile telephony offers easier communication than does instant messaging since anyone can correspond from anywhere at every hour of the day, contrary to its computer equivalent. But the space used to communicate is much more restricted on the screen of a mobile phone and there is a larger choice of emoticons in instant messaging. The keyboard of a computer is furthermore composed of keys made for fingertips whereas their size is much more little on mobile phones, which explains why we use the thumb (or both in the case of androids) to write SMS and our ten fingers to produce an instant message (Ling & Baron, 2007). Linguistic consequences include a larger ability in writing a message on computer rather than mobile phones.

Fukking and Hermanns (2009) chose to compare two media of communication, which are landline phone and instant messaging, to know which one would be favored in an association of help for youth. They showed that adolescents chose to communicate through instant messaging rather than landline phone. This medium has been prefered to SMS, less appropriate to long exchanges. Indeed, corpuses of SMS consist mainly of a few sentences whereas corpuses of ENMI are much longer. Pétillon (2006) mentioned that the average number of character per SMS was approaching 160 (unlimited SMS package was not as usual as it is today). In DWIM, the average is not based on the number of characters but on the number of lines (one per exchange, Anis, 2003). Indeed, SMS were not free at the very beginning, contrary to instant

messaging, which is free since ever. Something unexpected happened then: some SMS users reduced the quantity of characters, even if it means they have to distort spelling of words. If the economical aspect was first put forward, it does not seem to be the only reason why there has been a considerable increase of distorted spelling in digital language. Indeed, if it had been the only reason, this characteristic would have never appeared in free NICTs media as instant messaging.

But the abbreviations and other components of SMS language do appear on any screen, which means that calling up the economic criteria is not anymore sufficient. It then seems necessary to rely on the work done in the SMS area to help defining DWIM since research on this field is rarer and focuses on the quality of spelling production and much less on lexical, syntactic or grammatical aspects.

1.3 Corpuses and typologies

Many authors tried to make corpuses and typologies in digital writing. Such is the case of Panckhurst (2009) who invited French-speaking Master students to list the different kinds of written forms – or modifications – that appear in digital writing (such as abbreviations, "consonant skeletons", Anis, 2003, etc.). She worked on an "eSMS" website, which allows users to send SMS online. She showed that DWIM is a combination of both traditional and digital writing that her participants have deliberately produced. We could infer that using both kinds of written forms would have an impact on literacy skills. This is the statement of Plester, Wood and Joshi (2009) who showed that English-speaking students aged from 10 to 12 were more able to develop their phonological awareness or to choose the correct synonym when they were in possession of a mobile phone, lent for the occasion, than others who do not. To do so, they tested all their literacy skills, where Bouillaud et al. (2007) focused on the spelling performances of French students enrolled in 4th, 7th and 9th grades. Participants were invited to write two dictations down in traditional writing first and then in digital writing.

The authors established a correlation between digital world knowledge (thanks to a questionnaire) and the quantity of oralisations[1] used. It means that the more children use digital writing, the more he uses modifications in DWIM. But these results are only valid for 4th grade students since other results were not significant. Furthermore, the authors had difficulties in knowing if some particular forms were real modifications or misspellings (e.g., is *to* a misspelling or a modification in the sentence *I lov u to*?).

---

[1]We prefer the term of *modifications* already mentioned since all the linguistic items that appear in digital writing are not only a matter of oralisation process.

Febvrel and Hureau (2008) had to deal with the same difficulty. Concerning their methodology, they also invited their participants to write two dictations down and to answer to a 30 items questionnaire categorized in four axes: "the use of computers, blogs and MSN®; general use of mobile phones; reading habits; writing habits". Although their results were found to be unusable, Febvrel and Hureau (2008) noticed that students with a correct spelling level did not pay enough attention to the quality of their spelling production.

This is also true for Plester, Wood and Bell (2008) since these authors reported the existence of child metalinguistic skills in every kind of written production. To do so, Plester et al. (2008) measured the relationship between the ability to use modifications and traditional writing by asking their English-speaking participants aged from 10 to 12 to provide a translation from traditional writing to digital writing and vice versa. Results showed that the students who use SMS modifications the most were those who achieved the best scores to the spelling test. Furthermore, Febvrel and Hureau (2008) concluded that their 10 to 11 years old participants were able to have appropriate adaptation when producing digital and then traditional writing. They implied that children who were able to make such a distinction the most were those who get the best scores in reading and writing.

Whereas latest research focused on the impact of the use of digital writing on spelling, studies on grammar production in NICTs are rarer. Stark (2012) analyzed a corpus of SMS and showed that the inflection of some words disappeared. As a consequence, written forms such as *tu a* in French (i.e., *you have* in English), where the morpheme *–s* is removed, showed that subject-verb agreement in SMS are not very different from their traditional equivalent. Her work focused on inter-words agreement whereas previous studies targeted intra-words analyses.

All this research helped to find out that digital writing has a different structure from traditional writing. We have to explore this idea by comparing its structure to other traditional systems.

1.4 Oral code and written code

In order to provide a definition to digital writing, we have to rely on its grammatical structure and on the different characteristics of traditional oral and written codes. Crystal (2001) mentioned that traditional writing needs a physical medium to subsist, which implies it has space constraints. Thanks to a system of abstract and linguistic signs, we can use a wide variety of words and of spelling, grammatical, syntactic,

lexical and semantic rules. The context has to be defined clearly to lead to understanding. That is why the quantity of information has to be more important and that the identification of protagonists is required (Crystal, 2001). Syntax is complex in traditional writing and is organized around the common noun (Cellier, 2003a; Halliday, 1985). Nouns must pertain to grammatical gender and number, they can be either animate or inanimate (e.g., *a horse* in the first case, *a watch* in the second) or related to a state (e.g., anxiety), a characteristic (e.g., patience), a connection (e.g., complementarity), etc. (for further details, see Riegel, Pellat & Rioul, 1994).

The information provided is in limited number in oral since the definition of the context is not mandatory (which includes words are in limited number too). Information is indeed inferred and the protagonists are clearly identified, which leads to a massive use of deixises. These words find their meaning in a context of enunciation, which has previously been defined (e.g., the use of *I* or *tomorrow* depends on who produces the sentence, or when). The determination of their referents is not mandatory in oral either, as the description of gestures and facial expressions.

