Empirical re-assessment of stimulus poverty arguments[1]

JULIE ANNE LEGATE AND CHARLES D. YANG

Abstract

It is a fact that the child learner does not entertain logically possible but empirically impossible linguistic hypotheses, despite the absence of sufficient disconfirming evidence. While Pullum & Scholz claim to have shown the existence of disconfirming evidence, they fail to demonstrate its sufficiency. By situating the acquisition problem in a quantitative and comparative framework, we show that the evidence is, after all, insufficient. Hence the argument from the poverty of the stimulus, and the innateness of linguistic knowledge, stand unchallenged.

1. Introduction

The so-called Innateness Hypothesis, which claims that crucial components of our tacit linguistic knowledge are not learned through experience but are given by our biological/genetic specifications, is not really a hypothesis. Rather, it is an empirical conclusion mainly based on observations of child language acquisition, one of which is now known as the *Argument from the Poverty of Stimulus* (APS).

---

1. A bibliographical note. This article started as a paper by the first author (Legate 1999). At that time, we were not aware of Pullum's (1996) paper; rather, the first author was interested in an empirical evaluation of the standard APS using a quantitative model of acquisition being developed by the second author. Subsequently, the second author discovered Pullum's paper and discussed it in Yang (2000). We thank Noam Chomsky for his comments on both documents. We would also like to thank the audience at the 25th Penn Linguistics Colloquium for comments and discussion. This work was partially funded by SSHRC #752-97-2087 to the first author.

The Linguistic Review 19 (2002), 151–162 0167–6318/02/019-0151

In the discussion article in this issue, Geoffrey Pullum and Barbara Scholz (2002; henceforth P&S) seek to undermine the Innateness Hypothesis by challenging the APS. However, we will argue here (in Section 3) that there are both logical and empirical flaws in P&S's spirited discussion. To begin, Section 2 will review the relevant facts.

2. The case

The logic of the APS is simple: if you know X, and X is underdetermined by learning experience, then the knowledge of X must be innate. The best known example of the APS concerns the knowledge of *structure dependency* in question inversion (Chomsky 1975; Crain and Nakayama 1987).

This instantiation of the APS has the following structure:

(1) given language data *D*, and a simple but incorrect hypothesis of *D*, *H*,
   a. the child behaves as though he/she does not entertain *H*
   b. the evidence necessary to rule out *H* is not available to the child
   ∴ the child possesses innate knowledge excluding *H* from the hypothesis space

Forming a question in English involves inversion of the main clause auxiliary verb and the subject:

(2) *The data*
   a. Is Alex *e* singing a song?
   b. Has Robin *e* finished reading?
   c. Are the boys *e* back yet?

Exposure to such sentences underdetermines the correct operation for question formation, as there are many possible hypotheses capable of generating the surface strings in (2) (Lasnik and Uriagereka's article, this volume, makes a similar point):

(3) *The hypotheses*
   a. Linear Ones
      (i) front the first auxiliary
      (ii) front the last auxiliary
      (iii) . . .
   b. Linear + hierarchical ones
      (i) front the first auxiliary following the first NP
      (ii) front the first auxiliary preceding some VP
      (iii) . . .

c. Creative ones
   (i) front the auxiliary whose position in the sentence is a prime number
   (ii) front the auxiliary which most closely follows a noun
   (iii) ...

Note that although some possible hypotheses may appear less plausible than others, if all learning is to be data-driven induction, the child must eliminate all competing hypotheses. The correct operation for question formation is, of course, structure dependent: it involves parsing the sentence into structurally organized phrases, and fronting the auxiliary that follows *the subject NP*, which can be arbitrarily long:

(4) a. Is [the woman who is singing] *e* happy?
    b. Has [the man that is reading a book] *e* eaten supper?

Of particular interest is hypothesis (3a-i) – call it the first auxiliary hypothesis – which can be argued to involve simpler mental computation than the correct generalization. Although the first auxiliary hypothesis correctly generates the data in (2), it yields erroneous predictions for the data in (4):

(5) a. *Is [the woman who *e* singing] is happy?
    b. *Is [the man that *e* reading a book] has eaten supper?

Yet children don't go astray like the inductive learner in (3a-i). They stick to the correct operation from very early on, as Crain and Nakayama (1987) showed using elicitation tasks. The children, aged between 3;2 and 5;11, were instructed to "Ask Jabba if the boy who is watching Mickey Mouse is happy", and no error like those in (5) was found.

Surely, if children hear enough sentences like those in (4), then they could reject the first auxiliary hypothesis. But if such evidence is virtually absent from the linguistic data, one can not but conclude that children do not entertain the first auxiliary hypothesis, because the knowledge of structure dependency is innate. And so goes the standard version of the APS.

3. The challenge

P&S do not challenge the logic of the APS – that much seems indefeasible. Rather, they challenge its empirical content:[2] they claim that children do encounter disconfirming evidence which serves to rule out the incorrect, structure independent, hypotheses.

2. P&S's article contains more than the usual amount of rhetoric, which strikes us as largely empty. To take one example, P&S survey the arguments for the innateness hypothesis in the

P&S examined four case studies in the spirit of the APS. Since all their arguments share the same structure, we will only consider the best known case, the inversion of auxiliary verbs.

P&S (implicitly) assume that there is only *one* alternative hypothesis to be ruled out, namely, the first auxiliary hypothesis (3a-i). As pointed out earlier, this is incorrect: the learner in fact has to rule out *all*, in principle infinitely many, hypotheses, such as those enumerated in (3), all of which are compatible with the simple auxiliary inversion examples that children most frequently encounter.[3] But for the sake of argument, let's grant that the learner has only a binary choice to make, while keeping in mind that if the learner did not have prior knowledge of structure dependency, the task of ruling out all possible hypotheses can only be more difficult than the task of ruling out the first auxiliary hypothesis.

Following Sampson (1989), P&S claim that auxiliary inversion in yes-no questions such as (6) is not the only type of sentence that rules out the first auxiliary hypothesis:

(6) Is [the boy who is in the corner]$_{NP}$ *e* smiling?

Although it is not necessarily the case, it is likely that wh-questions with an inverted auxiliary over a complex NP are also informative:

(7) How could [anyone that was awake]$_{NP}$ *e* not hear that?

P&S then proceed to count the frequency of critical evidence exemplified in (6) and (7), using a *Wall Street Journal* corpus. They find that in the first 500 sentences examined, 5, or 1 %, are of the two types. Some examples are given below:

---

literature, identifying 13 separate arguments, and observing that authors choose to focus on a subset of these arguments. The natural conclusion from this would seem to be that the innateness hypothesis is convergently and thus strongly supported. However, P&S's conclusion is rather different. They complain that "each successive writer on this topic shakes together an idiosyncratic cocktail of claims about children's learning of language, and concludes that nativism is thereby supported" (P&S Section 2.1: 12), and try to convince us that this failure of exhaustivity on the part of each author somehow weakens the arguments themselves: given this "mix-and-match presentation", "what is not clear at all is the structure of the reasoning that is supposed to get us to this [nativist] conclusion" (P&S Section 2.1: 12). Then they attempt to justify their silence on 12 of the 13 nativist arguments by declaring that "there is little chance of a comprehensive critical discussion of the supposed argument in anything less than a full-length book. It would be like fighting the hydra to tackle such a many-headed beast" (P&S Section 2.1: 14). Given this kind of reasoning, we decided to ignore the rhetoric and focus instead on P&S's main substantive claim, that in acquisition, disconfirming evidence is available to rule out logically possible but empirically impossible hypotheses.

3. See Lasnik and Uriagereka (2002) for discussion of this point.

(8) a. How fundamental are the changes these events portend? (P&S (28a))
b. Is what I'm doing in the shareholders' best interest? (P&S (30a))
c. Is a young professional who lives in a bachelor condo as much a part of the middle class as a family in the suburbs? (P&S (29a))

As they themselves note, the *Wall Street Journal* hardly reflects the composition of realistic acquisition data. After citing some anecdotal examples, they report three relevant sentences in one recorded session of conversations between a child, Nina, and her caretakers (CHILDES file NINA05.CHA):

(9) a. Where's the little blue crib that was in the house before?
b. Where's the other dolly that was in here?
c. Where's the other doll that goes in there? (P&S (34a–c))

Thus, critical evidence does exist, contrary to the previous claims P&S identified in generative linguistics and adjacent fields. According to their estimate, no less than 0.1 to 1 percent of all input sentences are of the forms (6) and (7). [4] However, an extra logical step is required: the *existence* of disconfirming evidence says nothing about its *sufficiency* to rule out the competing first auxiliary hypothesis; only the proof of the latter undermines the APS. This crucial step, as we shall show, is missing from P&S's argument.

