

Lexical availability

Chris Payne analyses *available* words as opposed to *frequent* words.

The study of lexical availability is not new. It originated in France in the 1950s and 60s. Since the seminal work in French by Georges Gougenheim *et al* in 1964, numerous studies have been carried out, notably the work of Naum Dimitrijevic in English in 1969 and, more recently, the research conducted in Spanish by Rosa María Jiménez and her colleagues.

Defining lexical availability

Lexical availability (LA) can be defined as the words that spring to mind in response to 'word prompts' – topics related to daily life, such as *food and drink, health and medicine, the countryside, the city, animals, parts of the body, garden, clothes*, and many more.

LA is measured by means of a test that reflects the participants' spontaneous vocabulary production.

In the 1950s, great importance was attached to vocabulary teaching in France, so an organising principle was needed for the vast French lexicon. In order to select words for inclusion in the *élémentaire* level, it was decided that priority should be given to the most frequent words because they were the most useful. It is interesting that frequency lists were being compiled so long before the advent of computer corpora. However, it soon transpired that there was a recurrent problem with using frequency as the sole selection

criterion: some well-known words which were deemed to be common by French speakers did not necessarily feature in the frequency lists.

LA studies emerged as a way of compensating for the perceived shortcomings in word frequency data, as they served to foreground useful native-speaker vocabulary that was excluded from frequency counts.

Considering word frequency

Today, word frequency information obtained from computer corpora provides lexicographers, coursebook writers and teachers with an invaluable source of reference for vocabulary selection, and a large number of materials for language teaching and learning are now corpus-informed.

Despite the importance of selecting and teaching the most frequent words in English, contemporary corpus linguists acknowledge that an adherence to frequency data alone would lead to some strange teaching practices. Some important words, especially those belonging to closed lexical sets, fall outside the list of the 2,000 most frequent words in English. Words like *Tuesday*, *Wednesday* and *autumn*, to name just a few, are relatively low-frequency.

But no teacher would omit *Tuesday* and *Wednesday* and *autumn* when teaching the days of the week and the seasons, simply because these words are

statistically less frequent than their co-hyponyms. As the authors of *From Corpus to Classroom* state: 'Corpus statistics need to be combined with a notion of psycholinguistic usefulness and the availability (*disponibilité*) of items in the mental lexicon.'

To date, although substantial research into lexical availability has focused on Spanish, studies have also sought to compare English as an L1 and L2. Here, we will consider LA within the context of English as an L1.

Testing lexical availability

Let us look at an example of an LA test and how it is implemented.

This paper-based LA test is designed to last about 20 minutes:

- Fifty native-speaker participants are presented with ten topics in no particular order. For example, *health and medicine*, *the garden*, etc.
- Each topic is displayed at the top of a different page, and the participants are asked to write as many words as possible from the given topic in two minutes.
- Space is provided for 50 words or lexical items consisting of more than one word, eg *mow the lawn*.
- The participants are not allowed to move on to the next topic until the two-minute time limit has ended, and they are asked to hand in the test immediately after 20 minutes, when all the topics have been completed.

There is no single standard model of an LA test, so there is considerable leeway as regards the design of the test:

- The topics used to elicit responses can differ.
- The number of topics can be changed, thus lengthening or shortening the test.
- An open list can be used, so that the participants have a time limit (also variable) but not a cap on the number of words they produce.

Measuring lexical availability

The test results are interpreted using two criteria:

- 1 The frequency with which a word is produced under a certain topic.

- 2 The number or position a word occupies on the page for a given topic.

Words produced by a large number of participants and which appear early on the page are given a high lexical availability value, whereas words produced by few people or appearing at the bottom of the list are assigned a low lexical availability value.

We ought to bear in mind that lexical availability is not a *fixed* property of words, and that the availability of a word can vary according to the topic for which it is produced. Thus, the word *car* has a high LA value for the topic of *transport*, a medium LA value for the topic of *the city*, and a low one for *the countryside*.

Here are the ten words with the highest LA value which were produced by British participants in 2006 for the topic of *health and medicine*:

- | | |
|---------------|-------------|
| 1 doctor | 6 injection |
| 2 hospital | 7 NHS |
| 3 nurse | 8 death |
| 4 cancer | 9 surgery |
| 5 paracetamol | 10 disease |

These ten most elicited words are all a good representation of vocabulary that people actually use in conversation, regardless of whether they are statistically 'frequent' or not.

Generally, the words that are produced in LA tests are useful for our learners to know. But in the event that a participant produces 'strange' words in a test, they are unlikely to be written by other participants, so these words end up at the bottom of the final list and they never obtain a high LA value.

The following guidelines are also used to edit the results of a test:

- Words that are misspelled are accepted and counted.
- Words that are repeated under the same topic are counted once.
- Illegible words are not counted, neither are words which clearly don't belong in the topic in which they were written.

The implications for teaching

Teachers can address the claim that word frequency lists do not adequately represent certain vocabulary used by native speakers in everyday life: by distinguishing between *frequent* words – those with a high statistical occurrence –

and *available* words – those that come to our memory first in response to a stimulus or prompt. However, as you may have noticed in our *health and medicine* example, *available* words can also be *frequent* words, but although there is sometimes overlap between *frequent* and *available* words, frequency is not a reliable predictor of lexical availability.