Oral intonation gives the best clue to punctuation and people perform faster communication. The oral code is organized around the verb (Cellier, 2003a; Halliday, 1985), which is the most meaningful part of the verbal phrase. It then provides a lot of information as tense, mood, voice, aspect, and agrees with the person, the grammatical number and gender. It also helps to define a state of being, an occurrence or an action (see Riegel et al., 1994 for further details).

Other words provide a different kind of information; phatic expressions as *hello* or *you know what I mean* help to start, maintain or close contact to ensure successful communication. If these words are used both in oral and written languages, the use of emphatic expressions for instance (which has a completely different meaning in French) do not appears to be close to the *subject-verb-adjective* pattern (e.g., *super, ce film!,* which could be translated *great, this movie!*). Indeed, a French native speaker can put some expressions at the beginning of a sentence to highlight an important point.

Furthermore, function words (or free morphemes) are most widely produced in oral whereas it is content words (or root words), which are more commonly used in writing (Cellier, 2003a) (for further details, see Table 1).

New form located between the oral and the written appeared with digital writing (Fairon, Klein & Paumier, 2006), which gave rise to a *hybrid* (Anis, 1998; Jalabert, 2006; Mourlhon-Dallies, Rakotonoelina & Reboul-Touré, 2004; Pétillon, 2006;

Tagliamonte & Denis, 2008). It means syntactic, spelling, grammatical, etc. rules must no longer be as rigid as they are in traditional codes.

	Oral code	**Written code**
Lexical categories	Basic vocabulary (3 to 5000 words)	Large vocabulary (more than 10000 words)
	General and imprecise vocabulary	Varied vocabulary
	Phatic expressions without informational content (e.g. *hello*)	*Full* words
	Truncated words, which are specific reductions (i.e. *tomoz* for *tomorrow* etc.)	
	High creativity	Creating neologisms does not occur frequently
	High quantity of function words	High quantity of content words
	The determination of deixises and proper nouns are not mandatory	The determination of deixises and proper nouns are mandatory
	Nonverbal communication: visible gestures and facial expressions	Description of nonverbal communication
Syntactic categories	Syntax organized around the verb	Syntax organized around the common noun
	Dynamic word set	Static and dense word set
	Utterances can be very incomplete	Complete sentences
	Sentences sometimes incoherent	Coherent sentences
	The French negation adverb *ne* is often deleted[2] (it could be compared to *isn't, hasn't, wouldn't*, etc.)	Complete negation *ne … pas/ jamais/ rien, etc.* (it could be compared to *is not, has not, would not*, etc.)
	Interrogative intonation (e.g., *You want something?*)	Subject-verb inversion in interrogative sentences (i.e. *When will you leave?*)
	Duplication of pronouns (e.g. *Me, I am not* …); introduction of information (i.e. *This is me and my son*); emphatic expressions (as the term is understood in French) (e.g., *super / ce film*; *great / this movie*)	Canonical sentence order, with rules

---

[2]In French, two adverbs are needed to express negation (e.g., *ne/pas* as in *je ne veux pas écrire une letter,* which means *I do not want to write a letter*).

Intonation can disambiguate sentences	Sentences are disambiguated by punctuation
Clustering of utterances (e.g. *I have something to tell you well you said to me yesterday uh I was late whereas I wasn't now I will always show up a few minutes early*)	
Repetitions	Repetitions are avoided as much as possible

Table 1. Lexical and Syntactic characteristics of the oral and written codes (Cellier, 2003a; Riegel et al., 1994).

However, Tagliamonte and Denis (2008) proved that DWIM and traditional writing were not as different as it is usually thought. They showed that both kinds go through the same grammatical transformations. To do so, the authors collected data from traditional and digital writing productions of English-speaking adolescents between the ages of 15 and 20. But other works underline several differences between both kinds of writing. Baron (2010) indeed mentioned that the use of words is reduced as much as possible in digital writing. Furthermore, he proved that its users write their instant messages by breaking sentences down thanks to the "Enter" key. They proceeded that way as if they were speaking, which explains why sometimes the message could be misinterpreted (Grinter & Eldridge, 2001). To avoid misunderstanding, users created new written form such as initials (e.g., *omg* instead of *oh my god*) or acronyms (e.g., *lol,* which means *laughing out loud*).

As a consequence, the grammatical definition of digital writing had to be done. Panckhurst (2009) built a corpus of SMS to provide an analysis organized in four parts of speech that included verbs, common nouns, adjectives and adverbs. She was able to get a complete picture of the emails, SMS and eSMS[3] syntactic structure. She found out that the formers were organized around the common noun and the latters around the verb. We can deduce that emails have a written-based structure whereas SMS/eSMS have an oral-based structure. A work still needs to be done on a corpus of DWIM to define its structure since we showed the two systems were not as similar as it is commonly considered. Such a corpus would help gathering data from language users in a more or less natural context of enunciation.

---

[3] ESMS are little messages sent online.

As a result, we have to provide a definition of DWIM by constructing two typologies. The former will allow a better understanding of the grammatical structure of DWIM and the latter of its code organization thanks to the description of its traditional equivalents (see Table 1). This work will be as specific as possible since the corpus must be produced by adolescents who get their literacy skills in permanent contact with both traditional and digital writing and with no language impairment. They indeed have to be born after 1992, which corresponds to the advent of SMS (Simoës et al., 2012). The research that has been done until know concerned adults, or children with language disorders' written production (Anis, 2003; Fairon et al., 2006; Kobus, Yvon & Damnati, 2008; Liénard, 2008; Panckhurst, 2009; Simoës et al. 2012; Véronis & Guimier de Neef, 2006; Yvon, 1998).

It is hypothesized that the DWIM code is a hybrid closer to the oral code than to the written code. Furthermore, Febvrel and Hureau's (2008) directions for further research help us to assume that participants who reported reading and writing on traditional media would be those who use the DWIM code the most, compared to participants who do not have such habits.

2. Method

We used a two-stage process. The former part helped to define DWIM grammatically, lexically and syntactically, and the latter to gather data with regard to reading and writing habits to see if it is positively linked to the use of the DWIM code.

Thanks to the works of several authors, we are able to hypothesize that:

- 1st hypothesis: The DWIM code is closer to the oral code than to the written code.

