

2 *The lexical plight in second language reading*

Words you don't know, words you think you know, and words you can't guess

Batia Laufer

Introduction

No text comprehension is possible, either in one's native language or in a foreign language, without understanding the text's vocabulary. This is not to say that reading comprehension and vocabulary comprehension are the same, or that reading quality is determined by vocabulary alone. Reading comprehension (both in L1 and in L2) is also affected by textually relevant background knowledge and the application of general reading strategies, such as predicting the content of the text, guessing unknown words in context, making inferences, recognizing the type of text and text structure, and grasping the main idea of the paragraph. And yet, it has been consistently demonstrated that reading comprehension is strongly related to vocabulary knowledge, more strongly than to the other components of reading. Anderson and Freebody (1981) survey various studies (correlational, factor analyses, readability analyses) that show that the word variable is more highly predictive of comprehension than the sentence variable, the inferencing ability, and the ability to grasp main ideas. Beck, Perfetti, and McKeown (1982), Kameenui, Carnine, and Freschi (1982), and Stahl (1983) have demonstrated that an improvement in reading comprehension can be attributed to an increase in vocabulary knowledge.

A similar picture of vocabulary as a good predictor of reading success emerges from second language studies. Laufer (1991a) found good and significant correlations between two different vocabulary tests (the *Vocabulary Levels Test* by Nation [1983a] and the *Eurocentres Vocabulary Test* by Meara and Jones [1989]) and reading scores of L2 learners. The correlations were .5, significant at the level of $p < .0001$, and .75, significant at the level of $p < .0001$, respectively. Even higher correlations are reported by Koda (1989) between vocabulary (tested by a self-made test) and two reading measures, cloze and paragraph comprehension. These correlations are .69, $p < .0002$ and .74, $p < .0001$. Coady, Magoto, Hubbard, Graney, and Mokhtari (1993) conducted two experiments that showed that increased proficiency in high-frequency vocabulary also led

to an increase in reading proficiency. Vocabulary materials were so successful that in the second experiment, no control group, which was to be taught without the materials, could be set up, since all the students wanted to use them.

Syntactic complexity, on the other hand, was found not to affect the level of reading comprehension. Ulijn and Strother (1990) claim that 'while a complete conceptual and lexical analysis may be necessary for reading comprehension, a thorough syntactic analysis is not' (p. 38). As for the effect of general reading strategies on L2 reading, the current view is that since reading in L2 is both a reading problem and a language problem, 'some sort of threshold or competence ceiling has to be attained before existing abilities in the first language can begin to transfer' (Alderson, 1984, p. 20). In other words, even if a reader has good metacognitive strategies, which he or she uses in L1, these will not be of much help in L2 before a solid language base has been reached. This conclusion is borne out by empirical evidence. Perkins, Brutton, and Pohlmann (1989) tested Japanese learners of English on various components of reading in L1 and L2 and then related the correlations (between L1 reading and L2 reading scores) to the TOEFL scores of the learners. The authors suggest, albeit cautiously, that the threshold of L2 competence is reflected in the score interval of 375–429 on the TOEFL test. In other words, this interval indicates the lowest level at which reading in L2 can be expected to start resembling reading in L1.

Since language threshold is essential for comprehension, an important question is, therefore, what the nature of this threshold is. According to Ulijn (1984), Ostyn and Godin (1985), Ostyn, Vandecasteele, Deville, and Kelly (1987), and Ulijn and Strother (1990), indispensable conditions for reading in L2 are understanding the text's words and the knowledge of its subject matter. Laufer and Sim (1985a, 1985b) found that, in interpreting texts, students tend to regard words as main landmarks of meaning. Background knowledge is relied on to a lesser extent, and syntax is almost disregarded. Haynes and Baker (1993) too came to the conclusion that the most significant handicap for L2 readers is not lack of reading strategies but insufficient vocabulary in English. What these studies indicate is that the threshold for reading comprehension is, to a large extent, lexical. Lexical problems will, therefore, hinder successful comprehension. Intuitively, most people associate lexical problems in comprehension with nonfamiliarity with words in the text. This is certainly a major obstacle, which I will discuss and refer to as 'words you don't know'. However, in addition to the phenomenon of noncomprehension, the learner will often experience miscomprehension of certain words that look familiar even though they are unknown. Since meaning is, to some extent, in the eye of the beholder, the problem of 'pseudofamiliar' words will be discussed as well. It will be referred to as 'words you think you