The results of LA tests provide us with an alternative vocabulary source to consult, and can complement the use of word frequency data by also helping us to expand and enrich our learners' vocabulary. We can draw attention to the most *available* lexis and highlight different aspects of words, depending on the level and needs of our learners.

For example, from our previous list of ten words we could focus on the following:

- **hospital** – teach collocations like *be admitted to hospital* and *be discharged from hospital*.
- **NHS** – explain what it stands for, which is particularly relevant for learners living in the UK.
- **injection** – focus on its collocates, like *give, have*, etc. (Spanish learners might be tempted to say *put an injection*.)
- **surgery** – explore the polysemous nature of this word in British English.
- **disease** – contrast its use and collocations with those of illness, with which it is sometimes confused.

Anyone who is interested in this area of research can devise their own test and give it to native English speakers in their community. A paper-based LA test can be taken almost anywhere, as administering it doesn't rely on technology of any sort: all you need to do is find willing participants!

Obviously, this is easier in countries where English is the language of the community, such as the UK and the USA, but there is often a local population of native speakers living in places where English is an L2. Although it is not feasible for busy teachers to set up tests for 50, 100 or more people at a time, we can focus on small groups or even individuals, staggering our research over a period of time that suits us.

Small-scale teacher research on LA can be fruitful, and a piecemeal approach enables us to build up a picture of *available* words which are potentially useful in the classroom.

Lexical availability

Several factors can influence the results of an LA test, which I have grouped under two separate headings: the participants or 'testees', and the topics or word prompts.

The testees

The number of people who are tested is important – the larger the sample of language production that is gathered, the more representative it will be of current English. Naum Dimitrijevic tested 185 secondary school students in Edinburgh, but some tests, like my chosen example, are conducted with 50 people or fewer.

The type of testee should also be considered. Ideally, I suggest selecting a cross-section of people in the community that takes into account age, gender and specialised knowledge. People of different ages use different words, so testing teenagers, young and middle-aged adults and senior citizens will offer a more complete picture of available words.

Dimitrijevic looked at the effect of gender on learners' LA with the subjects of his study in Scotland in 1969, but his results and those of most other research since then generally conclude that gender has no significant influence on lexical availability. Even so, for variety's sake, it makes sense to choose a mix of male and female testees.

It is also worth finding out in advance whether test participants have specialised knowledge of the topics to be tested. For example, the words produced for the topic of *health and medicine* by 30 doctors or health professionals will no doubt be quite different from those written by 30 laypeople.

The word prompts

If we take another look at the ten most available words from our earlier example of *health and medicine*, we can see immediately that all the words elicited were nouns. This is understandable: after all, we use nouns to name things. The topic prompts in most LA tests are overwhelmingly nouns, and many are superordinates, such as *professions*, *animals*, *clothes*, etc. The relevance of this is that the word prompts tend to trigger responses that belong to the same word

class, ie nouns have a propensity to elicit other nouns. Of course, there are some examples that buck this trend, like *mow the lawn* under the topic of *garden*.


Another observation is that some topic prompts lead to more production than others: *food and drink*, *animals* and *the town* are a particularly fertile ground for available words.

Hopefully, future LA studies in English will include prompts that elicit more adjective and verb responses, with the aim of obtaining a broader range of lexis.



I have barely scratched the surface of research into lexical availability; I have taken just a cursory glance at some factors that can impact the outcome of LA tests. There are, of course, others, like cognitive factors such as *imageability* and *typicality*:

- 'Imageability' is the ease with which you can evoke a mental image of a word – concrete nouns like *apple* tend to be easier to envisage than abstract nouns like *intelligence*.
- 'Typicality' refers to prototypes – the words that are the most characteristic and representative of a given category. For instance, a *robin* is considered to be a more typical example of a bird than an *ostrich* or a *penguin*.

Nor have I touched on the usefulness of the tests for comparing the available words of L1 speakers and L2 learners. My wish is that, in the future, there will be more extensive research in English into this approach to vocabulary selection – and that the results will be made widely available. 

Jiménez, R M *Lexical Availability in English and Spanish as a Second Language* Springer 2014

O'Keefe, A, McCarthy, M and Carter, R *From Corpus to Classroom* CUP 2007



Chris Payne is the owner of Paddington School of English, Linares, Spain. He is an experienced teacher and a former Cambridge English oral examiner and oral examiner trainer. He has written several published articles on ELT.

paddingtonschool@outlook.com

ENGLISH TEACHING professional

This is your magazine.
We want to hear from you!



IT WORKS IN PRACTICE

Do you have ideas you'd like to share with colleagues around the world?

Tips, techniques and activities; simple or sophisticated; well-tryed or innovative; something that has worked well for you? All published contributions receive a prize!

TALKBACK!

Do you have something to say about an article in the current issue of ETp?

This is your magazine and we would really like to hear from you.

It really worked for me!

Did you get inspired by something you read in ETp? Did you do something similar with your students?

Did it really work in practice?

Do share it with us ...

Reviewing for ETp

Would you like to review books or other teaching materials for ETp?

We are always looking for people who are interested in writing reviews for us.

For guidelines and advice, write to us or email:

helena.gomm@pavpub.com

ENGLISH TEACHING professional
Pavilion Publishing and Media Ltd,
Rayford House, School Road,
Hove BN3 5HX, UK

Fax: +44 (0)1273 227308

Email: info@etprofessional.com