- 2nd hypothesis: Having reading and writing habits helps users to produce more linguistic items related to the DWIM code than those who do not have such habits.

2.1 Participants

Forty 8th grade students participated to the study. As we had to constitute two equal groups (a first one with reading and writing habits and the other without such habits), we decided not to take into account the results of 8 adolescents who reported that thay never write on traditional media. Children of this school level were chosen since they (a) are supposed to have reached grammatical, lexical and syntactic construction

basis and (b) can legally have an account on instant messaging website, in the US at least (Leloup, 2012).

Two tools helped in assessing the grammatical, lexical, syntaxic and digital knowledge needed.

A French spelling test (Doutriaux & Lepez, 1994)

This multiple choice test assesses spelling level and focuses on the spelling of both function and content words. Participants have 30 minutes to answer the questions. The French spelling test (FST) helped to control the spelling level of each student and to ensure none of them had language disorders (cf., Table 2).

A four-axed questionnaire (cf., Appendix A)

A questionnaire was developed in four axes, which were "the use of computers, blogs and MSN®; general use of mobile phones; reading habits; writing habits" (Febvrel & Hureau, 2008). Participants had 10 minutes to reply to the questionnaire. It helped to assess the use of digital and traditional writing, and more specifically the reading and writing habits of participants. It also revealed global computer and Internet access at home wich both are needed to often read and write digital writing. Analyzing the results helped to constitute both groups according to the answers to question 13 and 15, which were "*You often read books, novels or comics in your free time.*" and *"When you write outside school, are you used to do it on a piece of paper or on a notebook for instance?*" respectively (see Appendix B).

Groups	G1	G2
Gender	8 girls/ 8 boys	7 girls/ 9 boys
Mean age	13.36	13.56
Standard deviation for age	0.5	0.51
Average FST score	36.31	35.12
Standard deviation for FST score	5.25	4.82

Table 2. Gender, mean age and standard deviation for age, average FST score and standard deviation for FST according to each group.

The first group (G1) included students who answered they have traditional reading and writing habits and the second (G2) those who do not. There is a significant difference between groups, t (30) = 4.392, p < .001 (see Table 2 and Appendix B for the questionnaire results).

2.2 Material and procedure

Participants were invited to write a semi-structured written production on an instant messaging website in the computer room of their school. Instant messaging accounts were created so that students were able to write their digital production. Participants had to open a window part so they were able to see their production received in the middle. It has previously been written in the taskbar at the bottom window where participants were able to select an emoticon (e.g., ☺, ☹).

Participants were asked to produce DWIM for an hour (which is the time needed to bring out automatic reflexes). They were free to choose the other student with whom they would like to chat but do not have to sit next to him/her to avoid oral conversation.

In case the adolescents would lack inspiration, two topics of conversation had been created; (a) the former was about their conception of friendship, (b) the latter about their career prospects.

We explained that we were ready to speak with participants in any case, especially if they encounter a problem and would prefer to leave the study. As participants asked several times if they were assessed during the experimentation, we informed the students that only the analysis of words got our interest and not the whole textual content of their production.

We ensured ongoing informed consent from participants and their relatives (anonymity has been guaranteed). Actually, the headteacher of the secondary school left a note in their student's liaison book (that is how we got adolescents' and their parents' consent). We used pseudonyms to ensure anonymity and invited the whole group to read our results once completed.

We selected an e-mail platform, which was allowed by the French board of education. Its firewall is indeed very effective, which meant we were able to guarentee the security of the students. Participants were not allowed to speak with strangers but only with the other adolescents of the class group. We chose topics of conversation, which were ethically acceptable. The experimenter watched the screens to ensure no one would insult one of their peers.

3. Results

The objectives of this study were (a) to provide a definition of the DWIM code based on grammatical, lexical and syntactical analysis of 8th grade students written

production (see Appendix C) (Cellier, 2003a), that would have helped to show that the DWIM code is closer to the oral code than to the written code, as mentioned in the 1st hypothesis; and (b) to assess the impact of traditional reading and writing habits on the use of the DWIM code. We then wanted to prove that the students who have traditional reading and writing habits are more likely to use the DWIM code than those who do not have such habits, as mentioned in the 2nd hypothesis.

3.1 Définition of DWIM

Both tools (or typologies) listed below correspond to the first objective (data are rounded to the nearest hundredth). The first typology focused on parts of speech and the second on traditional oral and written codes. Both were submitted to an interrater reliability calculation (with two of the authors), which showed acceptable tolerance in both cases (the kappa coefficient was 0.7 for typology 1; 0.93 for typology 2[4]).
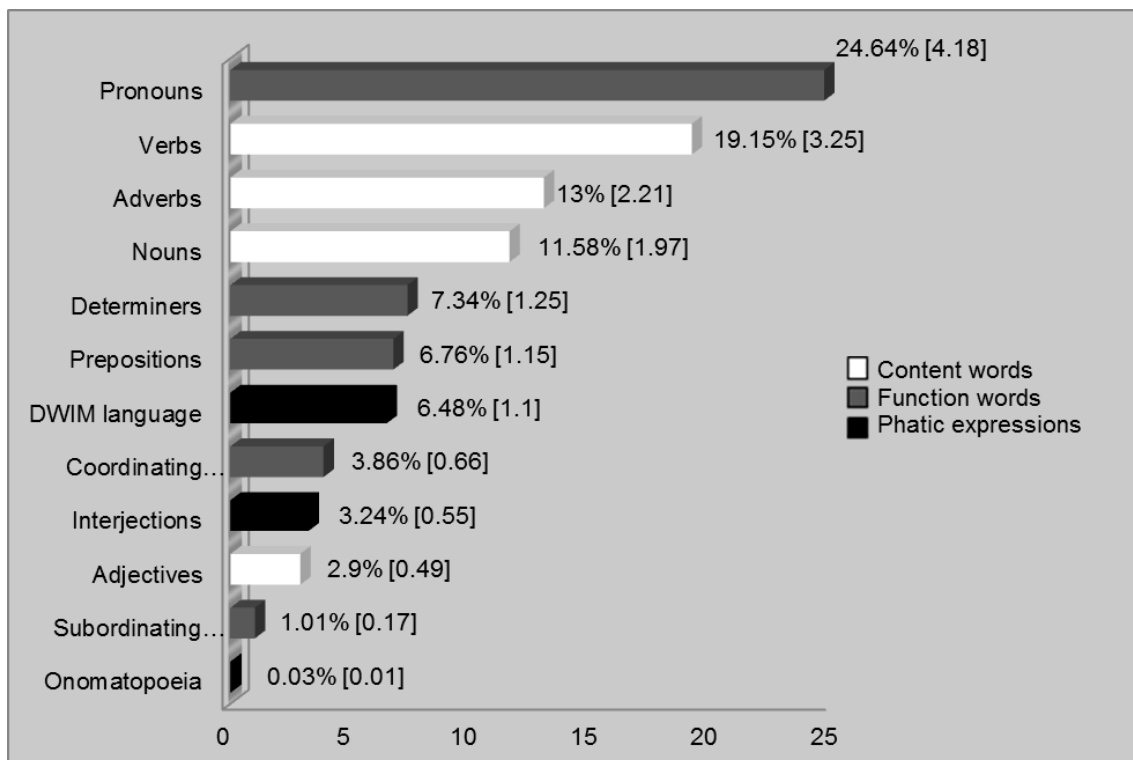

