

# Steps Toward Deep Lexical Acquisition

**Sourabh Niyogi**

Massachusetts Institute of Technology

niyogi@mit.edu

## Abstract

I describe steps toward “deep lexical acquisition” based on naive theories, motivated by modern results of developmental psychology. I argue that today’s machine learning paradigm is inappropriate to take these steps. Instead we must develop computational accounts of naive theory representations, mechanisms of theory acquisition, and the mapping of naive theories to lexicalizable concepts. This will enable our theories to describe the flexibility of the human conceptual apparatus.

## 1 Where We Are Now

### The present Machine Learning Paradigm

Much of computational linguistics has converged onto a machine learning paradigm that provides us soothing clarity. The machine learning approach defines a problem as a mapping problem – map some acoustic stream onto a list of word tokens, map a list of word tokens onto a parse tree, map a parse tree onto a set of semantic roles or “logical form”, map each word in a tree onto its best sense, and so on. We then develop a learning algorithm to accomplish the desired mapping. Multiple groups describe how well their algorithm maps various test sets given various training sets, and describe a “result” to improve upon. The clarity provided by this paradigm is so soothing, one gets the sense we can turn a crank, and indeed, in many cases, progress has been made proceeding precisely along these lines.

Turning the crank on deep lexical acquisition, however, we might feel something is missing. What is it? Underlying any model of deep lexical acquisition is a theory of the *human conceptual apparatus*. Unlike our handle on acoustic streams, word lists, and parse trees, our handle on a suitable “output” for the space of word meanings is remarkably poor. Somehow, via experience (of some kind or another), children acquire a mapping from a space of vocabulary items to a space of lexicalizable concepts – the lexicon; our task as modelers is to figure out how this mapping can occur. Many models for the space of lexicalizable concepts exist: concepts are points in  $R^n$ , concepts are Jackendoff’s lexical conceptual structures, concepts are FrameNet’s frame elements, concepts are Schankian script activators, concepts are distributions over syntactic frames, concepts are grounded in sensorimotor statistics, or all of the above. Almost everyone nowadays reports how their algorithm accomplished some mapping to one or more of these models of concepts. They have to, because today’s de facto idea of what constitutes a “result” according to the machine learning paradigm today is to do exactly this.

### The Golden Oldies formed our concept models

Our models of conceptual spaces did *not* originate from computational linguists following the machine learning paradigm. They were proposed from linguists, psychologists and philosophers back in earlier eras - what we will call Golden Oldies – when the idea of a “result” was somewhat different. There are too many to recall: Quine (1960) argued that the linguist watching the natives uttering *Gavagai!* in the context of a rabbit would nec-

essarily require far more constraints than met the eye. Brown (1957) showed that children used syntactic cues to disambiguate between possible meanings; Landau and Gleitman (1985) followed on these insights, showing just how deep it could be, that even blind children could learn *look* and *see*, basing their mapping on syntactic constraints. Chomsky’s (1965) notion of “deep structure” – proposed to account for commonplace syntactic phenomena – motivated many insights explored in Gruber (1965)’s thesis, Fillmore (1968)’s classical thematic roles, and Jackendoff (1983)’s Lexical conceptual structures. Hale and Keyser and many linguists labored under the MIT Lexicon project in the 1980s to determine the fundamental features of the lexicon; many of these hard-earned observations appear in Levin (1993). Schank (1972)’s Conceptual dependency theory, Minsky (1975)’s Frames were proposed for the broader goals of capturing commonsense knowledge. Quillian’s (1968) and Miller et al (1990)’s WordNet were not intended for models of lexical acquisition or databases to be used in computational linguistics but as models of human semantic memory. Many other Golden Oldies exist, and our debt to them is quite large. Ask what motivates our collection of subcategorization statistics or what drives the quest for semantic roles, and the roots are found in the science questions of the Golden Oldies.

### The present Myopic Learning Paradigm

It would have been extremely myopic to take any one of these classical results and accuse their authors of not demonstrating a learning algorithm, not evaluating them on large corpora, and not getting together in workshops to share the results on test sets. The standard for what constituted a result back then consisted of none of these things, because today’s machine learning paradigm was just not present then. The questions were:

- Question (1): What is a lexicalizable concept?
- Question (2): How can a word-concept mapping be learned from evidence?

But for reasons that no one really talks about, somehow, the standard of what constitutes a result changed from some balance of Question (1) and (2) to a machine learning paradigm essentially focused

on Question (2). The dependency between Question (1) and (2) is quite well-understood, but do we have an adequate answer to (1)? We tell ourselves: *We’ve gotta build better parsers, speech recognizers, search engines, machine translation systems, so...* let’s take shortcuts on Question (1) so as to make progress on Question (2). For many, that shortcut consists of semantic role labels and learning from frame distributions. These shortcuts don’t answer Question (1), unfortunately.

