Simulating Language Games of the Two Word Stage using Unification and Substitutions on a Corpus of Exemplars with a focus on Semantics

. . . being an endeavor in cognitive simulation to parsimoniously re-enact verbal interactions of a toddler through translation and reckoning with pragmatic and semantic annotations of its linguistic history.

Andreas van Cranenburgh*

June 26, 2009

March - June 2009 Bachelor thesis Artificial Intelligence
Supervisors: Remko Scha & Jos de Bruin,
University of Amsterdam

*acranenb@science.uva.nl, 0440949

Abstract

An exemplar-based model of language acquisition is presented, based on a minimalist meaning representation. The model demonstrates that semantics and pragmatics in combination with unification and substitution mechanisms for recombination can account for both comprehension and production in the two word stage, as well as mimicking the discrepancy in performance between language comprehension and production of children in that stage. The model is evaluated by comparing its reaction to data from the Childes corpus, as well as by demonstrating the interpretation of novel utterances. Results seem to indicate that sensible utterances can be interpreted correctly, whereas non-sensible utterance get rejected as interpretation fails.

Try to attach a meaning
To words that you've heard

Stumbling through the dark
Seems I'm stumbling through the dark
Everybody's stumbling through the dark

The men who proceeded us here
Left only questions and fears
[...]

— from the album *Rainy day music*, The Jayhawks (2003)

Acknowledgements

First and foremost I have to acknowledge the crucial commentary of my two supervisors, Remko Scha and Jos de Bruin. Our discussions have been fruitful and inspiring. Many shortcomings they have pointed out have been addressed. Any remaining errors are surely my own. I would also like to thank Wouter van Atteveldt for reading and commenting on drafts.

Contents

1 Introduction

General linguistics has been dominated by Chomskian generative linguistics (e.g., Chomsky 1975) for several decades. The focus is on rules and their creativity, viz. systematicity and productivity. The central dogma is that an in-born, Universal Grammar is necessary to adequately explain these phenomena. It holds on to the continuity assumption[1], which states that language as used and understood by children is qualitatively equal to that of adults (for criticism, see Tomasello 2005).

However, from a developmental psychology angle, several empirical findings (Tomasello, 2000; 2005) shed doubt on whether this approach is applicable to language acquisition by children. It rather appears that language learning is bootstrapped in a haphazard fashion, learning constructions[2] here and there, which can only later be synthesized to form a coherent grammar.[3]

Rather than trying to resolve this age-old debate between rationalism and empiricism along theoretical lines, it might be fruitful to try to model the behavior of early language users, and demonstrate in this way that a universal grammar is in fact not necessary to explain the phenomena observed. This strategy echoes a suggestion made by Turing (1950):

[1] Two kinds of continuity should probably be distinguished, the first being a static kind of continuity where a set of rules is present at birth and continues to be employed unchanged throughout the life of an individual, the second being a dynamic continuity where the same general-purpose mechanisms are used to actively develop one's grasp on language throughout the life of an individual (de Kreek, 2003). It is only the former that is criticized here.

[2] A construction consists of collocation of two or more words paired with a schematized meaning, which can be used to interpret and produce a class of utterances.

[3] Note that this grammar might remain implicit (except for what is taught explicitly in grammar school, that is), because language is not essentially a game of rule-following, but rather of successful expression and communication. More on this in section 2.4.

> "Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? [...] Presumably the child-brain is something like a note-book as one buys it from the stationers. Rather little mechanism, and lots of blank sheets."