Figure 1. Proportion of each part of speech

---

[4]This tool includes the *Punctuation* characteristic which comprises three subsections, which were (a) capital letter, (b) comma and (c) period, question and exclamation marks (Fayol, 1997). It was also submitted to an interrater reliability calculation, which showed acceptable tolerance (the kappa coefficient was 0.84).

Parts of speech in DWIM

The results are presented in the following order, according to: (a) function words that help to link information; (b) content words that include information; and (c) phatic elements that help to start, maintain or end a conversation, to describe emotions, or to use polite phrase (Riegel et al., 1994).

The largest category of function words (and of the total amount of words also) is pronouns (24.64% [4.18]). The average of verbs is higher than all the other categories of content words (19.15% [3.25]) such as nouns for instance (11.58% [1.97]). Lastly, DWIM language such as *lol* or ☺ has the highest rate of phatic elements (6.48% [1.1]) (see Figure 1).

Comparison between DWIM, oral and written codes

Lexical and syntactic characteristics were used as a basis for the analysis (see Table 3 and 4).

Few elements that should not be present in traditional writing (except perhaps for direct, indirect and free indirect speech) do appear in DWIM and are related to the oral code. It concerns phatic expressions (9.76% [0.18]), truncated words (1.29% [0.04]) and the determination of proper nouns that is not mandatory (proper nouns represent 3.52% [0.13] of words).

Generally speaking, the results from the statistical analysis showed no significant difference between function and content words, but do show significant differences between content words and phatic expression, and function words and phatic expressions (t (31) = 7.77, p < .001; t (31) = 7.04, p < .001 respectively). The proportion of phatic expressions is therefore significantly lower than those of function and content words.

One characteristic gathers function and content words into one. Indeed, a large proportion of deixises, such as the personal pronoun *I*, has been found (undefined deixises share 22.19% [4.74] of the total amount of words; determined deixises, 0.02% [0.01]). We note that these results has been characterized as being statistically significant, t (31) = 26.451, p < .001.

If we consider phatic expressions, we found that 11.58% [0.25] of words are related to DWIM language. Its proportion is therefore significantly higher than the visible and nonverbal communication characteristic, t (31) = -8.501, p < .001.

Lexical characteristics	Oral code		Written code
Phatic expressions	9.76% [0.18]	*Full* words (information content)	90.24%
Truncated words	1.29% [0.04]	Untruncated words	98.71%
Function words	43.61% [0.89]	Content words	46.63% [0.97]
The determination of deixises is not mandatory	23.28% [0.54]* (21.48% [0.54] of determiners and pronouns, and 1.8% [0.05] of adverbs)	The determination of deixises is mandatory	0.02% [0.01]
The determination of proper nouns is not mandatory	3.52% [0.13] *	The determination of proper nouns is mandatory	0%
Nonverbal communication	ø[5]	Description of nonverbal communication (e.g., *I am kidding,* ☺ [6]*, lol,* etc.)	11.58% [0.25] *

Table 3. Lexical analysis.

Syntactic chategories	Oral code		Written code
Syntax organized around the verb	19.15% [3.25] *	Syntax organized around the verb	11.58% [1.97]
Deletion of the French negative adverb *ne*	1.78% [0.06]	Complete negation	0%
Interrogative intonation	2.66% [0.06]*	Subject-verb inversion in interrogative sentences	0.05% [0.01]
Duplication of pronouns (e.g., *Moi, je – Me, I*); introduction of	1.15% [0.04]	Canonical sentence order, with rules	98.85%

---

[5] Indeed, emoticons are glyphs.
[6] We chose to consider expressions such as *I am kidding* as a part of this syntactic characteristic since emoticons, initials and acronyms have the same phatic role.

information (e.g., *This is*); emphatic expressions (e.g., *great / this movie!*)			
Clustering of utterances	0.81% [0.04]		
Repetitions	1.1% [0.04]		
Intonation can disambiguate sentences	Cf., analysis infra	Punctuation can disambiguate sentences	Cf., analysis infra

Table 4. Syntactic analysis.

Firstly, we notice that syntax is organized around the verb (with 19.15% [3.25] of words of the corpus, compared with only 11.58% [1.97] of nouns). Difference between these two characteristics is statistically significant, t (31) = -5.9, p < .001, which shows that verbs are much more used than nouns.

Furthermore, other syntactic characteristics that should not be present in traditional writing also appear in DWIM. The deletion of the word *ne* has been constant and applies to 1.78% [0.06]) of the total amount of words (compared with 0% of usual two-part negative adverb). Difference between these two characteristics is significant, t (31) = -5.9, p < .001.

We have made similar observations with interrogative intonation, which means that no subject-verb inversion has been used (i.e., *her name is?*, example taken from the corpus). This linguistic process is not used in 2.66% [0.06] of cases whereas it is in 0.05% [0.01], which means that the user does not produce a canonical subject-verb inversion most of the time. It also showed that the difference between interrogative intonation and subject-verb inversion was statistically significant, t (31) = 7.683, p < .001. It is important to notice that this latter proportion is due to the fact that one of the participants wrote a wordplay in French (e.g., Question: *Mister and Madam Net(1) have a girl. What is her name?* Answer: *Jessica(2),* which is becoming *J'ai six ca(2) (r)nets(1)*, or in English, *I have six notebooks*; example taken from the corpus).

