The Language of Text Messaging:

“Linguistic Ruin” or Resource?

Francesca Farina & Fiona Lyddy

The Irish Psychologist
Volume 37, Issue 6
April 2011
rst in messin,
mi headn/me englis,
try2rite essays,
they all come out txtis,
gron not plesd w/letters shes getn,
swears i wrote better
b4 comin2uni.
&she's african

This is the winning entry in the first year (2001) of the Guardian newspaper's text-message poetry competition, in which entrants were required to write a poem within a 160-character limit. Written by Hetty Hughes, a 22-year-old undergraduate studying peace studies, it was selected for first prize from over 7,500 competition entries. The poem displays many of the stereotypical features of text messages: lack of capitalisation, running together of words, abbreviation and non-standard spelling, with only 10 of the 27 words in standard spelling (Crystal, 2008). But how typical is this of text messages?

Media reports over the past decade have cast text language and online communication as a form of “youth communication” consisting predominantly of non-standard language forms. Cited examples of text and computer-mediated messages characteristically include heavily abbreviated, symbolically rich and occasionally undecipherable sentences. Journalist John Humphrys, in an article titled IMHO Txt MsGs: How Texting is Ruining our Language, provides as an example: “IMHO U R Gr8” - “in my humble opinion you are great” (Humphrys, 2007). Media accounts suggest that such stylisations are universal in everyday text language, or, as Thurlow (2006) puts it: “Mst f d tym dey usd ds kind f lang ge’t: “Most of the time they use this kind of language”.

Text messaging, short messaging service (SMS), or “texting” refers to the transmission of text between mobile phones. The majority of such messages involve one-to-one, person-to-person SMS, with interlocutors known to each other, but texting can also involve communication from a messaging service (e.g., competition entry, advertisements or services). Texting can be considered as a type of computer mediated communication (CMC) – although strictly speaking it is not computer-based, it is technologically mediated and text-based and it shares many of the features of other types of CMC (Blieswanger, 2008). CMC can involve the exchange of messages synchronously or asynchronously using applications such as e-mail, instant messaging (IM) and social networking sites (Baron, 2004). IM is generally assumed to involve synchronous communication, as both or all users are present online at the same time and rapid responses are the norm. E-mail and texts are typically asynchronous, but both might be considered as near-synchronous. While texts do not necessitate an immediate response, they can be sent and responded to quickly such that a text “conversation” is possible. Dialogic exchanges are increasingly apparent in sociolinguistic analyses of young people’s text messages (e.g., Kasesniemi & Raatulainen, 2002). Therefore, it has been
suggested that synchrony and asynchrony might be more appropriately viewed as a continuum, with text messaging providing near-synchronous communication (Rettie, 2009). The near-synchronous and brief nature of text communications – along with restrictions posed by the mobile phone keypad – produce distinctive conversation patterns and language use.

Because there are distinctive features to text language, the tendency is to overestimate the degree to which it is non-standard. But, as Crystal (2008) notes, the most striking feature of text messages is the combination of standard and non-standard forms. The language used in young people’s CMC has been labelled “teen-talk”, or more specifically “textisms”, “textspeak”, “textese”, “textspeak” (in the case of SMS), “netspeak”, “netlingo”, and “weblish” (in the case of computer-based communication). Such terms support the notion of a distinct (and, it is generally assumed, deviant) language. It has even been suggested that there may be a link between CMC language patterns (while texting, IMing, social networking and so forth) and a perceived decline in literacy standards in children and young adults (Thurlow, 2006), who are the largest user groups of texting and CMC worldwide (Ling, 2005).

