Vocabulary Learning and Instruction
Volume 3, Issue 1, December 2014
http://vli-journal.org

Commentary

Reaction Time Methodologies and Lexical Access in Applied Linguistics

John P. Racine
Dokkyo University
doi: http://dx.doi.org/10.7820/vli.v03.1.racine

1 Introduction

The first issue of this journal featured a paper by Iso (2012) in which the author described research conducted to validate his Lexical Access Time Test (LEXATT2). While the details of the test procedure are scant in the write-up, it appears that there are a number of methodological issues that require thorough examination before this test can be considered a valid measure of lexical access. Notable among these issues are the accuracy of the reaction time (RT) measurements and the manner in which the RTss are interpreted. Other aspects of the study–including its relation to prior research and theory–also deserve scrutiny.

As pointed out by Mochizuki (2012) in a discussion of four vocabulary test studies which included Iso's, very few researchers in applied linguistics are conducting research on lexical access. Indeed, lexical research involving RT measurement of any kind is particularly rare in our field, despite the abundance of such studies in cognitive psychology and psycholinguistics circles. The comments below are thus not intended as criticism of Iso's paper in particular. Rather, this commentary is intended to highlight some of the general principles central to cognitive approaches to the mental lexicon, particularly where RT is to be measured. It is hoped that the issues raised below will serve as an introduction to some key elements of this kind of research and encourage lexical researchers who wish to undertake further studies in this area.

2 The Original Study

Iso's LEXATT2 program attempts to test lexical access by examining the difference in RTs between nonverbal and verbal recognition tasks. Nonverbal task trials began with subjects clicking and holding down a mouse button. A stimulus (in this case, a circular shape) then appeared on screen and subjects were required to release the button when they saw it. When they released the button, the stimulus disappeared. The RTs (i.e., the time between presentation of the stimulus and when the button was released) from five such trials were recorded and the average time was calculated. Iso calls this the "mean reaction time for nonverbal stimulus" (MRT (NV); p. 78). In the verbal task trials, subjects once again began by holding down a mouse button. In this case a word appeared (either four, six, or eight letters in length). After reading the word, subjects again released the button and the RT

Please cite this article as: Racine, J.P. (2014). Reaction time methodologies and lexical access in applied linguistics. *Vocabulary Learning and Instruction*, *3* (1), 66–70. doi: 10.7820/vli.v03.1.racine

was recorded (RT for verbal stimulus, RT (VL)). Lexical access time was thus calculated for each word as:

$$RT\,(VL) - MRT\,(NV)$$

Verbal task trials included a second stage in which subjects were then given a test of receptive knowledge in which they had to choose the meaning or a synonym of the word from a number of options. For trials in which subjects failed to recognize the meaning of the word in this stage, RTs were discarded as subjects had either not known the word or had presumably failed to read it correctly.

3 Technical Considerations in RT Research

One of the fundamental issues that researchers in applied linguistics should realize about RT measurements, such as the one employed here, is that they are extremely delicate. In any experiment involving RT, many trials fail to record what they are intended to. Accidental keystrokes, lapses in subject attention, and distractions in the laboratory environment are common factors. In research where more than one key are used to trigger the timer (e.g., yes/no responses), slower left-hand reactions (for right-hand dominant participants) need to be accounted for and counterbalanced (Shen & Franz, 2005). Some researchers have attempted to circumvent these problems with voice-activated triggers, only to find that hems and haws, coughing, and extraneous noises in the lab can lead to just as many lost trials. There are more technical issues to consider as well. Screen and keyboard refresh rates can vary widely and can contribute dozens of milliseconds to RTs. Insensitive or oversensitive microphones regularly lead to lost data in voice-triggered RT studies as well. While researchers hope that these issues occur randomly across conditions, this may not always be the case and thus certain measures should be taken to limit the influence these factors have on resulting data.

Common methods of preventing undue influence on RTs from the factors listed above include allowing subjects ample practice on task and the inclusion of an abundance of experimental trials. Iso makes no mention of practice at all and bases the MRT (NV) measure on the mean from only five nonverbal trials. Compare this to the eight practice trials followed by 160 experimental trials in a study of attribution (Bassili & Racine, 1990; Expt. 1). Lexical decision studies similar to Iso's invariably include such practice sessions followed by an abundance of experimental trials. Fitzpatrick and Izura's (2011) study, for example, included 20 practice trials followed by 72 experimental ones. Balota and Chumbley (1984) utilized as many as 60 practice trials before subjects received experimental trials. This study also employed five buffer trials at the beginning of each experimental block. A buffer trial is one that appears to subjects as any other experimental trial, but the results from which are not included in the data. The decision to discard this data is made by the researcher a priori, and is an excellent way to allow participants to become accustomed to experimental procedures without adversely affecting the final results.

Another common procedure for avoiding undue influence on RT scores is to normalize data via the elimination of outliers. In experiments where average

response latencies may only be 200 or 300 ms in duration, a distraction of 5 or 6 seconds can have a huge impact on experimental cells. Iso does not report utilizing any measures to adjust or omit such erroneous data. As can be seen in the sampling of lexical decision task (LDT) studies listed in Table 1, however, it is very commonplace for researchers to employ such adjustments. In these studies, outliers were either adjusted (e.g., a score exceeding 2.5 standard deviations (*SD*s) above a subject's mean RT was replaced with a score of precisely 2.5 *SD*s above the mean) or they were eliminated from the data entirely.

