Learning Words in Context: a Survey

Raluca Budiu

August 27, 1997

Abstract

This paper surveys a number of theories of word meaning acquisition from context; it discusses possible (text-dependent or text-independent) constraints or biases that might help in the process, and also factors like student ability and the strength of contextual support that affect the learning task. It also reviews incidental learning from context during normal reading.

1 Introduction

The literature on word meaning learning in context seems to agree (more or less tacitly) in seeing this process as a search in a huge space of possible hypotheses. What differentiates various approaches is, grosso modo, the commitment to distinct types of cues (heuristics) — some times called constraints — that presumably help people to prune this space. Are the same type of constraints (or cues) used by high verbal ability and low verbal ability individuals? How are the context and the word properties related to such cues? Are there any innate constraints or any constraints that are constant across individuals? These types of questions are usually asked by papers in the field of word learning in context.

2 Theories of Learning Words In Context

We discuss in this section several papers that address questions related to the nature of search and/or the nature of heuristics used in word learning in context. We also included here, although they do not quite satisfy this criterion, the study by Werner and Kaplan, which pioneered the research in this field; the study highlights some types of error made by children in the learning process.

- van Daalen-Kapteijns and Elshout-Mohr [Van Daalen-Kapteijns et. al 1981] propose a theory of word learning in context for high and low verbals[1]. College students are asked to read five sentences referring to one unknown word and to write down their hypotheses about what the word meaning could be. The words denoted plausible, but inexistent concepts.

  The authors found out that both low and high verbals start by attributing to the unknown word a meaning familiar to them (the so-called *model*) denoted by a known word. The difference between low and high verbals lies in the way they process the model subsequently, after seeing other contexts. High verbals proceed *analytically*, i.e. they adapt their model by

[1] "High/low verbal" stands for "high/low verbal ability individual".

replacing some of its features with features of the unknown meaning as deduced from the new contexts. On the other hand, low verbals tend to use the model *holistically*: rather than refining one and the same model, they define one local model for each sentence, and finally, when asked to decide for a meaning, they try to integrate the local models through some weighting process.

High and low verbals differ also in their ability to extract information (features of the model) from the context (*decontextualization*). High verbals use their model to guide the decontextualization process — certain slots of the model are modified according to the new information. Low verbals too use models to solve ambiguities, but to a lesser extent. The model delimits only the domain of the new word meaning. If it is possible to find a reformulation of the context that is not contrary to the model, then the domain is better defined; if it is not, then the current model is abandoned and a new one takes its place. Therefore, in low verbals, the model-based search is overruled by sentence-based processes.

The surprising thing is that, although the decision problems of the low verbals are more difficult than the ones of the high verbals, the van Daalen-Kapteijns and Elshout-Mohr's data show that both groups do as well in the learning task.

The authors explain the different use of the models by high and low verbals by hypothesizing that high verbals have a better-structured knowledge organization. They are able utilize word meaning as aggregates built of meaning components that can be acted on independently.

- The study by McKeown [McKeown 1985] also offers a model for the meaning acquisition process:

  1. the person recognizes the unknown word in a context and selects from the context concepts that constrain the meaning of the word
  2. the person searches and tests the possible hypotheses (obtained from known concepts that fit the constraints)
  3. a hypothesis about the word is formed
  4. if the formed hypothesis does not decide that the word is known, then on a subsequent encounter of the word, the whole process is repeated, but, between steps 1 and 2, the previous constraints are coordinated with those met at present.

  Starting from this theory, McKeown designed an intricate experiment that measured how well low and high verbal (the degree of expertise being measured in terms of vocabulary knowledge and reading ability) fifth-graders performed each of the steps and substeps described above. The children were given some possible initial word meanings.

  Three results came out of the study.

  1. First, the low verbals exhibited a misunderstanding of the relationship between word and context (step 1). They seemed to create a story behind the context (behavior common for both groups), but relied more on that story than on the context when it comes to assigning a meaning for the unknown word.
  2. Second, when multiple contexts were used, low verbals tended to transfer the meaning obtained with one context to the other, and, thus, poorly evaluated the second context. This behavior was exhibited by high verbals, too, but to a lesser degree.