2 Theory

2.1 Literature review

One of the foremost proponents of the developmental take on language acquisition is Tomasello (2005). He argues that linguistic abilities are acquired gradually, in an incremental fashion. Linguistic forms are memorized in conjunction with their communicative functions or meanings. These constructions are then generalized so that language use becomes ever more expressive and productive. Aspects which distinguish this approach from that of generative linguistics (Chomsky, 1999; Cook, 1988) is the rejection of the autonomy of syntax and the consequential focus on semantic and pragmatic influences on learning. Aside from that the idiomatic and figurative dimension of language presents problems for purely formal accounts of semantics and syntax. [4]

The formal nature of traditional theories goes back to American (Bloomfieldian) structuralism and the supposed arbitrariness of the sign (Bloomfield, 1914). A counter-argument to the arbitrariness of the sign is that derived (e.g., figurative) meanings are relatively systematically related to their canonical meanings. For example, the verb 'to come' has the canonical interpretation of spatial movement, but it can also be applied to an event which is temporally approaching: "Christmas is coming" (Lakoff and Johnson, 1999). Notice how 'approaching' is also a spatial verb, and can be analogously applied with a temporal interpretation. On top of this, analyses of semantic networks such as Wordnet indicate that semantics is scale-free, i.e. it exhibits the small-world phenomenon and a fractal-like self-similarity on all scales. This entails that related words form local clusters sparsely interconnected by hubs with short average path length. This discovery suggests that lexical semantics is a bipartite, scale-free graph connecting words and concepts (Steyvers and Tenenbaum, 2005), rather than an arbitrary mapping between the unordered sets of words and concepts.

The work of van Kampen (2003) on children's' use of languages in the two word stage indicates that their (proto-)grammar employs pragmatic operators and content signs, instead of distinguishing all the syntactic categories present in adult language. Verbs are not yet inflected, and determiners are absent.

Chang and Gurevich (2004) demonstrate a computational model of Embodied Construction Grammar that combines constructions to interpret new constructions. Their semantic representation could serve as an inspiration. Also, the use of Minimum Description Length learning provides a good way to prune the database of learned constructions.

Steels (2004) describes his experiments with situated agents (robots fitted with cameras) that employ language games as a learning strategy. An example of a language game is the description game: one agent describes an event that has just happened, and the other responds by agreeing if the description matches its own experience. These experiments simulate language genesis and grammaticalization *ab initio.*

[4] Instead of enumerating problems with formal treatments of language, one can also hold an apprehensive attitude towards formal ratiocination altogether:

> "I look upon logical proofs the way a well-bred girl looks upon a love letter"
> — Johann Georg Hamann

```
1. "ball gone"  la score = 1
LINGUISTIC ABSTRACTION:
        WORDORDER: VAR:gone
        FRAME: action
                ID: action:move
                FRAME: object
                        ID: VAR
                        ABSTR: object:toy
```

Figure 1: Example abstraction produced by a previous model, showing the semantic annotation of a generalization from 'ball gone' towards 'X gone.'

Van Kampen and Scha (2007) discuss the modeling of early syntax acquisition using the Data Oriented Parsing framework (Bod and Scha, 1996). This means that all input is stored in memory in the form of exemplars which are a pairing of form and structure [5], (be it syntactic, semantic, or otherwise), and made available for recombination in the recognition of novel utterances.

2.2 Motivation

A previous project (van Cranenburgh et al., 2007) attempted to model the acquisition of constructions in the two word stage of early child language. The model used a corpus of utterances spoken to children, annotated with semantic representations of the context. The aim was for this model to be able to generalize over the sentences to discover the correct associations between words and their semantic representations, and to be able to combine sentence fragments into novel utterances. This model did not consider syntax and semantics separately, in the style of construction grammar (as employed in e.g., Tomasello 2000; 2005). Although indeed correct associations were found, and novel utterances could be recognized, most of the former were incorrect, and most of the latter non-sensical (although in part this was due to the first issue worsening the second). Figure 1 illustrates an example of an utterance as it was interpreted (in this case correctly) by this model.

In this sentence the construction "X gone" was applied to "ball", because it matched the condition of being a toy. The construction was apparently previously encountered when a toy was being moved. In this case the result was satisfactory, but unfortunately most other abstractions were spurious or at least infelicitous.