## 2 Where We Need to Go

While the Golden Oldies were used as the foundations of today’s lexical acquisition, psychology began to sing a new tune, still balancing Questions (1) and (2).

### Children have naive theories

Developmental psychology after the Golden Oldies has shown just how deep our “deep lexical acquisition” theories have to be. On this view, word meanings are couched in changing *naive theories* of how the world works. The model of the child is that the child possesses a naive theory  $T^*$  changing state from  $T1$  to  $T2$ , and that there is a space of concepts accessible from  $T1$  that substantively different from the space of concepts accessible from  $T2$ . A learner undergoes *radical conceptual change*. Developmental psychology has not been explicit about the precise form of  $T^*$ , nor have they characterized how  $T^*$  relates to lexicalizable concepts. But their contributions inform us about the fundamental ingredients of concepts (Question (1)) and inform us what deep lexical acquisition must consist of (Question (2)).

A few examples must suffice in place of a review (c.f. Gopnik and Meltzoff (1997)). Keil (1989)’s transformation studies illustrate theory change in the domain of biology. First, children are shown a picture of a skunk; then, are told a story – that the animal received either (A) surgery or (B) a shot in infancy – and then are shown a picture of a raccoon. Young preschool children judge that the animal is a raccoon, as if they base their judgements on superficial features. Children between 7 and 9 ( $T2$ ) on the other hand, judge that the raccoon-looking figure in (A) is still a skunk. Adults ( $T3$ ) judge that the raccoon-looking figure in both conditions is still a skunk. Apparently, preschoolers’ theory  $T1$

lacks the belief that an animal's kind is determined at birth, but this becomes part of the adult's  $T3$ .

Similarly, preschool children at  $T1$  have concept of *death* involving a belief in a continued existence in an alternate location (like sleep); When asked whether dead people dream, eat, defecate, and move, 4 to 6 year olds will say that dead people do all of these, except move (Slaughter et al, 2001). Missing in  $T1$  are the causes of death (a total breakdown of bodily functions) and that death is an irreversible, inevitable end. Between 4 and 6, children become superficially aware of the general function of various body parts (e.g. "You need a heart to live"). Other phenomena serve the same point: the child at  $T1$  thinks *uncle* means friendly middle-aged man, and at  $T2$  thinks it means parent's brother. The child at  $T1$  thinks *island* means a beachy territory and at  $T2$  thinks it means body of land surrounded by water (Keil 1989). And, "theory of mind" concepts/words such as *belief*, *desire*, *wonder*, *pretend* (Wellman and Bartsch 1995, Leslie 2000) are similarly situated.

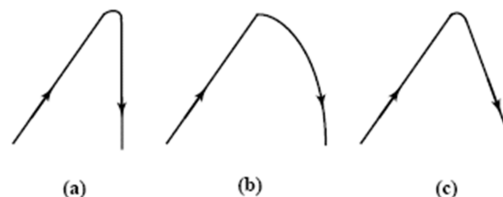
How "theory-like"  $T1$  and  $T2$  are is subject to considerable debate (diSessa 1993, Leslie 2000). disessa (1993) describes a large number of causal "p-prims" that are highly context specific and considerably larger in number than what Carey (1985) describes; these are shown to apply to everyday physical phenomena – "force as mover", "vacuums impel", "overcoming", "springiness", "bigger means lower pitch (or slower)", to name a few. Each of these have a FrameNet-like causal syntax, of some unknown mapping to vocabulary items. Similarly, Rozenblit and Keil (2003) show that non-expert adults have a remarkably superficial notion of how common mechanisms work – such as how a helicopter changes from hovering to forward flight. Theories may be suspiciously weak.

### Students have alternative frameworks

Educational psychologists have characterized  $T^*$  by asking a different, more practical question: why is it difficult for science students to learn certain scientific concepts (*weight*, *density*, *force*, *heat*, ...) when they come to class? The broad insight is this: students come to class not as blank slates but with alternative *pre*-conceptions that must be understood. Data on their pre-conceptions yields clues as to con-

tents of  $T^*$ , well before they walk into science class. Again, a few examples illustrate the point.

Many studies on physics misconceptions have observed deeply held views on the motion of projectiles (McCloskey 1983, Halloun and Hestenes 1985). Ask students to predict what happens when a projectile is thrown upward at an angle, and their answers will typically be consistent with one of (a-c)



These answers are consistent with an "impetus" theory of motion, where an object's motion is exclusively dominated by whatever "impetus" the thrower provides it. Medieval scientists such as Buridan also held similar beliefs; Newtonian mechanics, of course, shows that the answer is a parabola. disessa (1993) report a wider array of these types of physics misconceptions in a theoretical framework.

