Factors Affecting the Learnability of Technical Vocabulary: Findings from a Specialized Corpus

Simon FRASER
Institute for Foreign Language Research and Education
Hiroshima University

The "difficult" lexis of a medical discipline such as pharmacology is probably the biggest challenge facing the learner of English for Specific Academic Purposes (ESAP). But what exactly makes this type of vocabulary so problematic? Morphological complexity is clearly a source of difficulty, but there are several other less obvious factors which need to be considered. These form the primary focus of this paper, which draws on my research into the lexical characteristics of pharmacology texts (Fraser, 2007; 2009), in which I used frequency counts to compile lists of the most useful words from a corpus of pharmacology research articles.

The paper will first of all give a brief account of the work I have done in establishing new categories of specialized vocabulary and creating word lists. Then, focusing on the words in the lists, it will investigate the potential difficulties faced by learners in the acquisition of technical vocabulary. We will see how intralexical factors such as *synformy* and *deceptive transparency* (Laufer, 1997) and the notion of *technicality as compression* (Ward, 2007) can be helpful when investigating specialized vocabulary. Evidence will be provided, too, of the influence of the learners' L1 (Japanese, in this case) on word learnability. Finally, the paper will look at the ways in which the findings might inform the creation of pedagogical word lists and other teaching materials for pharmacology and related medical disciplines.

LISTING AND CATEGORIZING SPECIALIZED VOCABULARY

My investigations began with a replication of Chung and Nation's (2003) analysis of the size and importance of a technical vocabulary. Chung and Nation found that 31% of words in an anatomy textbook, and 21% in an applied linguistics textbook, were technical. In my study (Fraser, 2005), a pharmacology textbook replaced the anatomy text used in the original research. I found that as many as 36% of words were technical, which supports Chung and Nation's assertion that a technical vocabulary, particularly for a medical discipline, comprises a much larger proportion of the lexis of specialized texts than had generally been supposed; Nation (2001), for example, had suggested a figure of only 5%.

Establishing New Categories

Carrying out a replicative study also presented the opportunity to devise a new way of classifying vocabulary that would be more helpful than the usual technical/subtechnical distinction, and it was suggested that the vocabulary of specialized texts might better be divided into *fully technical, lay technical, cryptotechnical*, and *academic* categories.

Fully technical vocabulary consists of words with meanings which are clearly technical; they are specific to the field and not likely to be known in general language. Examples taken from my pharmacology corpus are *fluconazole, sulphydril*, and *tricyclic antidepressant*. Typically these are the names of drugs, but they may also be terms which are found in closely related fields such as physiology or molecular biology. As we might expect, almost all of the fully technical words are of Greco-Latin origin: words like *cardiac, cardiomyopathy, oedema, angiotensin*, and *ventricular*. This is not exclusively the case, however, as we can also find words like *afterload* and *preload*.

Cryptotechnical vocabulary consists of polysemous words such as *transmitter, dependence*, and *relaxation*; these words could be said to be "cryptic" in that they have a technical meaning which may be obscure to a non-specialist. Lay-technical words, on the other hand, are those terms which are obviously technical, but whose basic meaning would, nevertheless, be understood by someone without specialist knowledge in the field.

Academic vocabulary contains a large number of discourse-structuring words and words with an analytic or evaluative role. In pharmacology, items of this type include *adequate, maintain, insufficient, involved*, and *volume*. At present, despite its limitations, the best source we have of this kind of vocabulary is Coxhead's (2000) Academic Word List (AWL), after removing the words that can be considered technical (i.e., cryptotechnical words). However, there are also words from the high frequency categories that would seem to belong here, many of which are verbs involved in analytical description (e.g., *cause, develop, result in/from, lead to*), or anaphoric nouns such as *problem, findings*, and *study*.

Creating Word Lists

The categorization system described above provided a good basis on which to proceed to the next stage of creating a list of words that would be of optimum use for learners. The methodology employed in the construction of the word list drew upon the assumption that learners first learn general vocabulary (e.g., the GSL or equivalent lists), followed by academic vocabulary (the AWL), and only then move on to the specialized terms of their particular field.

In order to know how useful a word list is likely to be, coverage — the percentage of running words (tokens) necessary to ensure reasonable comprehension of a text — is an important concept. Laufer (1989a) calculated that 95% is the absolute minimum percentage of the words in a text needed to be known for comfortable comprehension, and Hu and Nation (2000) have placed the target at an even higher 98%. While supporting this figure, Schmitt et al. (2011) have evidence to suggest there is no actual "threshold" percentage, which means that it is well worth pushing for as high a coverage as possible, although 98%, or even 95%, may be beyond reach.

Using an initial corpus of 180,000 words (50 research articles), with frequency as the primary criterion, a 600-word list of specialized pharmacology words (the Pharmacology Word List) was compiled. This list gave 13% coverage of the corpus, which was significantly better than that provided by the similarly-sized Academic Word List. When combined with words in the GSL and AWL, a total coverage of 88% was achieved — falling short of the 95% target, but

respectable, nonetheless.

The next stage of the research attempted to address the following issues: the relatively small size of the pharmacology corpus, the fact that general and academic words can be used with specialized meanings, and the realization that it may not be necessary, or even desirable, for learners to master these more "basic" categories of words before starting to learn specialized vocabulary. A larger (360,000-word) corpus was created which better represented the various areas of pharmacology (e.g., cardiovascular pharmacology, endocrine pharmacology, and toxicology). From this corpus, a single, 2,000-word list was created which provided coverage of almost 90%. To achieve this, although range across the articles in the corpus was taken into account, frequency was again the primary criterion. This meant that there were some words in the list which, although they occurred with high overall frequency, were found in only two or three articles; such words will obviously be less useful for learners to know than those occurring in more than half of the articles, for instance.

The final stage was the creation of a 570-word "Essential Pharmacology Word List" (EPWL). With this list, the aim of creating a list of manageable size that will provide learners with the most important words in pharmacology has been achieved. The EPWL will be particularly useful for learners for whom classroom time is limited, and who may have difficulty in learning as many as 2,000 words; according to Milton and Meara (1998), L2 vocabularies typically grow by about 500 words for every year of university study. The list includes words which not only occur with the highest frequency, but are also found in the widest range of articles. At 570 words, it is the same size as the AWL, and it provides coverage of 27% of the Pharmacology Corpus. This is considerably better than the 9.5% coverage that the AWL gives of the same corpus, or the 12% that Wang et al.'s (2008) substantially larger Medical Academic Word List gives of medical corpora. An additional bonus is that the EPWL performs almost as well on texts taken from the wider field of general medicine as it does on pharmacology corpora.

POTENTIAL DIFFICULTIES FACED BY LEARNERS IN THE ACQUISITION OF TECHNICAL WORDS

Having identified the words that will be most useful to learners, we need to think about what we can do to ensure that these words will be acquired in the most efficient manner. This entails understanding what exactly it is about the words that we might anticipate will cause problems for learners. To this end, it is useful to look at word difficulty in terms of the categories suggested by Laufer (1997), who lists several features inherent in the word itself that could affect the ease or difficulty with which it is learned. The ones which are of particular interest to us, and which I feel are most likely to present learners of specialized language with difficulties, are: *pronounceability, orthography, length, morphology, inflexional and derivational complexity, synformy*, and *semantic features*.

Let us, then, take Laufer's criteria for word difficulty and illustrate them with words from the Essential Pharmacology Word List. As we proceed, we will see the importance of what Lado (1955: 31) considers to be "the most powerful factor in acquiring the vocabulary of a foreign

language": the vocabulary of the learner's L1. Meara and Bell (2001) point out that we cannot draw up a list of "difficult" words without taking into account the language learning context and background of the learners, something which has been investigated by Swan (1997), who warns of the dangers of learners constructing unrealistic "equivalence hypotheses". Ryan (1997: 94), too, stresses the need for a "greater understanding of the influence of some first language reading systems on second-language learning", an observation which most certainly holds true for the language teaching situation in Japan.

Pronounceability, Orthography, and Length

Problems with pronunciation are most likely to result from the presence of foreign phonemes, phonotactic irregularity, and variable stress and vowel change. The extent of any difficulty will, of course, be determined by the learner's L1 system. Some examples of potential problems for Japanese learners are given below.

1. Confusion between /s/ and /ʃ/ sounds: words like *shift, sigma, syndrome*, and *synaptic* will be difficult to articulate correctly.
2. The lack of the /θ/ phoneme: *therapy, ethanol, hypothesis*, and *synthesize*, for instance, will often be pronounced with an "s" sound replacing the "th".
3. Confusion between /b/ and /v/ sounds (/v/ is rare in Japanese): *variability, volume, survival*, and *vivo* may cause problems, with /v/ often being mispronounced as /b/.
4. Confusion between /r/ and /l/ (the sound in Japanese is somewhere between these two phonemes): words such as *cerebral, clearance, laboratory*, and *relevant* are likely to be difficult for many learners to get their tongues around, as well as to spell.
5. For Japanese learners, the way in which the syllabary of the language is rendered in the Roman alphabet is also often a source of confusion: in Japanese, the letters "a", "e", "i", "o", and "u" are always pronounced /ɑ/, /e/, /i/, /o/, and /u/. Because of this, the vowels in words like *ion, variable, diabetes*, and *domain* will frequently be mispronounced.

Laufer (1997: 144) draws our attention to the fact that a great many English words provide no clues to their pronunciation (she invites us to consider, for example, the different ways in which the letter "o" is sounded in *love, chose, woman, women*, and *odd*). Words characterized by such sound-script incongruence are prime candidates for not only errors in articulation, but also misspellings. Pharmacology words that could trip up learners include the following pairs which, although spelt in a similar way, are pronounced quite differently from each other: *protein/vein, macrophage/image*, and *figure/failure*.

Other words for which we would anticipate both pronunciation and spelling difficulties might include *fluorescence, withdrawal, simultaneous*, and *phosphorylation*. The length of these words is obviously a factor, although as Laufer (1997: 145) points out, it is not necessarily length *per se* that is the cause of difficulty; long words such as *pharmacological, administration*, and *predominantly* should be relatively straightforward for most learners to pronounce and spell.

For Japanese learners, the fact that there are several items in the pharmacology lists which are essentially the same in both English and Japanese is important. Japanese often uses loan words in preference to the original Japanese terms, and sometimes the loan word is used exclusively. This phenomenon is welcome to a certain extent, but problems arise because the pronunciation (and spelling in the Roman alphabet) of these words is altered, often quite markedly, to conform to the Japanese phonetic system. Words falling into this category include *alcohol* (rendered as アルコール, or "arukoru", in Japanese); *syndrome* (シンドローム, or "shindoromu"); *methane* (メタン, or "metan"); and *liver* (レバー, or "reba").

Complicating matters for Japanese learners of medical English is the fact that many words are false cognates (often borrowed from German rather than English), and these can be a source of much confusion. Examples include レントゲン ("rentogen": *X-ray*); カルテ ("karute": *medical chart*); ホルモン ("horumon": *hormone*); クランケ("kuranke": *patient*); and ノイローゼ "noiroze": *neurosis*). We also find terms such as ナトリウム ("natoriumu": *sodium*, and カリウム ("kariumu": *potassium*), which have entered the language from Latin via German.

Morphological Complexity

Laufer (1997) suggests that the learning load of a word with multiple inflexional forms will be greater than that of an item with no such complexity. Therefore, a family such as *withdrawal, withdraw, withdrew, withdrawn* might be expected to cause more problems than one with more regular inflexion (e.g., *transport, transported, transporter, transportation*). Irregular plurals, too, as we see with *datum/data* and *criterion/criteria*, will also require more attention.

The ease with which a word can be decomposed into its constituent morphemes has been termed its "derivational complexity" by Laufer. The difficulties faced by learners will obviously depend on their familiarity with the most frequently occurring affixes, and if a word is formed in a regular way, it may not be as difficult as it first appears. Take *depolarization*, for example: if learners know *polarity*, which occurs quite frequently, and are acquainted with the familiar morphemes *de-*, *-ize* and *-ation*, then they should be fairly easily able to work out the meaning. As we know, many technical terms are of Greco-Latin origin, and knowledge of prefixes and suffixes such as *hyper-, patho-*, and *-itis* will be important for pharmacology students.

Lack of regularity, however, can be problematic, and difficulties will arise when morphemes combine irregularly to create meanings, or when it is possible for multiple meanings to result. Of particular interest is Laufer's category of "deceptively transparent" words (Laufer, 1989b): words which are made up of morphemes with a familiar meaning, and appear to be easily understandable, but which in fact may have a sense that is quite different from the one assumed by a learner.

To illustrate the idea of deceptive transparency, we can consider the words *intake* and *uptake*, both of which are common in pharmacology. The meaning of *intake* can be gleaned from its constituent parts *in* and *take* as "the action of taking something in, or the amount taken in". The same is not true, however, for *uptake*, which in fact means "the action of taking in or absorbing of a substance"; the meaning of the word has nothing to do with moving an object in an upwards direction. Another example from the Pharmacology Word List is *outcome*, which

of course does not mean "come out", or "exit from", as the unwary learner might imagine. Also, *clearance* could be confused with *clarity*, and *onset* does not have the meaning of placing something on something else. *Withdraw* is another word which might fall into the category of deceptively transparent words.

Synformy

Similarity of lexical forms is another potential source of confusion identified by Laufer (1988). As the name suggests, "synforms" are pairs of words which differ in meaning but sound alike ("synophones"; see Laufer, 1981), look similar, or both. Laufer cites studies carried out by Henning (1973) and Meara (1982) which indicate that form interference from an already known word can lead to difficulties in retaining the correct form of a new word. The findings from Laufer's own studies enabled her to propose ten categories of synforms, with pairs differing, for example, in terms of a single phoneme, length, or a prefix or suffix (see Laufer, 1988, Appendix 1, for details of all of these categories). She found that the most problematic synforms were those which differed according to suffixes (e.g., *industrial/industrious; comprehensive/comprehensible*) and synforms with identical consonants but different vowels (e.g., *conceal/cancel; adopt/adapt; proceed/precede*).

Table 1 lists pairs of synforms, each consisting of words found in the Pharmacology Word List, which look as though they might cause problems for pharmacology learners:

Table 1. Candidates for Synformic Confusion in Pharmacology

affect/effect	insulated/isolated
assay/assess	internal/interval
base/bias	mode/model
constant/consistent	mediate/medicate
contact/contract	modify/modulate
efficacy/efficiency	regimen/region
efficient/sufficient	simulate/stimulate
formation/formulation	section/secretion
identical/identified	state/status

Semantic Features

Laufer (1997) identifies *specificity* and *multiple meaning* as two of the semantic properties of words that affect their learnability.

The notion of specificity obviously concerns us in our investigation into the lexis of specialized texts. When Laufer states that words with high specificity are more difficult for learners, what she means is that words which cover a large area of meaning and fit a variety of contexts are preferred to those which are restricted to a very narrow area of use. Of course, one reason why we are creating specialized word lists is to help learners acquire these more "difficult" words. However, the situation is more complicated than can be summed up by "general words are easy and specialized words are difficult" (as we have seen with cryptotechnical

words, which can be words found in general word lists but used with specialized meanings). Indeed, as Strevens (1973: 228) points out, for the ESAP student who is already familiar with the scientific field, fully technical words may actually be easier to learn than general words.

Another way of looking at the issue as it relates to the learning of a specialized discipline is to consider that there will be increasing degrees of specificity even within an already specialized field. Thus, there will be words which are found in all areas of pharmacology (*receptor, channel, agonist*, and *channel*, for example), and words with higher specificity, such as *cardiomyocyte* (a type of muscle cell found only in the heart) or *toxicosis* (in toxicology, a diseased condition resulting from poisoning). We might expect the former category of words to be more readily acquired by the majority of pharmacology learners.

The fact that many words have multiple meanings is one of the most important considerations when dealing with the difficulty of specialized lexis. Laufer (1997) suggests that whether these words are polysemes or homonyms is irrelevant; the challenge will be the same, with learners needing to be able to discriminate between the different senses of the same form, and to use these senses correctly. However, it would seem that this is something that most learners find far from easy to do. In a study of lexical guessing, Laufer found that those of her students who knew one of the meanings of a polyseme/homonym persisted with that meaning even though it made no sense in context.