The problem was that sentences were being learned as isolated fragments, without any notion of discourse or pragmatics. Also, the semantic representation did not fit well with all the words to be learned: it was only good at representing actions and objects; prepositions and demonstratives and other abstract words were not being learned. Instead of merely focusing on semantically describing a situation, the learner should consider the total communicative function of an utterance. The learning was implemented as making associations between words and each part of the semantic representation, and counting how often these associations occurred. This meant that a lot of incorrect associations were made. Unfortunately the model did not make use of pruning, as there was no way to know which associations had been incorrect.

Last year another project (Odolphi, 2008) developed a formal grammar for the two word stage, based on empirical work on child language (e.g., van Kampen

[5]This definition is similar to the definition of a construction; the contrast is that an exemplar contains a complete utterance, whereas constructions contain schematized fragments.

2003). This grammar does not make use of adult-like syntactic categories such as verb and noun, but groups expressions as topics, comments and operators. Using this grammar it is possible to parse and produce child utterances, because it turns out that almost all of the two word utterances follow the pattern of this formal grammar.

These projects focused on children's own utterances. However, it appears that children can already comprehend more complicated sentences than they produce themselves, as suggested by such exchanges as:

```
*MOT:  wanna [: want to] put (th)em [= crayons]
        back in the box ?
%act:   <4-7> MOT taps the box with her finger
%gpx:   MOT looks at CHI
*CHI:  no .
%gpx:   <bef> CHI looks up at the box .
        CHI looks down at the chair
```

— Childes,[6] New England corpus[7], Liam, November 30th, 1984

Van Turnhout (2007) implemented a conversational agent for the two word stage using Discourse Representation Structures. His model was capable of learning a semantic grammar to interpret and generate utterances. Two problems he reported were overgeneration and difficulties with elliptical sentences, both of which I shall try to address in the present thesis.

2.3 Research question

Can an exemplar-based model of language acquisition account for the discrepancy in performance between language comprehension and production of children in the two word stage? Can this model facilitate the simulation of simple language games of parent and child?

These questions will be addressed by attempting to implement a model of linguistic comprehension and production using an exmplar-based model of language. I will evaluate this model by comparing its reaction to data from the Childes corpus with actual reactions from children, as well as by testing how well it copes with novel utterances requiring generalization.

2.4 Some philosophical considerations

2.4.1 Limitations of symbolic AI and cognitivism

Since the cognitive revolution cognition has been conceived as symbol manipulation. This idea has kept researchers in both cognitive psychology and artificial intelligence (AI) in business. The more or less official ideology of 'Methodological Solipsism' (Fodor, 1980) secures research grants by asserting that other fields such as neuroscience and biology can have no bearing on the subjects of the so-called 'special sciences.'[8]

This doctrine has come under fire from different directions. Within cognitive science itself there is talk of a second generation (Lakoff and Johnson, 1999) putting forth embodiment as vital; as well as a revival of connectionist systems with sub-symbolic, distributed representations. But long before that there has been vocal criticism from philosophy. Dreyfus (1972) correctly predicted the failure of the

[6] MacWhinney and Snow (1995)

[7] `http://childes.psy.cmu.edu/data/Eng-USA/NewEngland.zip`

[8] Natural science, the dismal science, special sciences; personally I believe natural science is science enough, and that the rest is a mere sideshow.

bombastic ambitions of early AI, and basically claimed that this was due to the symbol manipulation metaphor being a pipe dream:

> "Philosophers have thought of man as a contemplative mind passively receiving data about the world and then ordering the elements."
> — Dreyfus (1972)

In short, Dreyfus warned that artificial intelligence is rather like alchemy: suffering from unwarranted optimism and badly in need of re-evaluating its dogmas. I shall now proceed with the latter.

