Extending the English LexTALE to assess vocabulary knowledge in native speakers and advanced L2 learners

Robert Fiorentino[a], Münir Özturhan[a], Andrew Collins[a], Xiao Yang[b], Braden Swaim[a], Manuel Carreiras[c], Simona Mancini[c], Lesa Hoffman[d], & Alison Gabriele[a]
University of Kansas[a], Northeastern University[b], BCBL[c], University of Iowa[d]

In studies of second language (L2) acquisition, it has become increasingly common to use the LexTALE, developed by Lemhöfer and Broersma ([1]), which is a short, easy to administer measure of L2 English proficiency that has been shown to be correlated with standardized proficiency measures. In studies of native language (L1) processing, vocabulary skills have been shown to be related to language comprehension (e.g., [2]) and it has become increasingly common to include vocabulary tests as measures of language experience (e.g., [3]). We examined whether an 'extended' version of the English LexTALE, including lower-frequency words, could be used as an assessment of proficiency in both English native speakers and L1 Spanish learners of English (see [1, 4]).

The original LexTALE included 40 real words (CELEX frequency 1–26 per million, mean 6.8), and 20 orthographically legal and pronounceable nonwords (4–12 letters in length, mean 7.3) ([1]). We added 20 additional challenging, low-frequency words (CELEX frequency <1 per million, mean 0), and 10 additional orthographically legal, pronounceable nonwords (4–12 letters in length, mean 7.9) Participants (N=72 native English, N=77 native Spanish-speaking adult learners of English) were instructed to indicate whether the stimulus was an existing English word or not, and to select 'no' if they were unsure, as in [1]. The task took about 3.5-5 minutes to complete. Participants also took the ACTFL Reading Proficiency test in English, an adaptive measure of reading ability that can assess proficiency from novice to superior levels of proficiency. The Author Recognition Task ([5]) was also included as a measure of language experience; L1 English participants were given an ART based on Acheson et al. [6] with new items included to replace items marked as an ineffective by Moore and Gordon [7]. L1 Spanish participants took an ART developed to target both classic and bestselling authors in Spain.

Following [1], we examined correlations between the LexTALE and measures of language proficiency/experience (ACTFL RPT, and ART). Results (Table 1) revealed significant correlations for both the original and extended LexTALE; however, only the extended LexTALE showed a significant correlation with language experience (ART) for native speakers, and correlation coefficients were larger for the extended LexTALE in all cases. We also conducted analyses to assess the reliability of the extended test. In examining the pool of 90 items, 42 items with accuracy rates > .95 in either group were removed. Psychometric assessment for the extent to which a single latent trait could predict the associations among 48 binary items was conducted using Item Factor Analysis (IFA) in M*plus* v 8.8 ([8]). Analyses were conducted for each group separately to avoid assuming measurement invariance across groups. Models specifying separate latent traits for accuracy in detecting words and nonwords had better fit than single-factor models; one item that did not relate significantly to its latent trait in either group was removed. The remaining 47 items (18 extended and 9 original words, 5 extended and 15 original nonwords) were then used to examine reliability of measurement across the latent traits. Figure 1 compares reliability for accuracy in detecting words when using only 9 original items versus also using 18 extended items within each group. It shows a benefit in both samples to adding the extended items, more so for higher trait abilities in native speakers. There was a correspondingly wider distribution of higher-level latent trait estimates for the L1 English sample. A much smaller benefit to reliability was found in each sample for the extended nonwords.

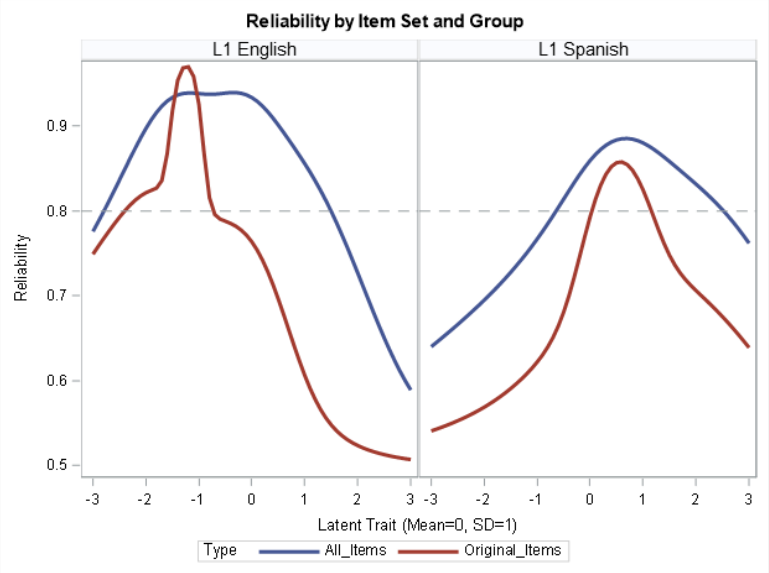
These results together suggest the utility of the newly extended LexTALE, an efficient, easy-to-administer test, for capturing variability in vocabulary knowledge in both English native speakers and L2 learners, including native speakers with relatively high language skills.

Table 1. Correlations between LexTALE (original and extended) and measures of language proficiency/experience

	L1 English speakers				L1 Spanish speakers			
	LexTALE Extended†	LexTALE Original†	RPT English	ART U.S.	LexTALE Extended	LexTALE Original	RPT English	ART Spain
LexTALE Extended	1	.95****	.45****	.38**	1	.91****	.49****	.42****
LexTALE Original	.95****	1	.32***	.22	.91****	1	.37***	.28*
RPT English	.45****	.32***	1	.4***	.49****	.37***	1	.25*
ART U.S.	.38**	.22	.4***	1	NA	NA	NA	NA
ART Spain	NA	NA	NA	NA	.42****	.28*	.25*	1

* $p < .05$, ** $p < .01$, ***$p < .005$, ****$p < .001$
†Original LexTALE scoring formula: ((number of words correct/40*100) + (number of nonwords correct/20*100))/2; Extended LexTALE scoring formula: ((number of words correct/60*100) + (number of nonwords correct/30*100))/2

Figure 1.



References

1. Lemhöfer, K. and Broersma, M. (2012). *Behavior Research Methods*, *44*(2): 325-343.
2. Van Dyke, J.A., Johns, C.L., & Kukona, A. (2014). *Cognition*, *131*(3): 373-403.
3. James, A. N., Minnihan, C. J., & Watson, D. G. (2023). *Journal of Cognition*, *6*(1): 30, pp. 1–29.
4. Izura, C., Cuetos, F., & Brysbaert, M. (2014). *Psicológica*, *35*,1, 49-66.
5. Stanovich, K.E. and West, R.F. (1989). *Reading Research Quarterly*, *24*(4): 402-433.
6. Acheson, D.J., Wells, J.B., & MacDonald, M.C. (2008). *Behavior Research Methods*, *40*(1): 278-289.
7. Moore, M., and Gordon, P. C. (2015). *Behavior Research Methods, 47*(4), 1095–1109.
8. Muthén, L.K. and Muthén, B.O. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén.