Alternative Methods for Measuring TXTMSG Frequency

Missouri Western State University | Mary L. Still | Alison L. Morris | Iowa State University

Overview

With increased use of text-messaging abbreviations comes an increased need to understand how they are recognized. Unfortunately, large-scale normative data are not available for text-messaging abbreviations. We present a case in which the absence of normative data (frequency) could lead to questionable conclusions. In doing so, we consider potential measures of abbreviation frequency.

Recognition of Abbreviations

Abbreviations as Lexical Units

Common abbreviations can facilitate lexical decisions to related words (Brysbaert, Speybroeck, & Vanderelst, 2009) and elicit an N400 (Laszlo & Federmeier, 2007) providing evidence for the lexical status of abbreviations.. ***ABS*** can prime ***BRAKES***

Emotion and Text-Messaging Abbreviations

Even with lexical status, abbreviations may not be functionally equivalent to the words they represent. Still, Morris, and Jones (2011) used an emotional Stroop task to examine emotional responses to text-messaging abbreviations and their translations.

I hate it Emotional Stroop Effect **IH8IT** No Effect

These results suggest that emotion is not automatically elicited when abbreviations are recognized. Perhaps a more sensitive measure would reveal an effect.

Perceptual Identification of Abbreviations

Method

A perceptual identification task was used to investigate processing of text-messaging abbreviations. Each trial began with a fixation cross followed by a briefly displayed letter string (90 msec) and pattern mask (####; 500 msec). Participants typed what they had seen. Abbreviation status (abbreviation vs. pseudo-abbreviation) was counterbalanced across participants.

	Negative (13 each)	Neutral (25 each)	Positive (13 each)
Abbreviation	**WTF MOFO**	**ASAP TTYL**	**LMAO GR8**
Pseudo-Abbreviation	**WRF DOFO**	**ASEP TFYL**	**LMEO GT8**

Participants

26 students (12 male) participated. A post-experiment survey revealed that participants knew the meaning of 86% of the abbreviations and 88% text daily and use abbreviations.

Results

Main Effect of Abbreviation Status

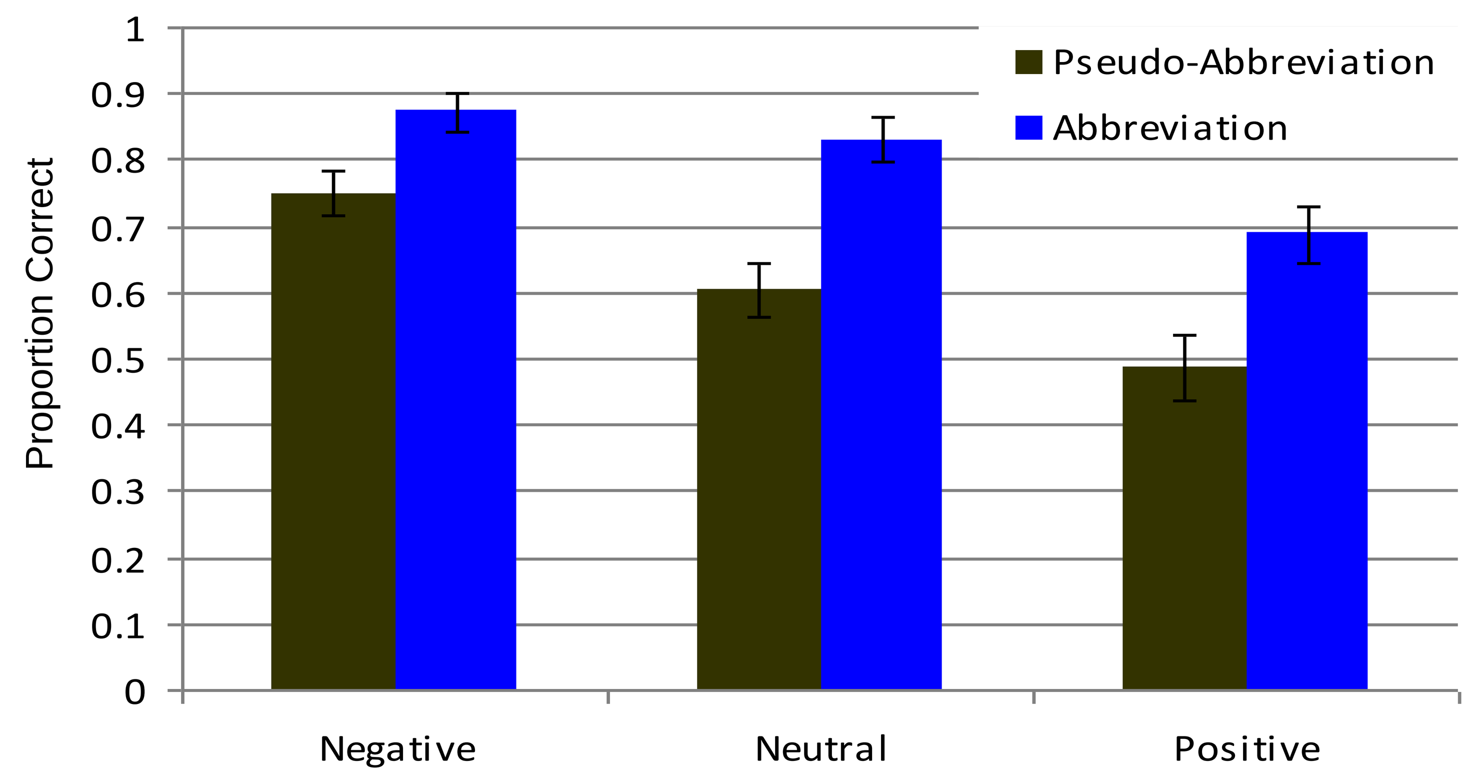
$F(2,50) = 39.90$, $p < .001$

Higher accuracy for abbreviations.
Evidence for lexical status of abbreviations.