Bogaards (2001), on the other hand, has shown that if learners are aware that a word has multiple meanings, the presence of already established meanings does not hinder the learning of new senses, and as many as four new meanings can be added to known forms without causing confusion. Bogaards also found that when the unknown sense of a known form is related to the meaning already known, it can be advantageous in at least the initial learning of the new sense.

Technicality

For Laufer (1997), the more specialized a word is, the harder it is to acquire, and so by definition, a technical word will be difficult. Obviously, technical words will be difficult for most learners to acquire because they occur infrequently in general language, meaning that there will be few opportunities to encounter them. However, this will be far less of a problem for ESAP learners, who will already have a knowledge of the discipline and thus be familiar with many of the concepts; they will be helped by the fixed meanings of most technical words, and by cross-linguistic equivalence.

Still, there are innate properties of these words, in addition to the categories suggested by Laufer, which may affect their ease of acquisition. Ward (2007), drawing upon the work of Halliday (1985) and Pueyo and Val (1996), considers the question of why a technical word should be inherently difficult. He argues that there are two aspects of technicality that are important in this regard: 1) *compression*, and 2) *precision*.

1. *Technicality as compression*

This refers to the process of nominalization which allows complex phenomena to be summarized in a few words. Ward (2007: 24) exemplifies this with the following definition of the term *enthalpy*:

> The total heat content of a system, expressed as a thermodynamic quantity obtained by adding its free energy to the product of its pressure and volume.

Ward explains that this definition, or "explanatory nominalization", represents an "unpacking" of the term into three further nominal phrases (*total heat content, thermodynamic quantity*, and *free energy*), which themselves require additional unpacking. The reverse of this process, *compression*, involves the nominalization of the original propositions, and noun phrases become increasingly difficult to disambiguate with each stage of the process (the "unrecoverability" problem; see Quirk et al., 1985). In the above example, compression culminates in the creation of a new single word, *enthalpy*. A single technical word such as this, then, can be expected to be more difficult than a phrase at an earlier stage of technicalization, because the burden of unpacking and unrecoverability is greater.

2. *Technicality as precision*

What is meant here is that when a word is used with its technical sense, it is used in a totally unambiguous, or precise, manner. For Ward, it is the complex relationship that words enter into that determines whether they are technical or not. This is illustrated with the definition of *enthalpy*, which can only be understood if the terms *free energy* and *heat content* (themselves interrelated) are known, as well as the technical concept of thermodynamics. This property of technical words can also be seen to apply to basic, everyday words. A word like *gas*, for instance, is technical in many disciplines because of its precise definition relating it to other terms. The technical definition of *gas*, "a species at above boiling point", assumes an accurate understanding of the concepts *species* and *boiling point*.

DEALING WITH THE DIFFERENT TYPES OF SPECIALIZED VOCABULARY IN THE CLASSROOM

Having identified a number of factors which might influence the learning of technical words, let us now look at how these potential sources of difficulty are likely to manifest themselves in terms of the distinct categories of specialized vocabulary that have been established: fully technical, lay-technical, and cryptotechnical words. The observations we make will be followed by some suggestions for dealing with the different types of words in the classroom. Particular reference will be made to the context of the Japanese university ESAP classroom, but many of my recommendations will apply equally to a variety of learning situations.

It has often been stated that it is not the job of the language teacher to teach technical words (see, e.g., Higgins, 1966; Cowan, 1974; and Hutchinson and Waters, 1987), and that in any

case these words should not actually be difficult for ESP learners. However, it should be clear from the above discussion that words can be difficult or easy along different dimensions: They can, for instance, be semantically easy but difficult in form (phonologically, orthographically, or in grammar). The teacher needs to be aware of these different aspects of difficulty, and be prepared to give the appropriate help when necessary.

Fully Technical Words

Fully technical pharmacology words are those which are used almost exclusively in pharmacological or related medical contexts. Table 2 shows all of the fully technical words contained in the Essential Pharmacology Word List, together with their frequencies of occurrence. Two immediate observations can be made: 1) There are not actually that many fully technical words (only 71 in a 570-word list, or 12.5%); and 2) even a non-specialist will recognize and be able to pronounce a good number of them.

Table 2. Fully Technical Words in the EPWL

Word family	Frequency	Word family	Frequency
1. RECEPTOR	1303	37. RENAL	112
2. AGONIST	353	38. NUCLEUS	110
3. NEURON	348	39. DNA	108
4. ANTAGONIST	327	40. AORTIC	104
5. PLASMA	297	41. ANAESTHETIC	103
6. OXIDATION	247	42. HEPATIC	100
7. ENZYME	238	43. DEPOLARIZATION	97
8. ASSAY	224	44. mRNA (MESSENGER RNA)	95
9. MEMBRANE	210	45. ION	85
10. SYNAPTIC	199	46. SIGMA	82
11. CARDIOVASCULAR	198	47. CEREBRAL	80
12. KINETIC	194	48. AMINO	76
13. PHARMACOLOGICAL	194	49. LIPID	74
14. PEPTIDE	192	50. IC50	65
15. CL (CHLORIDE)	191	51. PHOSPHORYLATION	58
16. IN VIVO	182	52. CATALYTIC	56
17. LIGAND	181	53. SUBTYPE	53
18. PH	173	54. INTRAVENOUS	52
19. MOLECULAR	170	55. PATHOLOGY	49
20. OXIDE	160	56. CO (CARBON MONOXIDE)	48
21. SERUM	156	57. ADRENERGIC	47
22. VASCULAR	152	58. GASTROINTESTINAL	46
23. ENDOTHELIUM	151	59. NITRIC	46
24. IN VITRO	151	60. SUBCUTANEOUS	45
25. ATP	149	61. WILD-TYPE	45
26. CYTOKINE	147	62. KCL	43
27. ACH (ACETYLCHOLINE)	145	63. NACL	43
28. MACROPHAGE	140	64. EPITHELIAL	42
29. HYPERTENSION	131	65. HOMEOSTASIS	42
30. PHYSIOLOGICAL	129	66. CACL2	36
31. MOLAR	125	67. INTRAPERITONEAL	33
32. SODIUM	121	68. BIOCHEMICAL	31
33. SUBSTRATE	120	69. CENTRIFUGE	27
34. ANTIBODY	116	70. ALIQUOT	25
35. SUBUNIT	114	71. PHENOTYPE	25
36. KINASE	113		