Thurlow’s (2006) analysis of 101 media reports on text and CMC language found that the vast majority of media reports portrayed the language used in a negative light. Teenagers have been labelled “generation text”, “generation grunt” and the “Net generation”, while descriptions such as “bleak, bald, sad shorthand” have been attributed to SMS (Sutherland, 2002, para. 8). Humphrys reinforces this: writing in the Daily Mail in 2007, he described SMS as “abused”, “grotesque”, and a “barrier” to communication and even described texters as “vandals who are trying to do to the language what Genghis Khan did to his neighbours eight hundred years ago.” This suggestion that SMS is a “barrier” reflects the stereotype of the texter as an inferior communicator; it has been suggested that an over-dependency on technology has culminated in a youth generation with deficient communication skills (e.g., Blair, 2004), causing a “dumbing down” of language and a “lowering of standards” (Thurlow, 2006, p. 11).

However, empirical research does not support this negative appraisal of text language nor of texters’ language skills. The data show that the majority of text language is standard form, and the non-standard forms used are often creative, serve an obvious communicative function and reflect a skilled command of language (e.g., Tagliamonte & Denis, 2008). Research analysing genuine examples of the types of textism and netspeak which appear in CMC – such as non-conventional spellings (fone/phone) and shortenings (goin/going) – has allowed for an examination of the frequency of such linguistic forms, and of their distinctiveness compared to “standard” language. This review considers the research relating to language use in texts and discusses implications for literacy.

Language in Texts and Other CMC

The popular perception is that texts and other forms of CMC contain an abundance of emoticons, obscure abbreviations and acronyms, often to the point that they are indecipherable to the uninitiated. Text messages do tend to involve short forms. Texts are limited to 160 characters and writing a text message requires the use of complex multiple keystrokes on most phones. These factors, combined with the time pressure of near-synchronous responses, make shortcuts likely. However, many of the shortcuts are found in other written communication. For example, shortened words (e.g., mon in place of Monday), removed letters (e.g., goin for going) and accent stylisations (e.g., gonna going to) are commonplace, while “cos”, “luv”, “wot” and “ya” appear in the Oxford English Dictionary.

Text also tends to contain more abbreviations than IM or other CMC (Ling & Baron, 2007), and apostrophes are more commonly omitted in texts, as on most mobile phones (which do not have a QWERTY keyboard) they require four key presses when texting, compared to a single keystroke when typing on a keyboard. In the next section of this article, we look at the features of language typically found in texts, before considering the frequency with which they occur.

Features of Text Language

Emoticons and Typographic Symbols

The use of emoticons and typographic symbols is generally overestimated, but they do occur. Emoticons are a type of pictogram that typically convey an emotion or a facial expression, though objects can also be represented (Crystal, 2008). Examples provided by Dresner and Herring (2010) include :-) (smile) and <3 (a heart). Glossaries of emoticons include a multitude of forms, most of which are not found in analyses of actual texts (e.g., :joy:); glossed “baseball cap”; see Crystal, 2008). Baron (2004) considers emoticons as text modifiers, while Neviarouskaya, Prendinger and Ishizuka (2010) have likened them to visual cues in face-to-face conversation. Typographic symbols, on the other hand, are single or multiple characters which represent whole words (Bieswanger, 2008). A popular example is one or several “x’s” used to symbolize a kiss, or “zzz’s” to suggest sleep, tiredness or boredom. Multiple punctuation is also often found in SMS and IM, i.e., “!!!” or “???” for emphasis, and “...” to express contemplation.

Letter/Number Homophones

Also termed logograms, phonetic reductions, or syllabograms, letter/number homophones use a letter or number to represent a word or part thereof (Thurlow & Poff, in press). It is the pronunciation of the individual letters/numbers which is significant, as opposed to the overall appearance of a homophone. A commonly cited example in both SMS and IM, “cu” is “cur/see you later”. This example illustrates both types of homophone, where “cu/see you” denotes a letter homophone and “cur/cur” denotes a number homophone. Other examples reported from actual text and IM messages include “wu2/what you up to”, “ne1/1/anyone” and “BCNU/be seeing you” (e.g., Plester, Wood & Bell, 2008; Thurlow & Brown, 2003).

