

## Learning words in L1 and L2

*Paul Meara*

The study of second language acquisition has been influenced in many ways by developments in the study of first language acquisition. In fact, many of the more interesting developments in second language research have come about as a direct result of people applying to second language acquisition theoretical models and research tools which were originally developed in connection with L1 acquisition research. This phenomenon is particularly noticeable in the areas of morphology and syntax, and to a lesser extent in pragmatics. What this paper is about is why there is no evidence of a similar phenomenon in the area of lexical acquisition, and what we can learn from this.

The research literature on vocabulary acquisition in L1 contains two main types of study: diary studies and experimental studies.

Diary studies have a long and venerable history in the study of L1 acquisition. They date at least as far back as the Egyptian Pharaoh who is reputed to have had two children brought up by silent guardians in an attempt to find out what the first words of untutored children would be (Campbell and Greaves 1982). The issues of *The Pedagogical Seminary* which appeared between 1900 and 1925 contain a very extensive collection of more modern, the less dramatic, studies on vocabulary developments in English, and the basic methodology is still in evidence in the pages of the more specialised journals.

Diary studies have not been widely used in studies of vocabulary acquisition in an L2. There are a few studies which make explicit comparisons between lexical acquisition in an L1 and an L2. Most of these are case studies, however, which look at individual children, usually very young children, in very special circumstances, usually learning the L2 by living in the country where it is spoken. Most people will be familiar with the discussion of vocabulary development in Leopold (1949). More recently, Celce-Murcia (1977) provides an account of the vocabulary of a single child, aged 2;4, becoming bilingual in English and French. She identifies four different types of words in these two languages, and argues that the child systematically avoids words which cause phonological problems. A similar study is to be found in Yoshida (1978), a detailed categorisation of the vocabulary acquired by a 3;6 year-old Japanese child learning English. Kinzel (1964) is a study of another, slightly older, bilingual child, aged 6;3-7;2. One chapter of the study is devoted to lexical acquisition, and especially to lexical interference.

The principal difficulty with diary studies is that once the subject's vocabulary gets beyond critical size, it becomes effectively impossible to keep track of the new items that appear in speech, and this means that the technique works well only with very young children. Even then, it might be

argued that diary studies are essentially data driven studies, not model driven studies, and this limits their value to some extent.

For L2 acquisition, however, this type of study is of extremely limited value. There are always difficulties in generalising from a case study to a wider population. In the case of the diary studies of L2 acquisition quoted above these difficulties are immediately apparent: they deal with young children, so that issues of cognitive development are mixed up with issues of vocabulary acquisition; they deal with the special case of bilinguals, so that issues of vocabulary acquisition are mixed up with issues relating to linguistic dominance; and they deal with vocabulary acquisition in naturalistic settings, so that issues of presentation and exposure are difficult to separate out from more general theoretical issues.

However, not all studies of L1 vocabulary acquisition have relied on the relatively crude technique of the diary study. Although the number of studies in this area is not large, a sizeable collection of empirical studies is also to be found, and these have led to the development of some straightforward, readily testable models of how children acquire word meanings in an L1. This is obviously not the place to provide a complete survey of his work. Suffice it to say that one of the principal characteristics of this work is that it has concentrated attention on small subsystems in the child's emergent vocabulary, often on pairs of words with closely related meanings, or even individual words. Usually, the word set under observation is chosen because of intrinsic linguistic interest, so we have studies of MORE and LESS (Donaldson and Balfour 1968), a number of studies on dimensional terms - BIG, LITTLE, SHORT, THICK and so on (e.g. Maratsos 1973), studies of relational conjunctions like BEFORE and AFTER (Clark 1971), and studies of colours (Heider 1971).

An extreme example of a study of this sort is to be found in Carey (1978). Carey attempted to study how nursery school children learn to handle the colour **olive**, a colour which was not part of their normal repertoire at the time the study. Carey introduced two objects, an olive cup and an olive tray, into the things the children normally played with. These objects were referred to by the label **chromium** in case the children were already familiar with olives in a culinary context. Two very deliberate presentations of the word were made over a week with each child, and then a series of tests were given over a period of 18 weeks in order to ascertain how the children were internalising the new word. Carey records that most children quickly learned that **chromium** was a colour word, but that few children had managed to match the term stably by the end of the study. Two main strategies developed: some children showed an "odd colour, odd name" strategy, associating the label **chromium** with any colour that was not already well labelled; others developed a "false synonym" strategy, and used **chromium** as a variant of **green**.

