Students’ use of text messaging language:

Effects of place and collection method

Abbie Grace
Nenagh Kemp
Frances Heritage Martin
School of Psychology, University of Tasmania, Australia

Rauno Parrila
Department of Educational Psychology, University of Alberta, Canada

Corresponding author:
Nenagh Kemp, School of Psychology, University of Tasmania, Locked Bag 30, Hobart 7001, Tasmania, Australia. Phone +61 3 6226 7534, Email Nenagh.Kemp@utas.edu.au
Abstract

Studies of mobile phone text messaging have reported widely varying proportions of textisms (e.g., u for you, 2 for to). We investigated whether conclusions about textism use are influenced by participant country, text message collection method, and categorisation method. Questionnaire data were collected from 241 undergraduate students in Australia and Canada, who also provided text messages via three methods used in previous research: translation from conventional English, writing a message in response to a scenario, and providing naturalistic messages. Significantly higher proportions of textisms were observed in messages written by Australians than Canadians, and in messages collected experimentally than naturalistically. A novel way of categorising textism forms as ‘contractive’ versus ‘expressive’ was explored and overall implications for text-message collection are discussed.
Mobile phone text messaging has continually increased in popularity since its inception in 1992 (Hillebrand, 2010). Global text messaging volume has risen from an estimated 1.8 trillion messages in 2007 to 6.1 trillion in 2010, equating to close to 200,000 messages sent per second (International Telecommunication Union (ITU), 2010). The text messaging ‘language’ that has been adopted and developed by users of texting (as well as of Instant Messaging) has been referred to as textese, with abbreviations sometimes called textisms. Textisms are often described in terms of categories such as ‘single homophones’ (e.g., u for you, c for see) and ‘initialisms’ (e.g., lol for laugh out loud) (e.g., Plester & Wood, 2009; Thurlow & Brown, 2003). However, previous research describing textese has reported widely varying estimates of mean textism density (proportion of textisms out of all words written). In the current study, we aimed to investigate factors which might explain the range of estimates regarding the number of textisms found in text messages.

Reported textism densities have ranged from five percent in American adults’ naturalistic messages (Ling & Baron, 2007), to 59% in English children’s translated messages (Plester, Wood, & Bell, 2008). The reasons for this variation have not been directly explored, although we assume that much of the variability in estimates of textism use does represent genuine differences between samples of participants. Patterns of textese use could have developed differently in different age groups and countries. Young adults are the primary users of texting technology (Ling, 2010), and for this reason, first-year undergraduate students were invited to participate in the current study. Even across countries with the same dominant language, variations in accent, as well as variable patterns of mobile phone ownership and text messaging uptake by residents, may lead to group differences. To our knowledge, no previous study has compared, in the one study, the texting behaviour of individuals from two countries, in this case, Australia and Canada.
A second possible explanation for variations in the use of textese by different samples lies in differences in the research methodologies employed to collect messages. Some researchers have relied solely on students’ self-reports of textism use (e.g., Massengill Shaw, Carlson, & Waxman, 2007). Others have collected examples of messages by asking participants to translate written or dictated messages (e.g., Neville, 2003), or to write a text message for a specified scenario (e.g., Plester, Wood, & Joshi, 2009). A third, related, issue concerns the way that textisms have been categorised in past research. There is some variability between the types of transformations counted within each textism category, and whether some words are counted as textisms at all. This variability could also lead to divergent conclusions about the type and extent of textism use in different participant groups.

It is important to investigate the possible effects of place, collection methods, and categorisation processes to establish accurate measurements of the prevalence and nature of textisms in messages. The conclusions drawn regarding students’ use of textese can have implications for larger societal views and educational issues. This study provides an important first step in differentiating experimental and individual factors that could explain wide variance in estimates of textism use.

**Sample differences: Effects of Place**

Textese has been described as a hybrid of spoken and written English (Plester & Wood, 2009) because of its frequent representation of features of spoken language. These include spellings which represent pronunciation (e.g., havin for having, cangrads for congrat[ulation]s), and the inclusion of pause fillers (e.g., ummm...), which are not normally represented in formal written language. For this reason, the use of textisms may be influenced by differences in accent and word choice between English-speaking countries. In a study of accent markers in internet homepages, Shaw (2008) observed accent-based differences in the spelling of numerous words. For example, with was most often respelled as wit in the US and
in England. However, these kinds of differences may not affect textism counts (wit and wiv are still both textisms), and are unlikely to explain international variation in textism density. One of the aims of this study was thus to investigate the extent to which textese differs between users of English in two countries. In order to explore the factors which differ between countries but which could also affect textism use (e.g., length of phone ownership, text input type), information about participants’ phone use and behaviour was also collected.

**Texting culture and technology.** The uptake of mobile phone technology has varied greatly even between developed countries. In 2003, the number of mobile phones per 100 people was approximately 73 in Australia, but only 41 in Canada. By 2009 these numbers had increased to approximately 114 and 69, respectively (CIA, 2004, 2010). Between 2003 and 2009, the iPhone and other ‘smartphones’ were introduced, and other forms of full keyboards became increasingly available alongside or in place of traditional alpha-numeric keypads. The constraints of relying on alpha-numeric keypads for text entry have been cited by many researchers as having encouraged the development of textese (e.g., Crystal, 2008; Hillebrand, 2010). In developed countries with a relatively later or slower uptake of mobile phone technology (such as Canada), fewer phone users may have experienced the constraints of keypads that require multiple key-presses to select letters and punctuation. Kemp and Bushnell (2011) confirmed that text entry system type can have a significant effect on textism use. In their study, children who used multi-press entry on an alpha-numeric keypad composed messages more slowly, and included more textisms, than children who instead used a predictive entry system on an alpha-numeric keypad.

The financial cost of sending multiple messages has also been cited as fuelling students’ use of textism abbreviations (Anis, 2007; Ling, 2004). It is possible that differences between countries in the price of text messages versus mobile phone calls may have contributed to international differences in texting behaviour (Baron & Ling, 2010). In recent
years, changes to messaging plans to include, for example, free text messages on weekends, or multiple messages for a set cost, may have further eased pragmatic pressures on current students to depend on abbreviated forms of spelling. As noted above, Australians have had higher levels of mobile phone ownership than Canadians throughout these changes, and this difference may have shaped the development of local texting cultures, in particular the popularity of using textisms.