Shortenings, Contractions and Clippings

Shortenings are words with missing end letters (Thurlow & Brown, 2003). Days and months are commonly shortened in SMS and IM, for example, “sun/Sunday”, and “feb/February”. Contractions are words with omitted middle letters, usually vowels, as, in English, consonants provide greater information than vowels (Crystal, 2008). Reported contractions include “txt/text” and “hmwrk/homework” (Plester, Wood & Joshi, 2009). Contractions can alternatively be classified as the short form of words (e.g., using “don’t” instead of “do not”). Clippings can be sub-categorised as G-clippings and other clippings. The former are words for which the final “g” has been omitted, for example, “goin/going”. The latter represents other final letter omissions, typically final consonants, for example, “wil/will”, and silent vowels, for example, “hav/have” (Crystal, 2008).
Acronyms and Initialisms
In general terms, acronyms and initialisms involve shortening words to their initial letters (Crystal, 2008). Acronyms are sometimes considered as formal shortenings such as “North Atlantic Treaty Organisation/NATO” or “Radio detection and ranging/radar”, while initialisms are more informal, for example, “omg/oh my God”, “bf/boyfriend”, “IMHO/ in my humble opinion”, and “ttyl/talk to you later”. However, Bieswanger (2008) distinguishes between acronyms (letters pronounced as one word, as above) and alphabetisms (pronounced letter by letter), for example “TV”, “FBI” or “BBC”. Many such forms are not unique to CMC, and many have been in general use for considerable time, and are now standard (Crystal, 2008).

Non-conventional Spellings and Accent Stylizations
Non-conventional or non-standard spellings follow legitimate letter-sound correspondences in a language, but they are not the conventional spelling for that particular word, for example, “sum/some”, “thnx/thanks”, and the much publicized use of “k” in “skool!” (Thurlow & Brown, 2003; Thurlow & Poff, in press). Accent stylizations, categorized broadly as “youth code” (Piester et al., 2008), refer to words that are spelled in accordance with informal/regional speech, for example, “wanna/want to”, “gonnagoing/to” and “dat/that” (Crystal, 2008). Accent stylisations are also often employed as humorous alternatives (Thelwall, 2009).

Other Forms
Other documented text language categories reported throughout the literature include onomatopoeic spellings (e.g., “haha” or “zzzz”), omitted apostrophes (e.g., “cant”), misspellings/typos (e.g., “are” for “our”) and hybrids, which are a word or utterance using two or more of the categories outlined above. Crystal (2008, p. 54) offers the following example; “iowan2bwu/l” only want to be with you” which contains a full word, an initialism, a clipping, a number homophone, an initialism and a letter homophone.

The Frequency of Textisms
The majority of language in text messages is standard form. Ling (2005) found that only 6% of words in texts provided by a Norwegian sample were abbreviated, while Ling and Baron’s (2007) American texting corpus contained less than 5% abbreviated words. However, textisms tend to be higher in English than in some languages (e.g., Bieswanger, 2008) and texting is relatively underused in the United States, compared to IM. Indeed, Thurlow and Brown’s (2003) sample from Wales contained a higher number of abbreviations, with abbreviations accounting for 19% of message content (see also Thurlow & Poff, in press). Similarly, we have recently collected and analysed a corpus of 1,200 text messages provided by young adults in Ireland. We found that 75% of message content involved standard spelling, with 17% of the messages containing no “textisms” at all. Only 10% of messages contained more textisms than standard spelling; most of these were short responses (Lyddy & Farina, in preparation). However, in a study investigating textism use in children aged between 10 and 12 years, Piester et al. (2009) found a higher rate of 34%. In this study, the children were given written scenarios (e.g., texting a friend about a birthday party) and were instructed to compose a text message. This, and other experimental methods of text generation, may not accurately reflect real world data, however.

