Lexical facility:
Developing vocabulary knowledge as a skill -it's what you know and when you know it-

Michael Harrington
School of Languages & Comparative Cultural Studies
m.harrington@uq.edu.au

関西英語教育学会 2013年度（第18回）
Talk outline

Introducing lexical facility
The importance of vocabulary size.
The importance of vocabulary speed.
Measuring lexical facility
Relating lexical facility to English performance
A lexical facility approach to vocabulary instruction.
Introducing lexical facility (LF)

*Lexical facility* refers to having both vocabulary knowledge of adequate knowledge and being able to access that knowledge in a timely manner.

Lexical facility brings together these two key aspects of vocabulary knowledge in a unitary construct to assess how the two develop, interact, and serve to measure or predict performance in English as a second language.

Lexical facility is a critical bottleneck in discourse processing

“Successful text comprehension requires that lower level linguistic processes be *efficient*, in that they are available in a timely manner, and *effective*, in that they provide information of an adequate quality to the higher level processes (Just & Carpenter, 1992).”

– This requires a vocabulary of *adequate size* and the ability to *access this knowledge* in a way appropriate to the context.
Vocabulary size as element of LF


Benchmarked by use frequency of occurrence statistics obtained from corpora, e.g. British National Corpus (BNC) and readily available on-line, e.g. [http://www.lextutor.ca/](http://www.lextutor.ca/).

Discrete, largely context-free approach to measuring L2 vocabulary.
Data-driven learning: Lexical frequency profiles

The likelihood of knowing a word depends on how frequently the word is encountered in the language.
How much vocabulary is needed to function in English?

Everyday conversation/80% text written text coverage

(Schonell, Meddleton & Shaw, 1956)

> 2000 (most frequent) word families

Threshold for initial access to authentic reading = 95% text coverage.

(Nation, 2001)

> 2000-3000 word families

Reading English university textbooks = 98-99% text coverage.

> 8000-9000 word families

Native speaker vocabulary size:

15,000-20,000 word families

Nation (2006); Adolphis & Schmitt (2003)
### Table 5.1 Number of unfamiliar tokens and the number of lines of text containing one unfamiliar word

<table>
<thead>
<tr>
<th>% text coverage</th>
<th>Number of unfamiliar tokens per 100 tokens</th>
<th>Number of text lines per 1 unfamiliar word</th>
</tr>
</thead>
<tbody>
<tr>
<td>99</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>98</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>95</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>90</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>80</td>
<td>20</td>
<td>0.5</td>
</tr>
</tbody>
</table>
Vocabulary speed as element of LF

• Speed of lexical retrieval measured by mean response time & standard deviation (measure of response consistency)


• In addition to differences in mean response time, response stability, or consistency, can be measured by the coefficient of variation (Segalowitz, Segalowitz, & Wood, 1998; Hulstijn, Van Gelderen, & Schoonen, 2009).

• Experiment-based research
LF: Bringing size and speed together

- Vocabulary size researchers (Laufer, Nation, et al) have had a primary focus on vocabulary assessment in the context of formal language instruction, while speed/automaticity researchers (Segalowitz et al) have been concerned with understanding basic psycholinguistic mechanism responsible for fluent language performance.

- The lexical facility construct brings these two research perspectives together, with a primary focus on what the inclusion of speed in our models of L2 proficiency will do for SLA theory and L2 assessment and teaching practice.
Measuring lexical facility

The Yes-No test (Meara, 1989; Eyckmans, 2004, Harrington, 2006) measures receptive L2 vocabulary knowledge by eliciting a simple judgement as to whether a presented item is known or not.

Test item selection is based on frequency-of-occurrence statistics. These are used to inferring the size of the individual’s receptive vocabulary (Meara, 1996). Size is indexed using performance on items sampled from frequency of occurrence bands

1-1000\(^{th}\) most frequent words = 1K band
1001-2000\(^{th}\) = 2K band
2001-3000\(^{th}\) = 3K band
Etc.
Adding a timed component

The Timed Yes/No Test measures both vocabulary size and the speed with which the judgement is made.

Both elements contribute to proficiency measurement

The use of timed responses also adds an additional task demand.
Timed Yes/No Test format

- Base on lexical decision task widely used in cognitive psychology

- Uses word and nonword (or pseudowords) to assess vocabulary size.

- Nonword performance used to adjust overall score.

- Speed-accuracy trade-off potential problem
‘Guessing’ on the part of the test-taker is reflected in nonword performance.

