Vocabulary targets and language proficiency. A framework for learning, teaching, and assessing vocabulary

Dr Veronica Benigno (Pearson, London, UK)

13-14 October 2016, New Directions 2016
Outline of talk

• **Section I**
Overview of vocabulary research in linguistics and related fields

• **Section II**
Introduction to the Pearson *Global Scale of English (GSE)* Vocabulary

• **Section III**
Discussion and new directions
Section I:
Overview of vocabulary research in linguistics and related fields
Why is vocabulary important in assessment?

• Vocabulary is regarded as a crucial component of communicative competence and as one of the strongest indicators of language performance:

  “Vocabulary load is the most significant predictor of text difficulty”; Alderson (2000, p.73)

• A positive relationship exists between vocabulary knowledge & proficiency, reading in particular (Anderson and Freebody; Laufer, 1996; Perfetti, 2007)
A multidimensional construct

- Vocabulary knowledge has been commonly interpreted as a multidimensional construct (Read, 2004; Meara, 2005; Daller et al., 2007; Schmitt et al., 2010)

- **Size** refers to how many (single) words a speaker knows – in relation to the mapping word-form, usually without any implication about degree of knowledge

- **Depth** refers to how well a speaker knows a word – in terms of associational knowledge, collocational knowledge, inflectional and derivational knowledge, knowledge of concepts and referents, and knowledge of constraints on use (Read, 2000)
**What is involved in knowing a word** by Nation, 2001

<table>
<thead>
<tr>
<th>Form</th>
<th>spoken</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R  What does the word sound like?</td>
</tr>
<tr>
<td></td>
<td>P  How is the word pronounced?</td>
</tr>
<tr>
<td>written</td>
<td>R  What does the word look like?</td>
</tr>
<tr>
<td></td>
<td>P  How is the word written and spelled?</td>
</tr>
<tr>
<td>word parts</td>
<td>R  What parts are recognisable in this word?</td>
</tr>
<tr>
<td></td>
<td>P  What word parts are needed to express the meaning?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Meaning</th>
<th>form and meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R  What meaning does this word form signal?</td>
</tr>
<tr>
<td></td>
<td>P  What word form can be used to express this meaning?</td>
</tr>
<tr>
<td>concepts and referents</td>
<td>R  What is included in the concept?</td>
</tr>
<tr>
<td></td>
<td>P  What items can the concept refer to?</td>
</tr>
<tr>
<td>associations</td>
<td>R  What other words does this make us think of?</td>
</tr>
<tr>
<td></td>
<td>P  What other words could we use instead of this one?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Use</th>
<th>grammatical functions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R  In what patterns does the word occur?</td>
</tr>
<tr>
<td></td>
<td>P  In what patterns must we use this word?</td>
</tr>
<tr>
<td>collocations</td>
<td>R  What words or types of words occur with this one?</td>
</tr>
<tr>
<td></td>
<td>P  What words or types of words must we use with this one?</td>
</tr>
<tr>
<td>constraints on use</td>
<td>R  Where, when, and how often would we expect to meet this word?</td>
</tr>
<tr>
<td>(register, frequency)</td>
<td>P  Where, when, and how often can we use this word?</td>
</tr>
</tbody>
</table>

R = receptive knowledge, P = productive knowledge
Measuring vocabulary

- Vocabulary can be tested as an **independent construct**, e.g. in a test of vocabulary size, or as an **embedded construct**, e.g. in a multiple choice item in a test of general proficiency (*Read & Chapelle, 2004*)

- In either case, the decision on what vocabulary corresponds to what level of achievement should be made using a common **framework or standard**

- But isn’t such a standard missing? What is the relationship between vocabulary and proficiency levels?
Evidence from the CEFR (2001)