Likewise, ask students for their knowledge of how their eyes work, and they reveal an "extramission" belief: something somehow shoots out from the eye and reaches the objects (Winer et al 2002); they also say that eye is the sole organ in the body responsible for vision. Plato and da Vinci shared these same beliefs. Systematic catalogues of these sorts of observations have been compiled for just about every domain – e.g. megaphones create sounds, heat is a substance, eggs are not alive, the moon and sun are the same size, and so forth (AAAS 1993).

### 3 What Steps We Must Take

Consider this fascinating phenomena from the Best of Today and the comfort of the grammar-generates-sentence relation will be replaced by queasiness: the terms *theory*, *concept*, and *change* are most unclear, as many developmental psychologists freely admit. But computational linguists may contribute significantly to rendering new clarity: If the Golden Oldies drove the efforts on today's shallow lexical acquisition, the Best of Today's Psychology may drive the results of tomorrow's progress in deep lexical acquisition.

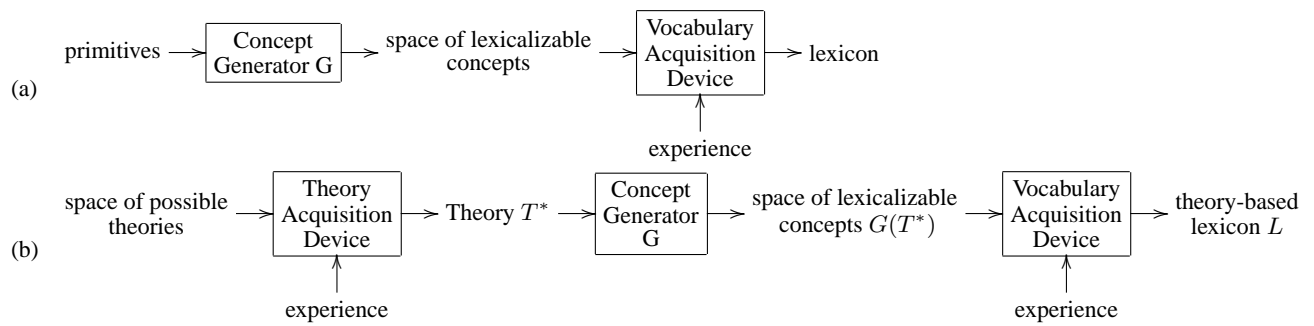


Figure 1: (a) The Model of Concepts from the Golden Oldies: used in the present Machine Learning Paradigm; (b) The Universal Theory Model of Concepts: necessary for deep lexical acquisition

### The new framework: Universal Theory

We have much progress to make: We can *describe* naive theories precisely; we can *describe* how theory acquisition occurs; we can *describe* the map from naive theories to a set of lexicalizable concepts. We can *describe* how vocabulary acquisition occurs. Figure 1(a) shows the Golden Oldies model of concepts that we must abandon: a Vocabulary Acquisition Device receives a fixed hypothesis space of possible concepts completely determined by a fixed set of primitives; Figure 1(b) shows the *Universal Theory Model of Concepts* that we must take steps towards: A *Theory Acquisition Device* (TAD) outputs a state  $T^*$  that describes a learners’s naive theory; A *Concept Generator*  $G$  maps  $T^*$  to a set of lexicalizable concepts  $G(T^*)$ . A *Vocabulary Acquisition Device* (VAD) uses  $G(T^*)$  to learn a lexicon. The theory of the TAD states is *Universal Theory* (UT); a UT metalanguage enables an abstract characterization of *possible theories* – each possible theory describes a system of kinds, attributes, relations, part-whole relations, and causal mechanisms. Within this *Universal Theory Model of Concepts*, we can begin to answer the following core questions:

1. what is the initial state of the TAD?
2. what are possible final states of the TAD?
3. how can the TAD change state?
4. how can the TAD use  $T^*$  to parse experience?
5. how does the concept generator  $G$  map  $T^*$  onto a set of lexicalizable concepts  $G(T^*)$ ?
6. how can the VAD use  $G(T^*)$ ?

### We have made progress on these core questions

Many of these questions have been addressed already in computational models where a candidate

UT metalanguage and theory  $T^*$  is latent. diSessa (1993) catalogs sets of p-prims in naive physics. Atran (1995) describes a theory of family structure. Gopnik et al (2004) uses Bayesian networks to model preschooler’s causal reasoning about *blickets*. McClelland and Rogers (2004) describe connectionist models of some of Carey (1985)’s classic results.