In contrast to the media representations of text language, letter/number homophones, contractions and accent

stylisations are infrequently recorded in text message analyses (e.g., Ling & Baron, 2007; Thurlow, 2006). The variety and complexity of emotions has, in particular, been exaggerated. Entries such as “:-(” (glossed as “a raspberry”) and “0:-)” (an angel; see Crystal, 2008) are not encountered in text analyses, with the “smiley” and “frown face” ( ; ) or “:O and “:D or “:L) being the main emoticons used. Even then, they account for a tiny proportion of message content across corpora (e.g., Ling & Baron, 2007; Thurlow & Brown, 2003). Similarly, the main typographic symbol used in texts is an “<” to signal affection, a convention found in traditional writing also (e.g., Thurlow & Brown, 2003). Thurlow and Brown’s data (2003) show a low frequency of emoticons (< -), typographic symbols (xxx) and letter/number homophones (“gr8/great”) in comparison to non-conventional spellings (“nite/night”), accent stylisations (“ello/hello”) and onomatopoeic spellings (”yayy!” “haha”).

The phonological approximations of non-conventional spellings (“fone/phone”) demonstrate knowledge of language rules. Piester and colleagues (2009) suggest that such spellings in children’s texts reflect an understanding of letter-sound rules of English. For example the letter “f” is used in “fone” because it shares the same sound as the letters “ph” together, not because it is a complete deviation from standard spelling. This borrowing of sounds can also be seen in letter/number homophones where symbols along with sounds are substituted. For example the number “2” shares sounds with the words “to/too”.

Similarly, word contractions preserve information within a language, and knowledge of orthography informs conventions that are adopted. In English, consonants are retained (e.g., “yr/your”, “mt/meet”, “bt/but”), as these are meaning rich. For example, a contraction of the word “tomorrow” that omits all vowels (“tmmrw”) is relatively easy to decipher whereas contracting by omitting consonants (“ooo”) eliminates meaning. So, although some features of text language diverge from traditional written language, the deviation is not an illogical one. Thurlow and Brown (2003) note that young people’s text messages serve the “sociolinguistic maxims of (a) brevity and speed, (b) paralinguistic restitution and (c) phonological approximation” making them “both linguistically unremarkable and communicatively adept” (p.1).

Another misperception of text language that is not upheld by empirical data concerns punctuation. Although punctuation is certainly reduced in texts, there is still much use of punctuation, particularly within the message. In their study, Ling and Baron (2007) found that transmission-final marks were often omitted, but punctuation within messages, and particularly use of question marks, was evident. They found
that punctuation was less frequent in text messages than in IM; in the case of IM, the use of a computer keyboard supports the inclusion of punctuation marks that would require multiple keystrokes when texting. The data also show that the apostrophe is retained in many cases, and particularly where interpretation benefits from its inclusion. Thurlow and Brown (2003) reported the use of apostrophes in one out of every three messages.

The most prolific users of SMS are women and teens/young adults, and those that text a lot, and in preference to a phone conversation, tend to be young and female (e.g., Ling, 2005; Reid & Reid, 2003). Textisms are also more common in these groups (e.g., Ling, 2005; Plester et al., 2009) and there are a number of gender differences in the language forms used that may need to be taken into account when considering the effect of textism use on literacy.

As Thurlow (2006) points out, popular discourse regarding new technology usually treats it as “all good” or “all bad”. A more measured approach would seem to be required. As new technology has been adopted, overall literacy has risen. Massey, Elliott and Johnson (2005), for example, examined samples of examinations taken by UK-based 16 year-olds between 1980 and 2004 and concluded that their standard of literacy had increased across time. However, it is nonetheless important to monitor use of non-standard language in inappropriate contexts (e.g., a formal school assignment), and, similarly, the use of textisms by children, and particularly by weaker readers, requires attention. For the majority of texters, the use of text language should not be any cause for concern. Drouin and Davis (2009) found no difference in standardised literacy scores between texters and non-texters in an American young adult sample. Kemp (2010) assessed the effects of textisms on literacy in Australian university students, noting neutral and positive relationships between scores on linguistic tasks and reading and writing accuracy for both textism and standard text. However, Rosen, Chang, Erwin, Carrier and Cheever (2010) noted a negative correlation between textism use and formal writing, an effect moderated by gender and level of education. Rosen et al. also noted a positive association between textisms and informal writing. Their data suggest that the precise type of textism used might be informative as regards the texter’s writing skill.