4. The rebuttal

So how much data is sufficient? It would surely be nice to give some absolute figures, e.g., "250 of these examples will set this parameter right", but we are far from that level of understanding in language acquisition. However, there are indirect but equally revealing ways of testing for data sufficiency, which P&S fail to recognize.

Suppose we have two independent problems of acquisition, P_1 and P_2, each of which involves a binary decision. For P_1, let F_1 be the frequency of the data that can settle P_1 one way or another, and for P_2, F_2. Suppose further that children successfully acquire P_1 and P_2 at roughly the same developmental stage. Then, under any theory that makes quantitative predictions of language development, we expect F_1 and F_2 to be roughly the same. Conversely, if F_1 and F_2 turn out significantly different, then P_1 and P_2 must represent qualitatively different learning problems.

Now let P_1 be the auxiliary inversion problem. The two choices are the structure-dependent hypothesis (3b-i) and the first auxiliary hypothesis (3a-i).

4. Which, as we shall see momentarily, is a gross overestimation.

Let F_1 be the frequency of the disconfirming evidence against the latter hypothesis, and hence the frequency of sentences like those in (6) and (7). Children apparently have "learned" P_1, assuming there is inductive learning and hypothesis disambiguation, by 3;2, the age of the youngest children in Crain and Nakayama's (1987) experiment. The conclusion that the APS fails is only valid if the problem, P_1, is situated in a *comparative* setting of language acquisition. That is, we need an independent yardstick to quantitatively relate the amount of relevant linguistic experience to the outcome of language acquisition: we need to find P_2, which children also acquire at around 3;2, and F_2, the frequency of critical evidence for P_2. If P&S are right, F_1 and F_2 must be comparable.

The well-known subject drop phenomenon is a perfect candidate for P_2. There is strong longitudinal evidence that English children's subject drop stage ends at around the 36th month (Valian 1991): this is comparable to 3;2. Now, in both P_1 and P_2, the learner will make a binary choice: Valian's children have to determine whether the language uses overt subjects, and Crain and Nakayama's children would have to rule out that the language uses the structure-independent rule, "invert first auxiliary". Following the generalization that the use of *there*-type expletives correlates with obligatory subjects, *there*-type expletives have been argued to be the evidence disconfirming an optional subject grammar (Hyams 1986; Jaeggli and Safir 1989; among many others). Thus, we need only to count the frequency of *there* expletive sentences to get F_2. Based on a random sample of 11,214 adult sentences in CHILDES, we estimate F_2 to be around 1.2 percent (140/11214).

The subject drop case is not the only P_2 that comes to mind. Clahsen (1986) reports that German children start producing V2 sentences at adult level frequencies by the 36–39th month. Yang (2000), using the longitudinal data in Haegeman (1995), finds that the acquisition of V2 in Dutch also seems to be successful by 3;0–3;2, according to the standard 90 percent criterion of Brown (1973). In both cases, the critical evidence for the V2 grammar is the OVS pattern (Yang 2000), which, according to corpus statistics cited in Lightfoot (1997) as well as our own CHILDES counts, is also around 1.2 percent. [5]

5. An anonymous reviewer wonders whether 1.2 percent could actually be *more* than the required amount of evidence to learn P_2. However, if this were the case, acquisition of P_2 would occur earlier, at the point when the required amount of evidence was obtained. Thus, other problems of acquisition which have a much higher frequency of critical evidence are learned much earlier. One such case, P_2', is verb raising in French. Pierce (1992) observes adult-like behavior in verb raising at 1;6. Following the standard analysis that verb raising is evidenced by the pattern [… V Adv/Neg … O], we estimate F_2' at 7 percent (based on French CHILDES).

Returning to the APS, if P&S are right, then F_1, the frequency of (6) and (7) sentences that allegedly rule out the first-auxiliary hypothesis, should also be approximately 1.2 percent.