In my own work, I have been situating the elements of the Universal Theory Model of Concepts in a microgenesis study, where adult subjects undergo a  $T1$  to  $T2$  transition (Niyogi 2005). The transition can be understood with a minimal UT metalanguage needed to characterize a set of possible theories:  $T^*$  is characterized by a interrelated sets of kinds, attributes, relations, and causal laws.  $T1$  and  $T2$  are described in that UT metalanguage, and the simplest concept generator  $G$  is described that mechanically maps  $T1$  and  $T2$  onto  $G(T1)$  and  $G(T2)$ . Subjects undergo theory change in a Blocksworld universe (see Figure 2(a)) while learning 3 verbs (*gorp*, *pilk*, *seb*) that refer to the causal mechanisms governing the universe. Subjects interact with a set of 29 blocks, some of which activate other blocks on contact. On activation, subjects are shown a transitive verb frame (“Z is *gorping* L”, “U is *sebbing* F”, “D is *pilking* Y”) in a Word Cue Area. Unbeknownst to subjects, each block belongs to 1 of 4 kinds (A, B, C or D) and 3 activation mechanisms exist between them: *lawab*: As activate Bs, *lawc*: Cs activate Cs, and *lawd*: Ds activate Ds; each of the 3 verbs refers to one the 3 mechanisms. Subjects are probed for the naming conditions on each of the 3 verbs.

Subjects’ responses indicate that their TAD state changes from  $T^* = T1$  (there is 1 kind of block governed by 1 causal mechanism *lawq*) to  $T^* = T2$

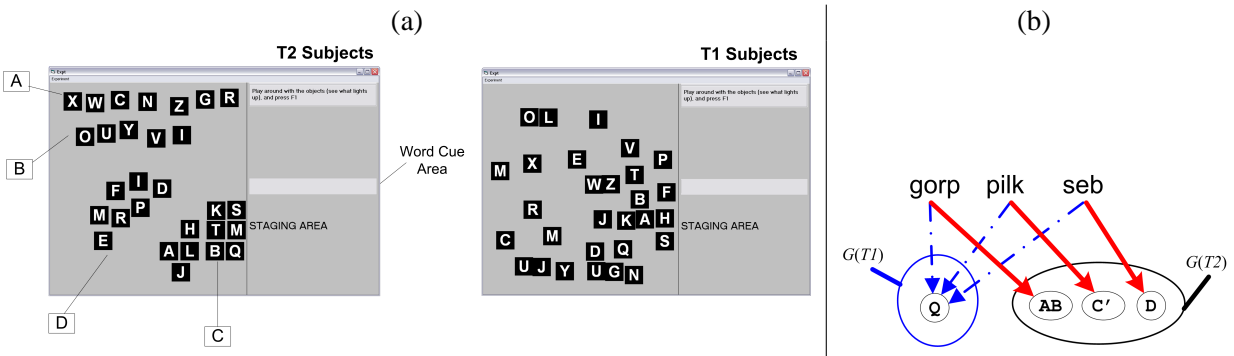


Figure 2: (a) Subjects try to learn the laws and word meanings in a “Causal Blocksworld” computer application by dragging and dropping blocks onto each other. Cues to the meaning of 3 verbs (*gorp*, *pilk* and *seb*) are given in a Word Cue Area. Shown is how two kinds of subjects –  $T2$  Subjects and  $T1$  Subjects – clustered the blocks; the clusters for the kinds A, B, C and D (boxed) are clear for  $T2$  Subjects but no such differentiation is apparent for  $T1$  subjects; (b) When  $T^* = T1$ , all 3 verbs can only be mapped to a single concept in  $G(T1) = \{Q\}$  (dashed arrows); When  $T^* = T2$ , *gorp*, *pilk* and *seb* can be mapped to 3 new concepts AB,  $C'$  and D in  $G(T2)$  (solid arrows).

(there are 4 kinds of blocks governed by 3 distinct causal mechanisms, *lawab*, *lawc'* and *lawd*). But this is not true for all subjects: some remain “ $T1$  subjects” while others move onto become “ $T2$  subjects”. Critically, when  $T^* = T1$ , the verbs can only be mapped to a single concept in  $G(T1) = \{Q\}$ ; When  $T^* = T2$ , the verbs can be mapped to 3 distinct concepts in  $G(T2) = \{AB, C', D\}$  (See Figure 2(b)). Once  $T^* = T2$ , subjects can “parse” the activation and infer the hidden kind and causal mechanism involved. Critically, subjects cannot learn to distinguish the 3 verbs until  $T^* = T2$ , when the 3 new concepts emerge in  $G(T^*)$ . Then *gorp*, *pilk* and *jeb* may be mapped onto those 3 new concepts. These verbs are thus theory-laden in the same way as *death*, *uncle* and *island*.

This UT architecture concretely *dissolves* the Puzzle of Concept Acquisition (Laurence and Margolis 2002): how can a person ever acquire a “new” concept, when a fixed set of primitives exhaustively span the space of possible concepts? Taking the viewpoint of the learner’s VAD at a specific moment in time with a *specific*  $T^*$ , it has access to *just* those concepts in  $G(T^*)$  – acquisition of a new concept is possible if  $T^*$  changes. Taking the viewpoint of the learner’s *species* across all possible times, the species has access to the *union* of  $G(T^*)$  over all possible TAD states – thus a “new” concept for the species is impossible. Which viewpoint one takes is a matter of perspective. Critically, the Golden Oldies model of concepts does not expose the TAD state re-

vealed in the UT model of concepts (Fig. 1a,b).