Carey uses this data to argue that learning new words involves a two-stage process. The first stage, which she calls "fast mapping", is rapid and efficient, and locates a word roughly in the appropriate semantic area, but no more. In this case, fast mapping would locate **chromium** as a colour word, but fail to specify what sort of colour word it was. Her data suggests that the fast mapping stage can be achieved with minimal exposure; in some cases, only a single exposure to the word was required. "Extended mapping" is the process by which the meaning of a word is made more precise and exact. Carey sees this process as essentially one in which the semantic features that provide a full specification of the meaning of word are gradually filled in. This idea is

not very different from Clark's (1973) Missing Feature Theory, though Carey herself prefers a modified version of Clark's model, which she calls the Missing Feature and Haphazard Example model. The details of this discussion will not concern us here. The important point for this paper is that lexical acquisition is slow and laborious. We will return to this point later.

There are several reasons why studies of this sort do not provide a paradigm which easily transfers to L2 learners. To a large extent these reasons derive from the fact that the majority of studies of vocabulary acquisition in an L1 deal with relatively young children. This focus has three important consequences.

Firstly, it means that almost all the studies one finds discussed in the literature are principally concerned with cognitive development in the L1, rather than with the acquisition of vocabulary *per se*. In young children, these two issues are inextricably linked together in a way which may not apply to older learners and second languages. In Carey's case, for example, what she is really interested in is how the child learns to differentiate between **chromium** and other colour words. A study like Carey's only makes sense if we assume that the cognitive task involved here is a difficult one, and in a sense, the fact that she uses linguistic tools to explore this effect is of secondary importance. This is not to suggest that the problem Carey is investigating does not exist for second language learners. On the contrary, there are plenty of examples where pairs of languages carve up reality in different ways, and these mismatches do cause some difficulty for learners. Nevertheless, instances of this sort are not so plentiful that they are a major source of difficulty in an L2. In Spanish, for example, you have a colour word **pardo** which is used to name a range of colours not unlike **chromium** which do not have a single name in English. **Pardo** covers what English speakers would label **gray**, **olive**, **dun**, **rust**, **khaki**, and so on. Now one could imagine setting up an experiment in which English-speaking learners in Spanish were introduced to **pardo** and their understanding of this word was monitored systematically over eight months. The results would presumably show that the learners identified **pardo** as a colour word fairly easily, and that as long as they didn't forget it altogether, their understanding of the niceties of its meaning developed perceptibly over a period of months. However, it would take a brave researcher to generalise from this one instance to more sweeping claims about the nature of vocabulary acquisition in an L2.

The basic problem here is that the L1 vocabulary studies often exhibit what can only be described as a perverse fascination with what children find difficult, but they don't provide a lot of insight into what children find easy. Everyone knows for example, that very young children have difficulty deciding who can be referred to as **dada**, that older children don't always get **teach/learn** or **borrow/lend**. On the other hand it takes some ingenuity to show that **ask/tell** is still a problem for nine-year-olds, and no one has made serious case for the view that such items form a significant proportion of the vocabulary of children at age. If anything, in fact, quite the contrary is true. Many L1 words seem to get learned almost instantaneously, and do not pose problems of the sort that we typically investigate. Similar arguments apply to L2 learners. Obviously, L2 learners do have some difficulties with "stubborn" words which don't map one-to-one onto English, but these difficulties are not major. Most learners of French eventually sort out the difference between **connaître** and **savoir**, just as most babies eventually get to recognise their fathers and avoid bearded strangers. This type of model, then, does not serve L2 acquisition well from a theoretical

point of view. Worse than this, it reinforces assumptions and prejudices about L2 vocabulary acquisition being difficult and demanding.

The second reason why it is difficult to generalise from L1 studies to L2 is that the range of learning strategies available to young children is much narrower than what is available to an older L2 learner. L2 learners can use dictionaries; they can quiz their teachers on points of difficulty; they can systematically set about increasing their vocabulary in a whole variety of ways; they have access to printed matter or broadcast material that is not available to young children. This means that vocabulary acquisition in an L2 is likely to be less uniform than it is an L1, and less constrained by the limitations of the input channel.

The third reason why studies of L1 may not generalise easily to L2 situations is the children have small vocabularies. This may not seem to be an important point, and my hunch is that adding new words to a small vocabulary is qualitatively different from adding words to a large vocabulary. Let me illustrate this point with some simple examples using graph theory.