Effects of methodology

Collection methods. Inconsistent results regarding textism use may also be explained by the varying ways in which text messaging data have been gathered by researchers. Some studies have relied on data from students’ self-reports of, for example, the words and phrases they most often abbreviate (e.g., Massengill Shaw et al., 2007) or the frequency with which they use textisms for given words (e.g., Rosen, Chang, Erwin, Carrier, & Cheever, 2010). Other researchers have used translation/dictation tasks (e.g., How would you write the following sentence as a text?), scenario-based exercises (e.g., What would you text a friend in the following situation?) and the naturalistic method of participants copying actual messages from their mobile phones (e.g., Ling & Baron, 2007; Thurlow & Brown, 2003). Only occasionally have researchers used more than one collection method. For example, Plester, Lerkkanen, Linjama, Rasku-Puttonen, and Littleton (2011) gathered both naturalistic and elicited messages, Drouin and Davis (2009) gathered both self-reports and translated messages, and Bodomo (2010) gathered both self-reports and naturalistic messages. However, there has been no systematic comparison of the possible effects of different collection methods on textism use.

In studies based on translation tasks, textism densities (proportion of textisms per message) of 43-57% have been reported in university students’ messages (Kemp, 2010), and 57-59% in children’s messages (Plester et al., 2008). Plester and colleagues (2009) developed
scenario-based tasks to elicit more spontaneous messages from children when naturalistic message collection proved difficult, and observed a textism density of 34% in the messages of 10- to 12-year-olds. Naturalistic messages from university students have contained considerably lower textism densities, ranging from 23% (Grace, Kemp, & Martin, 2010) down to five percent (Ling & Baron, 2007). Accurate estimations of textism type and density are crucial to discussions of students’ use of textese, and this study was designed to measure the influence of collection method on textism use by collecting messages from each participant via three methods: translation, elicitation, and naturalistic collection.

**Categorisation.** Finally, inconsistency also exists in the way that textisms have been categorised in previous research. For example, Shaw (2008) categorised *Fri* for *Friday* as a ‘clipping’, and *bein* for *being* as a ‘representation of a spoken form’, whereas Thurlow and Brown (2003) categorised *mon* for *Monday* as a ‘shortening’, and *goin* for *going* as a ‘g-clipping’. Similarly, *u* for *you* has been categorised by Ling and Baron (2007) as an ‘abbreviation’; Thurlow and Brown (2003) as a ‘letter/number homophone’, Rosen et al. (2010) as a ‘shortening’, and De Jonge and Kemp (2010) as a ‘logogram’. However, it should be noted that in all of these cases, these variations in coding would not result in variation in the overall counts of textisms produced.

In other cases, however, the total number of textisms has been affected by differences in the types of transformation that different researchers have counted. Spelling errors and typographical errors have been counted as textisms by some researchers (e.g., Thurlow & Brown, 2003; De Jonge & Kemp, 2010), but not by others (e.g., Ling & Baron, 2007; Plester et al., 2009). Similarly, lowercase *i* for *I* has been counted by some (e.g., Rosen et al., 2010; De Jonge & Kemp, 2010) and not others (e.g., Durkin et al. 2011; Plester et al., 2009). Finally, only in rare cases have ‘contractions’ been defined as including, for example, *can’t*
for *cannot* (Ling & Baron, 2007), or have ‘acronyms’ been defined as including standard acronyms such as *UK* for *United Kingdom* (Plester et al., 2008).

New categories have also been instituted to capture the use of new expressive devices. These include ‘emotion punctuation’, such as ........................, ‘emotion word’, such as *soooooo* (Varnhagen et al., 2009), and ‘all capitals’, such as *I AM ANGRY* (Rosen et al., 2010).

Textisms have also been grouped more broadly into ‘linguistic’ and ‘contextual’ categories, with the latter including emoticons, emotion punctuation, and all capitals (Rosen et al., 2010). The inclusion of such expressive devices allows recognition of the more creative strategies that texters incorporate in their messages to enhance communication (see Shortis, 2007).

Statistics describing textism use have inevitably been affected by these types of variation in categorisation. Despite the diversity of category names and definitions, a growing number of researchers are employing a fairly consistent categorisation of textisms based on those employed by Thurlow and Brown in 2003 (e.g., Durkin et al., 2011; Kemp, 2010; Plester et al., 2009), and we also based our initial categorisation on this scheme. However, we also considered textisms in terms of a broader division based on abbreviation of standard spellings, compared to expressive devices such as emoticons, emotion punctuation, and repeated letters or words.

A second categorisation issue concerns the counting of words which have been transformed in two or more ways. For example the textism *im* for *I’m* could be classified as belonging to the category of ‘missing capital’ and/or ‘missing apostrophe’. Some researchers (e.g., Plester et al., 2009) have categorised textisms only on the basis of the first transformation in the word, while others have considered all transformations (De Jonge & Kemp, 2010; Varnhagen et al., 2009). Others have not identified their procedure for coding these cases. Categorising only the first transformations within words potentially conceals important detail about the use of each transformation type.
Rationale

The present study is the first, to our knowledge, to investigate the potential effects of country, collection method, and categorisation method on the number and types of textisms observed in text-messages. Participants were students at two universities in two English-speaking countries: Australia and Canada. As noted earlier, mobile phones have been in more widespread use, and for longer, in Australia than in Canada. Thus we assume that in general, Australians would have had relatively more experience with text-messaging via the older technology of alpha-numeric keypads and multi-press entry, which encouraged the use of textisms. Accordingly, we hypothesised that Australian participants would produce significantly more textisms than their Canadian counterparts. In terms of collection method, we hypothesised that overall, students would produce significantly more textisms when translating messages than when writing them from a scenario, with the smallest proportion of textisms observed in participants’ naturalistic text-messages. Finally, we aimed to explore the use of different textism categories, both those used previously in the literature, and in terms of contracted versus expanded (expressive) forms.