be $3,000 \times 1.6 = 4,800$ (for the conversion formula, see Nation [1983a]). The level at which good L1 readers can be expected to transfer their reading strategies to L2 is 3,000 word families, or about 5,000 lexical items. Until they have reached this level, such transfer will be hampered by an insufficient knowledge of vocabulary. A regression analysis was carried out to check how an increase in vocabulary quantity related to the improvement in reading. The analysis showed that an increase in 1,000 words resulted in an increase of 7% on a comprehension test. Thus, if we tried to predict the relationship between vocabulary increase and comprehension increase, we might expect a knowledge of 3,000 word families (5,000 lexical items) to result in a reading score of 56% (the minimum passing grade on reading tests in the students' institution at the time of the experiment), a knowledge of 4,000 (6,400 lexical items) in 63%, 5,000 (8,000 lexical items) in 70%, 6,000 (9,600 lexical items) in 77%, and so on. These figures are correct if the progress in reading *vis-à-vis* vocabulary size is always linear. It is possible, however, that when the learner reaches a certain vocabulary level, progress in the reading score will decrease and finally level off.

Even if the results are not conclusive for all vocabulary levels, they provide, nevertheless, a general idea of how reading progresses above the threshold level of 3,000 word families and what vocabulary size should be aimed at for different reading levels. If the optimal reading score is considered to be, for example, 70%, then the vocabulary size to aim for will be 5,000 word families; if 63% is taken as a passing score, then 4,000 will suffice.

Further support for 3,000 word families as the lexical threshold is found in Laufer (1992). Adult learners were compared on their vocabulary level, reading comprehension in EFL, and general academic ability, which included a reading in L1 component. (The General Academic Ability score and the EFL reading score were taken from the university psychometric entrance test.) It was shown that learners below the 3,000 vocabulary level (5,000 lexical items) did poorly on the reading test regardless of how high their academic ability was. In other words, even the more intelligent students who are good readers in their native language cannot read well in their L2 if their vocabulary is below the threshold.

In terms of text coverage, i.e., the actual percentage of word tokens in a given text understood by a reader, the 3,000 word families are reported to provide a coverage of between 90% and 95% of any text (for surveys of frequency counts, see Nation, 1983a and 1990). The figure claimed to be necessary for text comprehension is 95% of text coverage (Deville, 1985; Laufer, 1989b). Thus, both earlier frequency counts and later empirical studies of L2 vocabulary and reading suggest a similar vocabulary minimum, which is 3,000 word families, or 5,000 lexical items. As mentioned

before, the higher the comprehension level expected, the larger the vocabulary should be.

Words you think you know

Deceptive transparency

So far we have discussed one aspect of the lexical plight in L2 reading – words one does not know, or, more precisely, words the reader knows he or she does not know. In addition to these, there are other words that are not even recognized as unfamiliar by the reader. These words are deceptively transparent, i.e., they look as if they provided clues to their meaning. For example, *infallible* looks as if it were composed of *in-fall-ible* and meant 'something that cannot fall'; *shortcomings* looks like a compound of 'short' and 'comings', meaning 'short visits' (these are actual misinterpretations provided by students). Huckin and Bloch (1993) found similar types of misinterpretations. They refer to the deceptively transparent words as cases of 'mistaken ID'. The deceptively transparent words (hence DT words) seem to fall into one of five distinct categories.

Words with a deceptive morphological structure

These are words that look as if they were composed of meaningful morphemes. Thus, *outline* was misinterpreted as 'out of the line', *nevertheless* as 'never less', *discourse* as 'without direction'. The learner's assumption here was that the meaning of a word equalled the sum of meanings of its components. This assumption is correct in the case of genuinely transparent words, but not when the 'components' are not real morphemes.

Idioms

Hit and miss, *sit on the fence*, *a shot in the dark*, *miss the boat* were translated literally, word by word. The learner's assumption in the case of idioms was similar to that of deceptive transparency, i.e., the meaning of the whole was the sum of the meanings of its parts.