### Universal Theory and the Linguistic Analogy

Computational linguists can progress on these questions, because naive theories are like grammars. Just as a grammar generates a set of possible sentences, a theory  $T^*$  generates a set of possible worlds. Just as the space of possible grammars is restricted, so is the space of possible theories. Just as learning a grammar consists of picking a point from a space of possible grammars, learning a theory consists of picking a point from the space of possible theories. The task of writing a naive theory is like writing a grammar. The task of characterizing the space of possible theories requires a theory metalanguage just as characterizing the space of possible grammars requires a grammar metalanguage.

Moreover, research into naive theories does not proceed *separately* from the program of research in grammar. The two programs are bridged by the concept generator  $G$ :  $T^*$  generates  $G(T^*)$ , a set of lexicalizable concepts. An adequate account of  $G$  would generate concepts present in a particular language, for every language, and for every possible  $T^*$ .

Miller et al (1990) distinguish between a constructive and a differential lexicon. In a *differential* theory of the lexicon, meanings can be represented by any symbols that enable a theorist to distinguish among them; In a *constructive theory* of the lexicon, the representation should “contain sufficient information to support an accurate construction of the concept (by either a person or a machine)”.

The conceptual analyst who desires to produce a *constructive* theory of the lexicon has four kinds of accounts to provide: (see Niyogi 2005)

- an explanatory account of the space of possible theories, for all persons P
- an explanatory account of the space of possible concepts, for all persons P, for all possible theories
- a descriptive account of a specific theory  $T^*$  held by a representative person P (e.g. of a 3-year old or of a 10-year old)
- a descriptive account of a specific lexicon  $L$  held by a representative person P (e.g. a 3-year old Chinese speaker, 3-year old English speaker, 10-year old Chinese speaker, 10-year old Chinese speaker)

We may envision a “theory-based lexicon” that would capture the *two* key state variables in Figure 1(b), the two descriptive accounts above: (1)  $T^*$  for an idealized human; (2) a set of vocabulary items mapped to points in  $G(T^*)$ . Very limited instances of a theory-based lexicon can be constructed already for subjects at the end of the experiment – such a theory-based lexicon has (1)  $T2$  in the UT metalanguage; (2) the mapping in  $L$  to  $G(T2)$ : *gorp* = AB, *pilk* = C', *seb* = D. This *constructive* theory-based lexicon would be in stark contrast to *differential* lexicons such as WordNet and FrameNet.

### Grounding language in perception is insufficient

Many have proposed deep lexical acquisition by “grounding language in perception” (Siskind 1996, Regier 1996, Roy and Pentland 2002, Yu and Ballard 2004), constructing systems that can learn to utter, e.g. *red*, *banana*, *hit* and *triangle* in contexts where there are, e.g., three triangles hitting red bananas. Such systems also propose a space of possible concepts exhausted by a *fixed* set of primitives, as in the Golden Oldies model. The initial state of the TAD ( $T^*(t = 0)$ ) can explicitly incorporate all these attributes and relations (contact, luminance, . . .); but then, the TAD can *further* change state to yield new kinds, attributes, relations, and causal mechanisms not present in the initial state, but motivated by the data (see Gopnik and Meltzoff 1997). As such, vague appeal to grounding is insufficient; associative processes that may work on *red*, *hit*, *ba-*

*nana*, *eye*, *three* are extremely challenging to generalize to *color*, *kind*, *wonder*, *pilk*, *seb*, *telescope*, *maybe* and uninvented *groobles* that cannot be perceived. Again, developmental psychology provides some insight on what theoretical innovations would be required for a suitable interface to sensorimotor apparatus (c.f. Mandler 2004).

### Commonsense AI gives UT foundations

Primitives well beyond the sensory apparatus have been developed to describe physical systems qualitatively (Regier 1975, Forbus 1984). They show us some of the possibilities of what  $T^*$  and candidate UT metalanguages may look like (quantity spaces, kinds, attributes, relations, part-whole relations, and causal mechanisms that interrelate these sets). Regier (1975)’s description of a *toilet* appears particularly close to Rozenblit and Keil (2003)’s *helicopter*. Later qualitative AI frameworks of Forbus (1984) and Kuipers (1994) may be applied to McCloskey (1982)’s intuitive physics and disessa’s (1993) p-prims. Except for the work of Hobbs, Pustejovsky and their colleagues, few have mapped commonsense theories onto the lexicon. Similar domain-general elements of naive math and causality are present in the works of Hobbs et al (1987), Kennedy and McNally (2002)’s degree representations for gradable predicates, Talmy (1988)’s force dynamics, and the quantity spaces of Kuipers (1994) and Forbus (1984). These disparate frameworks provide foundational elements for a UT metalanguage.