Method

Participants

Participants were students in first-year Psychology and Education classes at two urban universities, one in South-Eastern Australia, and one in Western Canada. There were 97 Australian students (70 female), mean age 23.3 years ($SD = 7.84$), and 170 Canadian students (132 female), mean age 22.7 years ($SD = 4.24$). All but 24 had English as their first language (15 in Australia, nine in Canada). All students received course credit for their participation, and provided informed consent. The study had ethical approval from both universities.

Materials
**Questionnaires.** All participants completed a questionnaire requesting demographic information and information regarding their mobile phone’s text entry systems (e.g., use of predictive text messaging), their use of text messaging (e.g., an estimate of messages sent per day), and their opinion of the appropriateness of textism use. The questionnaire is shown in Appendix 1. Data collection was completed in late 2010.

**Text Messaging Tasks.** Three tasks were developed to assess the three main methods that have been used in previous research for collecting text messages. The sentences used in the Translation and Scenario tasks are listed in Appendix 2.

*Translation.* In the translation task, students were given five written sentences (e.g., *Are you coming over today, honey? No problem if you can’t*.), and asked to rewrite each one “as you would if you were going to send them as text messages”, writing the messages “exactly as you would type them into your phone, with any abbreviations, emoticons etc., that you would normally use.” No direct reference to ‘translation’ was made.

*Scenario.* In the scenario task, students were given five written scenarios (e.g., *Write down what you would text to set up a meeting between you and your best friend somewhere in town on Saturday night.*), and asked to “write a message for each of the following scenarios as you would if you were going to send them as text messages”, with the same further instructions as above.

*Naturalistic.* Students were then asked to copy down the last five messages they had sent from their phones (and were comfortable sharing), including more messages as necessary to total at least 50 words overall. This number was chosen so that there would be approximately the same number of words elicited in each of the collection methods. Participants were asked to write messages “exactly as they appear on your phone, including punctuation and emoticons”.

**Procedure**
Participants completed the questionnaires and text message tasks individually in a quiet room in their university. This took approximately 30 minutes, and was completed within a 45-75 minute session that included several other tasks not reported in this study. Each participant recorded messages via all three collection methods, with the order of the translation and scenario tasks counterbalanced and the naturalistic task completed third in each case.

**Categorisation of textisms.** The textism categorisation system used in the present study was based on that of Thurlow and Brown (2003), with several minor additions and alterations designed to capture further detail, as shown in Table 1. As noted earlier, some researchers have counted the number of words containing textisms, while others have counted the number of transformations made, even within words. To allow comparisons with previous research, textisms were initially counted via both methods, referred to as ‘textism words’ where each transformed word/phrase (e.g., lolllll) was counted as a single textism, and as ‘textism events’, where each transformation was counted separately (e.g., lolllll was counted as two textisms: ‘initialism’ and ‘extra letters’). To capture the full detail of textism types used in messages, we chose to report textism events for each analysis. Textism density was calculated for each participant by dividing the number of textisms produced by that participant by the number of words included in his or her messages.

(Table 1 about here)

**Results**

Students’ demographic data were collated, and their messages in all three collection methods were checked for textisms, which were coded according to the categories listed in Table 1. An initial independent-samples $t$-test was calculated to check that the students without English as a first language did not create a significantly different overall proportion of textism events than those with English as a first language. However, we found that the
messages of non-native English speakers (15 Australian, and nine Canadian students) included a significantly higher proportion of textism events ($M = .307, SD = .159$) than the native English speakers ($M = .195, SD = .101$), $t(265) = -3.40, p = .002$, although they did not differ significantly on other demographic variables (age, phone ownership, or number of sent/received messages). This relatively large difference in textism use between the two groups, as well as the linguistic disparity of the non-native group (reporting 16 different first languages) led to our decision to restrict our analyses to the native speakers of English who also texted in English. Thus two further Australian participants who reported that they texted in other languages most of the time were also excluded.

The final sample therefore included 86 Australian (63 female; mean age 23.3 years, $SD = 8.2$), and 155 Canadian (118 female; mean age 22.6 years, $SD = 4.2$) participants. In total, the Australian students provided 1337 messages (approximately 16,000 words, mean message length 11.9 words, $SD = 6.3$) and the larger sample of Canadian students provided 2367 messages (approximately 27,300 words, mean message length 11.3 words, $SD = 5.8$). This provided a mean of approximately 180 words per student. Textism data were combined across collection methods to calculate an overall density for both textism words and textism events, as reported in Table 2.

(Table 2 about here)

The overall .015 difference in textism density between the two textism counting methods may appear small. However, it represents approximately 660 textism transformations (or 8% of all textism events) that may have otherwise been ignored (e.g., the missing apostrophe in *im* for *I’m*). Independent-samples $t$-tests revealed that Australian students used significantly more textisms than Canadian students whether textism words, $t(239) = 3.13, p = .002$, or textism events, $t(239) = 3.06, p = .003$, were calculated. All further analyses were conducted using the textism events data.
Effects of Place: Attitudes, Texting Experience and Technology

The questionnaire included a 5-point Likert scale item designed to capture students’ opinions regarding the appropriateness of using text message-type abbreviations when texting a friend (where 1 = not at all appropriate and 5 = entirely appropriate). An independent-samples t-test revealed that the ratings did not differ significantly between Australia ($M = 4.48$, $SD = 0.86$) and Canada ($M = 4.54$, $SD = 0.78$), suggesting that any differences in textism use between the countries cannot be explained by differences in students’ attitudes regarding the appropriateness of textism use.

Demographic data regarding students’ phone use are summarised in Table 3. The table also displays the results of t-tests which confirmed that the Australian students had owned phones for significantly longer than the Canadian students, although the Canadian students gave significantly higher mean estimates of messages sent and received per day. (The score for one (Canadian) outlier, who reported sending 1000 messages per day, was excluded from the analysis.)