False friends

Sympathetic was interpreted as 'nice' (Hebrew 'simpati'); *tramp* as 'lift' (Hebrew 'trem'); *novel* as 'short story' (Hebrew 'novela'). The mistaken assumption of the learner in this case was that if the form of the word in L2 resembled that in L1, the meaning did so too.

Words with multiple meanings

It often happens that students know one meaning of a polyseme or a homonym and are reluctant to abandon it even when, in a particular context, its meaning is different. For example, *since* was interpreted as 'from the time when' though it meant 'because'; *abstract* as 'not concrete' instead of 'summary'; *state* as 'country' instead of 'situation'. The mistaken assumption of the learner in this case was that the familiar meaning was the *only* meaning.

'Synforms' (similar lexical forms)

The largest category of DT words is that of 'synforms', pairs/groups of words that are similar in form. (For criteria of synform similarity, classification of synforms, and discussion of the problems they raise, see Laufer, 1988 and 1991b.) Generally speaking, some synforms are similar in sound (cute/acute, available/valuable, conceal/cancel, price/prize); some are morphologically similar (economic/economical, industrious/industrial, reduce/deduct/induce).

Synformic confusions may have two sources: The learner might have learnt one word of the pair/group, but since its representation in the memory is insecure or defective, a similar word that shares most of its formal features might look identical to it. Or, the learner might have studied both synforms but since the knowledge of both is insecure, he or she is not sure which word form is associated with which meaning. Whatever the reason, the result is to misinterpret one synform as its counterpart.

Deceptive transparency and reading comprehension

Laufer (1989a) studied students' interpretations of various DT words in text context. The results were as follows: (a) errors of word understanding were more frequent with DT words than with non-DT ones; (b) students were less aware of their ignorance with DT words than with non-DT ones; (c) there was a significant correlation between reading comprehension and learners' awareness of unknown DT words. Correlations do not necessarily show cause-effect relationships between the variables. Nevertheless, I will try to argue for such a relationship between reading and awareness of DT words.

When a foreign learner does not understand a word in a text, he or she has the following options: ignore it (if it is considered unimportant), look it up in a dictionary, ask someone who knows its meaning, or try to guess it from context. But an attempt to guess (regardless of whether it is successful or not) presupposes awareness on the part of learners that they

are facing an unknown word. If such awareness is not there, no attempt is made to infer the missing meaning. This is precisely the case with deceptively transparent words. (See also Huckin and Bloch, 1993.) Learners think they know them and assign the wrong meaning to them, distorting the immediate context in this way. But this may not be the end of the distortion process. The misinterpreted words will sometimes serve as clues for guessing words that the learner recognizes as unknown, which may lead to larger distortions. Graphically, the process can be represented in the following manner: unawareness of ignorance of DT words → misinterpretation of DT words → distortion of immediate context → using distorted context for further interpretation → distortion of larger context.

Here is an example of a distorted sentence resulting from misinterpretation of three words. The original sentence was: 'This nurturing behaviour, this fending for females instead of leaving them to fend for themselves, may take many different forms'. *Nurturing* was confused with 'natural', *fending* with 'finding', *leaving* with 'living'. The result was the following: 'Instead of living natural life, natural behaviour, females and children find many different forms of life'. For more examples of similar misinterpretation, see Laufer and Sim (1985b).

The suggested cause-effect relationship between awareness of DT words and reading comprehension can therefore be explained as follows: A better awareness of DT words is necessary for attempting to find their meaning. Such an attempt will result in a larger number of correctly interpreted words. These will in turn reduce the density of unknown words. Such reduction will result in an increase in contextual clues that are necessary for understanding additional new words. This understanding will increase the total number of correctly interpreted words. A larger number of known words will be an asset to global comprehension of the text.

Words you can't guess

The third lexical problem in L2 reading can occur when the reader is trying to guess unknown words, more precisely words that have been recognized as unfamiliar (as opposed to DT words, which are not recognized as such). The importance of guessing has been widely discussed (Liu Na & Nation, 1985; Nation & Coady, 1988; Van Parreren & Schouten-van Parreren, 1981, and many others). I do not contest the value of the guessing activity as such or the fact that it takes place. What I find hard to accept is taking for granted that guessing in L2 is indeed possible with most unknown words and that successful guessing depends mainly on the learner's guessing strategies. This seems to be a naive belief

since a variety of factors will interfere with the guessing attempts of the reader.

