

Introduction

The purpose of the English Language Learner's Lexicon is to provide teachers and learners with a rough guide to the relative importance of the words of the English language for the learner. With an estimated lexicon of 500,000 words, the acquisition of the vocabulary of English is arguably the most daunting task facing the learner. Of course, one does not need to know the total lexicon of English, and it is quite possible to live comfortably in the English-speaking world with a vocabulary of only a few thousand words. However, for the learner it is important to spend one's energy on words that will be necessary and useful. For the teacher, this lexicon will be helpful for designing appropriate language materials and it can be a useful reference for checking on what lexical items have and have not been covered in a course.

It is obvious that there are many variables to be considered as a learner's lexicon is compiled. Of considerable importance is the **learning context** of the learners. Are they self-studying? in an environment where English is not spoken? in a formal class? taking night classes while attempting to survive in an English-speaking land? preparing to travel or already traveling in the US? Canada? Great Britain? Ireland? Australia? New Zealand? South Africa? Jamaica? any other country where English is the lingua franca?

Another critical variable is **the learners**. The **age** of the learners certainly makes a difference. The **native language and culture** of the learners must be considered, along with the learners' **previous contact with English** and **educational background**.

Also for consideration is the **needs of the learners**, as implied above — for tourism/travel? for academic preparation? for economic and social survival? for international business? for professional needs and advancement? for using the World Wide Web? Obviously, the lexical requirements of the learners will vary with the learners' needs, especially as the learner gains in proficiency and requires more specific or technical language.

Of course, yet another fundamental but complex variable is **What is a Word?** At a very simple level, do we consider, for example, *an* to be simply a variant of *a*, and therefore these two forms are one word, or are they two? How do we deal with inflections, such as *learn*, *learned*, or *know*, *knew*, or *child*, *children* and prefixes: *unimportant* and *important* — two words or one? What about compound words such as *printout*? Can we assume that if the learner already knows *print* and *out*, they will therefore know *printout*? And how do we deal with derivations formed by affixing, for example, *different*, *difference*. Do we count them as one word or as two closely related forms of the same word?

In compiling this lexicon, it has been necessary to deal with the questions posed above and to make a number of assumptions that will determine what is included in the list, and at what level.

The first and foremost assumption is that the learner intends to use English for **social-interactional communication** with other speakers of English in a face-to-face situation. In other words, functioning in the spoken language is more important than functioning in the written language through reading and writing (although email correspondence may be blurring this distinction somewhat). Therefore, words such as *hello* or *excuse*, as in *Excuse me*, will appear early in this list.

The second assumption concerns the learning context. We have assumed that **the learner is engaged in some kind of formal teaching/learning situation**, as opposed to self-study. Therefore, especially at the survival/beginner stage/level, many classroom words such as *pen*, *listen*, *copy* and *assignment* are included in the list. We have further assumed that very basic metalinguistic terms such as *noun*, *verb* and *sentence* will be useful, even necessary, for the learner at the beginner level.

Although there are many varieties of English, this list is based on **North American English** and of course reflects the lexicon of this variety of English. Nevertheless, the lexical variation from one variety of English to another is not a significant problem (although pronunciation of the words can be a troublesome matter). The North American variety has been used because it is nowadays probably the most widespread and influential variety, and because the compiler is a North American.

As for the assumptions about the learner, the first is that the learner is **not a child learning the language in a true bilingual context**, i.e., as a child acquiring two native languages. Furthermore, we have assumed that the learner is **at least in elementary school**, and so some high frequency words used by very young children, for example, *Mommy*, *doll*, *blocks*, *doggie*, will not appear in our list.

Because of the tremendous variation among learners and their backgrounds and needs, this lexicon, limited to 2400 words, is based on **a generalized list of words that virtually all learners** need to know, whether child or adult, immigrant or traveller. A close examination of the list will reveal that the 0-300 level is heavy with function words and very basic verbs and adjectives, whereas the number of nouns is more limited. As one progresses through the list, the proportion of nouns increases as do the number of derivations, as the learners' ability to express themselves becomes more sophisticated, and the topics of communication become more numerous and wider. This lexicon purposefully goes only to the 2400-word level, because beyond this level, the predictability of the learner's requirements becomes less and less certain, and more and more diffuse.