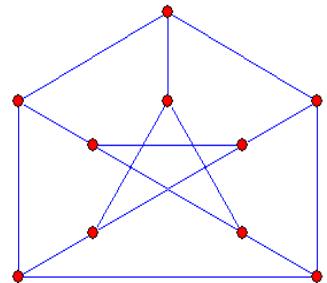
Imagine, for the sake of argument, that a vocabulary can be described as a set of points linked by arcs colour and that each point has a limited number of arcs attached to it. Call the number of points in the vocabulary its **size**, and call the number of arcs attached to a point the **valency** of the point. Now these two variables, size of the graph and the valency of the points, affect how closely the individual points on the graph are linked. The **diameter** of the graph increases if the size of the graph is bigger but the valency of the points remains the same. This is shown in figure 1. In figure 1a we have a graph of 10 points, and each point has a valency at 3. The longest path between any two points in the graph is three steps. Increasing the size of the graph by 20 points to 30, while preserving the valency of each point at 3 increases the diameter of the graph appreciably (figure 1b). In this case, the longest path goes up to four steps. The increase could be offset by increasing the valency of points. Figure 1c shows a graph of 30 points where each point has a valency of 5. In this graph, the longest path is three steps, but most points are joined by paths that are only one or two steps in length.

Now obviously, this type of model does not mirror exactly what goes on in the mental lexicon, but the analogy may be one worth exploring further. In particular let us imagine that human brains put a premium on developing lexicons with moderately small diameters, say in the region of 6 to 8. This will clearly have implications for the valency of each point. The smaller the permitted diameter, the larger the valency has to be. Let us also assume that children start off with low valencies. This would mean that lexical acquisition in children will be characterised by a sequence of catastrophic developments where the size of the vocabulary combined with low valencies to produce an unacceptably large diameter. At this point the lexicon would need to undergo some form of restructuring before further growth could occur.

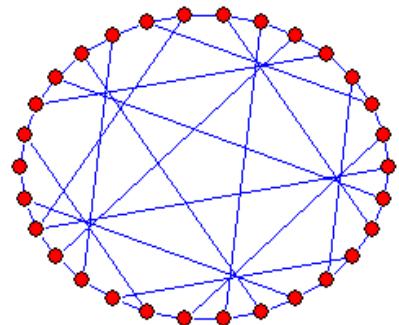
The exact mathematics of all this is not clear to me. For one thing, the graphs I have been discussing are all very simple ones, and as you move to graphs with a more realistic number of points, it becomes increasingly difficult to work out the distribution of path lengths. Secondly, the graphs that I have described are efficient in the sense that they are symmetrical and this minimises the path lengths. Real vocabularies are likely to be much less obliging. Nevertheless, it is easy to see

**Figure 1:**  
Graphs of varying sizes and varying valencies

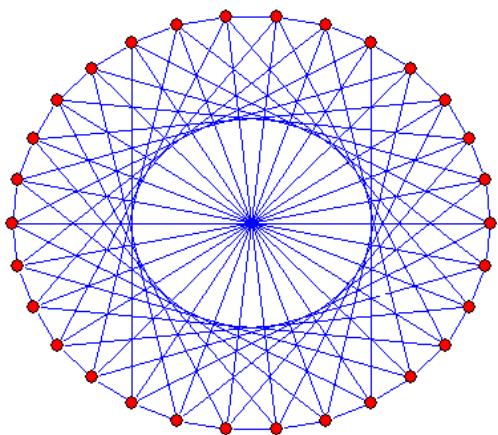
**Figure 1a**



**Figure 1b**



**Figure 1c**



that restructuring will occur very frequently at first, and then, as the vocabulary size increases and the valencies increase to compensate for this the necessity for restructuring will occur less frequently. Large numbers - hundreds - of words can be added to the vocabulary without incurring any penalty. Adding six words to vocabulary of 20,000 will certainly not have the same dramatic effects as adding six words to vocabulary which only contains six words to start with.

These arguments suggest to me that there is not a lot to be gained in trying to find point of contact between studies of early vocabulary acquisition in an L1 and an L2. On the other hand, they do suggest that there might be points of contact between L2 acquisition and studies of vocabulary acquisition in older L1 speakers, particularly adolescents.