Nonexistent contextual clues

Looking for contextual clues for the unknown word will not help if the clues are not there to be exploited. One cannot as a rule rely on contextual redundancy since there is no guarantee that a given context is redundant enough to provide clues to precisely those words that are unknown to the reader. Kelly (1990) collected a random sample of unknown words from two Italian books and tried to guess their meanings from context. His lack of success, together with a similar experience with students, led him to believe that 'unless the context is very constrained, which is a relatively rare occurrence, or unless there is a relationship with a known word identifiable on the basis of form and supported by context, there is little chance of guessing the correct meaning' (p. 203). Bensoussan and Laufer (1984) found that in a fairly standard passage of academic prose, out of seventy words that the learners were asked to guess, clear contextual clues could be found only for thirteen words. This may not be the case with other texts. Some may provide more clues, others even less. But the basic assumption, that a text will provide clues to unknown words, is overoptimistic.

Unusable contextual clues

The word guessability problem can occur in spite of the existence of clues. If the clues to the unknown word happen to be in words that are themselves unfamiliar to the reader, then, as far as the reader is concerned, the clues do not exist since they simply cannot be used. Let us look at the example mentioned previously: 'This nurturing behaviour, this fending for females instead of leaving them to fend for themselves, may take many different forms'. If 'nurture' is unknown and we are looking for a clue to its meaning, it is right there in the sentence - 'fending for someone', 'not leaving someone to fend for himself'. But what if 'fend' is also unknown? The learner may be lucky enough to recognize *nurture* and *fend* as unknown words (instead of confusing them with 'nature' and 'find', as described earlier), but since each of the words serves as a clue for the other, the two words will remain unguessable from context. Here is another example from our study: 'unless they accept, by specification or by implication, that the nature of man . . .' A reader trying to guess *implication* can notice that it is contrasted with *specification* and is therefore its opposite. This strategy will lead to a correct guess only if the meaning of *specification* is known. If it is not, the meaning of its opposite will remain a mystery.

It is precisely because clues that appear in unfamiliar words cannot be exploited that the density of unknown words in a text is of crucial importance to successful guessing. We do not have a definite answer as to the optimal ratio of unknown to known words in a text. I suggested earlier that a minimum of 95% text coverage, which is achieved by 3,000 word families, will lead to the transfer of reading strategies (including successful guessing). One should not forget that learners with such coverage received the minimum passing grade on the comprehension test. Hirsh and Nation (1992) discuss a different type of reading, which is easier and more fluent - reading for pleasure. Such reading, they argue, requires a higher lexical text coverage - a coverage of 98%. The remaining 2% can easily be guessed. To achieve a coverage of 98%, the learner needs to know about 5,000 word families, or 8,000 lexical items, according to Nation (1990). Whether one reads for pleasure (with 98% of coverage), or for bare necessity (with 95% of coverage), familiarity with a large number of words is a prerequisite for successful guessing.

Misleading and partial clues

We saw earlier that the deceptively transparent words tend to look familiar to the reader even though they make no sense in context. Familiar form makes it look right, thus overriding contextual clues (see also Haynes, 1984; Huckin and Bloch, 1993). But there are words recognized as unfamiliar by the reader and guessed incorrectly, or at best imprecisely, because the clues are misleading. Some of Kelly's (1990) Italian examples illustrate the point. The unknown words are in italics, the English translations in parentheses.

E dappertutto si beveva, si cantava, si ballava, si rissava. (And everywhere people were drinking, singing, dancing, brawling.) The unknown word was understood as 'enjoying themselves' instead of 'brawling' as the interpretation fitted the context of drinking, singing, dancing.

. . . scaldava le sue membra *intirizite* (he warmed his numb limbs). *Intirizite* was interpreted as 'frozen/cold' instead of 'numb' as it seemed sensible that someone would warm his 'cold' limbs. For more inadequacies of contextual clues, see Stein (1993).