2.4.2 Compositionality

Compositionality[9] is possible, but not necessary, in the model that will be presented. This is because the usage of exemplars allows the proliferation of arbitrary exceptions to the *implicit* regularities in language. Such exceptions include, but are not limited to, vernacular (e.g., "believe you me"), idiom (e.g., "to kick the bucket") and sentence context coordinating word meaning (e.g., "a bank *deposit*" is most likely not performed on a sofa). Debate continues about whether these present actual problems for compositionality, because some of these items can simply be added to the lexicon (and this might even be plausible). However it is clear that the strong, concatenative version of compositionality can not fully describe all language use, which requires the addition of exceptions to the otherwise strict and elegant rules. Using exemplars with domain-general combination mechanisms sidesteps this issue, because there is no need to pay special attention to problematic cases.

In the most specific case, a sentence is fully described by a single exemplar; in the most general case, a sentence is interpreted word for word, one exemplar each. However, it would appear to be optimal to employ a sort of 'basic-level constructions,' (cf. basic-level categories; Rosch et al. 1976) corresponding to stable collocations that describe a large number of sentences using a small number of multi-word fragments from exemplars. This should be optimal because it reduces the memory load, since not every sentence has to be stored, and because it allows the meaning of words to be dependent on sentence context.

2.4.3 The poverty of syntax

The dominant trend in linguistics is syntacto-centrism, as observed by Jackendoff (1983). What makes syntax so interesting is rarely made explicit,[10] but a desire for immediate and rigorous results probably favors the systematic nature of syntax, at the expense of the more elusive and sometimes vague nature of semantics. Sometimes semantics is given up from the outset, perhaps assuming it to be impervious to scientific scrutiny. Even accounts that explicitly focus on semantics and pragmatics are often syntactic in nature; a case in point is formal semantics. But whether these accounts actually describe semantics or merely mimic parts of it is a difficult question. Obviously planets do not need to be able to solve differential equations in order to orbit as we have come to expect. What we can be reasonably sure of is that semantics is residing (or perhaps presiding) in the human

[9] "The meaning of a compound expression is a function of the meanings of its parts"
— Janssen (1996)

A subtly more specific definition is often employed, adding the involvement of rules:

"The meaning of a complex expression is a function of the meanings of its immediate syntactic parts and the way in which they are combined." — Krifka (1999)

[10] There are execptions, for example Fodor has claimed that symbol manipulation is 'the only game in town'; this argument is dated, however, because connectionism has proved to be a serious alternative. The jury is still out, though.

brain, but this is rather like predicting that it will rain somewhere, tomorrow. The useful question is whether there is some higher-order abstraction of semantics, and whether it can be mechanized or otherwise reproduced in certain systems. This entails that cognition is not just a projection or construction, but a valid abstraction over neural (and possibly other) details. This thesis makes the assumption that such an abstraction should exist, and that an approximation can be attempted and evaluated in a model of language use. Furthermore there shall be no treatment of syntax in isolation from other aspects of language use.

2.4.4 Mentalism

Most accounts of cognition and language in particular are mentalistic. That is, they posit a mental entity which manipulates explicit representations corresponding to external states of affairs. Representations, however, are problematic, because representations have to come from somewhere, either learned or innate, and should, serendipitously or otherwise, faithfully describe, i.e., be isomorphic to, both distal external events and subjective experience. The most dramatic example is the 'Language of Thought' hypothesis (Fodor, 1975), which posits that cognition must operate on first-order logic predicates, which are taken to be universal and in-born. These predicates form the so-called semantic primitives, which can be composed to give rise to an apparent infinity of meanings – including such artefacts as door knobs and scissors which were certainly not part of our humble ancestors' inventories. It is safe to say such theories are far from parsimonious or even empirically responsible.