(Peason correlation coefficient were calculated to explore whether factors of texting experience (phone ownership and sent message volumes) are related to textism use. Overall textism density (collapsed across collection methods) was not significantly related to the length of time that students had owned mobile phones in either country (Australia: $n = 86$, $r = .11$, $p = .31$, Canada: $n = 155$, $r = -.029$, $p = .72$). The correlation between textism events and number of messages that students sent daily was not significant for Canadian students ($n = 154$, $r = .051$, $p = .53$), but it was for Australian students ($n = 86$, $r = -.22$, $p = .047$). Contrary to our hypothesis, in the Australian sample, more frequent senders of messages used fewer textisms. Students also noted their text messaging hardware (e.g., whether their mobile phone has an alpha-numeric keypad), and software (e.g., whether they use predictive software), as
reported in Table 4. The percentages shown in Table 4 regarding software use reflect the text entry strategy used *most* of the time. However, it should be noted that approximately 38% of the Australian and Canadian students reported using multiple entry systems at least some of the time (e.g., switching between multi-press and single-press predictive texting). Australian students reported using phones with alpha-numeric keypads and with full keyboards approximately equally, whereas for Canadians, full keyboards were much more widely used. Reflecting these differences, Australians reported using predictive entry most commonly, whereas Canadians used non-predictive entry most commonly.

We conducted two two-way analyses of variance (ANOVAs) to examine the effect of hardware and software use on the density of textism events in students’ text messages. (The few students who reported using both hardware systems, or using several software systems approximately equally, were excluded.) The ANOVA testing the potential influence of hardware choice on textism use had one between-subjects factor, Country (Australia, Canada), and one within-subjects factor, Hardware Type (keypad, keyboard). There was a significant main effect of Country, with more textisms produced by Australian than Canadian students, $F(1, 232) = 7.08, p = .008, \eta^2_p = .030$, and a significant main effect of Hardware Type, with more textisms produced by users of alpha-numeric keypads than full keyboards, $F(1, 232) = 6.25, p = .013, \eta^2_p = .027$. The interaction between Country and Hardware Type was not significant, $F(1, 232) = 0.75, p = .39, \eta^2_p = .003$.

The ANOVA for students’ use of textisms according to mobile phone software had one between-subjects factor, Country (Australia, Canada) and one within-subjects factor, Software Type (multi-press, predictive, non-predictive). Significantly more textisms were produced by Australian than Canadian students, $F(1, 233) = 11.31, p = .001, \eta^2_p = .047$, and the main effect of Software Type was also significant, $F(2, 233) = 7.98, p < .001, \eta^2_p = .065$. 
Tukey post-hoc tests showed that messages written by students who relied predominantly on multi-press entry contained more textisms than messages written by those who used single-press predictive ($p = .006$), or non-predictive ($p < .001$) systems. Messages written by students who used predictive systems contained more textisms than those written by students who used non-predictive systems, but this difference was not significant ($p = .18$). The interaction between Country and Software Type was not significant, $F(2, 233) = 2.06, p = .13, \eta_p^2 = .018$.

Some previous research has asked participants to report on their typical use of textisms. In the current study, only three students reported using textisms “most of the time” in their everyday messages, and approximately half of the participants in both countries rated their textism use as occurring “some of the time”, and half as “none of the time”. As the participants who selected “none of the time” used textisms for about 13% of the words in the naturalistic messages that they provided, it seems that students’ ability to judge their own textism use was not necessarily accurate.

**Effects of Collection Method**

We then considered the important question of whether the proportion of textisms produced by students in the two countries varied with collection method, as seen in Table 5.

(Table 5 about here)

As shown in the table, Australians used more textisms than Canadians in each method, and students in both countries used more textisms in translation than scenario messages, and in turn, than in naturalistic messages. A repeated-measures ANOVA was conducted with one between-subjects factor, Country (Australia, Canada), and one within-subjects factor, Method (translation, scenario, naturalistic), with mean proportion of textism events as the dependent variable. The order in which the translation and scenario tasks were administered did not make a significant difference to proportion of textisms produced, $F(1, 229) = 0.21, p = .65$. 
\[ \eta_p^2 = .001, \] and thus was not included in our analyses. After Greenhouse-Geisser corrections, there was a significant main effect for Country, \( F(1, 231) = 10.03, p = .002, \eta_p^2 = .042, \) with Australians producing more textisms than Canadians. A significant main effect of Method was also found, \( F(2, 462) = 38.45, p < .001, \eta_p^2 = .14. \) Tukey post-hoc tests confirmed that, as predicted, messages produced via the translation method contained more textisms than via the scenario method \( (p < .01), \) which in turn contained more textisms than the naturalistic messages \( (p < .01). \) The interaction between Country and Method was not significant, \( F(2, 462) = 2.47, p < .094, \eta_p^2 = .011. \) Thus, textism collection method does appear to influence textism use, but this does not vary significantly across the two countries.

**Categorisation of Textisms**

Textisms were coded according to the 19 categories used in previous research, as noted earlier, so that we could determine the potential effects of collection method on types of textisms produced. The means and standard deviations for textism events within these categories are reported in Table 6. Categories which occurred less than 0.01 of the time were combined into ‘other categories’ \( (g\text{-clippings, other symbols, extra punctuation/letters/words, spelling errors, other textisms).} \)

(\textit{Table 6 about here})

As can be seen in Table 6, overall the highest proportions of textisms were youth stylisations, omitted capitals, and homophones, followed by shortenings, contractions and omitted apostrophes. Patterns were similar between Australia and Canada, with the exceptions of homophones, omitted apostrophes, and expressive symbols, which occurred much more frequently in the Australian messages. However, consistent differences can be seen across collection methods. It appears from the means that for both countries, there was a higher mean proportion of textism transformations in the categories of single homophones, combined homophones, and contractions in the translation condition than the scenario
condition, and in the scenario condition than the naturalistic condition. These categories contain words that are shortened in some way (e.g., 2day for today, @ for at). Conversely, proportions of transformations in three categories involving the addition of characters (expressive symbols, extra letters and extra words), were found in the highest proportions in the naturalistic condition, and in the lowest proportions in the translation condition. In order to examine such broader patterns, we grouped transformation types into two new, larger categories: ‘contractive’ and ‘expressive’ textisms.

**Contractive versus expressive textism types.** Contractive textisms were defined as transformations in which characters were omitted from standard words, and therefore included all textisms previously coded as contractions, shortenings, g-clippings, omitted apostrophes, initialisms, other symbols, combined and single homophones, and other textisms. Expressive textisms were defined as transformations in which extra characters were added to words or messages, and therefore included all textisms previously coded as expressive symbols, extra words, extra letters, and extra punctuation. All of these transformations appear to add expressive intent to the message. We note that this method of categorisation excluded categories that contained a mixture of contractive and expressive forms (youth stylisations and nonstandard spellings), and categories that could not easily be associated with either broader category (e.g., spelling errors); approximately 44% of textism events. A summary of the resulting textism transformation proportions included in each of these novel categories is reported in Table 7.