Context may not necessarily be misleading. It can provide partial clues that will help the reader to arrive at a general word meaning. Here is an example from Clarke and Nation (1980). 'Typhoon Vera killed or injured 28 people and *crippled* the seaport city of Kellung'. *Crippled* can be understood as 'damaged, destroyed' since a typhoon will have some kind of negative influence upon a city. In some texts it may be sufficient to arrive at the general or approximate meaning of words in order to understand the passage. But how safe is it to rely on approximations in a legal or medical text? It is all too easy and tempting to say, 'The word looks

right in context, so it presumably means what I think it means'. What looks right may be wrong; and reliance on what is more or less right may sometimes produce an irresponsible interpretation.

Suppressed clues

One of the factors that contribute to successful guessing is the reader's background knowledge of the subject matter of the text, or content schemata. Inferences are drawn from the text on the basis of the reader's expectations of certain content. Since 'reading is a psycholinguistic guessing game' (Goodman, 1967), the successful players should, among other things, draw on their experiences and concepts. This strategy may work quite well, except when the reader's expectations and concepts are different from those of the author of the text. Readers tend to disregard information that, according to their world view, seems unimportant, add information that 'should' be there, and focus their attention on what, in their opinion, is essential (Steffensen & Joag-Dev, 1984). So strong is the effect of background knowledge that it overrides lexical and syntactic clues. In an experiment by Laufer and Sim (1985b), students were given a passage where the author (Margaret Mead) discussed biological differences between men and women, and clearly implied that boys and girls should get a different education. Some learners, however, insisted that the author was advocating the same education for both sexes. From interviews it became clear that the students were using their knowledge of the world, which was that nobody today would dare to suggest different education for men and women, certainly not a woman author. When a biased opinion of this kind is introduced into the interpretation, individual unknown words will be taken to mean whatever suits the reader's own notion of what the text says. If there are clues in context that would suggest a different interpretation, they can easily be suppressed.

Summary and conclusion

In this essay, I have discussed three lexical problems that may seriously impede reading comprehension in L2: (1) the problem of insufficient vocabulary, (2) misinterpretations of deceptively transparent words, and (3) inability to guess unknown words correctly. I will recapitulate my arguments starting with the third issue. A learner who has been taught guessing strategies will not automatically produce correct guesses. The following factors, which are beyond the reader's control, will affect guessing.

1. Availability of clues. It cannot be taken for granted that clues are present in the text and need only to be discovered by the reader. They may be there, or they may not be there.
2. Familiarity with the clue words. Mere availability of clues does not mean that they can be used by the reader. To be used, the clue-containing words have to be understood. A high density of unknown words will reduce the usability of clues as there will be a higher probability that words that explain each other will be unfamiliar.
3. Presence of misleading clues. Not everything that looks right in context is necessarily right. Yet the learner who has been taught that there is no need to understand the precise meaning of words may remain satisfied with whatever makes sense in the context, whether it is right or wrong.
4. Compatibility between the reader's schemata and the text content. If the two are different, the reader may impose his or her interpretation on the text and try to understand individual words in a way that will fit the global meaning, suppressing the clues that suggest a different interpretation.

One of the factors that overrides contextual clues is deceptive transparency of words. Words that look familiar will be interpreted to mean what the learner thinks they mean. These are words with a deceptive morphological structure, idioms, false friends, words with multiple meanings, and synforms. Since learners are unaware of their ignorance of the DT words, they will stick to the false meanings and may use them as clues to guessing other words. When this happens, both the immediate context of the DT word and the larger context are distorted. Readers with lower awareness of DT words also score lower on reading comprehension tests.

By far the greatest lexical obstacle to good reading is insufficient number of words in the learner's lexicon. Lexis was found to be the best predictor of success in reading, better than syntax or general reading ability. Whatever the effect of reading strategies is, it is short-circuited if the vocabulary is below the threshold, i.e., below the minimum of 3,000 word families, or 5,000 lexical items.