3. Third, even if the constraints were correctly identified, low verbals weren't always able to test the meaning choices within constraints. Also, they had more difficulties in identifying the correct meaning, and even if given the correct definition, they did worse than the high verbals at testing the meaning in new contexts.

- Some of the van Daalen-Kapteijns and Elshout-Mohr's and also McKeown's findings were actually present in a much earlier study on word acquisition in context, due to Werner and Kaplan [Werner et al. 1950a, Werner et al. 1950b]. They monitored the learning process in children from five age groups (between 8 and 13 years old). The subjects were given six sentences containing an artificial word and, after each sentence, they were asked to define the word.

  Werner and Kaplan observed two patterns in the errors the children made. First, the younger children failed to differentiate between the meaning of the new word and the overall meaning of the sentence. They either transferred part of the sentence meaning to the word meaning, or interpreted a new context as the same with another one encountered previously (this error was exhibited by the adult subjects of McKeown [McKeown 1985], too, as discussed before), or combined two local different meanings, or included parts of the sentence in the word meaning. The second class of errors (also, some of them encountered in the McKeown's study) referred to the fusion between the word meaning and some broader situational connotation (imagined by the child, a kind of "story") unstated in the context. The concept might change from sentence to sentence, new features being added/removed. Sometimes a unique scenario was used, a kind of "nucleus" that was enriched on each new context encountered (this is similar to the van Daalen-Kapteijns and Elshout-Mohr's model, but at the scenario level rather than at word meaning level). Also, a concept found valid for a previous context was kept subsequently if it was in some relation (causality, spatial contiguity) with another concept that was fit at present.

- Fischer's study [Fischer 1994] on second language students, claims that learning words from context involves a substitution strategy (same as for definitions). Contextual information guides the substitution and overgeneralization is avoided by relating the new word to the type of contexts in which it was first met. This strategy is similar with that of subjects using only dictionaries for word learning. In the same paper, the author finds that second language learners do not seem to benefit from the presence of both context and dictionaries, but they prefer to rely mostly on dictionaries when these two sources are present.

- Sternberg and Powell [Sternberg et al. 1983] identify the context features (contextual cues) and the individual-dependent features (mediating variables) which influence word learning from context. The contextual cues

  > "determine the quality of a definition that can theoretically be inferred for a word from a given context."

  The authors propose eight types of contextual cues ("temporal", "spatial", "value", "stative descriptive", "functional descriptive", "causal/enablement", "class membership" and "equivalence").

  The mediating variables specify the factors that can affect the application of the contextual cues in a given instance. They are: the number of occurrences of the unknown word, the variability of the context in which multiple occurrences of an unknown word appear, the density of unknown words, the relevance of the unknown word for passage comprehension, how helpful the context is perceived for the understanding of the new word, how abstract the target word is, how useful the prior knowledge is in cue utilization.

Sternberg and Powell extend this theory at the internal, word-level context; the contextual cues are in this case morphological (stem, prefix, suffix, interactive) and the mediating variables are the ones that may prime the usage of the knowledge of the morphology to infer what a word means.

3 External Constraints Guiding Word Learning

In this section we examine some papers that investigate the existence of some non-contextual constraints on word meanings. There is much controversy on the hypotheses of innate, universal constraints on what constitutes a possible word meaning. The studies that we discuss try to asses the idea that implicit prior knowledge (gained by the speakers during their experience with the language) constitutes a useful bias for the learning of words in context.

- [Nagy et al. 1990b] assume that a word learner is helped in his understanding of a new words by the previous experience he had with the language and with what may constitute a possible word meaning. The authors use the term *word schema* to refer to the prior background knowledge about words. The background knowledge can be metacognitive and metalinguistic (i.e. what students know about the task of word learning), morphological (e.g. affixes, roots) or related to typical patterns of word meaning (e.g. verbs of motion in English usually convey information about motion rather than about path or the object of the verb). In [Nagy et al. 1990b] an experiment to test this assumption is described; the results show that high school and college students are sensitive to overall plausibility of definitions (with increased sensitivity for verbs rather than nouns) and that college students rely on morphological cues to decide if a word is plausible or not.