Surprisingly, this is an area which appears to have received very little coverage, and one which is barely mentioned in textbooks on language acquisition. Dale (1976), for example, has an extensive chapter on the acquisition of word meaning, but none of the studies he mentions uses subjects aged more than eight years. He does quote some very old data from Smith (1926) via McCarthy (1954) which shows how vocabulary grows from 0-6 years, but even here his discussion of the data is cursory in the extreme. The data is mentioned principally so that Dale can point out the difficulties of undertaking research with older children. The entire question is dismissed in one page.

The assumptions underlying Dale's argument seems to be that the size of older children's vocabulary makes it impossible to do interesting qualitative research. When you have vocabularies of 20,000 or 30,000 words, a detailed investigation of a handful of individual words is not likely to produce interesting generalisations. This is probably true, but it overlooks the fact that purely quantitative aspects of vocabulary growth can in themselves provide us with interesting questions and objects of wonder.

The most impressive body of work in this area that I know of is a series of studies by Nagy, Anderson and sundry colleagues at the University of Illinois (see bibliography). Nagy and Anderson argue that children's L1 vocabularies are indeed very large - they suggest 50,000 words using conservative estimates of what counts as a word – and that the very size of the task involved in learning vocabularies as large as this imposes severe constraints on the way in which they are learned. There is, for example, a lot of research suggesting that formal vocabulary tuition does improve the ability of children to handle individual words and that formal tuition is considerably more effective than more haphazard informal methods of acquisition. Unfortunately, the amount of time required for formal tuition makes it an unrealistic option in real-life situations. You may be able to teach 10 or 20 words in this way, but not 10,000 or 20,000. Nagy and Anderson argue that school children's vocabularies increased by something like 5000 words a year, and they present some very elegant evidence to show that a large proportion of this could be achieved by exposure from reading. The argument is basically as follows. In normal reading texts, there is a probability of about between 0.05 and 0.1 that the context provides enough clues for a new word to be acquired from a single exposure. The number of unknown words that school children meet in a representative 1000 word text is between 15 and 55. The number of words that the average school child meets in print in the year is about one million. Combining these three figures suggests that children could learn large numbers of words from single exposures during reading, the maximum

and minimum suggested by these figures are 5500 and 750. Similar arguments presumably apply to spoken input as well, though here the total number of words children are exposed to is likely to be greater than what they experience in reading. On the other hand, the proportion of new, unknown words encountered is likely to be much less than in written text, and the probability of acquiring a word's meaning at a single exposure from speech is likely to be considerably lower than the figures quoted above. Nagy and Anderson, then, believe that the key element in the acquisition of vocabulary is exposure to large quantities of natural text. There is a mechanism which makes it possible to acquire the meanings of some new words from single exposure, and though this mechanism is relatively inefficient, it will nevertheless ensure that a large vocabulary develops, given enough input to work on.

Obviously, Nagy and Anderson's figures do not apply directly to typical L2 learning situations, principally because their assumptions about exposure to language input do not fit what normally happens to L2 learners. Few L2 learners will be exposed to one million words of running L2 text in a year; the figure is more likely to be in the region of a few thousand. Few learners are exposed to more than a few hours L2 speech in a week. Indeed if you take the not untypical case of an adult following a BBC language course, the total exposure available is in the region of one hour a week, plus some reading in the form of course notes. Much of this exposure will typically be in English anyway - explanatory notes, translations, instructions and so on are rarely provided in the L2. Speech rate in materials of this sort will generally be slower than normal language, and much of it is repetitious. Suppose, however, that a steady rate of one word a second is kept up throughout the broadcast material (a generous assumption, probably) then the total exposure to running text in a week is going to be not much more than three or four thousand words with perhaps 200 new unknown words. Nagy and Anderson's data suggest that a base of this size is quite insufficient for building even moderately sized vocabulary.

However, if we ignore this fundamental difference between L1 acquisition and L2 acquisition, Nagy and Anderson's data are suggested in three important ways.

Firstly, their work suggests that the key to vocabulary acquisition is exposure. Mental processes exist which enable us to learn new words from minimal exposure. These processes are not very efficient, but they are good enough to ensure that large vocabularies can be acquired as long as the language we are exposed to is extensive enough. This might lead us to expect fundamental differences in the vocabularies of learners taught in classroom situations, and those who have the benefits of extended residence in a country where the target language is spoken. Surprisingly again there is very little research on this issue. Folklore asserts that residence abroad is indeed beneficial, but most of the research which might have been expected to show material gains in specific linguistic skills (e.g. Willis et al. 1977) has actually concentrated on measures of attitude and motivation, with a few tests of general language competence as support. This is obviously an area which merits closer investigation.