### Shortcuts on UT foundations will not work

We must resist the urge to take shortcuts on these foundations. Simply creating slots for foundational phenomena will impede progress. Pustejovsky (1995)’s observations for co-composition have clearly illustrated how much flexibility our interpretation systems must have, e.g. in *He enjoyed the beer/movie*. But specifying the telic role of *beer* and *movie* to be *drink* and *watch* does not constitute an adequate theory – we require constraints that relate to the state space of the human conceptual apparatus. Pustejovsky (1995)’s telic, formal, constitutive, agentive roles may be mapped onto  $T^*$ ’s characterization of artifacts, materials, and so on. We require nothing less than absolute conceptual transparency.

## We must bridge UT to analogy

Lakoff and Johnson (1980) and subsequent cognitive linguistics work have catalogued a stunning level of metaphoric usage of language. Lexical extension of items such as *illuminate* in, e.g. *Analogies illuminate us on theory acquisition* are couched in terms of conceptual metaphors such as “ideas are light”. Significant steps have been taken to model analogical mapping (c.f. Falkenhainer et al 1989, Bailey et al 1997) and conceptual blending (Fauconnier and Turner 1998). These processes may motivate TAD state changes. In most cases, the underlying predicates in the source and target domains are ad hocly constructed; a natural source of these predicates may be the sets internal to  $T^*$  (kinds, attributes, relations, causal mechanisms); similarity between domains may be determined by the structural properties of the UT metalanguage and  $G$ . If  $T^*$  incorporates the common causal mechanisms behind ideas and light transmission, for example, then one may strive for a shorter lexicon where the vocabulary item *illuminate* happens to be used in both domains with “one” core entry. An adequate theory of this process would obviously reduce the number of so called “senses” in word sense disambiguation.

## 4 What We Assumed Wrong

Modern computational linguistics appears to have made a set of assumptions that deserve reanalysis, given the availability of other options.

### **Assumption: A fixed alphabet of meaning components exists, and we know what it is**

A key assumption dating to the Golden Oldies is that the meaning of a sentence is adequately captured by a “logical form” (LF) characterized by a *fixed alphabet* of meaning components (e.g. thematic roles, lexical semantic primitives, conceptual dependency primitives). Today’s computational linguistics program uses this assumption to demonstrate systems that answer “who did what to whom, where, why, . . .” questions, given sentences like:

*John saw the man with the telescope.*

*John hit the man with the umbrella.*

Is the computational linguist expected to be satisfied when systems can answer *Who saw the man with the telescope?* or *Who did John hit with the umbrella?* This year’s CoNLL Shared Task, mapping

sentences onto semantic roles, assumes the above. But try these: Does John have eyes? Were they ever open when he was looking through the telescope? Could John know whether the man was wearing underwear? Did the umbrella move? Did John move? Did the man feel anything when he was hit? Was John alive? Was the man alive? Why would John need a telescope to see the man, when he has eyes? Why would John use an umbrella when his hands would do? Something is missing in these systems.

We should be more accountable. Developmental psychology showed that theory change and conceptual change is possible, proving this assumption is *wrong*: the alphabet behind sentence meaning is a *varying* set of lexicalizable concepts  $G(T^*)$ . Missing in today’s systems attaching AGENT (or FrameNet’s Perceiver\_passive, or Impactor) to *John* and INSTRUMENT to *umbrella* and *telescope* is  $T^*$ , and a mapping of the lexical items to  $G(T^*)$ . What  $T^*$  must contain, in some as yet unknown form, is a T of physics described by McCloskey and disessa (1993), a T of vision studied by Landau and Gleitman (1985) and Winer et al (2002), a T of body studied by Carey (1985), a T of materials and artifacts studied by Hobbs et al (1987) and Pustejovsky (1995). This  $T^*$ , when mapped via  $G$ , forms the alphabet of the above 2 sentences.

### **Assumption: The machine learning paradigm can treat deep lexical acquisition.**

If we reject the assumption that there is some “meaning” of a sentence spanned by a set of meaning primitives, the soothing clarity of the machine learning paradigm is no longer available. We cannot map parse trees onto sentence meanings. The possibility of “Putting Meaning in Your Trees” (Palmer 2004) completely disappears. We may still use the machine learning paradigm to parse, disambiguate and recognize speech. But these results are of little use to model theory, concept and lexical acquisition, because there is no output representation where a suitable training set could be collected. The human conceptual apparatus is not that simple: the VAD requires  $G(T^*)$  (which changes, as  $T^*$  changes), and for *that* we need explanatory accounts of UT and  $G$ , and must recognize the diverse ways the TAD may change state.