- Since the CEFR is not tied to any particular language, the **framework is vague** about what vocabulary is needed at different proficiency levels
<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2</td>
<td>Has a good command of a <strong>very broad lexical repertoire</strong> including idiomatic expressions and colloquialisms; shows awareness of connotative levels of meaning.</td>
</tr>
<tr>
<td>C1</td>
<td>Has a <strong>good command of a broad lexical repertoire</strong> allowing gaps to be readily overcome with circumlocutions; little obvious searching for expressions or avoidance strategies. Good command of idiomatic expressions and colloquialisms.</td>
</tr>
<tr>
<td>B2</td>
<td>Has a <strong>good range of vocabulary</strong> for matters connected to his field and most general topics? Can vary formulation to avoid frequent repetition, but lexical gaps can still cause hesitation and circumlocution.</td>
</tr>
<tr>
<td>B1</td>
<td>Has a <strong>sufficient vocabulary</strong> to express him/herself with some circumlocutions on most topics pertinent to his everyday life such as family, hobbies and interests, work, travel, and current events.</td>
</tr>
</tbody>
</table>
| A2    | Has **sufficient vocabulary** to conduct routine, everyday transactions involving familiar situations and topics.  
Has a sufficient vocabulary for the expression of basic communicative needs.  
Has a sufficient vocabulary for coping with simple survival needs. |
| A1    | Has a **basic vocabulary repertoire** of isolated words and phrases related to particular concrete situations. |
Vocabulary guidelines before the CEFR

• “Specifications” of language abilities (Threshold\(^1\); Waystage\(^2\); Vantage\(^3\)) were published in the 70’s

• They included a description of the linguistic exponents (words + grammar) needed to carry out the communicative activities outlined in the framework

• The vocabulary specifications distinguished between:
  ✓ Specific Notions
  ✓ General Notions
  ✓ Functions

\(^1\) Van Ek & Trim, 1975; 1991
\(^2\) Van Ek & Trim, 1990
\(^3\) Van Ek & Trim, 2001
Specific Notions – Vantage (Council of Europe, 2001)
“[…] those which deal with more concrete details of the here-and-now.”

10 Food and drink

10.1 types of food and drink
(See also General Notions 5.1.6.)
food, meal, snack
to eat, to drink
names of meals e.g.
breakfast, lunch, tea, dinner,
supper
to have breakfast, etc.
course (as in: What’s the main course?)
vegetables
names of vegetables e.g.
beans, cabbage, carrots,
cauliflower, mushrooms,
peas, potatoes, spinach
chips

banana, grape, lemon, nut,
orange, raisin, pear,
strawberry, raspberry
cereals
crisps
ice-cubes
spice(s), to spice
names of spices e.g. curry,
garlic, mustard, pepper, salt,
oil, vinegar
flavour
spaghetti, macaroni
flour
bread
slice, piece (as in: Would you like a piece of cake?)
bread roll
toast, sandwich
bun
General Notions - Vantage (Council of Europe, 2001)

“[...] are largely concerned with abstract, relational concepts”

2 Spatial

2.1 location
noun: location, place, position, site, situation
adverbs: here, there, here and there, everywhere, somewhere, nowhere, (not) anywhere, where?, wherever; inside, outside; in the east/north/south/west
verbs: to be (as in: Harrods is in Knightsbridge.), to be situated, to lie
determiners: this, that, these, those

2.2 relative position
prepositions: above, against, among, at, at the end of, at the

2.5 direction
direction (as in: In which direction is Slough?), destination
direct (as in: Is there a direct route to London from here?)
adverbs: away, back, backward(s), (anti-)clockwise, there/here, across, over
Functions - Vantage (Council of Europe, 2001)
“[...] denoting in a general way what people do by means of language”

- Is anything wrong?
- What's the matter/problem?
- What's wrong?

2.5.24 complaining
- I have a complaint (to make).
- I'm sorry, but ...
- I don't want to complain, but ... 
  Then expressions and structures as per 2.5.22

2.5.25 expressing bad temper
- I'm in a bad mood/temper (just now/today).

- expressing interest
  - Really!
  - Fascinating!
  - Is that so!
  - Well, well.
  - (How)(very) interesting!
  - I am (very) interested in NP/VP gerund
    I am interested in old stamps.
  - NP interest(s) me (greatly)
    Greek men interest me greatly.
Limitations of current approaches (1)

- **Frequency is the main principle to identify vocabulary requirements**
  - It’s not all a matter of frequency - functional aspects of vocabulary such as *usefulness* to be considered (*Carter, 1987; Stubbs, 2009*)
  - Corpus-based frequency counts are ‘overall’ frequency counts (based on all meanings)!
Limitations of current approaches (2)

• The commonly used **unit of counting** is either the lemma or the word-family

• But do lemmas or word-families actually reflect the way we learn vocabulary? *(Bogaards, 2001; Schmitt & Zimmermann, 2002)*:
  ✓ **Word-families?** *Nation > national > nationwide > nationalism, etc.*
  ✓ **Lemmas?** *Date*
Our approach

- Combining **quantitative (frequency) and qualitative (usefulness)** to establish the relative importance of vocabulary
  - How useful is a word meaning to communicate efficiently in a particular context and for a particular purpose?