The other extreme is to reject mentalism and representations altogether, and stress the tight coupling of embodied agents with their surroundings; in a certain sense representations are made redundant by direct interaction with the world. In this conception there should be no need for the category of mind, and body and world are both necessary and sufficient ingredients for cognition. Rorty (1979) is a proponent of this view, and accuses the representationalist school of presuming language and mental events to be a "mirror of nature," as part of the foundationalist's program in epistemology. Instead, he argues, there is no need for a mirror, and language is merely a part of nature. The later Wittgenstein (1953) rejected mentalism on the grounds that a private (mental) language would be impossible, because language is a social phenomenon, useful only by virtue of being shared and understood by a speech community. He argued that language games are the fundamental building blocks of language, in which language use is the sole criterion of meaning. In general anti-mentalistic accounts see language as a way of skillful coping with the world and one's con-specifics, as opposed to the possibly conscious manipulation of explicit symbols. Language is not a conduit which encodes propositional and illocutionary content, but a tool by which we negotiate our ways in the world.

Although I highly sympathize with these views it is highly difficult to apply such a philosophy to artificial intelligence, because they reject abstract mental processes and representations *tout court*. This excludes the possibility of modeling aspects of language in isolation, since the situatedness of embodied agents is what cognition revolves about. Short of making a robot that catches up with millions of years of evolution, it would be impossible to responsibly model the cognition of language. To break this impasse I will make use of semantic and pragmatic representations, but without assuming them to be canonical and actually present in the minds of children. Instead they stand for or hint at a modicum of experience and the spreading activation of neurons, and the social conventions immanent in language use. Since the focus of this thesis is on language and not on modeling full-blown sensory-motor cognition, the linguistic model which will be developed will have to be provided with a ready-made corpus of utterances and their interpretations. How children arrive at such faithful interpretations is a non-trivial problem, but one which is beyond the scope of this project.

3 Practice

3.1 Empirical assumptions

Children always use language meaningfully (language use is never semantics-free, as it is intended to appear in Dada poetry). Pragmatics and semantics are the main pivots on which language use is based. The best-first interpolation of exemplars simulates the interpretation of utterances.

Semantics is a very difficult topic, and hard to get right. Contemporary approaches are Latent Semantic Analysis (LSA), Semantic Networks and formal semantics. LSA views semantics as a high-dimensional Euclidean space whose dimensionality can be reduced to make inferences. Although this approach is powerful and able to cope with vast amounts of information, it is not generally accepted that semantics is Euclidean. Also, the approach requires sufficient data to be of use. Griffiths and Steyvers (2002)

The second approach, Semantic Networks do not suffer from these constraints, but they often lack a rigorous formulation, as well as a clear interpretation (ibid.). The last approach, formal semantics (e.g., categorial grammar and Discourse Representation Theory), explicitly focuses on the formalization of fragments of natural language, because it appears to be very difficult to capture the intentions and significance of a conversational move in open-ended discourse. Because the present thesis does not make use of large amounts of data, and because I want to be able to model open-ended (children's) dialogues, I will settle for a kind of semantic network approach, taking the lack of clear interpretation and formalization for granted (in the scope of this thesis).

Exemplars are memorized by rote, but not declaratively; they are recognized but not necessarily recalled (MacWhinney, 1982). Perhaps rehearsal and self-talk provide the mechanisms to maintain and retain a useful corpus of exemplars. It has been suggested that dreams and imagination function in this way. Rehearsal is also related to the Vygotskyan notion of self-talk and egocentric speech, which differs from Piaget's usage of the term in that it develops from social speech and provides, as a precursor to inner speech, an important scaffolding for cognition, according to Vygotsky (1962):

> "The earliest speech of the child is [...] essentially social. [... A]t a certain age the social speech of the child is quite sharply divided into egocentric and communicative speech [...] Egocentric speech emerges when the child transfers social, collaborative forms of behaviour to the sphere of inner-personal psychic functions [... T]he child starts conversing with himself as he has been doing with others. [...] Egocentric speech, splintered off from general social speech, in time leads to inner speech, which serves both autistic and logical thinking. [...] the true direction of the development of thinking is not from the individual to the socialised, but from the social to the individual." (Op cit., ch. 2)

3.2 Exemplars and Semantics

The model works with a set of exemplars. Exemplars contain an utterance and its meaning representation. The minimalist meaning representation consists of:

1. operator: speech act, indicates the illocutionary force (Austin, 1962) of the utterance

2. a flat list of one or more clauses, reflecting the content of an utterance (insofar as it is understood) as well as the background of the context, ordered by salience (activation):
    (a) first predicate:
        - action: a fluent
        - category:[11] monotonic, atemporal
    (b) second predicate: concrete objects or icons[12], that which is ready-to-hand or '*zuhanden*' in Heideggerian terminology; alternatively, a variable (represented by an uppercase letter), when the information is missing from or asked in the utterance.