4 Conceptual issues

A more serious concern for the validity of this study can be seen in the way the RT scores are interpreted. Iso's verbal task, as explained above, consists of two stages: the RT measure and a meaning confirmation stage. Subjects were asked to read the word and release the button when they had done so. The word recognition task is a separate stage, Iso explains, "so that LEXATT2 is not measuring word recognition time" (p. 79). Then the question is: What is it measuring? What exactly are subjects doing before they release the button? Have they said the words to themselves? Have they thought about the meaning? What are the L2-learning subjects doing during this time if they do not know the word? More essentially, can a subject "read" a word without considering its meaning? And if that was possible, would subjects allow themselves to do so, knowing that they are about to be given a multiple choice recognition task? Iso would have us believe that the RT measure here is equal to fixation time on text as measured in experiments involving passage reading. Thus "lexical access" is something akin to perception of the orthography of a word before semantic processing has occurred. In fact, these two processes are not necessarily serial tasks and are often treated as inseparable. This is precisely how they are treated in the RT studies cited above. All of these LDTs measure RTs when subjects have completed judgments as to whether a string of letters constitutes a word. Researchers typically employ words and nonwords as stimuli in these studies. Semantic processing is inherent to the subject's task and that is

Table 1. Outliers in Lexical Decision Task Studies

LDT study	Outliers
Altarriba and Knickerbocker (2011)	2.5 *SD*s above or below a participant's mean RT
Fitzpatrick and Izura (2011)	3 *SD*s above or below a participant's mean RT
Hu and Jiang (2011)	2.5 *SD*s above or below a participant's mean RT Responses exceeding a 2500 ms time limit
Segalowitz and Segalowitz (1993)	2 *SD*s above or below a participant's mean RT Responses exceeding a 3000 ms time limit

what the RT scores are intended to measure. Iso should have designed his methodology similarly.

It is also interesting to note that Segalowitz and Segalowitz's (1993) research (listed in Table 1) is the only RT study cited in the Iso (2012) paper. This gives the impression that Iso's research is based upon, or is perhaps even a replication attempt, the former study. In fact, there are fundamental differences between the two studies. Some of these differences have already been explained above (i.e., Iso did not make adjustments for outliers and employed a lexical task for which the RT scores may not be interpretable). Another difference is that the Segalowitz and Segalowitz (1993) study includes a review of relevant prior research. While it may have been due to editorial considerations beyond Iso's control, his is the only study cited here that does not include a thorough literature review as background to the study. Another approach would be to cast the study as an attempt to explore the validity of a model or models of second-language lexical access (e.g., de Bot, 2010; de Bot, Paribakht, & Bingham Wesche, 1997; Kroll & Dijkstra, 2010). Without a broader literature review and a more detailed write-up, it is difficult to position Iso's work within the field.

Finally, as I stated in the introduction to this commentary, there are as yet very few applications of RT methodologies in applied linguistics. For that reason, Iso should be commended for his exploratory work. It is hoped that these comments will help to inspire further studies in the field and I welcome a LEXATT3 program that addresses the issues raised here.

References

Altarriba, J., & Knickerbocker, H. (2011). Acquiring second language vocabulary through the use of images and words. In P. Trofimovich & K. McDonough (Eds.), *Applying priming methods to L2 learning, teaching and research* (pp. 21–48). Amsterdam: Benjamins.

Balota, D.A., & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(3), 340–357. doi:10.1037/0096-1523.10.3.340

Bassili, J.N., & Racine, J.P. (1990). On the process relationship between person and situation judgments in attribution. *Journal of Personality and Social Psychology, 59*(5), 881–890. doi:10.1037/0022-3514.59.5.881

de Bot, K. (2010). Cognitive processing in bilinguals: From static to dynamic models. In R.B. Kaplan (Ed.), *The Oxford handbook of applied linguistics* (2nd ed., pp. 335–348). Oxford, UK: Oxford University Press.

de Bot, K., Paribakht, T.S., & Bingham Wesche, M. (1997). Toward a lexical processing model for the study of second language vocabulary acquisition. *Studies in Second Language Acquisition, 19*, 309–329.

Fitzpatrick, T., & Izura, C. (2011). Word association in L1 and L2: An exploratory study of response types, response times, and interlingual mediation. *Studies in*

Second Language Acquisition, *33*(03), 373–398. doi:10.1017/S0272263111000027

Hu, G.G., & Jiang, N. (2011). Semantic integration in listening comprehension in a second language. In P. Trofimovich & K. McDonough (Eds.), *Applying priming methods to L2 learning, teaching and research* (pp. 199–218). Amsterdam: Benjamins.

Iso, T. (2012). Examining the validity of the Lexical Access Time Test (LEXATT2). *Vocabulary Learning and Instruction, 01*(1), 78–82. doi:10.7820/vli.v01.1.iso

Kroll, J.F., & Dijkstra, T. (2010). The bilingual lexicon. In R.B. Kaplan (Ed.), *The Oxford handbook of applied linguistics* (2nd ed., pp. 349–370). Oxford, UK: Oxford University Press.

Mochizuki, M. (2012). Four empirical vocabulary test studies in the three dimensional framework. *Vocabulary Learning and Instruction*, *01*(1), 44–52. doi:10.7820/vli.v01.1.mochizuki

Segalowitz, N.S., & Segalowitz, S.J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics*, *14*(03), 369–385. doi:10.1017/S0142716400010845

Shen, Y.-C., & Franz, E.A. (2005). Hemispheric competition in left-handers on bimanual reaction time tasks. *Journal of Motor Behavior*, *37*(1), 3–9. doi:10.3200/JMBR.37.1.3-9