(Table 7 about here)

As shown in the table, there were more contractive than expressive textisms for both countries and across all three collection methods. Similar to results for overall textism density, Australians used more textisms than Canadians in each method, and students from both countries showed decreasing use of contractive textisms with increasingly realistic...
message types. However, the opposite trend was seen for expressive textisms, with decreasing use of expressive textisms in increasingly realistic text messages. A repeated-measures ANOVA was conducted with one between-subjects factor, Country (Australia, Canada) and two within-subjects factors, Method (translation, scenario, naturalistic) and Textism Type (contractive, expressive). To make the patterns more readily interpretable, the means are also shown in Figure 1.

(Figure 1 about here)

The contractive and expressive data were not normally distributed, and therefore an arcsine transformation was made to the data before the ANOVA was conducted. Significant main effects were found for Country, $F(1, 237) = 37.29, p < .001, \eta_p^2 = .14$; Method, $F(2, 474) = 29.93, p < .001, \eta_p^2 = .11$; and Textism Type, $F(1, 237) = 108.16, p < .001, \eta_p^2 = .31$. These main effects were modified by a significant interaction for Method by Textism Type, $F(2, 474) = 84.99, p < .001, \eta_p^2 = .26$. Interactions for Country by Method, $F(2, 474) = 1.05, p = .349, \eta_p^2 = .004$, and Country by Textism Type, $F(1, 237) = 1.75, p = .188, \eta_p^2 = .007$, were not significant. However, there was a significant three-way interaction between Country, Method and Textism Type, $F(2, 474) = 9.46, p < .001, \eta_p^2 = .038$.

A series of Tukey post-hoc tests were conducted on the three-way interaction. While in all cases Australians used more textisms than Canadians, Tukey post-hoc tests confirmed that this was significant for contractive textisms in the translated and scenario-based messages, $ps < .01$, and the naturalistic messages, $p < .05$, and for expressive textisms in the naturalistic messages only, $p < .01$. Similarly, there were more contractive than expressive textisms in all cases. This was significant for both countries in naturalistic and scenario-based messages, $ps < .01$, but only for Canadians in naturalistic messages, $p < .05$.

There were some differences in the ways that contractive and expressive textisms were used, depending on the collection method and the country. For contractive textisms,
proportions were higher for students in both countries in translation than scenario messages, and in turn, than in naturalistic messages. However, Tukey post-hoc tests confirmed this was significant for only translated compared to scenario-based and naturalistic messages for both countries, $ps < .01$, and for scenario-based compared to naturalistic messages for Australia, $p < .01$. For expressive textisms, Tukey post-hoc tests revealed that for Australians, proportions were significantly lower in translated than scenario-based messages, and in turn, than in naturalistic messages, $p < .01$, but for Canadians, no differences reached significance.

**Discussion**

The focus of this research was to investigate the potential effects of place and collection method on textism use. The results show that the number and type of textisms present in students’ messages were affected by where messages were collected (Australia or Canada), by the method by which messages were collected (translation, scenario, or naturalistic methods), and by the method by which textisms were categorised.

**Effects of Place**

The use of textisms differed significantly between countries, with Australian students using more textisms than Canadians in every analysis. These differences suggest that effects of place may explain some of the inconsistency in previously reported textism densities (e.g., Grace et al., 2010; Ling & Baron, 2007; Thurlow & Brown, 2003). There are various potential explanations for why textism use should differ between countries, but we hypothesised that differences in mobile phone technology and texting experience were likely to be at their hub. As outlined earlier, patterns of mobile phone ownership have varied between Australia and Canada (CIA, 2004, 2010), and the availability of technologies such as full keyboard entry systems and predictive texting also differed for our participants. In our sample, Australian students had owned a phone longer, whereas Canadian students sent more messages daily. However, neither of these variables was significantly correlated with
students’ textism use. Therefore the hypothesis that longer texting experience (i.e., during the more abbreviation-rich early development of textese) would have a measurable influence on textism density was not supported.

Students’ opinions regarding the appropriateness of textism use were almost identical between the countries, a finding which rules out local differences in attitudes as an alternative explanation for differences between countries. However, as also found by Kemp and Bushnell (2011), the current participants who used multi-press systems produced significantly more textisms overall than students who used single-press entry systems (either predictive or non-predictive). Further, higher proportions of textisms were used by students who relied on alpha-numeric keypads than full keyboards. Reflecting these results, Australian students used a higher proportion of textisms, and were more likely to use alpha-numeric and multi-press systems than Canadian students, who were more likely to use non-predictive and full keyboard entry systems. Therefore, the reason for the greater textism density observed in Australian than Canadian students’ messages seems to stem from current differences in hardware and software use, rather than from differences in overall texting experience. Continued advances in keyboard design and predictive technology may influence the proportion of textisms that students employ in the future.

Effects of Methodology

Inconsistencies in methodology employed in previous research, in both categorisation and collection of textisms, may also explain differences in reports of textism densities. In the current study, when all textism types were combined, participants produced significantly more textisms in the translation than scenario-based messages and in the scenario-based messages than in the naturalistic messages. These results support our hypothesis that artificial methodologies may elicit higher textism densities from students than are found in real-life messages. This finding may also explain some of the inconsistency in previous reports of
students’ textism use. Higher textism proportions have been reported in studies relying on translated messages than scenario-based messages, and lower proportions in studies involving naturalistic messages (e.g., Grace et al., 2010; Kemp, 2010; Plester et al., 2009). The current study also showed that students’ self-reports of textism use in a questionnaire item were not reliable, with about 13% textisms observed even in the naturalistic messages of those who reported that they normally used no textisms. Researchers would thus do well to interpret results in previous research with potential effects of collection method in mind, and to use naturalistic messages where possible.