- Choosing the **word meaning as unit of counting**
  - Learning of vocabulary takes place in meaningful contexts
  - Learning of word meanings is incremental, from basic to specialized (Brent, 2009; Nagy and Scott, 2000)

“An assessment of the number of meanings a reader knows enables a remarkably accurate prediction of this individual's ability to comprehend discourse” (Anderson & Freebody, 1979)
Section II:
Introduction to the *Global Scale of English (GSE) Vocabulary*
The GSE Vocabulary - overview

• A framework which describes vocabulary targets for adult learners of general English
• Aligned to the CEFR (A1 to C2) and the Global Scale of English (GSE – 10 to 90)
• Complements the functional guidance found in the CEFR (Council of Europe, 2001) by providing lexical exponents for English
• Primarily aimed at teachers and materials designers to help them select vocabulary
The GSE Vocabulary – key features

- Publicly available and searchable online by keyword, grammatical category, topic, subtopic, and proficiency level
- Organized by topics, e.g. Food and drinks, Health and body, Business and industry
- Very large in size: Words 20k+; Word meanings 37k+; Collocations 80k+ (e.g. clear the table); 7k+ phrases (e.g. Good point! I bet! Spot-on!)
Methodology

• STEP 1 – Corpus (frequency) analysis

• STEP 2 – Semantic annotation

• STEP 3 – Teacher ratings (usefulness)

• STEP 4 – Vocabulary scaling
Step 1– Corpus analysis

• Extraction of **frequency list (top 10k lemmas)** from a corpus of general English (L1 data; 2.5 billion words; Spoken + Written)

• Integration with the **Longman Active Study Dictionary of English (LASDE)** database (25k lemmas)

• **Final list: 20k lemmas with frequency information**
Reference corpora

• LCN (Longman Corpus Network)
  A balanced and representative corpus of British/American English

• UKWaC (Baroni et al., 2009)
  A web-based corpus crawled from the .uk domain, very large in size

• COCA (Corpus of Contemporary American English)
  Spoken component selected for this study
STEP 2– Semantic annotation

• The purpose of the semantic annotation was to organise the inventory by topics and subtopics, e.g. Food and Drink

• Over 37,000 word meanings were manually (!) annotated by our team of lexicographers

• We adapted the semantic categorization found in the Vantage Specifications (Council of Europe)
Specific Notions – Vantage (Council of Europe, 2001)
“[...] those which deal with more concrete details of the here-and-now.”

<table>
<thead>
<tr>
<th>10</th>
<th>Food and drink</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.1</td>
<td>types of food and drink</td>
</tr>
<tr>
<td></td>
<td>(See also General Notions 5.1.6.)</td>
</tr>
<tr>
<td></td>
<td>food, meal, snack</td>
</tr>
<tr>
<td></td>
<td>to eat, to drink</td>
</tr>
<tr>
<td></td>
<td>names of meals e.g.</td>
</tr>
<tr>
<td></td>
<td>breakfast, lunch, tea, dinner, supper</td>
</tr>
<tr>
<td></td>
<td>to have breakfast, etc.</td>
</tr>
<tr>
<td></td>
<td>course (as in: What’s the main course?)</td>
</tr>
<tr>
<td></td>
<td>vegetables</td>
</tr>
<tr>
<td></td>
<td>names of vegetables e.g.</td>
</tr>
<tr>
<td></td>
<td>beans, cabbage, carrots, cauliflower, mushrooms, peas, potatoes, spinach</td>
</tr>
<tr>
<td></td>
<td>chips</td>
</tr>
<tr>
<td></td>
<td>banana, grape, lemon, nut, orange, raisin, pear, strawberry, raspberry</td>
</tr>
<tr>
<td></td>
<td>cereals</td>
</tr>
<tr>
<td></td>
<td>crisps</td>
</tr>
<tr>
<td></td>
<td>ice-cubes</td>
</tr>
<tr>
<td></td>
<td>spice(s), to spice</td>
</tr>
<tr>
<td></td>
<td>names of spices e.g. curry, garlic, mustard, pepper, salt, oil, vinegar</td>
</tr>
<tr>
<td></td>
<td>flavour</td>
</tr>
<tr>
<td></td>
<td>spaghetti, macaroni</td>
</tr>
<tr>
<td></td>
<td>flour</td>
</tr>
<tr>
<td></td>
<td>bread</td>
</tr>
<tr>
<td></td>
<td>slice, piece (as in: Would you like a piece of cake?)</td>
</tr>
<tr>
<td></td>
<td>bread roll</td>
</tr>
<tr>
<td></td>
<td>toast, sandwich</td>
</tr>
<tr>
<td></td>
<td>bun</td>
</tr>
</tbody>
</table>
An example of semantic annotation