It expressly does not make use of recursion or nested frames that would require learning more detailed structures than I want to assume to be available.

The first clause will often contain a topic and a comment, while the rest might contain context, presuppositions and associated facts salient in the relevant situation:

```
"utterance"
operator: pred1(pred2) pred3(pred4) ...
```

For example:

```
"what does a bunny do ?"
whquestion: do(X) animal(bunny)
```

The most salient clause describes what information is asked. To answer the question the child also has to know what the question is about. In both clauses the first predicate describes the class of clauses to which they are compatible. For the first clause this is useful in answering the question, whereas for the second clause it is useful when hearing similar questions about other animals.

Another example:

```
"want some juice ?"
ynquestion: want(juice) food(juice)
```

This representation makes no hard-and-fast distinction between what is explicitly verbalized in the utterance, and that which is understood through context, because this distinction would amount to a fully context-free, introspectable understanding of each and every word in the utterance. Instead of precisely describing the semantic structure of the utterance, this style of representation views the utterance as an ellipsis glossing over parts which can reasonably be expected to be filled in by hearers. Since this filling in of contextual details is not necessarily a linguistic phenomenon, it is assumed to have been completed successfully, and to be present in the initial corpus of exemplars.

3.3 Language Use and exemplars

Adequate participation in a discourse context requires interpreting an utterance, transforming this interpretation into an appropriate response, and verbalizing this response. Interpretation consists of finding a minimal covering set of exemplars which are compatible under unification or constrained predicate substitution. The

[11] Basic-level categories which become implicitly activated in a situation.

[12] Icons denote something by virtue of their inherent similarity to what they represent (Peirce, 1902). They are invoked here because in the corpus employed in this model there are frequent occurrences of pointing at pictures of animals in children's books, followed by the request to name the intended animal.

constraint is that the model can only substitute an argument[13] in an exemplar for the argument of a matching (identical) predicate in the current interpretation. These constraints encode both linguistic and social (common sense) 'rules,' e.g. it makes no sense to throw an animal, because throw occurs only in conjunction with toys.

Response generation is finding a best fit exemplar according to an operator to operator mapping. This mapping is a set of adjacency pairs of speech act operators:

- `imperative` ⇒ { `acknowledgement, refusal` } [14]
- `ynquestion` ⇒ { `agreement, denial`}
- `whquestion` ⇒ `assertion`
- `assertion` ⇒ `assertion` (the child tries to mirror the parent)
- otherwise: respond with empty utterance (e.g., in case of confusion).

After picking the operator the rest of the meaning representation is concatenated to it and the same process of unification and constrained substitution against available exemplars results in the meaning representation of a response, which is then verbalized.

Verbalization is the mapping of instantiated clauses to lexical items inferred from multiple exemplar occurrences.

Reinforcement (e.g. when a parent reacts with "that's right") records an identifier linking the exemplars for the previous utterance and its response to strengthen their association.