- The studies conducted by Nagy and Gentner [Nagy et al. 1990a] investigate the same line of research. Their data shows that apparently speakers use a variety of constraints on word meanings, including some that are apparently language specific (e.g. for English a preferred method is not to include too contingent information into the meaning of a word, attaching manner rather than other determinants to a verb of motion, expressing events through verbs and objects through nouns etc.). Interestingly, subjects seemed to extract different properties of the context in order to attribute meaning to verbs or nouns.

4 Factors That Influence Word Learning in Context

Contextual Support [Carnine et al. 1984] shows that it is easier to determine the meaning of an unknown word if it is placed in a context rather than isolated. Contextual cues that are close to the target word make the process simpler. Also, synonyms rather than inferential cues help students derive the word meaning. Older students seem abler to derive meanings for unknown words.

Another study dealing with the influence of the context on the word acquisition task is [Li 1988], for second language learners. The author doesn't report the kind of cues he uses. However, his subjects are better at deriving and retaining word meanings in sentences that provide adequate contextual cues.

[Mondria et al. 1991] deals also with the effect of the pregnancy of context on word guessing and retention for second language learners. It also asks the question how correct and incorrect guessing affects the retention (after a learning stage). The data shows that a pregnant context (especially in sentences in which cues are given by the function words) does result in improved guessing, but not in improved retention (after guessing, there is a learning stage that uses the same context). Moreover, words that were correctly guessed in context were not remembered more often than the words incorrectly guessed.

Student Ability Shefelbine [Shefelbine 1990] shows that lower vocabulary students learn less word meanings from context, presumably because they have less available contextual information for figuring out the unknown words and because they have a less developed knowledge of known words. It looked like word learning primarily involved the same words for both groups; words already known by high verbals didn't seem easier to learn for the low verbals. Reasoning ability seemed to have no effect on the task. However, new concepts were harder to learn by low vocabulary subjects.

5 Incidental Learning

[Nagy et al. 1985] distinguishes between the ability to *derive* word meanings in context and the ability of incidental learning during normal reading. The first, the authors argue, is actually exhibited by subjects in studies which give them explicit instructions to figure out the meaning of unfamiliar words in a text, and it is rarely encountered in every-day situations. The claim of the paper is that incidental learning from context is the major source for vocabulary growth during school years, due to the volume of experience with written context. The hypothesis the authors make is that incidental learning occurs in small increments, and that new encounters with unknown words increase the partial knowledge of that word. They designed an experiment that simulated normal reading conditions and they found a small, but steady (across types of text, methods of measurement and levels of scoring) amount of learning from context.

[Nagy et al. 1987] analyzes the influence of word and text properties on incidental learning of word meaning in context during normal reading. The amount of learning is not very impressive in their study. However, their data shows that conceptual difficulty (i.e. how unfamiliar the concept denoted by a word is) is the only word property that affects learning (the other word properties considered were number of occurrences in the context, length of the word, part of speech, morphological transparency, word familiarity). On the other hand, the text properties that were found to interact with the learning were passage-level conceptual difficulty (the proportion of conceptually difficult words in the passage) and average length of unknown words in the passage. Unexpectedly, the strength of contextual support didn't seem to interact as much with the word acquisition, at least in narratives versus expositions.

Based on these findings, the authors argue for a schema theory of vocabulary acquisition. They consider that

"the most important factor in learning from context is the degree to which the readers can integrate the information in a passage into a coherent system consistent with his or her prior knowledge."

Indeed, the authors claim that what the conceptual difficulty of a word actually measures is the degree in which the word is part of a system of concepts and the familiarity of the learner with that system.