When we look at the words in the list, we see that, for reasons discussed earlier, the orthography or pronunciation of some of them (e.g., *acetylcholine, phosphorylation, intraperitoneal*, and *subcutaneous*) may be problematic, and this is especially true in the case of Japanese learners. There are also words, though, which should be relatively easy as far as spelling and pronunciation are concerned (*assay, subtype*, and *ATP*, for instance). However, as we have already noted, when considering difficulty we have to take into account the fact that a number of Japanese medical words have been borrowed from other languages: this means that the English word will often be easily recognized, but it can either help or hinder the learner with regard to pronunciation and spelling.

Table 3 gives some examples of medical/pharmacological words, found in the Pharmacology Word List, whose equivalents in Japanese have been borrowed from other languages.

Table 3. Fully Technical Words with Similar Japanese Equivalents

English	Japanese
neuron	ニューロン（NYURON）
synapse	シナプス（SHINAPUSU）
acetylcholine	アセチルコリン（ASECHIRUKORIN）
sodium	ナトリウム（NATORIUM）
sigma	シグマ（SHIGUMA）
amino	アミノ（AMINO）
peptide	ペプチド（PEPUCHIDO）
vivo	ヴィヴォ（VIVO）
vitro	ビトレ（BITORÉ）
adenosine	アデノシン（ADENOSHIN）
glutamate	グルタミン酸（GURUTAMINSAN）
kinase	キナーゼ（KINAZÉ）
renin	レニン（RENIN）

The existence of these cognate words, which share the same Latin or Greek origin with English, is certainly helpful, and some of them are pronounced very similarly in both English and Japanese. The fact that many of them have entered the Japanese language via other languages (primarily German), though, means that the English and Japanese pronunciations can often be very different. The Japanese pronunciation of *kinase*, for example, would probably be unrecognizable to a native English speaker. Adding to the problem is that, at least from my experience, it is often far from easy to convince students that the borrowed word may not in fact be the same in English.

The most conspicuously technical terms are perhaps not actually as difficult as we might expect, and we can make it easier for learners to acquire them by training them in strategies that will help them understand and remember the words. The teacher can, for instance, take advantage of the fact that a large number of technical words are of Greek or Latin origin. Learners should be encouraged to analyze such words wherever possible and relate the meanings of the word parts to the meaning of the word. Clearly, if learners can recognize the meaning of a

prefix or suffix, they will have a better chance of guessing the word's definition more accurately.

Table 4 shows some of the most commonly occurring prefixes, suffixes, and stems that make up technical pharmacology words. While this list is not exhaustive, if learners are familiar with these affixes they will gain access to many technical words, both in pharmacology and the wider field of medicine. Those words will include, importantly, a large number which are not in the lists of most frequent words. The fact that many unlisted words can be guessed from their etymology helps to address the criticism that the Pharmacology Word List falls short of providing 95% coverage of pharmacology corpora.

Table 4. Useful Affixes for the Learning of Pharmacology Words

a-, an-	-gram	-ology
amino-	-graph	-oma
angio-	haem-	-osis
anti-	hepat-	osteo-
-ase	hetero-	para-
-ate	histo-	patho-
auto-	homeo-	-pathy
bio-	homo-	pharmaco-
bi-	hyper-	photo-
brady-	hypo-	physio-
broncho-	immuno-	post-
cardio-	inter-	poly-
cerebro-	intra-	pre-
co-	iso-	pseudo-
counter-	-itis	psycho-
-crine	lipo-	re-
cross-	macro-	-scope
-cyte	micro-	-stasis
cyto-	mono-	self-
derm(o)-	multi-	sub-
-emia	myo-	super-
endo-	neo-	supra-
-ergic	neuro-	tachy-
exo-	nitro-	trans-
gastro-	non-	-tropic
-genic	-oid	

By way of example, two commonly occurring affixes in pharmacology are *exo-* and *cyto-*; with an awareness that *exo-* means "outside", or "external to", and *cyto-* means "cell", it should not be difficult to guess that *exocytotis* refers to the process of cellular secretion or excretion. Similarly, if learners know that the meaning of the widely used prefix *hyper-* is "extreme, or beyond normal", they will more easily understand and remember the meaning of a great many important words like *hypersensitivity* (abnormal susceptibility to a drug) or *hyperaemia* (an excess of blood in an organ). Another important prefix is *poly-* (more than one, many), and knowledge of this will make the meaning of numerous words transparent (e.g., *polydrug, polycyclic, polycationic*). The more learners are conscious of the way that words are made up, the more they will be able

to apply this knowledge to the understanding of new words. See Appendices 1 and 2 for sample glossaries of the most common affixes.