Secondly, the work suggests that it is basically wrong to compare L2 learners with young children from the point of view of vocabulary acquisition. Acquiring new words is not the prerogative of children; we all do it all the time, especially during adolescence, and there are no obvious reasons for assuming that acquiring L2 words is in any way as fundamentally different process from

acquiring new L1 words at this level. There may be some problems with acquiring the phonological and orthographic forms associated with new words, and these difficulties may differ somewhat from what happens in L1, but the process of acquiring a passive vocabulary is going to be essentially the same, whatever source of the new word. Interestingly, most of the commercial materials available make quite the opposite assumption. They seem to believe that explicit teaching of semantics is essential, but that is not necessary to help L2 learners overcoming difficulties they have in storing the **forms** of new L2 words (cf. For example Rudzka et al. 1981).

The third and most important idea that arises from Nagy and Anderson's work is that we may be severely underestimating the ability of learners to pick up new words. There is a generally held view that vocabulary acquisition is hard, and the only way of helping learners to surmount this difficulty is to limit the learning load by reducing the number of words learned to something of manageable proportions. This sort of idea underlies most attempts to limit the number of words learners are required to learn, whether by frequency, availability, interest centres, or whatever. Frequency analyses suggest that a vocabulary of 2000 words or so covers 80 percent of everyday language use, and the target in this order of magnitude is often taken as a reasonable first speech for a foreign language. O-Level examinations, for example, tend to work within a vocabulary of about this size, though there is a widespread belief among teachers that 2000 words is rather a lot for a five-year course, and that a lower figure would be more appropriate. 2000 words over five years works out at 400 words year, i.e. roughly 10 words week. This figure is actually quite large relative to the amount of imports L2 learners typically receive, but compared to the massive figures that Nagy and Anderson quote for L1 acquisition, this figure could only be described as puny. Nagy and Anderson's work suggests that the current practice of working with restricted vocabularies for as long as possible might be quite wrong, and that we ought perhaps to aim at much bigger vocabularies right from the very start of a language course. This might involve less emphasis on oral, communicative work, and more emphasis on reading, television, films, and other forms of "passive exposure" however, and I doubt whether such a suggestion would go down well in the current climate.

A corollary of the low targets we set for vocabulary acquisition in the beginners' foreign language courses is that we expect massive vocabulary expansion over very short periods of time in more advanced students. Anyone who has had the experience of doing an O-Level language followed by reading set texts for A-Level will remember the dreadful tedium of working through three or four hundred pages of novel, often looking up in a dictionary half of the words in each page, a process which continued for some years at university level. By the end of it, many learners emerge pretty fluent in the target language, but the process is far from easy. In Nagy and Anderson's terms, what is happening here is very rapid growth in the target language, from a vocabulary of one to two thousand words, to something more closely approximating the vocabulary of a native speaker, say 30,000 words at a conservative estimate over a period of five years. This is no mean feat. It represents something like 150 new words a week, or approximately half as much again as the upper limit Nagy and Anderson suggest for native speakers. However, even this figure turns out to be an underestimate, when you consider the language students at university level often acquire high levels of proficiency in two languages simultaneously, and at the same time continue to expand their L1 vocabularies. It looks then, as though highly verbal learners might be able to acquire four or five hundred new words a week in a variety of languages, and that they can

sustain this performance over a number of years. This is an extraordinary performance by any standards, but one which we have no real understanding of.

Susan Carey described her preschool children as "magnets for new words", estimating that they had learned about nine new words a day. She argued that her extended mapping process took some time to fix the meaning of new words, anything up to six months, and that meant that the child is working out the meanings of 1600 words at any one time. She finds this a staggering figure, and indeed it is, but it pales into insignificance compared with what older foreign language learners achieve. Unfortunately, the sheer size of the achievement of advanced L2 learners makes it very difficult to investigate processes involved. Certainly the single word approach, so much favoured in L1 studies, does not have the necessary power. What we really need is a reliable tool that would enable us to measure accurately the size of an individual L2 speaker's lexicon. Such a tool would open up large areas of research which are no-go areas at the moment. It would, for example, enable us to chart the rate at which L2 learners' vocabularies develop, and to discover whether this rate is steady or subject to fluctuations, and we could investigate the effects of L1 on vocabulary growth rates. We would also be able to discover whether it is possible to affect the rate of growth by particular types of instruction or particular types of exposure and to chart the rate at which vocabularies atrophy when they are not exercised. More importantly, the existence of such a measure would enable us to develop testable models of the acquisition process which take vocabulary size into account as a critical variable. It seems reasonable to suppose, for example, that learners with vocabularies of around 200 words behave differently from more advanced learners with vocabularies in the region of 2000 words, and that they in their turn behave differently from learners with 20,000 word vocabularies. If a good measuring tool existed, it would be possible to develop this sort of hunch into a set of specific hypotheses to describe levels of proficiency. At the moment, however, claims like this cannot really be investigated; we have to rely on crude measures such as length of time studying the language or crude measures of overall proficiency, which lump different types of learners together in extremely coarse groupings. This is clearly unsatisfactory.