A large sight vocabulary is also an answer to the other two problems: deceptive transparency and guessing ability. The larger the vocabulary, the fewer the words that will look deceptively transparent to the learner and the fewer the DT errors that will be made. A large vocabulary will provide a good lexical coverage of a text (3,000 word families, or 5,000 lexical items cover about 95% of text). The higher the coverage, the lower the density of unknown words. When the density is low, there is a good chance of finding clues to the unknown words. There is also less cognitive capacity involved in lower-level processing. As more clues are

made available, the better guessing becomes. As less cognitive capacity is invested in decoding words, more of it is available for higher-level processing.

In this discussion of the lexical plight in L2 reading I have tried to show how indispensable good vocabulary knowledge is to reading. Reading may well be a psycholinguistic guessing game. But words are the toys you need to play it right.

References

- Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem? In J. C. Alderson & A. H. Urquhart. London: Longman (Eds.), *Reading in a foreign language* (pp. 1-27).
- Anderson, R. C., & Freebody, P. (1981). Vocabulary and knowledge. In J. T. Gutrie (Ed.), *Comprehension and teaching: Research review* (pp. 77-117). Newark, DE: International Reading Association.
- Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of text construction and instructional procedures for teaching word meanings on comprehension and recall. *Journal of Educational Psychology*, 74, 506-521.
- Bensoussan, M., & Laufer, B. (1984). Lexical guessing in context in EFL reading comprehension. *Journal of Research in Reading*, 7, 15-32.
- Clarke, D. F., & Nation, I. S. P. (1980). Guessing the meanings of words from context: Strategy and techniques. *System*, 8, 211-220.
- Coady, J., Magoto, J., Hubbard, P., Graney, J., & Mokhtari, K. (1993). High frequency vocabulary and reading proficiency in ESL readers. In T. Huckin, M. Haynes, & J. Coady (Eds.), (pp. 217-228). Norwood, NJ: Ablex.
- Deville, G. (1985). Measuring a FL learner's lexical needs. Paper presented at the fifth LSP Symposium, Louvain, Belgium.
- Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist*, 6, 126-135.
- Haynes, M. (1984). Patterns and perils of guessing in second language reading. In J. Handscombe, R. A. Orem, & B. P. Taylor (Eds.), *On TESOL 83* (pp. 163-176). Washington, DC: TESOL.
- Haynes, M., & Baker, I. (1993). American and Chinese readers learning from lexical familiarization in English texts. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary acquisition* (pp. 130-152). Norwood, NJ: Ablex.
- Hirsh, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a foreign language*, 8(2), 689-696.
- Huckin, T., & Bloch, J. (1993). Strategies for inferring word-meanings in context: A cognitive model. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary acquisition* (pp. 153-180). Norwood, NJ: Ablex.
- Kameenui, E. J., Carnine, D. W., & Freschi, R. (1982). Effects of text construction and instructional procedures for teaching word meanings on comprehension and recall. *Reading Research Quarterly*, 17, 367-388.
- Kelly, P. (1990). Guessing: No substitute for systematic learning of lexis. *System*, 18, 199-207.

Koda, K. (1989). The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. *Foreign Language Annals*, 22, 529-540.

Laufer, B. (1988). The concept of 'synforms' (similar lexical forms) in vocabulary acquisition. *Language and Education*, 2(2), 113-132.

(1989a). A factor of difficulty in vocabulary learning: Deceptive transparency. *AILA Review*, 6 (*Vocabulary acquisition*), 10-20.

(1989b). What percentage of lexis is essential for comprehension. In C. Lauren & M. Nordman (Eds.), *From humans thinking to thinking machines* (pp. 316-323). Clevedon, UK: Multilingual Matters.

(1991a). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). Basingstoke: Macmillan.

(1991b). *Similar lexical forms in interlanguage*. Tübingen: Gunter Narr Verlag.

(1992). Reading in a foreign language: How does L2 lexical knowledge interact with the reader's general academic ability? *Journal of Research in Reading*, 15, 95-103.

Laufer, B., & Sim, D. D. (1985a). Measuring and explaining the threshold needed for English for Academic Purposes texts. *Foreign Language Annals*, 18, 405-413.

(1985b). Taking the easy way out: Nonuse and misuse of contextual clues in EFL reading comprehension. *English Teaching Forum*, 23(2), 7-10, 22.