A final point to make about technical words is that, of course, they do not occur in isolation, and they behave in very specific ways in specialized texts. Although we are primarily concerned here with the ways in which the innate properties of technical words affect their learning, it is worth mentioning that the text itself provides a number of clues that can help learners identify technical words and infer their meanings and behavior. For example, the words may be defined in the text, they may be written in bold or italics, or they may appear as a label in a diagram. Bramki and Williams (1984) and Chung and Nation (2003) give some further examples of the textual clues that learners need to be familiar with.

Lay-technical Words

We would not expect words such as *drug, liver, cell, heart*, or *patient* to present many problems for pharmacology students. However, there may be more to these "everyday" words than is at first apparent: when lay-technical words are used in specialized texts, it is often the way in which they combine and interact with other words that determines their technicality. It would, therefore, make sense to deal with a lay-technical word as part of a highly specialized multi-word unit.

To take an example, *blood* is most often found as part of a two or three-word multi-word unit: *blood count, blood pressure, blood glucose, blood sample, blood flow, blood clot, blood vessel, blood brain barrier*, and *cell blood count* are just some of the most frequently occurring units.

Another word, *cell*, frequently occurs as part of the following units: *endothelial cell, mast cell, muscle cell, T cell, red cells, white cells, vascular cell, inflammatory cell, endothelial cell, receptor cell*; also, *intracellular* and *extracellular* are common forms.

A further instance of a lay-technical word that behaves in this way is *heart*, which is commonly found in these combinations: *heart disease, heart failure, hypertensive heart disease,* and *heart rate*.

When lay-technical words are used in this highly specific way they are just as specialized, and as fundamental to the discipline, as fully technical words. However, the fact that these words will be familiar, even to low-level learners (and teachers without much knowledge of the field), means that they can be much more easily dealt with in the classroom. These are words, suggests Ward (2007: 25), that are at the threshold of technicality. In a sense, they provide a gateway into the murky world of incomprehensible technical terms: they can be defined, or unpacked into propositions, using language which does not require a high level of specialist knowledge. This attribute is clearly shown in the following definitions of *drug* and *blood* (taken from the Chambers Dictionary of Science and Technology). While these definitions are more obviously technical than those a layperson might offer, they are, nevertheless, relatively easy to interpret:

Drug: Any substance, natural or synthetic, which has a physiological action on a living body, either when used for the treatment of disease or the alleviation of pain or for recreation and self indulgence, leading in some cases to progressive addiction.

Blood: A fluid circulating through the tissues of the body, performing the functions of transporting oxygen, nutrients and hormones, and carrying waste products to the organs of excretion.

Cryptotechnical Words

It is this category of polysemous and homonymous words that I believe will provide learners with the greatest difficulty. The fundamental problem with these words, of course, is meaning-related: there is nothing about them that makes them intrinsically difficult, and it is the fact that learners (and ESAP teachers) may erroneously think they know them that is a source of concern. All learners will immediately know that they do not understand an overtly technical word like *cytokine* or *intraperitoneal*, but they may well pass over a commonplace word such as *medium* without realizing that they do not properly understand its meaning in pharmacology. Exacerbating the problem is the finding that cryptotechnical and lay-technical words are used at least as often as strictly technical words in pharmacology texts. It is, therefore, of critical importance to ensure that learners are aware of these words.

Of course, lists and glossaries of the most important items can be provided (for a sample glossary, see Appendix 3), but they will not cover all words that are likely to be problematic. Learners need to be trained to detect whether they do or do not understand how cryptotechnical words are being used in specialized texts, and that they are not being deceived by the common meanings of these words.

Academic Words

There are many non-technical, or perhaps we should say sub-technical, words and phrases such as *finding, result*, and *previous studies*, all of which play a vital part in structuring the text of academic articles. Some expressions, such as *in the absence of* or *in the presence of*, appear to be peculiar to pharmacology, where they have an important role in descriptions of experimental work. Most of these words are found in the GSL or AWL, and as with cryptotechnical words, learners should be taught to recognize how and when these apparently ordinary words are being used with a specialized function.

CONCLUSION

It is hoped that this paper has given some insight into the challenges that specialized lexis (and medical vocabulary in particular) poses to both teachers and learners of ESAP. The high proportion of technical words in pharmacology research articles obviously contributes to the difficulty of coping with these texts, but the causes of difficulties are various and not always obvious. Laufer's intralexical factors have proved helpful in the investigations, and it has been

shown that the learners' L1 cannot be ignored in any discussions on word difficulty. The different types of specialized words pose their own particular problems: Polysemous cryptotechnical words may be the most problematic due to the fact that they do not *appear* to be difficult; "familiar" lay-technical words may also not be as easy as we expect, with an important source of difficulty being the way in which collocation determines their technicality.

The claims that have been made made, however, are as yet untested, and the next logical step would be to test the words on learners. Carrying out case studies, for example, would make it possible to observe how learners deal with and process the different categories of vocabulary, and to note the actual, rather than imagined, problems that they encounter. It could, for instance, be established whether synforms or words with deceptive transparency really are as problematic as has been suggested for pharmacology learners. Analysis of learner corpus data would also be useful as a means of investigating the kinds of problems there might be in production; it could tell us which items are misused, or used idiosyncratically, for example. The different pharmacology word lists, too, need to be tested with students in order to determine just how well they help them improve their comprehension of research articles. We may find that we are overestimating the usefulness of the lists if we discover that learners are not sufficiently familiar with English morphology to benefit from them.