- **Fork**: [FOOD/GARDENING/TRANSPORT]
- **Somewhere**: [SPACE AND LOCATION]
- **I am sorry**: [APOLOGIZE; REGRET; SYMPATHY]
STEP 3 – Teacher ratings (usefulness)

- Usefulness of vocabulary (meanings) was rated by 19 teachers:
  - *plate*: a flat dish that you use for eating or serving food
  - *plate*: in baseball, the place where the person hitting the ball stands

1 = Essential: words learners would want to acquire first
2 = Important: words that become necessary at a next stage
3 = Useful: words enabling more detailed and specific language
4 = Nice to have: words to express concepts more accurately
5 = Extra: words some language users will use occasionally
99 “Escape” (high uncertainty judgement or no knowledge)
STEP 4 - Combine usefulness ratings and frequency

\[
\text{Combine} = \frac{R_a \times r_{\text{Rating}} + F_{\text{rank}} \times (1 - r_{\text{Rating}}) + F_{\text{rank}}}{2}
\]

Where:
- \(R_a\) is the Rating average
- \(r_{\text{Rating}}\) is the Reliability of rating data
- \(F_{\text{rank}}\) is the scaled Frequency rank.
A probabilistic model of receptive knowledge

A learner at 25 on GSE

Likelihood of Success

Word meaning GSE value

Boy, Father
Girl, Mother
Female
Bride
Bridegroom
GSE Teacher Toolkit

Let the Global Scale of English (GSE) save you time and improve the quality and relevance of your English classes. The GSE Teacher Toolkit gives you fast access to GSE learning objectives, grammar and vocabulary to help you plan lessons that are at the right level for your students.

Choose category

Learning Objectives  Grammar (beta)  Vocabulary

Who are you teaching?

Adult Learners

Choose a range on the GSE / CEFR

Choose topic  Grammatical Category

Filter search results with an exact word

Clear all filters  Show results
37 main topics – and hundreds of subtopics

Choose topic

- Body and health
- Business, industry, and the economy
- Clothes, fashion, and beauty
- Countries, cities, towns, and public places
- Cultures, customs, religion, and traditions
- Do or happen
- Education, study, and language
- Existence, presence, or availability
- Feelings, wishes, and opinions
- Food and drink
  - Buying and storing food
  - Diet and nutrition
  - Hunger and fullness
  - Language related to food and drink
  - Names of food and drinks
    - Names or parts of meals
      - Names of meals
      - Parts of a meal
      - Portions of food
  - People and food

[Choose button]
### A search by subtopic “banking”