This takes us to the empirical problem in P&S's argument: the procedure for estimating F_1. It is rather odd that, after indicting proponents of the APS for failing to be rigorous and empirically grounded, P&S themselves opt to cite two pages of anecdotes from TV shows, newspaper clips, and plays by literary pundits. What's more, the only realistic acquisition data they give, based on the Nina corpus in CHILDES, is curiously selective: they report counts from only one file, NINA05.CHA, which happens to be the file that has the most number of critical sentences, out of all 56 files. Even for this file alone, they don't give a denominator – how many adult sentences the file contains – to give us a sense of how robustly these critical sentences are attested.

Here are the real counts. In all 56 files in the Nina corpus, we found:

(10) 46,499 sentences, of which 20,651 are questions, of which

a. None were yes-no questions of the type in (6)

b. Fourteen were wh-questions of the type in (7), exhaustively listed below:

(i) Where's the little red duck that Nonna sent you? (NINA02.CHA)

(ii) Where are the kitty cats that Frank sent you? (NINA03.CHA)

(iii) What is the animal that says cockadoodledoo? (NINA04.CHA)

(iv) Where's the little blue crib that was in the house before? (NINA05.CHA)

(v) Where's the other dolly that was in here? (NINA05.CHA)

(vi) What's this one up here that's jumping? (NINA05.CHA)

(vii) Where's the other doll that goes in there? (NINA05.CHA)

(viii) What's the name of the man you were yesterday with? (NINA10.CHA)

(ix) What color was the other little kitty cat that came to visit? (NINA28.CHA)

(x) Where's the big card that Nonna brought you? (NINA38.CHA)

(xi) And what was the little girl that came who also had whiskers? (NINA41.CHA)

(xii) Where's the card that Maggie gave you for Halloween? (NINA41.CHA)

(xiii) Nina # where are the pants that daddy sent you? (NINA43.CHA)

(xiv) Where are the toys that Mrs. Wood told you you could bring home? (NINA46.CHA)

This puts F_1 at approximately 0.068 percent:[6] that is 40 times lower than 1.2 percent, the amount of evidence needed to settle on one of two binary choices by around the third birthday.

Just to confirm that the Nina statistics are no accident, we considered another corpus, that of Adam. In an earlier paper, Legate (1999) finds the following:

(11) In a total of 20,372 sentences, 8,889 were questions, of which
  a. None were yes-no questions of the type in (6)
  b. Four were wh-questions of the type in (7):[7]
    (i) Where's the part that goes in between? (ADAM43.CHA)
    (ii) What is the music it's playing? (ADAM37.CHA)
    (iii) What's that you're drawing? (ADAM47.CHA)
    (iv) What was that game you were playing that I heard downstairs? (ADAM52.CHA)

which gives a frequency of 0.045 percent.

Not only are those frequencies far below the magic figure of 1.2 percent required to learn the correct rule by the 36th month, it is also low enough to be considered negligible, that is, not reliably available for every human child. And interestingly, the canonical type of critical evidence, [aux [$_{\text{NP}}$ ... aux ...] *e* ...], appears not even once in all 66,871 adult sentences found in both the Nina

6. P&S also argue that sentences like those in (i) below disambiguate the correct rule from the first auxiliary hypothesis:

   (i) If you don't need this, can I have it?

   If the underlying representation of (i) is [*If you don't need this, I can have it*], the first auxiliary rule would front either *don't* or *can*, producing erroneous output. However, as acknowledged in their article, this line of reasoning doesn't work if children know where sentence boundaries are, i.e., the new clause signals a fresh start. There is, however, evidence that children do recognize sentence boundaries, for which even low level acoustic cues suffice (Fisher and Tokura 1996). In any case, we only found ten such sentences in the Nina corpus, four of which contain the special symbol "#", which encodes a significant pause separating the two clauses. Even including these examples would still give a frequency far lower than 1.2 percent.

7. Of these, it is not even clear whether the equative sentences (11b-iii) and (11b-iv) necessarily count as evidence against the first auxiliary hypothesis. The child might analyze them with the wh-word in the subject position and the complex NP in the object position (although this is arguably not the analysis ascribed to these questions in adult grammar). The Nina sentences in (10b-iii), (10b-vi), and (10b-viii) are of this type as well. For completeness, we note an additional wh-question containing a complex NP in the Adam files; however, the context reveals that it is unambiguously an echo question with the wh-word in subject position:
   Adam: Dat's de funniest bird I ever saw.
   Mother: What is the funniest bird you ever saw?

and Adam corpora – the standard statements of the APS are not hyperbole as P&S charged. Hence the original APS stands unchallenged: the knowledge of structure dependence in syntax, as far as we can test quantitatively and comparatively, is available to children in the absence of experience.[8] And the conclusion then seems to be Chomsky's (1975: 33): "the child's mind … contains the instruction: Construct a structure-dependent rule, ignoring all structure-independent rules. The principle of structure-dependence is not learned, but forms part of the conditions for language learning."