In terms of the type of textism categorisation used in previous studies (e.g., Plester et al., 2009; Thurlow & Brown, 2003), the patterns of textism use seen across all three collection methods in the present study were broadly similar to those reported by previous researchers (De Jonge & Kemp, 2010; Neville, 2003; Plester et al., 2009). However, textism transformations have been counted and categorised in slightly different ways, which could have an effect on the specific conclusions drawn. By counting all textism transformations, even when more than one appeared in a word, some textisms were analysed that would have otherwise been ignored. We also extended a commonly used categorisation schema (Thurlow & Brown, 2003) to include less commonly used categories that recognise the expressive nature of some textisms currently in use (e.g., extra letters). The combination of these strategies captured detail in the pattern of textism use that in turn led to a simplified, broader categorisation of textisms. The new categories of textism transformations were ‘contractives’ (with characters removed, e.g., pls for please) and ‘expressives’ (with characters added, e.g., pleeease for please). The term ‘expressive’ refers to the emotional intent and added communicative value contained in these textisms.

Analyses of the broader categories showed that while students overestimated their everyday use of contractive textisms in the experimental methods (translation and scenario),
they may have underestimated their use of expressive textisms in the same. Mean proportions of expressive textisms were higher in translation than scenario messages for both countries, and in scenario than naturalistic messages for Australia. Therefore, results based on experimental methods may lead not only to exaggerated estimates of textism use overall, but more specifically to exaggerated estimates of the abbreviated forms of textisms referred to in media concerns regarding texting effects on literacy. Conversely, it is possible that aspects of heightened communication used in naturalistic messages, such as representations of tone of voice (e.g., *pleeease*), and facial expression (e.g., :{-D}), have been underrepresented.

Several reasons may underlie the higher proportion of contractive textisms found in translated and scenario-based messages than in messages copied from students’ phones. It may be that perceived researcher expectations and the nature of the experimental tasks themselves prime students to use more textisms. For example, students might try to think of a textism they might use for each word of the translation task, and for each word of the potential message that comes to mind for each scenario, rather than focusing on the instruction to write messages as they would in real life. Researchers’ choice of particular ideas or words (which are often subject to text abbreviation) might stimulate the use of specific potential textisms, which may also contribute to increased textism use. For example, Kemp (2010) asked students to write lists of textisms, and then created messages that included these words for use in her experimental task. The experimental setting may also preclude the full expression of emotion, tone of voice, and facial expression that students try to represent in real-life messages (e.g., *xoxo, heeeelllllooo0000000000* for *hello*, and :{-p respectively). It is possible that any combination of these factors have affected previous efforts to examine links between textism use and literacy skills (e.g., Drouin & Davis, 2009; Kemp & Bushnell, 2011).
The present study shows that factors of both place and methodology may explain some differences in the density of textisms reported in previous research. It appears that Australians’ tendency to use more textisms than Canadians may be explained in terms of Australians’ greater use of older technology, the constraints of which encourage abbreviated messaging. Further, naturalistic text messages seem to show lower densities of textisms than do messages collected via translation or scenario-based elicitation tasks. Future researchers would do well to collect naturalistic messages where possible. Finally, texters continue to create a wide variety of textisms, and we recommend that future researchers consider the use of expressive devices in examinations of language use in text messages.
References


Hillebrand, F. (2010). *Short Message Service*. West Sussex, UK: John Wiley & Sons Ltd.


Table 1

*Categorisations Applied to Textism Transformations.*

<table>
<thead>
<tr>
<th>Category</th>
<th>Example</th>
<th>Category, continued</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single homophones</td>
<td><em>u</em> for <em>you</em>, <em>2</em> for <em>to</em></td>
<td>Symbols: Expressive</td>
<td><em>xxx, xox, :D, =P, ;_-;</em></td>
</tr>
<tr>
<td>Combined h’phones</td>
<td><em>2nite</em> for <em>tonight</em></td>
<td>Symbols: Other</td>
<td>*@ for <em>at</em>, &amp; for <em>and</em></td>
</tr>
<tr>
<td>Shortenings</td>
<td><em>mon</em> for <em>Monday</em></td>
<td>Omitted capitals</td>
<td><em>i</em> for <em>I</em>, <em>sarah</em> for <em>Sarah</em></td>
</tr>
<tr>
<td>Constructions</td>
<td><em>msg</em> for <em>message</em></td>
<td>Extra capitals</td>
<td><em>HAPPY BIRTHDAY</em></td>
</tr>
<tr>
<td>Omitted apostrophes</td>
<td><em>cant</em> for <em>can’t</em></td>
<td>Extra punctuation</td>
<td><em>???, !!!!!!, ??</em></td>
</tr>
<tr>
<td>g-clippings</td>
<td><em>goin</em> for <em>going</em></td>
<td>Extra letters</td>
<td><em>pleeease</em> for <em>please</em></td>
</tr>
<tr>
<td>Initialisms</td>
<td><em>btw</em> for <em>by the way</em></td>
<td>Extra words</td>
<td><em>cool cool</em></td>
</tr>
<tr>
<td>Nonstandard spelling</td>
<td><em>cos</em> for <em>because</em></td>
<td>Spelling errors</td>
<td><em>rekon</em> for <em>reckon</em></td>
</tr>
<tr>
<td>Youth stylisation</td>
<td><em>gonna</em> for <em>going to</em></td>
<td>Other textisms</td>
<td><em>abit</em> for <em>a bit</em></td>
</tr>
</tbody>
</table>
Table 2

*Mean Textism Densities (Standard Deviations in Parentheses) for ‘Textism Words’ and ‘Textism Events’.*

<table>
<thead>
<tr>
<th>Country</th>
<th>Textism words</th>
<th>Textism events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia (n = 86)</td>
<td>.203 (.103)</td>
<td>.221 (.117)</td>
</tr>
<tr>
<td>Canada (n = 155)</td>
<td>.164 (.073)</td>
<td>.178 (.082)</td>
</tr>
<tr>
<td>Overall (n = 241)</td>
<td>.178 (.087)</td>
<td>.193 (.098)</td>
</tr>
</tbody>
</table>

*Note:* For textisms words, each transformed word was counted once; for textism events, each transformation was counted, even if there was more than one within a word.
Table 3

*Means (Standard Deviations in Parentheses) for Mobile Phone Ownership and Text Messaging Use, with t-values Showing Differences between Countries.*