The final consideration is what counts as a word? The fundamental assumption is based on the unmeasurable, but somehow real criterion of **challenge to the learner**, i.e. how much energy is required to learn and use a word. In most cases the challenge is determined by the degree of semantic overlap. For example, if the learner knows the word *teacher*, it is not a significant challenge to the learner to know *teachers* (although some awareness of regular plural inflections, i.e. awareness of grammatical rules, may require some initial learning energy and challenge). *Teachers*, as a word, is not significantly different in meaning from *teacher*. Therefore, the two forms are considered to be one word.

Now the question is, where to draw the line? For example, *child* and *children* may represent a greater challenge because of the irregularity of the plural form, but because of the basic similarity in the form and meaning of the two words, they are not considered to be two words. Furthermore, *bring* and *brought* are considered as one word, despite the difference in form because the difference in meaning is grammatical, not lexical. However, the word *break* in *take a break*, despite the similarity in form with *break*, as in *did you break your arm* or *that was a lousy break*, is a separate word, different from the other two meanings of *break*. These three examples represent three different meanings and are therefore, three different words. Finally, the personal pronouns, although there are differences in form, are each considered to be one word, as in *I, me, my, mine*.

With all that in mind, this list of 2400 words is correlated with Pro Lingua's four stages of functional proficiency. These four stages represent the transition from a very limited ability to function in English to a level at which the learner is beginning to be bilingual and bicultural.

The stages of functional proficiency are correlated, in turn, with the commonly-used linguistic proficiency terms, beginner, intermediate, advanced, and each level includes the number of words that a learner could be expected to know at each stage and level. The following chart summarizes the correlations.

<u>FUNCTIONS</u>	<u>WORDS</u>	<u>LEVEL</u>
Limited Survival	0-300	Absolute Beginner
Surviving	300-600	Beginner
Adjusting / Settling in	600-1200	Low Intermediate
Participating	1200 - 2400	High Intermediate
Integrating	2400 +	Advanced

In summary, an absolute beginner knows fewer than 300 words and cannot really function well without a bilingual friend or possibly a bilingual phrase book. A beginner is a learner who can accomplish some basic tasks with some difficulty (survive) and knows up to 600 words. An intermediate learner can get around and knows up to 2400 words, and an advanced learner knows at least 2400 words and is beginning to function completely in English.

One final note on the dividing lines of 300, 600, etc. Of course there are no sharp lines in the language learning process and the decision to place a word at the 300, rather than 600 level, is based mostly on experience and intuition. However, it would be best to approach this list with the understanding that most of the words at the 0-300 level probably belong there, but any given word may or may not occur in a real teaching-learning situation at a given level, because real contexts vary considerably from one time and place to another.

This lexicon is divided into four parts, one for each level. Each list is cumulative, that is, the 301-600 list also contains all the words in the 0-300 level, as well as the new words at the 301-600 level. The new words at each level are in boldface.

So how does a teacher use this list? Cautiously and with awareness that no list will be the same for every teaching/learning situation, and this list is, after all, a rough guide to be used in deciding what and what not to teach learners at four levels of functional proficiency where the division between levels can never be a sharp dividing line.

User's Guide

The word lists are presented with a few conventions to help in understanding the nature of the word. They are summarized below.

Example(s)	Explanation
fall (<i>season</i>)	When there is more than one common meaning to a word, the intended meaning is italicized in parentheses.
begin/began/begun turn on/off	Irregular inflected forms and other closely related forms are considered one word, but the forms are listed and separated by a slash.
(un)important slow(ly)	Derivational affixes are indicated with the affix in parentheses when there is no other change in the form of the word.
(in) front (of) (air)plane	Compounds and collocations are also indicated with parentheses.
different/difference	Derivations that result in a spelling change are separated by a slash. The derived form is not considered a separate word.
-body anybody nobody somebody	High frequency compound words are shown with the base form preceded by a dash. The variations are indented, but not considered different words.

When there may be a question about which use of the word is intended, an abbreviation indicates the intended part of speech. The abbreviations are listed below:

n = noun	adj = adjective	excl = exclamation, interjection
pro = pronoun	adv = adverb	conj = conjunction
v = verb	prep = preposition	
m = modal verb	det = determiner	
av = auxiliary verb	int = intensifier	