[Herman et al. 1987] is another study concerning the effects of the text properties on word learning. The properties considered this time are the ones believed to influence reading comprehension: text macrostructure (i.e. features of global coherence, like the quality of titles, topic sentences, organization), text microstructure (i.e. logical and temporal relationships expressed in text) and the completeness of the description of the text concepts and of the relations among them. The study deals only with expository passages. Three effects were observed. First, an increase in the amount of word meanings known by middle-grade students who read a certain text (which confirmed again the incidental learning hypothesis). Second, high verbal ability students (in this study, the students who had high verbal comprehension skills) learned significantly more than their low verbal peers. And third, only the elaborated text version (which had good macrostructure, good microstructure and a complete explanation of the concepts in the text) seemed to affect significantly the acquisition of new words. Its benefits were greater for high verbals than for low verbals.

6 Partial Conclusion

Although this survey might not give this impression, it seems not very easy to state a conclusion about what is known on the topic of word learning in context. This is mainly because the experiments are done in different conditions and one should be very careful when comparing them. For instance, verbal ability is measured in different ways by different studies (although it is possible that these measures be correlated); certain settings may conduce to contextual support influencing the word acquisition and others may not (maybe because contextual cues or even text types are different); sometimes, low verbals may be as good as high verbals at learning, etc. However, the idea of a (more or less context-guided) search, at least for deriving word meanings, seems a constant.

References

[Carnine et al. 1984] Carnine D., Kameenui E. J., Coyle G., 1984. Utilization of Contextual Information in Determining the Meaning of Unfamiliar Words. *Reading Research Quarterly, 19*, pp. 188-204.

[Fischer 1994] Fischer U., 1994. Learning Words from Context and Dictionaries: an Experimental Comparison. *Applied Psycholinguistics, 15*, pp. 551-574.

[Herman et al. 1987] Herman P.A., Anderson R.C., Pearson P.E., Nagy W.E. 1987. Incidental Acquisition of Word Meaning from Expositions with Varied Text Features. *Reading Research Quarterly, 22*, pp. 263-284.

[Li 1988] Xiaolong L., 1988. Effects of Contextual Cues on Inferring and Remembering Meanings of Words. *Applied Linguistics, 9*, pp. 402-413.

[McKeown 1985] McKeown M.G. 1985. The acquisition of word meaning from context by children of high and low ability. *Reading Research Quarterly, 20*, pp. 482-496.

[Mondria et al. 1991] Mondria J., Wit-de Boer M., 1991. The Effects of Contextual Richness on the Guessability and the Retention of Words In a Foreign Language. *Applied Linguistics, 12*, pp. 249-267.

[Nagy et al. 1987] Nagy W.E., Anderson R.C, Herman P.A., 1987. Learning Word Meanings from Context during Normal Reading. *American Educational Research Journal, 24*, pp. 237-270.

[Nagy et al. 1985] Nagy W.E., Herman P.A., Anderson R.C., 1985. Learning words from context. *Reading Research Quarterly, 20*, pp. 233-253.

[Nagy et al. 1990a] Nagy W.E., Gentner D., 1990. Semantic Constraints on Lexical Categories. *Language and Cognition Processes, 5*, pp. 169-201.

[Nagy et al. 1990b] Nagy W.E., Scott J.A., 1990. Word Schemas: Expectations about the Form and Meaning of New Words. *Cognition and Instructions, 7*, pp. 105-127.

[Sternberg et al. 1983] Sternberg R.J., Powell J.S., 1981. Comprehending Verbal Comprehension. *American Psychologist, 38*, pp. 878-893.

[Van Daalen-Kapteijns et. al 1981] van Daalen-Kapteijns M.M., Elshout-Mohr M., 1981. The Acquisition of Word Meanings as a Cognitive Learning Process. *Journal of Verbal Learning and Verbal Behavior, 20*, pp. 386-399.

[Shefelbine 1990] Shefelbine J., 1990. Student Factors Related to Variability in Learning Word Meanings in Context. *Journal of Reading Behavior, 22*, pp. 71.

[Werner et al. 1950a] Werner H., Kaplan E., 1950. Development of Word Meaning through Verbal Context: an Experimental Study. *Journal of Psychology, 29*, pp. 251-257.

[Werner et al. 1950b] Werner H., Kaplan E., 1950. The Acquisition of Word Meanings: a Developmental Study. *Monographs of the Society for Research in Child Development, 15, No. 51.*