Perhaps most importantly, the concept of cryptotechnicality needs to be further validated, and it would be useful to test students to determine empirically their actual knowledge of cryptotechnical words. The extent to which the supposedly hidden technical senses of these words actually are inaccessible should be investigated, and determining their salience for learners in a pharmacological context will involve an understanding of which meanings are perceived to be dominant, and of how learners rate the semantic relatedness of the core and peripheral senses of a word.

REFERENCES

Bogaards, P. (2001). Lexical units and the learning of foreign language vocabulary. *Studies in Second Language Acquisition 23*, 321—343.

Bramki, D. and Williams, R. (1984). Lexical familiarization in economics text, and its pedagogic implications in reading comprehension. *Reading in a Foreign Language 2* (1), 169—181.

Chung, T. M. and Nation, I. S. P. (2003). Technical vocabulary in specialized texts. *Reading in a Foreign Language 15* (2), 102—116.

Cowan, J. R. (1974). Lexical and syntactic research for the design of reading material. *TESOL Quarterly 8* (4), 389—399.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly 34* (2), 213—238.

Fraser, S. (2005). The lexical characteristics of specialized texts. In K. Bradford-Watts, C. Ikeguchi, & M. Swanson (Eds.) *JALT 2004 Conference Proceedings*. Tokyo: JALT.

Fraser, S. (2006). The nature and role of specialized vocabulary: What do ESP teachers and learners need to know? *Hiroshima Studies in Language and Language Education 9*, 63—75.

Fraser, S. (2007). Providing ESP learners with the vocabulary they need: Corpora and the

creation of specialized word lists. *Hiroshima Studies in Language and Language Education 10*, 127—143.

Fraser, S. (2009). Breaking down the divisions between general, academic, and technical vocabulary: The establishment of a single, discipline-based word list for ESP learners. *Hiroshima Studies in Language and Language Education 12*, 151—167.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. London: Edward Arnold.

Henning, G. H. (1973). Remembering foreign language vocabulary: Acoustic and semantic parameters. *Language Learning 23* (2), 185—196).

Higgins, J. J. (1966). Hard facts. *ELT Journal 21*, 55—60.

Hu, M., and Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language 13* (1), 403—430.

Huchinson, T. and Waters, A. (1987). *English for specific purposes*. *Cambridge*: Cambridge University Press.

Lado, R. (1955). Patterns of difficulty in vocabulary. *Language Learning 6* (1), 23—41.

Laufer, B. (1981). A problem in vocabulary learning — synophones. *ELT Journal 35* (3), 294—300.

Laufer, B. (1988). The concept of 'synforms' (similar lexical forms) in L2 learning. *Language and Education 2,* 113—32.

Laufer, B. (1989a). What percentage of text-lexis is necessary for comprehension? In C. Lauren and M. Nordman (Eds.), *Special language: From humans thinking to thinking machines*, 316—323. Clevedon: Multilingual Matters.

Laufer, B. (1989b). A factor of difficulty in vocabulary learning: Deceptive transparency. *AILA Review 6*, 10—20.

Laufer, B. (1997). What's in a word that makes it hard or easy: Some intralexical factors that affect the learning of words. In: N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy*, 40—63. Cambridge: Cambridge University Press.

Meara, P. M. (1982). Word association in a foreign language. A report on the Birkbeck vocabulary project. *Nottingham Linguistic Circular 11* (2), 29—38.

Meara, P. M. and Bell, H. (2001). P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect 16* (3), 5—19.

Pueyo, I. G., and Val, S. (1996). The construction of technicality in the field of plastics: A functional approach towards teaching technical terminology. *English for Specific Purposes 15* (4), 251—278.

Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1985). *A Comprehensive grammar of the English language*. London, Longman.

Ryan, A. (1997). Learning the orthographical form of L2 vocabulary — a receptive and a productive process. In N. Schmitt and M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy*, 181—198. Cambridge: Cambridge University Press.

Schmitt, N., Jiang, X., and Grabe, W. (2011). The relationship between the amount of vocabulary known in a text and reading comprehension. *Modern Language Journal 95* (1), 26—43

Strevens, P. (1973). Technical, technological, and scientific English. *ELT Journal 27*, 223—234.

Swan (1997). The influence of mother tongue on second language vocabulary acquisition and use.

In N. Schmitt and M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy,* 156—180. Cambridge: Cambridge University Press.

Wang, J., Liang, S-l, and Ge, G-c. (2008). Establishment of a medical academic word list. *English for Specific Purposes 27* (4), 442—458.

Ward, J. (2007). Collocation and technicality in EAP engineering. *Journal of English for Academic Purposes 6*, 18—35.

West, M. (1953). *A general service list of English words.* London: Longman, Green & Co.