The duplication of pronouns, introduction of information and emphatic expressions should also not be present in traditional writing, but share 1.15% [0.04] of the total amount of words. The reader could encounter difficulties in decoding these kinds of phrases, and this reading task could become even more difficult when his interlocutor uses repetitions (with an average of 1.1% [0.04]) or clustering of utterances (which amounts 0.81% [0.04] of the corpus).

In addition, punctuation helps usually the reader, but is mainly unused in DWIM to the amount of 66.16% [1.42]. Only 30% [1.2]) of punctuation marks are written, which is significantly lower than the proportion of unused punctuation marks, t (31) = -3.866, p < .001. We noted furthermore that few emoticons were used to end a sentence in place of a period, question or exclamation marks, in 3.74% [0.15] of the cases. It is also significantly lower than the proportion of unused punctuation marks, t (31) = 7.981, p < .001 (difference between undeleted punctuation marks and emoticons used to end a sentence was significant, t (31) = 7.981, p < .001).

3.2 The use of DWIM and traditional reading and writing habits

No between-group differences were found to be significant. In other words, having traditional reading and writing habits did not show any quantitative of qualitative difference in using the DWIM code.

4. Discussion

The objective of this paper was twofold. We wanted (a) to provide a definition of DWIM based on grammatical, lexical and syntactic data to show that the DWIM code is closer to the oral code than to the written code (1st hypothesis) and (b) to know if having traditional reading and writing habits could constitute a predisposition to use the DWIM code (2nd hypothesis). Indeed, the studies available on the topic of digital writing mainly focus on SMS, but only a few were published in the field of DWIM. Although these two media of digital writing have many similarities, we must reiterate that the analysis of SMS involves analyzing a few sentences whereas the studies devoted to DWIM focus on text analysis.

Our 1st hypothesis helps us to consider DWIM as a *hybrid* (e.g., Anis, 1998; Jalabert, 2006; Mourlhon-Dallies et al., 2004; Pétillon, 2006; Tagliamonte & Denis, 2008), which means it is a combination of the traditional oral and written codes. Users of DWIM indeed produce as much function words, which help to link information as content words, which help to provide it.

But our results have confirmed and also completed the work of these authors. This 1st hypothesis was confirmed since the analysis showed that the DWIM code is much closer to the oral code than the written code. Keeping most function words in DWIM maybe includes an attempt on the part of the sender to facilitate the information

processing of the reciever as well as providing him lexical data, with the primary purpose of being understood. Our analysis indeed highlighted that no significant difference was found between the proportion of function words and content words (43.61% [0.89]) and 46.63% [0.97] respectively). These results complete what Tran et al. (2008) found about dyslexic children using function words in SMS. They took the sentence "ta ou blie (LE) rdv ou koi?" (which means "tu as oublié LE rendez-vous ou quoi?" in French, and "(did) you forget THE appointment?") as an example to show that function words could be more often used than deleted.

If Grinter and Eldridge (2001) mentioned that pressing the *Enter* key would help to disambiguate instant messages, it would also explain why punctuation marks are often unused. Our results showed that the use of emoticons helped to disambiguate instant messages too since the students produced smileys to end a sentence, as they would have done it with a punctuation mark. It confirms the conclusions of Dresner and Herring (2010) about the illocutionary force of emoticons. They indeed provided a definition to these linguistic signs by letting know it could be compared to punctuation marks, without showing that smileys could replace period, question and exclamation marks.

If emoticons could play such a role, they also help to describe nonverbal behavior, which includes other elements (i.e. *I am kidding*, *lol*). These written signs allow the user to add information to his message or to avoid misunderstanding, and correspond to the amount of 11.58% [0.25] of the corpus.

Although this kind of communication is undoubtedly available through writing, it seems that one of the lexical characteristics of the oral code are used in DWIM, which are (a) truncated words (1.29% [0.04]); and that some of its syntactic characteristics, (b) which are the deletion of the negative adverb *ne* (1.78% [0.06]), (c) the interrogative intonation (2.66% [0.06]), (d) the duplication of pronouns, the introduction of information, and the emphatic expressions (1.15% [0.04]), (e) the clustering of utterances (0.81% [0.04]), and (f) repetitions (1.1% [0.04]).

Other lexical characteristics of DWIM are included in the oral code as the determination of deictic, which is not mandatory. We have to add that only few deixises have been defined (0.02% [0.0], compared to 23.28% [0.54] of undefined deictics), and that no proper nouns have been defined whereas they are used in 3.52% [0.13] of cases.

Using more verbs (19.15% [3.25]) than nouns (11.58% [1.97]) allows saying that DWIM is syntactically closer to the oral code than the written code. This statement was

also demonstrated through the analysis of parts of speech. These results confirmed our 1st research hypothesis. The detailed analysis showed that the proportion of pronouns was the most important (24.64% [4.18]). According to Riegel et al. (1994), the etymology of the word *pronoun,* which means *a word that substitutes for a noun* does not define this part of speech in its entirety, but it helps to understand their use in DWIM. Pronouns do not only substitute for noun only but also for noun phrase and sometimes, they designate the referent directly. This particular use, that ensures a good understanding of information, is possible since the writers are clearly identified as students enrolled in 8th grade and indeed involves a high proportion of undefined personal pronouns (17.12% [2.91] out of a total of 24.64% [4.18] of pronouns).

The use of phatic expressions (interjections, onomatopoeia and other linguistic signs specific to DWIM) shows that it is very important for an adolescent to start, maintain and end a conversation, or to tell how he feels (see Appendix C). Their proportion equals to almost one tenth of the corpus (9.76% [0.18]), which is quite as much as the part of nouns (11.58% [1.97]).

To conclude, DWIM is a combination of both traditional oral and written codes (e.g., Crystal, 2001) but also remains close to the former (as it was mentioned in the 1st hypothesis). Thanks to this conclusion, we propose to define the DWIM code as follows (see Table 5).