**Assumption: Paths from shallow to deep lexical acquisition exist**

The Golden Oldies Models of concepts (Figure 1a) and the Universal Theory models of concepts (Figure 1b) are *incommensurable*. The path from the shallow to the deep cannot be declared to exist by fiat. Wishful thinking is inappropriate, because one architecture is more powerful than the other: the Golden Oldies model did not expose the TAD state space. Instead, lexical semantics results obtained under the Golden Oldies model require translation into the UT model: the privileged position syntactic positions that motivated thematic roles and lexical semantics primitives, the bi-partite event structure revealed through adverbial modification, and so on. This translation is mediated in *G*, and will not yield a notational variant of what we started with.

**Assumption: Verb classes determine meanings**

We must distinguish between a representation of verb meanings *determined* by the distribution of subcategorization frames and *cued* by these frames. Landau and Gleitman (1990) showed that verb’s participation in some frames but not others are *cues* that a child uses to constrain verb meaning. Levin and Rappaport-Hovav (1998) explicitly distinguish *structural* and *idiosyncratic* components of meaning. But neither claim that verb classes or statistical distributions of subcategorization frames *determine* verb meaning. Yet VerbNet maps verbs to predicates in precisely this way: (Kingsbury et al 2002).

<i>cure, rob, . . .</i> : Verbs of Inalienable Possession cause(Agent,E) location(start(E),Theme,Source)
<i>marry, divorce, . . .</i> : Verbs of Social Interaction social_interaction(. . .)

The distinction between *cure* and *rob*, or between *marry* and *divorce* is not astonishing to the English speaker. Causal mechanisms behind disease, possession, and the marital practices that were labeled *idiosyncratic* by the lexical semanticist must be captured in *T\**.

**Assumption: Language is separate from general systems of knowledge and belief**

This “defining” assumption helped for the Golden Oldies, but innovations in developmental psychology motivate dropping this assumption. The bridge is provided by the concept generator *G*: it maps a naive theory *T\** (general systems of knowledge and

belief) to *G(T\*)*, used by the VAD (language).

**Assumption: Real-world knowledge is Bad**

The absence of the soothing clarity of the machine learning paradigm and presence of real world knowledge in *T\** brings forth 2 associations:

Early Schank/Cyc = Much Knowledge = UT research = Bad  
Statistics = Little Knowledge = shallow semantics = Good

The associations lead to the inference that Universal Theory research will suffer a similar fate as the 70s Schankian program and the Cyc program (Schank 1972, Lenat and Guha 1990). However, this inference is incorrect. The 70s Schankian program and Cyc efforts did not carefully consider the constraints of syntactic phenomena or developmental psychology. Schank and his colleagues stimulated research in qualitative physics and explanation-based learning that addressed many of these deficiencies, but there is much work to be done to bridge today’s efforts in deep lexical acquisition to this.

**Assumption: Others will provide us the answers**

Lexical semanticists now rely on cognitive explanations far more heavily than ever before. Jackendoff (2002) concludes: “someone has to study all these subtle frameworks of meaning - so why not linguists?” Levin and Rappaport-Hovav (2003), addressing denominal verbs such as *mop* and *butter*, now freely point to “general cognitive principles” rather than situate knowledge in the lexicon. Rather than consume lexical semantics of the Golden Oldies, we can draw upon our toolbox to again answer Question (1): “what is a lexicalizable concept?”

**5 We Must Change Our Concepts**

Stop working with models of concepts from the Golden Oldies. Start questioning whether results under the machine learning paradigm are really *results*. Change your concept of a *result*. Learn how children do theory, concept and vocabulary acquisition. Expose the fundamental ingredients of concepts. Change your concept of *deep*. Change your concept of *computational linguistics*. Radical conceptual change is possible. Write some new songs, and sing some new tunes. We can have some Great Golden Oldies of Tomorrow.