<table>
<thead>
<tr>
<th>Country</th>
<th>Australia (n = 86)</th>
<th>Canada (n = 155)</th>
<th>t-values</th>
<th>Overall (n = 241)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phone ownership (years)</td>
<td>7.13 (2.44)</td>
<td>5.91 (2.37)</td>
<td>3.78***</td>
<td>6.35 (2.46)</td>
</tr>
<tr>
<td>Texts sent per day</td>
<td>23.93 (31.90)</td>
<td>39.12 (44.17)</td>
<td>3.06**</td>
<td>33.66 (40.78)</td>
</tr>
<tr>
<td>Texts received per day</td>
<td>23.97 (31.71)</td>
<td>39.27 (44.14)</td>
<td>3.10**</td>
<td>33.76 (40.72)</td>
</tr>
</tbody>
</table>

**p < 0.01, ***p < 0.001
Table 4

Percentages of All Students Regarding Their Self-Reported Phone and Textism Use, and Mean Proportions of Textism Events (Standard Deviations in Parentheses).

<table>
<thead>
<tr>
<th>Country</th>
<th>% of students</th>
<th>Textism Proportions</th>
<th>% of students</th>
<th>Textism Proportions</th>
<th>% of students</th>
<th>Textism Proportions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia (n = 86)</td>
<td></td>
<td></td>
<td>Canada (n = 155)</td>
<td></td>
<td>Overall (n = 241)</td>
<td></td>
</tr>
<tr>
<td>Text entry: hardware</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Keypad (alpha-numeric)</td>
<td>53</td>
<td>.242 (.115)</td>
<td>28</td>
<td>.195 (.088)</td>
<td>37</td>
<td>.219 (.105)</td>
</tr>
<tr>
<td>Keyboard (one letter/key)</td>
<td>47</td>
<td>.197 (.116)</td>
<td>66</td>
<td>.173 (.080)</td>
<td>59</td>
<td>.179 (.092)</td>
</tr>
<tr>
<td>Both systems used</td>
<td>0</td>
<td>-</td>
<td>5</td>
<td>.147 (.073)</td>
<td>3</td>
<td>.147 (.073)</td>
</tr>
<tr>
<td>Text entry: software*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multi-press</td>
<td>14</td>
<td>.319 (.146)</td>
<td>9</td>
<td>.214 (.104)</td>
<td>11</td>
<td>.262 (.134)</td>
</tr>
<tr>
<td>Single-press predictive</td>
<td>55</td>
<td>.211 (.092)</td>
<td>34</td>
<td>.187 (.090)</td>
<td>41</td>
<td>.199 (.091)</td>
</tr>
<tr>
<td>Single-press non-predictive</td>
<td>31</td>
<td>.196 (.124)</td>
<td>54</td>
<td>.169 (.071)</td>
<td>46</td>
<td>.175 (.087)</td>
</tr>
<tr>
<td>Multiple systems equally</td>
<td>0</td>
<td>-</td>
<td>3</td>
<td>.148 (.104)</td>
<td>2</td>
<td>.148 (.104)</td>
</tr>
</tbody>
</table>
Use of textisms in messages

<table>
<thead>
<tr>
<th></th>
<th>Most of the time</th>
<th>Some of the time</th>
<th>None of the time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>57</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>.278 (.117)</td>
<td>.146 (.064)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>46</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>.234 (.071)</td>
<td>.226 (.082)</td>
<td>.132 (.053)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>50</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>.234 (.071)</td>
<td>.247 (.100)</td>
<td>.137 (.057)</td>
</tr>
</tbody>
</table>

* System used most of the time

*Note*: Due to rounding, some percentages do not add to 100
Table 5

*Mean Proportions of Textism Events (Standard Deviations in Parentheses) for Each Collection Method.*

<table>
<thead>
<tr>
<th>Collection Method</th>
<th>Translation</th>
<th>Scenario</th>
<th>Naturalistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>86 (.268 .183)</td>
<td>86 (.218 .126)</td>
<td>86 (.188 .091)</td>
</tr>
<tr>
<td>Canada</td>
<td>152 (.204 .122)</td>
<td>155 (.183 .101)</td>
<td>150 (.156 .085)</td>
</tr>
<tr>
<td>Overall</td>
<td>238 (.227 .150)</td>
<td>241 (.195 .111)</td>
<td>236 (.167 .088)</td>
</tr>
</tbody>
</table>

*Note:* Means were calculated by averaging individual student means.
Textese: Place and collection method

Table 6

Mean Proportions (Standard Deviations in Parentheses) of Textism Events in Each Category for Both Countries and Each Collection Method.