Lexical categories	DWIM code
Hybrid characteristic	Function words and content words in same proportion
Written code characteristic	Description of nonverbal communication (e.g., emoticons)
Oral code characteristics	- Troncated words - The determination of deictic and proper nouns are not mandatory
Syntactic categories	**DWIM code**
Hybrid characteristic	Punctuation is mainly unused (emoticons replace period, question and exclamation marks in few cases)
Oral code characteristics	- Syntax organized around the verb - Deletion of the French negative adverb *ne* - Interrogative intonation - Duplication of pronouns; introduction of information; emphatic expressions - Clustering of utterances - Repetitions

Table 5. Lexical and syntactic structure of the DWIM code

If we were able to show that DWIM was closer to the oral code than the written code, we showed why all these phonetic forms appear in DWIM productions (e.g., *to* in place of *too*, or *aaaaaaah* instead of *ah*). Indeed, users of NICTs give priority to communication, and that is why they also try to simplify their spelling by using forms that are closer to the oral. Phonetic forms are written to save time and to make the written production easier.

Our 2nd hypothesis about traditional reading and writing habits was not confirmed. This result does not confirm the directions of Febvrel and Hureau (2008) since it seems that a regular use of a traditional tool is not a predisposition to write a digital message while respecting the DWIM code. However, an expert user of digital writing would be more able to produce sentences according to the DWIM code than a novice user, both quantitativally and qualitatively.

If we have benefited from the work of Cellier (2003a) and Riegel et al. (1994) on the traditional oral and written codes, we had to delete some of the characteristics of each code (such as its vocabulary) since we did not gather data from other media. Future research would be then devoted to a real comparison of these three codes by collecting three different corpuses of oral and written data. This eventual work would reinforce and complete the results of the present study.

5. References

Anis, J. (1998). *Texte et ordinateur: l'écriture réinventée?* Bruxelles: De Boeck.

Anis, J. (2003-May). Communication électronique scripturale et formes langagières. Paper presented at the *Actes des Quatrièmes Rencontres Réseaux humains / Réseaux technologiques*, Poitiers.

Baron, N. (2010). Discourse Structures in Instant Messaging: The Case of Utterance Breaks. *Language@Internet, 7*(4). Retrieved July 13th, 2011, from: http://www.languageatinternet.org/articles/2010/2651/Baron.pdf.

Bazerman, C. (2006). Ecriture, organisation sociale et cognitions: Repenser Goody. *Pratiques, 131/132*, 95-115.

Bekhuis, T. (2007). Misery and the Digital Divide. *PsycCRITIQUES, 52*(33).

Blanc, Q. (2012). Les ados envoient 2500 SMS par mois. *Le Figaro*. Retrieved October 22th, 2012, from:

http://www.lefigaro.fr/hightech/2012/10/05/01007-20121005ARTFIG00527-les-ados-envoient-2500-sms-par-mois.php.

Bonetti, L., Campbell, M. A., & Gilmore, L. (2010). The relationship of loneliness and social anxiety with children's and adolescents' online communication. *Cyberpsychology, Behavior and Social Networking, 13*(3), 279-285.

Bouillaud, C., Chanquoy, L., & Gombert, J.-E. (2007). Cyberlangage et orthographe: quels effets sur le niveau orthographique des élèves de CM2, 5ème et 3ème? *Bulletin de Psychologie, 60*(6), 553-565.

Burns, S., Espinosa, L., & Snow, C. E. (2003). Débuts de la littératie, langue et culture: perspective socioculturelle. *Revue des sciences de l'éducation, 29*(1), 75-100.

Cellier, M. (2003a). Un oral, des oraux. In M. Prouilhac (Ed.), *Maîtrise du langage* (pp. 17- 58). Toulouse: CNED.

Cellier, M. (2003b). Orthographe. In M. Prouilhac (Ed.), *Savoirs et culture* (pp. 193-247). Toulouse: CNED.

Crystal, D. (2001). *Language and the Internet*. Cambridge: Cambridge University Press.

Defays, J.-M., & Englebert, A. (2009). *Acteurs et Contextes des Discours universitaires.* Paris: L'Harmattan.

Demougin, P. (2003). Spécificités de l'écrit. In M. Prouilhac (Ed.), *Maîtrise du langage* (pp. 109-130). Toulouse: CNED.

De Saussure, F. (1979). *Cours de linguistique générale.* Paris: Payot.

Doutriaux, F., & Lepez, R. (1994). *Test de Niveau Orthographique*. Paris: Les Editions du Centre de Psychologie Appliquée.

Dresner, E., & Herring, S. C. (2010). Functions of the Nonverbal in CMC: Emoticons and Illocutionary Force. *Communication theory, 20*, 249–268.

Fairon, C., Klein, J. R., & Paumier, S. (2006). *Le langage SMS: étude d'un corpus informatisé à partir de l'enquête: Faîtes dons de vos SMS à la science.* Louvain: Presses Universitaires de Louvain.

Fayol, M. (1997). On acquiring and using punctuation: A study of written French. In J. Costermans & M. Fayol (Eds.), *Processing interclausal relationships: Studies in the production and comprehension of text* (pp. 157-177). Mahwah: Lawrence Erlbaum Associates.

Febvrel, A., & Hureau, C. (2008). "Msn c tro bi1?": Influence de la communication médiée par ordinateur sur les pratiques orthographiques: étude auprès d'une population de CM2. *Unpublished master's thesis*. Université de Lille 2.

Fischer, J.-P. (1997). L'apprentissage d'une grammaire artificielle par des enfants de 9 à 11 ans. *L'année psychologique, 97*(2), 207-236.

Fukking, R., & Hermanns, J. (2009). Children's experiences with chat support and telephone support. *Journal of Child Psychology and Psychiatry, 50*(6), 759-766.

Grinter, R. E., & Eldridge, M. A. (2001-September). Y du tngrs luv 2 txt msg ? Paper presented at the *European Conference on Computer-Supported Cooperative Work*, Dordrecht.

Halliday, M. A. K. (1985). *Spoken and Written Language*. Oxford: Oxford University Press.

Jalabert, R. (2006). MoliR, revi1 vit... il son 2vnu foo ! (« Molière, reviens vite… ils sont devenus fous ! »). *Dossier "Orthographe", 440*. Retrieved July 13th, 2011, from: http://www.cahiers-pedagogiques.com.