APPENDIX 1

Sample Glossary of Common Prefixes in Pharmacology

Prefix	Meaning	Prefix	Meaning
a-, an-	not, without	*macro-*	large, long
angio-	blood vessel	*micro-*	small
auto	self	*mono-*	one, single
bi-	both, double	*multi-*	many, multiple
brady-	slow	*myo-*	relating to muscle
bronchi-, broncho-	bronchus, bronchial	*neo-*	new
cardio-	pertaining to the heart	*neuro-*	nerves, nervous system
cerebro-	pertaining to the brain	*nitro-*	nitrogen, NO2
chemo-	chemical, chemicals	*non-*	not existing
co-	together, with	*osteo-*	bone
counter-	contrary, opposite	*para-*	beside, abnormal
cross-	across	*patho-*	suffering, disease
cyto-	cell	*pharmaco-*	indicating drugs
derm(o)-	skin	*photo-*	pertaining to light
endo-	inside, within	*physio-*	physical, natural
exo-	outside, external	*post-*	after, behind
gastro-	pertaining to the stomach	*poly-*	many, plurality
haem-	pertaining to blood	*pre-*	before (in position or time)
hepat-	pertaining to the liver	*pro-*	forward, before
hist(o)-	animal tissue	*pseudo-*	false, fake
homeo,	similar, like	*psycho-*	pertaining to the mind
homo	the same, common	*re-*	again, backward
hydro-	water	*self-*	oneself, itself
hyper-	excessive, beyond normal	*sub-*	below, beneath
hypo-	deficient, below normal	*super-*	in excess, above, superior
inter-	between, among	*supra-*	over, above, excessive
intra-	inside, within	*tachy-*	rapid, irregularly fast
lipo-	fat	*trans-*	across, through

APPENDIX 2

Sample Glossary of Common Suffixes in Pharmacology

Suffix	Meaning	Suffix	Meaning
-ase	enzyme	*-oid*	resemblance to
-ate	containing oxygen	*-ology*	study of
-crine	secretion	*-olysis*	breakdown
-cyte	cell	*-oma*	tumour, cancer
-emia	blood condition	*-osis*	disease, condition
-genic	origin, formative	*-pathy*	disease, disorder
-gram	record, picture	*-scope*	picture, inspection
-graph	record, picture	*-stasis*	stoppage
-itis	inflammation	*-tropic*	direction

APPENDIX 3

Sample Entries from a Glossary of Cryptotechnical Words

ACTIVITY

General

(noun) a situation in which many things are happening at the same time

Technical

1. (noun) the situation in which a biological agent is exerting an effect on something and causing change

Examples: *analgesic activity, inhibitory activity, [drug] induced activity*

2. (noun) the property or behaviour of a body system or organ

Examples: *brain activity, locomotor activity, neural activity*

BLOCK

General

(verb) to put an obstacle in the way of something

Technical

(verb) to prevent the action of a drug

Also *blocker* (noun); *blockade* (noun)

Examples: *blocked by [the antagonist], blockade of [receptors], channel blocker(s), beta blocker(s)*

CONTROL

General

1. (noun) some kind of regulation or check
2. (verb) to check; to regulate

Technical

Statistical: (noun) an experiment performed to afford a standard of comparison for other experiments

Examples: *control group, control conditions, control mice, control experiments*

DELIVERY

General

(noun) the transportation of goods

Technical

(noun) administration or transfer of a drug

Examples: *delivery of [drug], insulin delivery, synaptic delivery, delivery into synapses*

EXPRESSION

General

(noun) the process of making known one's thoughts or feelings

Technical

(noun) the detectable effect of certain protein molecules

Examples: *gene expression, protein expression, expression of [the protein]*

Also express (verb) to give a concise explanation of data

Examples: *expressed as a percentage, expressed as geometric means*

REACTION

General

(noun) a person's reaction to something is what he feels, says or does because of it

Technical

(noun) any change in the behaviour of an organism in response to a stimulus

Examples: *adverse drug reactions, reaction time, inflammatory reaction*

SAMPLE

General

(noun) a small portion of something to show the quality of the whole

Technical

1. (noun) a specimen taken for medical analysis

Examples: *blood sample, urine sample, saliva samples*

2. Statistical: (noun) a small but representative part of a population, used in a test or survey

Examples: *small sample size, control samples, a sample of general practices*

TREATMENT

General

(noun) the manner in which a person behaves towards or deals with someone or something

Technical

1. (noun) a session of medical care of a patient, or the administration of a drug

Examples: *treatment of [condition], treatment with [drug], course of treatment*

2. Scientific/statistical: (noun) referring to the factors or variables controlled by the researcher in an experiment

Also *treat* (verb)

Examples: *untreated controls, control and treated groups, treatment groups*

要　約

専門的語彙の習得における影響要因
―コーパスから分かること―

サイモン・フレイザー
広島大学外国語教育研究センター

薬理学等の医療分野における難解な語彙の習得は，特定学術目的の英語（ESAP）学習者にとって最大のチャレンジである。しかし具体的には何がこの分野の英語をそれほど難解にしているのであろうか。形態的複雑性が語彙学習を困難にしているのは確かであるが，それ以外にも考慮されるべき，さほど明白ではないいくつかの要因が存在する。

本研究では，これらの要因に着目し，薬理学のテキストの語彙の特徴について考察した。ここでは，先行研究ですでに作成した頻度カウントを利用し，薬理学の論文コーパスの中で最も有用だと思われる単語から構築した単語リストを利用した。

本稿においては，まず用語の新しい分類とワードリストの作成に関して簡単な説明を行い，次にリスト上の単語に着目し，学習者が専門用語を習得するに当たり，考えられる問題点について考察している。そして，“synformy”や“deceptive transparency”（Laufer, 1997）といった語彙内発生型要因と，“cryptotechnicality”（Fraser, 2006），“technicality as compression”（Ward, 2007）のような概念が，専門用語を見ていく上で有用である事を指摘している。また，学習者のL1（この場合は日本語）に起因する語彙習得における影響要因についても示唆を与える。最後に，薬理学とその関連分野における教材用ワードリストの構築や，教材開発において有用であると思われる点について言及する。