

# *Learnability and LIKELY/PROBABLE*

RICHARD HUDSON, AND ROSTA & NIKOLAS GISBORNE

---

One of the early contributions to the debate about learnability, Hudson (1972), pointed out the problem raised by the difference between LIKELY and PROBABLE in sentences like (1):

- (1) a It is likely/probable that John will be late.  
b John is likely/\*probable to be late.

Have we solved this problem in the past 21 years of work? We shall show that these words are still highly relevant, though not for the reasons originally given; and we shall introduce some more data which seem even more relevant.

One of the most influential theories of learnability is that a word's syntactic distribution is entirely predictable from its semantic structure, with various caveats about taking due care with the semantic analysis (e.g. Pesetsky 1982, Pinker 1989, Grimshaw 1990; but contrast Grimshaw 1979, Jackendoff 1985, 1993, Rothstein 1992 and Chomsky and Lasnik forthcoming). One very simple prediction of this theory is that true synonyms must have precisely the same subcategorization. So far as we can tell, PROBABLE and LIKELY are indeed true synonyms; but their subcategorization facts are different, as can be seen from (1). These facts therefore still stand as apparent evidence that at least some syntactic facts are not predictable from semantic structure.

There are other examples of syntactically different synonyms. Take the pair TRY and ATTEMPT, both of which allow a complement with *to*. This is optional with TRY, but obligatory with ATTEMPT (Gazdar et al, 1985:32):

- (2) a Please try/attempt to finish this work.  
b You may not be able to finish it, but please try/\*attempt.

Or consider the words ALMOST and NEARLY. Rather strikingly, VERY is possible with NEARLY, but not with ALMOST.

- (3) a I nearly/almost fell down.  
b I very nearly/\*almost fell down.

(This last example is especially interesting as a learnability problem because it came to our attention through a child's utterance containing *very almost*.) Either these pairs

are not true synonyms (contrary to first appearances), or syntactic structure is not predictable from semantic structure.

If these syntactic facts cannot be predicted from the semantics, how else might they be learned? It is not hard to see an alternative for the cases reviewed above: 'strict lexical conservatism' (Pinker 1989:17). The child extracts subcategorization facts from memory (of adult speech), without making any risky predictions at all. As far as LIKELY and PROBABLE are concerned, whatever other similarities there may be the child is guided simply by what it hears, and in the absence of precedents for PROBABLE in a raising structure like (1b), it assumes this is not possible. Similarly for TRY and ATTEMPT, where matters of optionality are directly observable, and many of the other examples.

One attraction of this explanation is that it generalises beyond complement/argument structures. Here the difference between *very nearly* and *\*very almost* is especially interesting, because VERY is not a complement, and therefore not obviously subject to subcategorization. The following explanation was suggested by Larry Trask. The words NEARLY and ALMOST are not adverbs, but members of a very small class which also contains JUST, ONLY and EVEN. (The name 'joe' has apparently been suggested for this class in recognition of these central members!) None of these words allows any modification at all, by VERY or by any other word, so if NEARLY and ALMOST belong to it as well, there is no reason to expect them to be different. From this perspective, then, the odd one out is not ALMOST, which rejects VERY, but NEARLY, which accepts it. This changes the learnability problem completely, because all the child needs is positive evidence that VERY is possible with NEARLY. We notice also that VERY is very nearly the **only** word that can modify NEARLY:

(4) I very/jolly/\*so/\*too/\*that nearly fell down.

The trouble with this conclusion is that it flies in the face of all the evidence (reviewed by Pinker) against lexical conservatism, which shows quite clearly that children do in fact generalise beyond what they actually hear. However all this evidence involves patterns for which the child receives quite overwhelming evidence, such as the various productive alternation patterns of English (passive, middle, inchoative, etc). Maybe then the correct generalisation is that children are lexically conservative until the data overwhelms them, at which point they jump with both feet and assume that all exceptions will be positive ones.

To illustrate a pattern which is below the 'overwhelm' threshold (whatever this may be), consider what we might call 'desiderative' verbs in English: WANT, LIKE,

PREFER, HATE, LOVE and a few more. These share a range of subcategorisation patterns as illustrated in (4).

- (5) a I want/like/prefer/hate/love food.  
b I want/like/prefer/hate/love to have fun.  
c I want/like/prefer/hate/love my food piping hot.

However when it comes to small clauses headed by a passive participle, HATE and LOVE seem to part company with the others.

- (6) I like/prefer/\*hate/\*love the dean told about any problems that arise.

The assumption is that although most of these patterns apply to the same range of verbs, each verb has to be observed in each pattern because the evidence for a generalisation is less than overwhelming.

However even this theory falls foul of the facts in at least two areas, both of which appear to involve overwhelmingly productive patterns which nevertheless contain lexical gaps.

At least one such gap exists in English word-formation. The addition of an agentive suffix *-er* is very productive, and can apply freely to newly-created verbs; but for some reason it is not possible with the agentive verb LOOK.

- (7) a Everyone was sitting around listening to the music, but one by one the listeners got up and went away.  
b Everyone was sitting around looking at the pictures, but one by the \*lookers got up and went away.

Our other example brings us right back to where we started, with the pair LIKELY and PROBABLE. They can both be used as the basis for an adverb, though only PROBABLE receives an extra *-ly*. (Adjectives which already end in *-ly* are often used as adverbs without any change - compare EARLY and DAILY.) However, when LIKELY is used as an adverb some speakers require it to be modified, usually by MOST (Quirk et al 1985:407).

- (8) a John will most probably/likely be late.  
b John will probably/%likely be late.

How do children learn that, in their dialect, LIKELY cannot be used as an adverb without a modifier, contrary to the general pattern for other adverbs?

## References

- Chomsky, Noam and Lasnik, Howard. forthcoming. Principles and Parameters Theory. To appear in J. Jacobs, A. von Stechow, W. Sternefeld and T. Vennemann (eds.) *Syntax: An International Handbook of Contemporary Research*, Berlin: Walter de Gruyter.
- Gazdar, Gerald; Klein, Ewan; Pullum, Geoffrey; and Sag, Ivan. 1985. *Generalized Phrase Structure Grammar*. Oxford: Blackwell.
- Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10, 279-326
- Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, MA: MIT Press.
- Hudson, Richard. 1972. Evidence for ungrammaticality. *Linguistic Inquiry* 3, 227.
- Jackendoff, Ray. 1985. Multiple subcategorization and the theta-criterion: the case of *climb*. *Natural Language and Linguistic Theory* 3, 271-295.
- Jackendoff, Ray. 1993. On the role of conceptual structure in argument selection: a reply to Emonds. *Natural Language and Linguistic Theory* 11, 279-312.
- Pesetsky, David. 1982. *Paths and Categories*. Cambridge, MA: MIT dissertation.
- Pinker, Steven. 1989. *Learnability and Cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
- Quirk, Randolph; Greenbaum, Sidney; Leech, Geoffrey; and Svartvik, Jan. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Rothstein, Susan. 1992. Case and NP licensing. *Natural Language and Linguistic Theory* 10, 119-139.