Kobus, C., Yvon, F., & Damnati, G. (2008). Transcrire les SMS comme on reconnaît la parole. *TALN 2008*. Retrieved July 11th, 2011, from: ftp://tlp.limsi.fr/public/013 _taln_1364.pdf.

Leloup, D. (2012). Mineurs et réseaux sociaux, une équation complexe. *Le Monde.fr Technologies*. Retrieved October 2nd, 2012, from: http://www.lemonde.fr/technologies/article/2012/06/06/mineurs-et-reseaux-sociaux-une-equation-complexe_1713216_651865.html.

Liénard, F. (2008). Analyse linguistique et sociopragmatique de l'écriture électronique: le cas du SMS tchaté. In J. Gerbault (Ed.), *La langue du cyberespace: de la diversité aux normes* (pp. 265-278). Paris: L'Harmattan.

Ling, R., & Baron, N. S. (2007). Text Messaging and IM: Linguistic Comparison of American College Data. *Journal of Language and Social Psychology, 26*(3), 291-298.

Mairesse, J., Cette, G., & Kocoglu, Y. (2000). Les technologies de l'information et de la communication en France : diffusion et contribution à la croissance. *Economie et statistique, 339*, 117-146.

Mourlhon-Dallies, F., Rakotonoelina, F., & Reboul-Touré, S. (2004). Les discours de l'internet: quels enjeux pour la recherche? *Les Carnets du Cediscor n°8: Les discours de l'internet: nouveaux corpus, nouveaux modèles?, 8*, 9-18.

Panckhurst, R. (2007). Discours électronique médié: quelle évolution depuis une décennie? In J. Gerbault (Ed.), *La langue du cyberespace: de la diversité aux normes* (pp. 121-136). Paris: L'Harmattan.

Pétillon, S. (2006). Le discours SMiste, un nouveau corpus? *Dossier "Orthographe", 440*.

Plester, B., Wood, C., & Bell, V. (2008). Txt msg n school literacy: does texting and knowledge of text abbreviations adversely affect children's literacy attainment? *Literacy, 42*(3), 137-144.

Plester, B., & Wood, C. (2009). Exploring relationships between traditional and new media literacies: British preteen texters at school. *Journal of Computer-Mediated Communication, 14*, 1108-1129.

Plester, B., Wood, C., & Joshi, P. (2009). Exploring the relationship between children's knowledge of text message abbreviations and school literacy outcomes. *The British Psychological Society, 27*, 145-161.

Privat, J.-M. (2006). Un habitus littératien? *Pratiques, 131/132*, 125-130.

Riegel, M., Pellat, J.-C., & Rioul, R. (1994). *Grammaire méthodique du français*. Paris: Presses Universitaires de France.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*, 855-863.

Simoës-Perlant, A., Thibault, M.-P., Lanchantin, T., Combes, C., Volckaert-Legrier, O., & Largy, P. (2012). How adolescents with dyslexia dysorthographia use texting. *Written Language and Literacy, 15*(1), 65-79.

Sjöberg, U. (2003). Making sense of screen-based media. The uses and readings of television, computer games, and Internet among Swedish young people. In I. Rydin (Ed.), *Media fascinations: Perspectives on young people's meaning making* (pp. 147-164). Goteborg: Nordicom.

Stark, E. (2012). L'analyse morphosyntaxique des SMS: variation et stabilité grammaticale. Centre de recherches sur la Cognition et l'Apprentissage. Retrieved October 2nd, 2012, from: http://www.google.fr/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0CC0QFjAB&url=http%3A%2F%2Fcerca.labo.univpoitiers.fr%2FIMG%2Fpdf_Resume_seminaire_Stark.pdf&ei=YnFUP7mC6Ol0QXz34GIBQ&usg=AFQjCNHHScYGIKwZDiPnd8LBzPkCNRu3YA.

Tagliamonte, S. A., & Denis, D. (2008). Linguistic Ruin? Lol! Instant Messaging and Teen Language. *American Speech, 83*(1), 3-34.

Tran, T. M., Trancart, M., & Servent, D. (2008). Littéracie, SMS et troubles spécifiques du langage écrit. Paper presented at the *Actes du Congrès mondial de linguistique française*, Cluj-Napoca.

Véronis, J., & Guimier de Neef, E. (2006). Le traitement des nouvelles formes de communication écrite. In G. Sabah (Ed.), *Compréhension automatique des langues et interaction* (pp. 227-248). Paris: Hermès Science.

Witt, A. (2010). L'apprentissage implicite d'une grammaire artificielle chez l'enfant avec et sans retard mental: rôle des propriétés du matériel et influence des instructions. *Unpublished Master's thesis.* Université de Bourgogne.

Yvon, F. (1998). Rewriting the orthography of SMS messages. *Natural Language Engeneering, 1*(1), 1-28.

6. Appendices

6.1 Appendix A. The four-axed questionnaire (translated from French)

Full name: **Age:**

Please circle your answer:

1/ You are enrolled in	6th grade 7th grade 8th grade 9th grade		
2/ Do you have computer access at home?		YES	NO
3/ Do you have internet access at home?		YES	NO
4/ Did you ever read a blog and/or leave a comment?		YES	NO
5/ Do you know and use instant messaging?		YES	NO
6/ Do you know how to use smileys on an instant messaging website?		YES	NO
7/ Leaving a comment on your instant messaging profile is usual for you. It helps you to give news to your friends, to say how you are going, to plan the day and other things like that.		YES	NO
8/ You can use your mobile phone or someone else's to send text messages?		YES	NO
9/ When you text, do you pay attention to spelling?		YES	NO

Please check the box that best applies to you:

10/ To send a message:	☐ you use your mobile phone the most
	☐ you use instant messaging the most
	☐ you have no preference.
11/ To write to people you know, your preference goes to:	☐ the keyboard of a mobile phone
	☐ the keyboard of a computer
	☐ you have no preference.

Please circle your answer:

12/ Do you have difficulties reading SMS or instant messages?	YES	NO
13/ You often read books, novels or comics in your free time.	YES	NO
14/ You are used to read articles or web pages online to do your homework or to search anything else in your free time.	YES	NO
15/ When you write outside school, are you used to do it on a piece of paper or a notebook for instance?	YES	NO
16/ When you leave school, are you used to write via the Internet?	YES	NO
17/ Do you play online?	YES	NO
If you do play online, do you spell every words correctly?	YES	NO

Table A1. The four-axed questionnaire (translated from French).