<table>
<thead>
<tr>
<th>Response</th>
<th>Item</th>
<th>Word</th>
<th>Nonword</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>correct</td>
<td>incorrect</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(‘hit’)</td>
<td>(‘false alarm’)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>incorrect</td>
<td>correct</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(‘miss’)</td>
<td>(‘correct reject’)</td>
<td></td>
</tr>
</tbody>
</table>
Evaluating the lexical facility construct

*Research questions*

1. How does vocabulary size and speed compare as stable measures of individual differences in English L2 performance? Do they correlate?

2. Do the two measures reliably predict performance differences in common domains of L2 performance?

3. Does a combination of size and speed provide a better picture of individual differences than size (or speed) alone?
The data set

<table>
<thead>
<tr>
<th>Data sets</th>
<th>N</th>
<th>Setting</th>
<th>Learner Level</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language program placement testing</td>
<td>88</td>
<td>ESL Australia</td>
<td>Low to Mid</td>
<td>Placement test</td>
</tr>
<tr>
<td>(Harrington &amp; Carey, 2009)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicting IELTS performance</td>
<td>310</td>
<td>ESL Australia</td>
<td>Low to High</td>
<td>IELTS overall bandscores</td>
</tr>
<tr>
<td>(Harrington, in preparation)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicting academic performance</td>
<td>70</td>
<td>EFL Oman</td>
<td>Low</td>
<td>GPA Academic skills test</td>
</tr>
<tr>
<td>for diagnostic purposes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Harrington &amp; Roche, forthcoming; Roche &amp; Harrington, 2012)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In addition to multiple sites in Australia, data also collected in Singapore and Kansas, USA.
Vocabulary size as stable indicator of L2 development

Accuracy Responses by Group and Level

![Graph showing accuracy responses by group and level](chart)

- Intermediate ESL
- L2 University
- L1 English

Accuracy (proportion)

Word frequency level

2000 | 3000 | 5000 | 10,000
Vocabulary speed as stable indicator of L2 development

Response Time by Group and Level.
Size and speed measures

Research question #1 Do item response times mirror vocabulary size differences as stable measures of individual differences in L2 lexical performance?

• Yes

• Accuracy (size) is better than RT

• RT shows more variability

• Both measures less sensitive at lower levels
Size and speed as predictors of performance

Research question 2: Do the two measures reliably predict performance differences in functional domains of L2 performance?

1. Language program placement
2. Predicting IELTS scores
3. Academic English and GPA in PELA setting
Size & speed as a placement measure

• Milton College: Sydney language school placement study

• *Aim:* To assess the effectiveness of the TYNT as a tool for placement decisions in a commercial language school. The study was carried out at an established English language school in Sydney, Australia.

• *Participants (n=88):* Ranged from 19 to 33 years (mean= 24.3, SD = 3.8) with many intending to continue on to university study in Australia and elsewhere at the end of language study. The largest number was from Korea (32) and Japan (18), with the remainder from 14 different first languages.

• Design. Performance on two versions of the TYN test was compared with inhouse grammar and listening tests on placement level decisions.
Vocabulary size and placement decisions

Percentage correct

Place of level

- Beginner
- Elementary
- Lower Intermediate
- Upper Intermediate
- Advanced
Vocabulary response speed and language program placement decisions

![Graph showing mean word response time (msec) across different placement levels for VLT and Milton.](image-url)
IELTS performance

Design. The TYN test was given to volunteers at the University of Queensland and the University of Queensland Foundation Program between 2008-2011. Tests were taken at entry to the program. IELTS scores were self-reported. Students took one or both versions.

Test consisted 80 items (66 words, 14 nonwords)

Participants. Approximately 75% of the test takers were Chinese L1, with most of the remainder (20%) from Southeast Asia and the remainder from the Middle East and elsewhere.

<table>
<thead>
<tr>
<th>IELTS</th>
<th>5</th>
<th>5.5</th>
<th>6</th>
<th>6.5</th>
<th>7</th>
<th>7.5</th>
<th>8</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version B</td>
<td>8</td>
<td>145</td>
<td>76</td>
<td>49</td>
<td>20</td>
<td>8</td>
<td>5</td>
<td>310</td>
</tr>
</tbody>
</table>
TYN Test performance and IELTS overall bandscores: Accuracy (N=310)
TYN Test performance and IELTS overall bandscores: Response times (N=310)
Lexical facility as a diagnostic in an EFL PELA setting

- English-medium college of education in the Sultanate of Oman, Arabic L1 users (N=70)
- Timed Y/N Test performance correlated with written Academic English Proficiency (AEP) and overall Academic Achievement (GPA)
- AEP was assessed using an academic writing test based on IELTS.
- Vocabulary size and speed correlated with both academic writing and GPA measures.