<table>
<thead>
<tr>
<th>Word</th>
<th>Subtopic</th>
<th>Part of Speech</th>
<th>Frequency</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>get</td>
<td>Banking</td>
<td>verb</td>
<td>15</td>
<td>&lt;A1 (10-21)</td>
</tr>
<tr>
<td>money</td>
<td>Banking</td>
<td>noun</td>
<td>15</td>
<td>&lt;A1 (10-21)</td>
</tr>
<tr>
<td>bank</td>
<td>Banking</td>
<td>noun</td>
<td>20</td>
<td>&lt;A1 (10-21)</td>
</tr>
<tr>
<td>save</td>
<td>Banking</td>
<td>verb</td>
<td>31</td>
<td>A2 (30-35)</td>
</tr>
<tr>
<td>card</td>
<td>Banking</td>
<td>noun</td>
<td>33</td>
<td>A2 (30-35)</td>
</tr>
<tr>
<td>account</td>
<td>Banking</td>
<td>noun</td>
<td>35</td>
<td>A2 (30-35)</td>
</tr>
<tr>
<td>credit card</td>
<td>Banking</td>
<td>noun</td>
<td>38</td>
<td>A2+ (36-42)</td>
</tr>
<tr>
<td>borrow</td>
<td>Banking</td>
<td>verb</td>
<td>40</td>
<td>A2+ (36-42)</td>
</tr>
<tr>
<td>PIN pin number</td>
<td>Banking</td>
<td>noun</td>
<td>44</td>
<td>B1 (43-50)</td>
</tr>
</tbody>
</table>

**DEFINITION**
a number you use to take money from a machine using a plastic card

**EXAMPLE**
A gang of up to five youths held her prisoner forcing her to reveal her PIN number
**A search for a specific word**

<table>
<thead>
<tr>
<th>Date</th>
<th>Definition</th>
<th>Grammatical Category</th>
<th>GSE</th>
<th>CEFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>date</td>
<td>A particular day of the month or of the year, shown by a number</td>
<td>noun</td>
<td>16</td>
<td>&lt;A1 (10-21)</td>
</tr>
<tr>
<td>Date</td>
<td>Point or period of time</td>
<td>noun</td>
<td>54</td>
<td>B1+ (51-58)</td>
</tr>
<tr>
<td>date</td>
<td>Marital status and relationships</td>
<td>noun</td>
<td>61</td>
<td>B2 (59-66)</td>
</tr>
<tr>
<td>date</td>
<td>Socializing and social arrangements</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>date</td>
<td>Names of food</td>
<td>noun</td>
<td></td>
<td></td>
</tr>
<tr>
<td>date</td>
<td>Plants as food</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Collocations of “ticket”

<table>
<thead>
<tr>
<th>VOCABULARY</th>
<th>TOPIC</th>
<th>GRAMMATICAL CATEGORY</th>
<th>GSE</th>
<th>CEFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ticket</td>
<td>Sports, hobbies, and interests</td>
<td>noun</td>
<td>21</td>
<td>&lt;A1 (10-21)</td>
</tr>
<tr>
<td></td>
<td>Countries, cities, towns, and public places</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Holidays, travel, and transportation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Media, arts, literature, and entertainment</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**DEFINITION**

A printed piece of paper that shows that you have paid to do something, for example travel on a train or watch a film.

**COLLOCATIONS**

- a ticket for sth
- buy a ticket
- a ticket price
- ticket sales

Show more collocations
Collocations of “ticket”

- a ticket for sth
- buy a ticket
- a ticket price
- ticket sales
- a free ticket
- an airline ticket
- a ticket holder
- book a ticket
- a plane ticket
- a ticket office
- an air ticket
- a ticket counter
- a bus ticket
- a return ticket
- a train ticket
- a concert ticket
- a ticket collector
- a one-way ticket
- a theatre ticket
- a ticket booth
- a ticket machine
- reserve a ticket
- a ticket agency
- a ticket inspector
- a coach ticket
- the ticket barrier
- a valid ticket
- a ticket to somewhere
More than 7k phrases: “Ways of…”

Choose a Topic

- Shopping, money, and personal finance
- Space and location
- Sports, hobbies, and interests
- Time and speed
- Ways of expressing feelings, wishes, attitudes, and opinions
- Ways of expressing obligation, necessity and probability
- Ways of interacting with other people
- Ways of organizing discourse
- Workplace and jobs

- Agreeing with a statement or opinion
- Asking about feelings
- Asking for an opinion
- Expressing amusement
- Expressing anger or bad mood
- Expressing approval
- Expressing belief
- Expressing certainty
- Expressing disagreement
- Expressing disappointment
## Phrases to “accept an apology”