6.2 Appendix B. Full questionnaire results expressed as percentages

Item	Results G1	Results G2	Total
A. "The use of computers, blogs and MSN®" (Febvrel & Hureau, 2008), affirmative response to a yes/no question			
2/ Computer acces	100%	100%	100%
3/ Internet access	100%	100%	100%
4/ Read a blog with/ without comments	93.75%	75%	84.375%
5/ The use of instant messaging	100%	100%	100%
6/ The use of emoticons	100%	100%	100%
7/ Profile comments	62.5%	56.25%	59.375%
B. "General use of mobile phones"			
8/ Sending SMS via mobile phones	93.75%	93.75%	93.75%
9/ Paying attention to spelling (affirmative response)	68.75%	43.75%	56.25%
10/ Preference for sending messages (tick boxes) :			
* Via mobile phones	87.5%	68.75%	78.125%
* Via computers	6.25%	12.5%	9.375%
* No preference	6.25%	12.5%	9.375%
* Abstention	0%	6.25%	6.25%
11/ Preference for writing (tick boxes) :			
* On mobile phones	6.25%	12.5%	9.375%
* On computers	37.5%	25%	31.25%
* No preference	56.25%	56.25%	56.25%

* Abstention	0%	6.25%	3.125%
C. "Reading habits", affirmative response			
12/ Reading SMS difficulties	6.25%	25%	15.625%
13/ Short story/book reading during free time	***100% ****	***56.25%***	***78.125%***
14/ Digital reading habits during free time	87.5%	87.5%	87.5%
D. "Writing habits", affirmative response			
15/ Writing on traditional media	***100% ****	***0%***	***50%***
16/ Writing online	68.75%	68.75%	68.75%
17/ Playing games online	75%	56.25%	65.625%
Use of correct spelling (among affirmative responses)	50%	22.22%	36.11%

Table B1. Full questionnaire results expressed as percentages.

6.3 Appendix C. Full description of French parts of speech (Riegel et al., 1994)

Function words	Content words
French determiners (they make common nouns usable in sentences by "marking" nouns): - Definite articles (e.g., *le, la, les* in French; *ø, the* in English) - Indefinite articles (e.g., *un, une, des* in French; *a, an* in English) - Partitive articles (e.g., *du, de la* in French; *some, ø* in English) - Possessive determiners (e.g., *mon, ton, ses, notre, leur* in French; *my, your, his/ her, our, their* in English) - Demonstrative determiners (e.g., *ce, cet, cette, ces* in French; *this, that, these* in English) - Indefinite quantifiers (e.g., *tous, chaque, aucun* in French; *every, any, none* in French) - Ordinal quantifiers (e.g., *premier, deuxième, troisième* in French; *the first, the second, the third* in English) - Cardinal numeral quantifiers (e.g., *un, deux, trois* in French; *one, two, three* in English)	Adjectives: - Qualifying adjective: it denotes some property of a noun (e.g., *black*) - Relational adjective: it helps to classify and categorize words. It can be replaced with a possessive phrase (e.g., *presidential → of the president*) - Adjective phrase (e.g., <u>*Nice*</u> *to meet you*)

- Exclamative determiners (e.g., *quel! quelle!* in French; *what!* in English) - Interrogative determiners (e.g., *quel? quelle?* in French; *what?, which?* in English)	
Preposition: it always introduces another word/ phrase (e.g., *à, de, par, pour, sans, chez, dans, en* in French; *at, for, without, in* in English). Prepositional phrase is included in this part of speech (it fulfil the same function). It is made up of multiple words (e.g., *in front of*), like all the other phrases.	Nouns (cf., 1.4): - Common nouns - Proper nouns (e.g., someone's name or the name of a location) - Nominal phrase (e.g., *tooth-brush*)
Pronouns (they are a word or phrase that substitute for another word or phrase. They help to avoid repetition. The identification of a referent is needed most of the time to understand the message): - Personal pronoun (e.g., *je, moi, vous, leur, eux* in French; *I, me, you, their, them*) - Relative pronoun (e.g. *qui, que* in French; *who, which* in English) - Demonstrative pronoun (e.g., *ce, celui, celle-ci* in French; *this, this one* in English) - Indefinite pronoun (e.g., *tous, aucun, chacun* in French; *every, none, no mone* in English) - Possessive pronoun (e.g., *le mien, le tien, le nôtre, les leurs* in French; *mine, yours, ours, theirs* in English) - Interrogative pronoun (e.g., *qui? que? lequel?* in French; *who, which, which one* in English) - Pronominal phrase (e.g., *à laquelle* in French; *to which* in English)	Verbs (cf., 1.4): - Inflected verb and infinitive - Verbal phrase (e.g. *to be fed up with*)
Coordinating conjunction: it is a word that connects two other words or phrases of equal	Adverb: it is a word that changes the meaning of other words such as verbs,

importance (e.g., *mais, ou, et, donc, ni-ni, néanmoins* in French; *but, where, and, neither-nor, nevertheless* in English). Conjunction phrase is a part of it (e.g., *ainsi que*)	qualifying adjective, etc. (e.g., *clearly, well,* etc.). Adverbial phrase is a part of it (e.g., *un peu* in French, *little* in English).
Subordinating conjunction: it is a word that connects two other words or phrases of unequal importance (e.g., *because*). It includes conjunction phrases (e.g., *as if, now that,* etc.)	

Table 8. Full description of French parts of speech (Riegel et al., 1994).

A third category of words is used in DWIM also. Phatic expressions include (a) interjections, which express an emotion, sentiment or polite phrase (composed of one word or phrase), and (b) onomatopoeias, which imitate the sound that they describe (e.g., *hello, aaaaah!* in the first case; *splash!* in the second). We chose to consider other form used specifically in digital writing as DWIM language. It includes (c) emoticons, initials and acronyms (e.g., *;), mdr, omg*), which plays a role of equal importance in DWIM sentences.

We also took into account that the word *lol* has been used as a verb twice (*I lol*) so it was classified as verbs. Emoticons were considered as punctuation marks to end a sentence (that is the way they have been identified).