5. The message

While in our opinion their challenge of the APS fails, P&S are correct in suggesting that a rigorous defense of the APS will "bring generative linguists into contact with two lines of work …: mathematical learning theory and corpus linguistics" (P&S Section 5: 46).[9] In fact, such effort is well under way, including our own work. In Yang (2000), we propose a model that views language acquisition as a probabilistic competition process among grammars circumscribed by UG. The "fitness" of competing grammars, e.g., the percentage of sentences in the primary linguistic data that each grammar can parse/analyze, can be extracted from corpus statistics. This allows one to make quantitative predictions of language development, such as the end of the subject drop stage and the successful acquisition of V2 noted earlier. In fact, our interest in the empirical status of the APS started out as a natural derivative of this line of work; see Legate (1999). However, it must be pointed out that our rebuttal of

8. In any case, the claim that children entertain the first auxiliary hypothesis for question formation is doubtlessly false. There is of course no way to prove that structure-independent hypotheses are *never* entertained. However, whatever data-driven hypothesis one conjures up, it had better crunch out some frequency counts not far from 1.2 percent.
9. Omitted from the original quote in the space of the "…" is "toward which they have traditionally been rather antipathetic" (Section 5: 46). This attribution is wildly off the mark. Mathematical learning theories, better known as learnability studies, investigate concrete models of language acquisition and their formal and empirical consequences. They have always had an important impact on linguistic theories: the Principles and Parameters framework, which constrains the child's learning space to a finite number of parameter choices, is a direct response to the learnability problem. As for corpus linguistics, one only need to flip through the pages of the generatively inclined journal *Language Acquisition*: virtually every paper includes statistics from corpora such as CHILDES. Or, go to the next Diachronics in Generative Syntax (DIGS) conference, or research the University of Pennsylvania's Middle English corpus, which has proven to be an indispensable tool for many of us in the business of language change.

P&S needn't require any specific model of acquisition: the comparability between F_1 and F_2 must hold in any adequate model.

Finally, some comments on "data-driven learning" (henceforth, DDL). What does it mean? In a strict sense, every acquisition model is data-driven; how could it be otherwise? But there surely is a world of difference between a UG-based model, which P&S call *innately primed learning*, and an inductive learner entertaining the hypotheses in (3). While P&S repeatedly pledge their neutrality on this issue, they seem to suggest that DDL, one which is devoid of innate knowledge of language, is a promising alternative to the innateness hypothesis. Yet it is important to separate promise from reality.

While linguists deal with quantitative and comparative problems in the acquisition of phonology, syntax, semantics, and pragmatics, DDL, often in the form of connectionist networks, chooses to deal with very rudimentary problems such as the learning of English irregular verbs: witness the exchanges in the past 15 years of *Cognition*. Yet, as far as we know, no connectionist model has even passed the famous Wug test (Prasada and Pinker 1993). Furthermore, DDL, proudly touted as innateness free, turns out to have many hidden assumptions – otherwise known as innate knowledge – built in; see, e.g., Marcus (1998) for a dissection of a class of connectionist models, and their inability to learn simple structural relations such as substitutability.