<table>
<thead>
<tr>
<th>Category</th>
<th>Australia Translation</th>
<th>Australia Scenario</th>
<th>Australia Naturalistic</th>
<th>Canada Translation</th>
<th>Canada Scenario</th>
<th>Canada Naturalistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sing. Homoph.</td>
<td>0.044 (.069)</td>
<td>0.042 (.068)</td>
<td>0.016 (.032)</td>
<td>0.024 (.044)</td>
<td>0.017 (.039)</td>
<td>0.011 (.022)</td>
</tr>
<tr>
<td>Comb. Homoph.</td>
<td>0.015 (.027)</td>
<td>0.005 (.013)</td>
<td>0.003 (.009)</td>
<td>0.004 (.010)</td>
<td>0.002 (.007)</td>
<td>0.001 (.003)</td>
</tr>
<tr>
<td>Shortenings</td>
<td>0.028 (.025)</td>
<td>0.020 (.022)</td>
<td>0.011 (.015)</td>
<td>0.022 (.023)</td>
<td>0.008 (.012)</td>
<td>0.012 (.017)</td>
</tr>
<tr>
<td>Contractions</td>
<td>0.020 (.022)</td>
<td>0.012 (.017)</td>
<td>0.005 (.011)</td>
<td>0.017 (.018)</td>
<td>0.013 (.013)</td>
<td>0.003 (.009)</td>
</tr>
<tr>
<td>Omitted Apost.</td>
<td>0.035 (.035)</td>
<td>0.005 (.010)</td>
<td>0.012 (.018)</td>
<td>0.002 (.006)</td>
<td>0.003 (.011)</td>
<td>0.001 (.006)</td>
</tr>
<tr>
<td>Initialisms</td>
<td>0.008 (.011)</td>
<td>0.003 (.009)</td>
<td>0.009 (.015)</td>
<td>0.010 (.012)</td>
<td>0.001 (.004)</td>
<td>0.009 (.015)</td>
</tr>
<tr>
<td>Nonst’d Spell.</td>
<td>0.018 (.025)</td>
<td>0.004 (.010)</td>
<td>0.005 (.011)</td>
<td>0.009 (.013)</td>
<td>0.003 (.010)</td>
<td>0.001 (.005)</td>
</tr>
<tr>
<td>Youth Styl’n</td>
<td>0.028 (.033)</td>
<td>0.052 (.031)</td>
<td>0.038 (.029)</td>
<td>0.020 (.022)</td>
<td>0.055 (.031)</td>
<td>0.038 (.032)</td>
</tr>
<tr>
<td>Symb. Express.</td>
<td>0.023 (.032)</td>
<td>0.031 (.034)</td>
<td>0.037 (.037)</td>
<td>0.011 (.015)</td>
<td>0.015 (.021)</td>
<td>0.009 (.014)</td>
</tr>
<tr>
<td>Omit. Capitals</td>
<td>0.025 (.028)</td>
<td>0.031 (.031)</td>
<td>0.037 (.027)</td>
<td>0.028 (.033)</td>
<td>0.037 (.041)</td>
<td>0.040 (.039)</td>
</tr>
<tr>
<td>Extra Capitals</td>
<td>0.011 (.017)</td>
<td>0.001 (.003)</td>
<td>0.002 (.010)</td>
<td>0.011 (.015)</td>
<td>0.005 (.017)</td>
<td>0.004 (.011)</td>
</tr>
<tr>
<td>Other Categories</td>
<td>0.011 (.016)</td>
<td>0.012 (.018)</td>
<td>0.011 (.013)</td>
<td>0.024 (.025)</td>
<td>0.019 (.026)</td>
<td>0.017 (.020)</td>
</tr>
</tbody>
</table>
Table 7

*Mean Proportions (Standard Deviations in Parentheses) of ‘Contractive’ and ‘Expressive’ Textism Transformations in Each Collection Method for Both Countries.*

<table>
<thead>
<tr>
<th>Country</th>
<th>Translation Contractive</th>
<th>Translation Expressive</th>
<th>Scenario Contractive</th>
<th>Scenario Expressive</th>
<th>Naturalistic Contractive</th>
<th>Naturalistic Expressive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>.157 (.142)</td>
<td>.026 (.037)</td>
<td>.092 (.105)</td>
<td>.035 (.037)</td>
<td>.061 (.055)</td>
<td>.043 (.039)</td>
</tr>
<tr>
<td>Canada</td>
<td>.090 (.089)</td>
<td>.017 (.023)</td>
<td>.050 (.060)</td>
<td>.025 (.033)</td>
<td>.040 (.045)</td>
<td>.018 (.023)</td>
</tr>
<tr>
<td>Overall</td>
<td>.114 (.115)</td>
<td>.020 (.029)</td>
<td>.065 (.082)</td>
<td>.029 (.035)</td>
<td>.047 (.049)</td>
<td>.027 (.032)</td>
</tr>
</tbody>
</table>
Figure 1

Mean Proportions of Contractive and Expressive Textism Transformations in Each Collection Method for Both Countries.
Appendix 1

Text messaging questionnaire

Age: _____ years and _____ months    Sex:    Male    Female

Is English your first language?    Yes    No    **If No:** What is your first language? ______

Do you always send and receive messages in English?    Yes    No    **If No:** What percent of the time do you send ______% and receive ______% text messages in English? What other languages do you text in? ______________

1. How long have you owned a mobile phone? _____ months     OR   _____ years

2. How often do you usually send or receive text messages? On average, about:
   
   **Send:** ___________ messages per day    **Receive:** ___________ messages per day

3. What type of keypad does your phone have?
   
   □ Alpha-numeric keypad (with letters written above number keys like on a landline phone)
   □ Qwerty keypad (like a computer keyboard)
   □ Touch-screen keypad (with separate touch-screens for numbers/letters/symbols)
   □ Other (please explain) ________________

4. About what percent of the time do you use each of the following text entry systems? % of the time
   
   Single-press predictive (press each key once and phone predicts likely word) ______
   Single-press non-predictive (press each key once without needing/using prediction) ______
   Multi-press (press each number key 1 to 4 times for the right letter) ______
   Other (please explain) ____________________________ ______

5. Which statement best describes the way that you write text messages? Tick one.

   **I use abbreviations and “text message” style spellings for…**
   
   □ Most of the words that I write.
   □ Only some of the words, such as “u” for “you”.
   □ Nothing at all; I try to write all my words in conventional English.
6. Which statement best describes the way you feel when you **read** text messages? Tick one.

- I find it easy to work out what the sender meant.
- I know the obvious abbreviations, but I sometimes struggle with the more obscure ones.
- I find it difficult to work out many of the abbreviations.

7. How **appropriate** do you think it is to use text-message-style abbreviations in each of the following type of message? (e.g., “u” for you, “lol” for laughing out loud, smiley faces)

**Texting a friend:** Not at all appropriate 1 2 3 4 5 Entirely appropriate

*[Other appropriateness items are not analysed in this paper.]*

**Appendix 2**

**Sentences used in the translation exercise:**

1. Are you coming over today, honey? No problem if you can’t.
2. Happy birthday! See you tonight at seven o’clock. It’s going to be awesome!
3. Yes, love, let’s meet at Lazenby’s on Friday, 11.45am.
4. Do you want to get your dad to pick us up, because Sarah can’t, okay?
5. By the way, did you get my text about tomorrow?

**Sentences used in the scenario exercise: Write down what you would text to...**

1. Set up a meeting between you and your best friend somewhere in town on Saturday night.
2. Your mum to ask her if she can pick something up from the supermarket for you on her way to your place.
3. A friend to tell them you are going to be five minutes late to meet them.
4. Congratulate your study partner on finishing a big essay and to ask them if they are going to attend a party that night.
5. Someone you know from uni to find out whether or not they are going to go to a lecture this afternoon, and whether they can take notes on your behalf.