Tagliamonte and Denis (2008) propose, based on their analysis of IM language, that the use of non-standard linguistic forms reflects a “skilled command” of language and the available linguistic systems. They argue that the manipulation of language evident from IM is possible due to the in-depth understanding of linguistic features, suggesting that this type of language signals “not the ruin of this generation at all, but an expansive new linguistic renaissance” (p.27). This idea that IM language is associated with a greater awareness of language properties has been supported by Clark and Dugdale (2009), who suggest that as long as children recognise which styles of writing are appropriate in a particular situation (e.g., texting, e-mailing or writing an exam), a positive association between CMC and literacy can occur. Thurlow and Poff (in press) also conclude that although some overlap between textisms and formal language undoubtedly occurs in a minority of users, the vast majority of texters understand the context specific nature of language use, a skill termed “metalinguistic competence” (Crystal, 2005, in Reid & Reid, 2005). Similarly, Plester et al. (2008), using a translation exercise, found that most children switched between standard spelling and textism proficiently.

Research examining children’s text use is currently limited, but available data suggest that concern about pervasive effects on literacy is unfounded. Plester et al. (2008) found that children’s knowledge and use of textisms was not related to written language outcomes in a sample of 11-12 year olds. Plester et al. (2009) found that use of textisms was positively related to word reading, vocabulary and phonological awareness. The last of these – greater phonological awareness – would seem logical given the inclination towards phonetic abbreviation such as non-conventional spellings (“cum/come”) reported in everyday CMC. The direction of the association is not clear however. Plester et al.’s (2009) study also demonstrates the importance of distinguishing between textisms and misspellings in assessing implications for literacy. Textisms such as g-clippings, symbols and accent stylizations showed positive associations with spelling ability, while texted misspellings were negatively associated with spelling ability, as might be expected. Plester and Wood (2009) concluded that “it is clear that [texting] does not contribute to the demise of pre-teen children’s literacy” (p.18).

However, it is worth noting that, because girls and women make more use of text messaging, many of the analyses to date have been predominantly based on data from these groups, with boys and men underrepresented. Female texters also seem to produce more textisms (Ling, 2005; Plester et al., 2009). Given existing gender differences in early literacy (e.g., Phillips, Norris, Osmond & Maynard, 2002; Ready, LoGerfo, Burkham & Lee, 2005), it would seem important to bear these differences in mind when monitoring effects on boys’ and girls’ literacy.

In summary, non-standard language accounts for a minority of words in text messages in the English language, with the majority of message content following conventional forms. Text language makes use of emoticons (:-)), typographic symbols (xxx), acronyms (BBC), initialisms (lol), letter/number homophones (lbr), shortenings (tues), contractions (wkd), g clippings (goin), other clippings (hav), non-conventional spellings (fone) and accent stylizations (gona). Additionally, onomatopoetic spellings (woohoo), omitted apostrophes (cant), and hybrids (two or more of the above) are found.

Of these, the most frequently occurring appear to be phonetically-based forms such as non-conventional spellings, accent stylizations and onomatopoetic spellings, which display language-specific conventions. Symbolically-based forms, such as emoticons, typographic symbols and homophones are less frequent in general, in contrast to the stereotypical treatment of texts in media reports. Use of emoticons and other symbolic forms is relatively infrequent, and a small range of such items appears in various text message corpora. In summary, the data suggest that text messaging language is not as deviant as media portrayals would have us believe. Furthermore, the use of textisms has been found to correlate positively with word reading, vocabulary and phonological awareness in children, and some aspects of language performance in young adults. This may reflect skilled use of metalinguistic knowledge, which allows the texter to switch between differing language systems. Thus, rather than signalling the demise of language, CMC and text language likely reflects the workings of a productive and flexible language system.
Acknowledgements

This review was supported by funding awarded by NUI Maynooth's Summer Programme for Undergraduate Researchers to the first author. The authors wish to thank Niamh Kelly O'Neill for assistance with editing.

References