(Roche & Harrington, 2013)
Oman study descriptive stats

Table 2: Descriptive Statistics (Means, Standard Deviations, Range) for Advanced and Basic word tests, Grade Point Average and IELTS Writing Scores, N = 70.

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advanced Word</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>35.96</td>
<td>12.05</td>
<td>8.93</td>
</tr>
<tr>
<td>Response time(^b)</td>
<td>1590</td>
<td>371</td>
<td>946</td>
</tr>
<tr>
<td>False alarm rate</td>
<td>23.70</td>
<td>13.46</td>
<td>3.57</td>
</tr>
<tr>
<td><strong>Basic Word</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>53.23</td>
<td>15.83</td>
<td>8.33</td>
</tr>
<tr>
<td>Response time</td>
<td>1571</td>
<td>372</td>
<td>909</td>
</tr>
<tr>
<td>False alarm rate</td>
<td>22.57</td>
<td>11.03</td>
<td>3.57</td>
</tr>
<tr>
<td>IELTS writing</td>
<td>5.18</td>
<td>.86</td>
<td>3.50</td>
</tr>
<tr>
<td>Grade point average</td>
<td>2.79</td>
<td>.55</td>
<td>1.00</td>
</tr>
</tbody>
</table>

\(^a\) Corrected for guessing score, proportion of yes responses to words (hits) minus proportion of yes responses to nonwords (false alarms).

\(^b\) Milliseconds (msc)
Research question #2

Do the two measures reliably predict performance differences in common domains of L2 performance?

• Yes

• Accuracy (size) is better than RT

• RT shows more variability

• Both measures less sensitive at lower levels
Combining vocabulary size and speed

Research question #3. Do size and speed together provide a better picture of individual differences than size (or speed) alone?
**Milton placement study Additional variance accounted for by RT**

Table 7
Hierarchical regression analyses of Yes/No test results with placement as criterion variable and accuracy and response time as ordered predictor variables.

<table>
<thead>
<tr>
<th>Test</th>
<th>n</th>
<th>variable</th>
<th>$R^2$</th>
<th>Standard error of estimate</th>
<th>$R^2$ change</th>
<th>$F$ change</th>
<th>Sig $F$ change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined</td>
<td>88</td>
<td>a</td>
<td>.386</td>
<td>1.017</td>
<td>.386</td>
<td>54.037</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b</td>
<td>.430</td>
<td>.968</td>
<td>.044</td>
<td>6.534</td>
<td>.012*</td>
</tr>
<tr>
<td>Program</td>
<td>88</td>
<td>a</td>
<td>.376</td>
<td>1.025</td>
<td>.376</td>
<td>51.875</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b</td>
<td>.427</td>
<td>.988</td>
<td>.051</td>
<td>7.554</td>
<td>.007*</td>
</tr>
<tr>
<td>General</td>
<td>86</td>
<td>a</td>
<td>.288</td>
<td>1.092</td>
<td>.288</td>
<td>34.00</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b</td>
<td>.342</td>
<td>1.057</td>
<td>.054</td>
<td>6.797</td>
<td>.011*</td>
</tr>
</tbody>
</table>

a = accuracy as 1st predictor variable  
b = response time as 2nd predictor variable
IELTS study

Additional variance accounted for by response time

<table>
<thead>
<tr>
<th>Variable</th>
<th>R²</th>
<th>Std Error Estimate</th>
<th>R² Change</th>
<th>F Change</th>
<th>df1</th>
<th>df2</th>
<th>Sig. F Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=310</td>
<td>Accuracy</td>
<td>0.459</td>
<td>0.447</td>
<td>0.459</td>
<td>261.085</td>
<td>1</td>
<td>309</td>
</tr>
<tr>
<td>RT</td>
<td>0.474</td>
<td>0.441</td>
<td>0.016</td>
<td>9.18</td>
<td>1</td>
<td>308</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Accuracy (size) as 1st predictor variable
RT = response time as 2nd predictor variable
Table 4: Hierarchical Regression Analyses of the Advanced Word and Writing Measures with GPA as Criterion Variable and Writing and Word Measure as Predictors.  N = 70