<table>
<thead>
<tr>
<th>VOCABULARY</th>
<th>TOPIC</th>
<th>GRAMMATICAL CATEGORY</th>
<th>GSE</th>
<th>CEFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>no problem</td>
<td>Accepting an apology</td>
<td>phrase</td>
<td>33</td>
<td>A2 (30-35)</td>
</tr>
<tr>
<td>that’s all right</td>
<td>Accepting an apology</td>
<td>phrase</td>
<td>36</td>
<td>A2+ (36-42)</td>
</tr>
<tr>
<td>it's fine</td>
<td>Accepting an apology</td>
<td>phrase</td>
<td>39</td>
<td>A2+ (36-42)</td>
</tr>
<tr>
<td>I understand</td>
<td>Accepting an apology</td>
<td>phrase</td>
<td>42</td>
<td>A2+ (36-42)</td>
</tr>
<tr>
<td>never mind</td>
<td>Accepting an apology</td>
<td>phrase</td>
<td>43</td>
<td>B1 (43-50)</td>
</tr>
<tr>
<td>don't worry (about it)</td>
<td>Accepting an apology</td>
<td>phrase</td>
<td>53</td>
<td>B1+ (51-58)</td>
</tr>
<tr>
<td>it's no big deal</td>
<td>Accepting an apology</td>
<td>phrase</td>
<td>57</td>
<td>B1+ (51-58)</td>
</tr>
<tr>
<td>forget it</td>
<td>Accepting an apology</td>
<td>phrase</td>
<td>57</td>
<td>B1+ (51-58)</td>
</tr>
<tr>
<td>no worries</td>
<td>Accepting an apology</td>
<td>phrase</td>
<td>58</td>
<td>B1+ (51-58)</td>
</tr>
</tbody>
</table>
Printable results

Vocabulary
Search Query: -
GSE level: 10-90
Grammatical Category: -
Topic: Accepting an apology

Scan the QR code to download the results on your mobile devices

Search results

<table>
<thead>
<tr>
<th>Vocabulary</th>
<th>Topic</th>
<th>Grammatical Category</th>
<th>GSE</th>
<th>CEFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>no problem</td>
<td>Accepting an apology</td>
<td>phrase</td>
<td>33</td>
<td>A2 (30-35)</td>
</tr>
<tr>
<td>Definition:</td>
<td>used when accepting an apology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>that's all right</td>
<td>Accepting an apology</td>
<td>phrase</td>
<td>36</td>
<td>A2+ (36-42)</td>
</tr>
<tr>
<td>Definition:</td>
<td>used to tell someone you are not angry when they say they are sorry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>it's fine</td>
<td>Accepting an apology</td>
<td>phrase</td>
<td>39</td>
<td>A2+ (36-42)</td>
</tr>
<tr>
<td>Definition:</td>
<td>used when accepting an apology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I understand</td>
<td>Accepting an apology</td>
<td>phrase</td>
<td>42</td>
<td>A2+ (36-42)</td>
</tr>
<tr>
<td>Definition:</td>
<td>used when someone has apologized, to show that you are not</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary

• A unitary model of lexical proficiency to help teachers and test developers produce vocabulary tasks which are meaningful (level-appropriate) and authentic to the students

• Integrating different dimensions (size and depth) of receptive vocabulary knowledge

• Combining the two principles of frequency and usefulness

• Using the word meaning as unit of counting and therefore taking into account the contextual dimensions of learning (polysemy, collocations, and phrases)
Section III:
Discussion and new directions
L1 data vs learner data

- Creating a standard means identifying a common learning goal and ensuring assessment is fair and consistent across different contexts and groups of learners.
- However, learner data are highly variable (depending on the learners’ L1 and cultural background) and their use to set a standard could be highly misleading.
New directions

• The GSE Vocabulary database (raw data) will be made freely available to external researchers for validation studies.
• We are currently conducting a study to identify the different weight that particular items have in the different varieties of English.
• Where there is a need for localized assessment, the GSE vocabulary can be adapted by teachers and test developers.
• We have recently started to develop a vocabulary framework for Young Learners.
Thank you!

• The GSE vocabulary is already **freely available at English.com**

• **For any information** about the project or if you would like to get access to the data for research purpose, please contact me: veronica.benigno@pearson.com
There’s so much more to learn

Find out more about us at English.com/gse