I am sorry to say, that I have not yet been able to develop such a test, though I do have a couple of pilot studies under way which rely on the application of signal detection theory to vocabulary testing - another idea which also derives from Anderson.

Some years ago it was fashionable to describe a vocabulary acquisition as a neglected area of L2 studies. Vivian Cook recently referred to the "massive" amounts of research which have appeared in the last few years. In some ways, however, it seems to me that the area is still in its infancy, and that it is unlikely to get out of this state until we move away from the study of how learners acquire tiny word sets, to the study of vocabulary acquisition on a much larger scale.

## References

### **Anderson, RC and P Freebody**

Vocabulary knowledge. In: J Guthrie (ed.) *Comprehension and Teaching: Research Reviews*. Newark, De.: International Reading Association. 1981.

**Anderson, RC and P Freebody**

Reading comprehension and the assessment and acquisition of word knowledge. In: **B Hutson** (ed.) *Advances in reading/language research*. Greenwich, Ct.: JAI Press. 1983.

**Campbell, R and R Greaves**

Royal investigations of the origins of language. *Historiographia Linguistica* 9(1982), 43-74.

**Celce-Murcia, M**

Phonological factors in vocabulary acquisition: a case study of a two-year-old, English-French bilingual. *Working Papers in Bilingualism* 13(1977), 27-41.

**Carey, S**

The child as word learner. In: **M Halle, J Bresnan and GA Miller** (eds.) *Linguistic Theory and Psychological Reality*. Cambridge Mass: MIT Press. 1978.

**Clark, EV**

On the acquisition of 'before' and 'after'. *Journal of Verbal Learning Verbal Behavior* 10(1971), 266-275.

**Clark, EV**

What's in a word? On the child's acquisition of semantics in his first language. In: **T Moore** (ed.) *Cognitive Development and the Acquisition Language*. New York: Academic Press. 1973.

**Dale, P**

*Language Development: structure and function*. New York: Holt, Reinhart and Winston. 1973.

**Donaldson, M and G Balfour**

Less is more: a study of language comprehension in children.

*British Journal of Psychology* 59(1968), 461-472.

**Heider, HE**

'Focal' colour areas and the development of colour names.

*Developmental Psychology* 4(1971), 447-453.

**Kinzel, P**

*Lexical and Grammatical Interference in the Speech of the Bilingual Child*. Seattle: University of Washington Press. 1964.

**Leopold, W**

*Speech Development of a Bilingual Child: a linguist's record*. Evanston: Northwestern University Press. 1949

**Maratsos, P**

Decrease in the understanding of the word 'big' in preschool children. *Child Development* 44(1973), 747-752.

**McCarthey, D**

Language development in children. In: **L Carmichael** (ed.) *Manual of Child Psychology*. New York: Wiley. 1954.

**Nagy, W and RC Anderson**

How many words are there in printed school English? *Reading Research Quarterly* XIX,3(1984), 305-330.

**Nagy, W, P Herman and RC Anderson**

Learning words from context. *Reading Research Quarterly* XX,2(1985), 223-253.

**Rudzka, B, J Channell, Y Putsyes and P Ostyn**

*The Words You Need*. London: Macmillan.1981.

**Smith, ME**

An investigation of the development of the sentence and the extent of vocabulary in young

children. *University of Iowa Studies in Child Welfare*. 1926.

**Willis, F, G Doble, U Sankarayya and A Smithers**

*Residence abroad and the student of modern languages*. University of Bradford. 1977.

**Yoshida, M**

The acquisition of English vocabulary by a Japanese speaking child. In: **E. Hatch** (ed.) *Second Language Acquisition: a book of readings*. Rowley, Mass.: Newbury House. 1973.

This discussion paper was first posted in 1987.