<table>
<thead>
<tr>
<th>Advanced Word Models</th>
<th>$R^2$</th>
<th>Adjusted $R^2$</th>
<th>$R^2$ change</th>
<th>B</th>
<th>SEB</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing</td>
<td>.160</td>
<td>.147</td>
<td><strong>.160</strong>*</td>
<td>.230</td>
<td>.077</td>
<td>.399</td>
</tr>
<tr>
<td>Advanced Accuracy</td>
<td>.160</td>
<td>.135</td>
<td>.000</td>
<td>.001</td>
<td>.005</td>
<td>.021</td>
</tr>
<tr>
<td>Advanced Response time</td>
<td>.198</td>
<td>.162</td>
<td><strong>.028</strong></td>
<td>-1.171</td>
<td>.659</td>
<td>-0.200</td>
</tr>
<tr>
<td>Advanced Accuracy</td>
<td>.02</td>
<td>.005</td>
<td>.02</td>
<td>.001</td>
<td>.005</td>
<td>.021</td>
</tr>
<tr>
<td>Advanced Response time</td>
<td>.091</td>
<td>.063</td>
<td><strong>.071</strong></td>
<td>-1.172</td>
<td>.659</td>
<td>-0.2</td>
</tr>
<tr>
<td>Writing</td>
<td>.198</td>
<td>.162</td>
<td><strong>.108</strong></td>
<td>.230</td>
<td>.077</td>
<td>.357</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Basic Word Models</th>
<th>$R^2$</th>
<th>Adjusted $R^2$</th>
<th>$R^2$ change</th>
<th>B</th>
<th>SEB</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing</td>
<td>.16</td>
<td>.147</td>
<td><strong>.160</strong>*</td>
<td>.215</td>
<td>.079</td>
<td>.334</td>
</tr>
<tr>
<td>Basic Accuracy</td>
<td>.184</td>
<td>.160</td>
<td>.024</td>
<td>.003</td>
<td>.004</td>
<td>.093</td>
</tr>
<tr>
<td>Basic Response time</td>
<td>.267</td>
<td>.234</td>
<td><strong>.083</strong></td>
<td>-1.708</td>
<td>.624</td>
<td>-0.299</td>
</tr>
<tr>
<td>Basic Accuracy</td>
<td>0.215</td>
<td>0.079</td>
<td>0.334</td>
<td>0.003</td>
<td>0.004</td>
<td>0.093</td>
</tr>
<tr>
<td>Basic Response time</td>
<td>0.003</td>
<td>0.004</td>
<td>0.093</td>
<td>-1.708</td>
<td>0.624</td>
<td>-0.299</td>
</tr>
<tr>
<td>Writing</td>
<td>-1.708</td>
<td>0.624</td>
<td>-0.299</td>
<td>0.215</td>
<td>0.079</td>
<td>0.334</td>
</tr>
</tbody>
</table>

B, Unstandardized coefficient; SEB, Standard Error B; $\beta$, Standardized coefficient. $F$ significant at * $p < .05$, ** $p < .01$, *** $p < .001$
Research question #3

Do size and speed together provide a better picture of individual differences than size (or speed) alone?

• In general yes, but RT variability can affect outcomes.

• Used as complementary measures or as composites.
Future research on lexical facility

• A normative model of temporal processing?

• Taking a closer look at response variability as a dimension of development.

• Response time variability is a central issue.

• Longitudinal development of size and speed in relation to performance.

• Response time and other domains of vocabulary knowledge.
Lexical facility and vocabulary instruction

The process of learning vocabulary is characterised by a set of features that distinguish it from other domains of language.

*Design features*

1. Word learning involves other words.

2. Vocabulary size is important.

3. Word meaning is open-ended.

4. Words must be available for use.

5. Words are things.
1 Word learning involves other words

A new word is not learned in isolation. Learning new words is a process of modifying the existing network of words in the learner’s head (the “mental lexicon”).

No word is an island.*

* Apologies to John Donne
2 Vocabulary size is important.

To be able to function in a language you must know a vast number of words. This *breadth* of knowledge continues to expand as your experience with the language continues.

The more the merrier.
3 Word meaning is open-ended.

Knowledge of a word is open-ended. The *depth* of word knowledge develops through experience.

Word knowledge is not black & white.
Do you know the meaning of ‘broke’?

The waves broke on the rock.
He broke his leg.
The cup broke.
After the incident, he was a broken man.
She broke his heart.
He broke his word.
The man broke his oath.
Which country has broken the cease-fire?
Some workers have broken the strike.
She broke the ice with a joke.
The crowd broke up as the police arrived.
His voice broke when he was 13.
Her fall was broken by a tree.
4 Words must be available for use.

For fluent performance individual words must be accessed quickly.

Having a word on the tip of your tongue doesn’t count.
Words are physical objects, whether in sound (phonological form) or sight (orthographic shape). Part of learning a word is learning the form and then practicing its recognition and production. Fast retrieval is important to both.

Vocabulary learning is partly perceptual
Conclusion

Lexical facility….

..it’s not just what you know,

*it’s when you know it.*
References