But a more serious problem with DDL, both present and future, has to do with the wild statistical disparities between what is presented to children and how children actually learn. As pointed out by Fodor and Pylyshyn (1988) and others, a DDL model without innate knowledge, or *learning priors*, can do nothing but recapitulate the statistical distributions of adult input. But children often learn their languages in ways that clearly defy such distributions. Take the subject drop phenomenon: while almost all adult English sentences contain a subject, children only acquire this aspect of the grammar at the much delayed age of 3:0. In contrast, as discussed in footnote 5, the placement of finite verbs over negation/adverbs is acquired by French children at a far earlier age (1;6), while this form only occurs at a frequency of 7 percent in adult French. Thus, we have one grammatical pattern that is heard frequently but learned late, and the other that is heard rarely but learned early: if there is no innate knowledge that primes children's language learning, how do we explain such statistical discrepancies? Or, consider the so-called Optional Infinitive problem (Weverink 1989). Children acquiring many languages go through an extended stage where infinitive verbs are used in co-existence with finite ones in root sentences. And yet, infinitive root sentences are vanishingly rare in adult language, for they are completely ungrammatical. Even the APS is a case in point. Why do children only entertain the hypothesis that has little clinching evidence but categorically reject a simpler one that is almost equally compatible with the adult data? Such examples are numerous in the empirical study of language

acquisition,[10] and from them it must be concluded that the innate knowledge of UG provides important learning priors to skew the distributional relations between adult language and child language.

Unless DDL models start to tackle the full range of acquisition problems with realistic acquisition data, and show that these statistical disparities can be accounted for (at least in principle), innately primed learning is, still, "the only game in town".

MIT
Yale University

References

Brown, Roger (1973). *A First Language*. Cambridge, MA: Harvard University Press.

Chomsky, Noam (1975). *Reflections on Language*. New York, New York: Pantheon.

Clahsen, Harald (1986). Verbal inflections in German child language: acquisition of agreement markings and the functions they encode. *Linguistics* 24: 79–121.

Crain, Stephen and Mineharu Nakayama (1987). Structure dependency in grammar formation. *Language* 63: 522–543.

Fisher, Cynthia and Hisayo Tokura (1996). Acoustic cues to grammatical structure in infant-directed speech: cross-linguistic evidence. *Child Development* 67: 3192–3218.

Fodor, Jerry A. and Zenon Pylyshyn (1988). Connectionist and cognitive architecture: a critical analysis. *Cognition* 28: 3–71.

Haegeman, Liliane (1994). *Introduction to Government and Binding Theory*. Malden, MA: Blackwell Publishers.

— (1995). Root infinitives, tense, and truncated structures. *Language Acquisition* 4: 205–255.

Hyams, Nina (1986). *Language Acquisition and the Theory of Parameters*. Dordrecht, the Netherlands: Reidel.

Jaeggli, Osvaldo and Kenneth Safir (1989). Introduction. In *The Null Subject Parameter*. Osvaldo Jaeggli and Kenneth Safir (eds.), 1–44. Boston, MA: Kluwer.

Lasnik, Howard and Juan Uriagereka (2002). On the poverty of the challenge. *The Linguistic Review* 19: 147–150.

Legate, Julie Anne (1999). Was the argument that made was empirical? Manuscript. MIT.

Lightfoot, David (1997). Shifting triggers and diachronic reanalysis. In *Parameters of Morphosyntactic Change*, Ans van Kemenade and Nigel Vincent (eds.), 253–273. New York: Oxford University Press.

Marcus, Gary F. (1998). Rethinking eliminative connectionism. *Cognitive Psychology* 37: 243–282.

Pierce, Amy (1992). *Language Acquisition and Syntactic Theory: A Comparative Analysis of French and English Child Grammars*. Boston, MA: Kluwer Academic.

Prasada, Sandeep and Steven Pinker (1993). Similarity-based and rule-based generalizations in inflectional morphology. *Language and Cognitive Processes* 8: 1–56.

Pullum, Geoffrey (1996). Learnability, hyperlearning, and the poverty of the stimulus. Paper presented at the Parasession on Learnability, 22nd Annual Meeting of the Berkeley Linguistics Society, Berkeley, CA.

10. And they convincingly rule out the possibility, raised by Sampson (2002), that such discrepancy may be attributed to a special statistical distribution in "Motherese".

Pullum, Geoffrey and Barbara Scholz (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 19: 8–50.

Sampson, Geoffrey (1989). Language acquisition: growth or learning? *Philosophical Papers* 18: 203–240.

Sampson, Geoffrey (2002). Exploring the richness of the stimulus. *The Linguistic Review* 19: 73–104.

Valian, Virginia (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition* 40: 21–82.

Weverink, Meika (1989). The subject in relation to inflection in child language. M.A. thesis, University of Utrecht, Utrecht, the Netherlands.

Yang, Charles D. (2000). Knowledge and learning in natural language. PhD. dissertation. MIT.