Main Effect of Emotion*

$F(1,25) = 28.33$, $p < .001$

Post-hoc comparisons revealed significantly higher accuracy for negative abbreviations than for neutral or positive abbreviations and higher accuracy for neutral than for positive abbreviations (all *p*s < .01).



*Conclusion or Confound?

One might be tempted to conclude that abbreviations automatically elicit emotion, capture attention, and thereby modulate PID accuracy. But stimulus frequency may influence perceptual identification. If negative, neutral, and positive abbreviations have different frequencies, the effect of emotion may be a spurious result.

Measures of Frequency

Traditional Corpora

Large collections of text are typically used to approximate printed word frequency. There are no freely available text-messaging corpora and developing corpora may be cost prohibitive (e.g., Underwood, Rosen, More, Ehrenreich, & Gentsch, 2011).

Search Engine Hits

Traditional corpus frequency counts are highly correlated with internet search engine hit rates (number of results for a given search) as long as the search engine relies on a large database (Blair, Urland, & Ma, 2002; Caldwell-Harris, Berant, & Edelman, 2012).

Familiarity Ratings

Participant subjective ratings of familiarity have been shown to predict stimulus frequency for word pairs (e.g., *legal fees*; Caldwell-Harris et al., 2012), but a similar relationship is not necessarily obtained for single words (Blair et al., 2002).

TXTMSG Abbreviation "Frequency"

1. Google was used to obtain hit rates for each abbreviation. These data were not normally distributed; thus, the log transformation of the total number of search results was used.
2. **Familiarity ratings*** were obtained using the following scale:

   1 – don't know what it means | 2 – seen it before, don't know what it means | 3 – seen it before, know what it means | 4 – I use it, know what it means

3. **Knowledge of abbreviation meanings*** may be a proxy for frequency - more likely to know the meanings of frequently encountered abbreviations. Values were obtained by having participants write out the meanings of abbreviations. After transformation, lower values represent abbreviations known by many participants, higher values represent abbreviations known by fewer participants.

*Values obtained from approx.100 participants *not* involved in the PID study.

Reconsidering the Results

Correlations between Perceptual Identification and Frequency Measures

	Google	Familiarity	Meaning
PID Accuracy	.748*	.171	-.038
Google		.225	-.011
Familiarity			-.717*

***p <.001**

Subjective familiarity and knowledge of abbreviations were highly correlated. Interestingly, Google frequency was the only measure predictive of PID performance.

Item Analyses for Abbreviation PID: Frequency Effects

	ANOVA Results			Post-hoc Analyses (*p*-values)		
Analysis w/ Frequency	Main Effect	*F*-ratio	*p*-value	Neg-Neu	Pos-Neu	Neg-Pos
No covariate	Emotion	2.5	.095	1.000	.187	.143
Google covariate	Emotion	1.3	.285	1.000	.520	.442
	Google	54.9	< .001			
Familiarity covariate	Emotion	5.1	.010	.390	.066	.009
	Familiarity	6.5	.014			
Meaning covariate	Emotion	2.8	.069	1.000	.167	.096
	Meaning	.8	.376			

Conclusions

- Interpretation of these PID results depend *critically* on which frequency estimate is used
  - Google counts best predict PID accuracy, but are they an accurate frequency measure for text-messaging abbreviations?
- In order to understand abbreviation processing, we must have reliable frequency estimates!

Blair, I.V., Urland, G.R., & Ma, J.E. (2002). Using Internet search engines to estimate word frequency. *Behavior Research Methods, Instruments, & Computers, 34*, 286-290.

Brysbaert, M., Speybroeck, S., & Vanderelst, D. (2009). Is there room for the BBC in the mental lexicon? On the recognition of acronyms. *The Quarterly Journal of Experimental Psychology, 62,* 1832-1842.

Caldwell-Harris, C.L., Berant, J.B., & Edelman, S. (2012). Measuring mental entrenchment of phrases with perceptual identification, familiarity ratings, and corpus frequency statistics. In S. T. Gries & D. Divjak (Eds.), *Frequency Effects in Cognitive Linguistics (Vol. 1): Statistical Effects in Learnability, Processing and Change.* The Hague, The Netherlands: De Gruyter Mouton.

Laszlo, S. & Federmeier, K.D. (2007). Better the DVL you know: Acronyms reveal the contribution of familiarity to single-word reading. *Psychological Science, 18,* 122-126.

Still, M.L., Morris, A.L., & Jones, K. (2011, November). *IDK if you Really H8IT Unless You Spell it Out.* Poster presented at the Annual Meeting of the Psychonomic Society, Seattle, WA.

Underwood, M.K., Rosen, L.H., More, D., Ehrenreich, S.E., & Gentsch, J.K. (2012). The BlackBerry project: capturing the content of adolescents' text messaging, *Developmental Psychology, 48,* 295-302.