Liu, Na, & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, 16, 35-42.

Meara, P., & Jones, G. (1989). *Eurocentres vocabulary test 10 KA*. Eurocentres, Zurich.

Nation, I. S. P. (1983a). *Learning and teaching vocabulary*. NZ: Wellington, Victoria University.

(1983b). Testing and teaching vocabulary. *Guidelines*, 5 (RELC supplement), 12-24.

(1990). *Learning and teaching vocabulary*. New York: Newbury House.

Nation, P., & Coady, J. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 97-110). London: Longman.

Ostyn, P., & Godin, P. (1985). RALEX: An alternative approach to language teaching. *Modern Language Journal*, 69, 346-353.

Ostyn, P., Vandecasteele, M., Deville, G., & Kelly, P. (1987). Towards an optimal programme of FL vocabulary acquisition. In A. M. Cornu, J. Vanparijs, M. Delaheye, & L. Baten (Eds.), *Beds or bracelets? How do we approach LSP?* (pp. 292-305). Oxford: Oxford University Press.

Perkins, K., Brutton, S. R., & Pohlmann, J. T. (1989). First and second reading comprehension. *RELC Journal*, 10(2), 1-9.

Schneider, W., & Shiffrin, H. M. (1977). Controlled and automatic human information processing: I. detection, search and attention. *Psychological Review*, 84, 1-66.

Stahl, S. (1983). Differential word knowledge and reading comprehension. *Journal of Reading Behaviour*, 15, 33-50.

Steffensen, M. S., & Joag-Dev, Ch. (1984). Cultural knowledge and reading. In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a foreign language* (pp. 48-61). London: Longman.

- Stein, M. J. (1993). The healthy inadequacy of contextual definition. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary acquisition* (pp. 203-216). Norwood, NJ: Ablex.
- Ulijn, J. M. (1984). Reading for professional purposes: Psycholinguistic evidence in a cross-linguistic perspective. In A. K. Pugh & J. M. Ulijn (Eds.), *Reading for professional purposes* (pp. 66-81). London: Heinemann.
- Ulijn, J. M., & Strother, J. B. (1990). The effect of syntactic simplification on reading EST texts as L1 and L2. *Journal of Research in Reading*, 13, 38-54.
- Van Parreren, C. F., & Schouten-van Parreren, M. C. (1981). Contextual guessing: a trainable reader strategy. *System*, 9, 235-241.

3 Orthographic knowledge in L2 lexical processing

A cross-linguistic perspective

Keiko Koda

Introduction

During the last decade, there has been a growing interest in vocabulary acquisition among second language (L2) researchers. The research stockpile has expanded in scope and complexity at a remarkable rate for the brief period of time. Despite the increasing number of empirical studies, however, relatively little attention has been given to orthographic considerations. In view of the major gains in L2 vocabulary research, the gap has become problematic.

Conceivably, the minimal concern with orthographic knowledge may stem from the predominance of top-down approaches to L2 reading research over the last decade. More than 60% of the empirical research between 1974 and 1988 reflected this perspective (Bernhardt, 1991). As a consequence, scant work had been done on lower-level verbal processing mechanisms until very recently, and the role and function of orthographic knowledge remain largely unexplored.

There are compelling reasons, however, to believe that orthographic knowledge plays a critical part in L2 reading, particularly in lexical processing. Current L2 vocabulary studies, for example, consistently demonstrate that the ability to utilize context in inferring the meaning of unknown words is highly correlated with reading proficiency (e.g., Chern 1993; Haynes, 1984). And, even more important, the failure to use context for lexical inference is, in many cases, attributable to word misidentification (Huckin & Bloch, 1993). When L2 learners mistakenly assume they know a word, they tend to ignore various contextual clues that highlight the semantic incongruity resulting from the misidentification. But of greatest significance, many identification errors result from insufficient information derived from orthographic processing (e.g., Holmes & Ramos, 1993; Huckin & Bloch, 1993; Laufer, 1988). Thus, inefficient orthographic processing can lead not only to inaccurate lexical retrieval, but to poor comprehension as well.

Further support for the importance of orthographic contribution can be found in the L1 word recognition research. Of late, a renewed interest