## References

- S. Atran. Classifying nature across cultures. In E. Smith and D. Osherson, editors, *Thinking: An invitation to cognitive science*, Cambridge, MA, 1995. MIT Press.
- D. Bailey, J. Feldman, S. Narayanan, and G. Lakoff. Modeling embodied lexical development. In *Proceedings of the Annual Cognitive Science Society*, 1997.
- K. Bartsch and H. Wellman. *Children Talk about the Mind*. Oxford University Press, New York, 1995.
- R. Brown. Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology*, 1957.
- S. Carey. *Conceptual Change in Childhood*. MIT Press, Cambridge, MA, 1985.
- N. Chomsky. *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA, 1965.
- A. M. Collins and M. R. Quillian. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8:240–247, 1969.
- B. Falkenhainer, K. Forbus, and D. Gentner. The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, 41:1–63, 1989.
- G. Fauconnier and M. Turner. Conceptual integration networks. *Cognitive Science*, 22(2):133–187, 1998.
- C. Fillmore. The case for case. In E. Bach and R. Harms, editors, *Universals in Linguistic Theory*, pages 1–90, New York, 1968. Holt, Rinehart and Winston.
- C. Fillmore, C. Wooters, and C. Baker. Building a large lexical databank which provides deep semantics. In *Proceedings of the Pacific Asian Conference on Language, Information and Computation*, Hong Kong, 2001.
- K. Forbus. Qualitative process theory. *Artificial Intelligence*, 24:85–168, 1984.
- D. Gentner. Why we're so smart. In D. Gentner and S. Goldin-Meadow, editors, *Language in mind: Advances in the study of language and thought*, pages 195–235, Cambridge, MA, 2003. MIT Press.
- A. Gopnik, C. Glymour, D. Sobel, L. Schultz, and T. Kushnir. Theory formation and causal learning in children: Causal maps and bayes nets. *Psychological Review*, in press.
- A. Gopnik and A. Meltzoff. *Words, thoughts and theories*. MIT Press, Cambridge, MA, 1997.
- J. Hobbs, W. Croft, T. Davies, D. Edwards, and K. Law. Commonsense metaphysics and lexical semantics. *Computational Linguistics*, 13:241–250, 1987.
- R. S. Jackendoff. *Semantics and Cognition*. MIT Press, Cambridge, MA, 1983.
- R. S. Jackendoff. *Foundations of Language*. Oxford University Press, Oxford, 2002.
- F. Keil. *Semantic and Conceptual Development: An Ontological Perspective*. Harvard University Press, Cambridge, MA, 1979.
- C. Kennedy and L. McNally. Scale structure and the semantic typology of gradable predicates. *Language*, 2002.
- P. Kingsbury, M. Palmer, and M. Marcus. Adding semantic annotation to the penn treebank. In *Proceedings of Human Language Technology Conference*, 2002.
- B. Kuipers. *Qualitative Reasoning*. MIT Press., Cambridge, MA, 1994.
- B. Landau and L. R. Gleitman. *Language and experience: Evidence from the blind child*. Harvard University Press, Cambridge, MA, 1985.
- S. Laurence and E. Margolis. Radical concept nativism. *Cognition*, 86:22–55, 2002.
- D. Lenat and D. Guha. *Building large knowledge-based systems: Representation and Inference in the Cyc Project*. Addison-Wesley, Reading, MA, 1990.
- A. Leslie. How to acquire a representational theory of mind. In D. Sperber, editor, *Metarepresentations: An Multidisciplinary perspective.*, pages 197–223, Oxford, 2000. Oxford Press.
- B. Levin. *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press, Chicago, IL, 1993.
- B. Levin and M. Rappaport-Hovav. Objecthood and object alternations. *ms*, 2003.
- J. Mandler. *Foundations of Mind: Origins of Conceptual Thought*. Oxford University Press, New York, 2004.
- M. McCloskey. Intuitive physics. *Scientific American*, 248:122–130, 1983.
- G. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. Miller. Five papers on wordnet. *International Journal of Lexicology*, 3(4), 1990.
- M. Minsky. A framework for representing knowledge. In P. Winston, editor, *The psychology of Computer Vision.*, pages 211–277, New York, 1975. McGraw-Hill.
- N. Nersessian. Comparing historical and intuitive explanations of motion: Does naive physics have a structure? In *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society*, pages 412–420, 1989.
- S. Niyogi. Aspects of the logical structure of conceptual analysis. *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*, 2005.
- S. Niyogi. The universal theory model of concepts and the dissolution of the puzzle of concept acquisition. *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*, 2005.
- M. Palmer. Putting meaning in your trees. In *CoNLL-2004*, 2004.
- J. Pustejovsky. *The Generative Lexicon*. MIT Press, Cambridge, MA, 1995.
- W. Quine. *Word and Object*. MIT Press, Cambridge, MA, 1960.
- T. Regier. *The Human Semantic Potential*. MIT Press, Cambridge, MA, 1996.
- T. Rogers and J. McClelland. *Semantic Cognition: A parallel distributed Processing approach*. MIT Press, Cambridge, MA, 2004.
- D. Roy and Pentland. Learning words from sights and sounds: A computational model. *Cognitive Science*, 26:113–146, 2002.
- L. Rozenblit and F. Keil. The misunderstood limits of folk science: an illusion of explanatory depth. *Cognitive Science*, 26:521–562, 2002.
- R. Schank. Conceptual dependency theory. *Cognitive Psychology*, 3:552–631, 1972.
- J. Siskind. A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, 61:39–91, 1996.
- V. Slaughter, R. Jaakola, and S. Carey. Constructing a coherent theory: children's biological understanding of life and death. In M. Siegel and C. Peterson, editors, *Children's understanding of biology and health*, Cambridge, 1999. Cambridge University.
- V. Slaughter, R. Jaakola, and S. Carey. Constructing a coherent theory: children's biological understanding of life and death. In M. Siegel and C. Peterson, editors, *Children's understanding of biology and health*, Cambridge, 1999. Cambridge University.
- C. Yu and Dana H. Ballard (2004) A Unified Model of Early Word Learning: Integrating Statistical and Social Cues. *Proceedings of the 3rd International Conference on Development and Learning*, 2004.