Several operations require knowledge of the connections between clauses and words. This lexical knowledge is derived from the corpus of exemplars by juxtaposing all exemplars containing a specific word, and picking the most salient clause they have in common as the meaning for that word. If that fails to produce a single clause, the procedure falls back to looking for links between clauses and words. These links reflect the words and concepts that have been acquired in the one word stage. The links are either explicit, as a word index to an utterance next to a predicate, or implicit, when a predicate and a word have the same form.[15] For example:

```
Exemplars with the word 'bunny':
 ('what does a bunny do ?', 'whquestion: do(X) animal(bunny)'),
 ('bunny .', 'assertion: animal(bunny)'),
 ('is that a bunny ?',
  'ynquestion: point(bunny) animal(bunny)')
Pair-wise intersection of clauses yields: animal(bunny)
```

This process is repeated until no new definitions can be gleaned from the corpus of exemplars. Content words are especially likely to receive correct definitions from this process. This bias is acceptable because they are already acquired in the one word stage, as opposed to function words. Function words do not necessarily carry meaning in isolation, but rather co-ordinate and decorate sentence meaning, which is adequately contained in exemplars.

Similarly, a corpus of constructions is derived from the exemplars. This corpus is created by iterating over all substrings of the utterances in the exemplars, and counting their frequencies. The most frequent collocation of each subset of words is kept and a meaning for it is sought for. If it can be obtained (in a manner similar to the lexical induction) the most frequent meaning is recorded as a form-meaning

[13] Henceforth I shall use the word 'argument' to refer to the second predicate of a clause, not to be confused with the usage of 'argument' in first-order logic.

[14] Since there is no refusal in the corpus used, it has been disabled in the model

[15] This is merely an implementation convencience. In reality the semantic concepts should not be strongly tied to a specific natural language.

pair, possibly abstracted (made variable) if the meaning is not present in the words according to the lexicon.

The corpus of exemplars that is available to the model has been taken from Childes[16] and annotated with meaning representations. It contains 70 utterances by Christopher (1 year and 6 months old), and 157 utterances from his mother. Christopher's mean utterance length is 1.15, his mother's is 3.27 (excluding punctuation).

The ten most frequent constructions in this corpus are:

```
(11, ('is that', 'ynquestion: pretty(X)')),
(7, ('is that a', 'assertion:')),
(6, ('this is a', 'assertion:')),
(6, ('me to', 'ynquestion:')),
(5, ('you want me to', 'ynquestion:')),
(5, ('you want me', 'ynquestion:')),
(5, ('want me to', 'ynquestion:')),
(5, ('want me', 'ynquestion:')),
(5, ('can you', 'imperative:')),
(3, ('you want another box', 'toy(box) ynquestion: want(box)'))
```

The construction 'is that a' is an assertion because it is derived from a few yes-no questions which have been annotated as assertions because they seem to be suggestions rather than questions.

3.4 The model

The first step in interpreting a novel utterance is finding the exemplar whose utterance is most similar to it. This is implemented by iterating over the ordered subsets of words occurring in a sentence, from long to short, and trying to find an exemplar containing these words. The meaning of the exemplar that is found is then used as a template to which other exemplars must conform if they are to be used in interpreting the rest of the utterance. An exemplar conforms to the current interpretation if it has a family resemblance with it, i.e., one of its clauses has a predicate in common with the current interpretation. If the matching clause has a variable argument, it is instantiated:

```
Parent:  throw the ball
initial exemplar:
         ('can you throw it [= ball] to Mommy ?',
          'imperative: throw(X) toy(X) family(mommy)')
instantiated (X) with (ball)
         'the' in 'is this the same as that [= jack+in+the+box] ?'
        and 'ynquestion: toy(jackinthebox)'
        matches 'imperative: throw(ball) toy(ball) family(mommy)'
        interpretation: imperative: throw(ball) toy(ball) family(mommy)
Child:  0
```

The utterance `'0'` means a response in the form of an action. This exemplar comes directly from the Childes data.

If the matching clause has a conflicting argument, it is substituted (see step 2 in figure 2). In order to curtail spurious instantiations and substitutions, only clauses describing the words being covered (as indicated by the derived lexicon) are considered open to modification.

After finding the first exemplar further exemplars are sought in order to cover the remaining words in the utterance. The words are covered in a greedy fashion,

[16] Childes, New England corpus, Christopher, July 19th, 1984

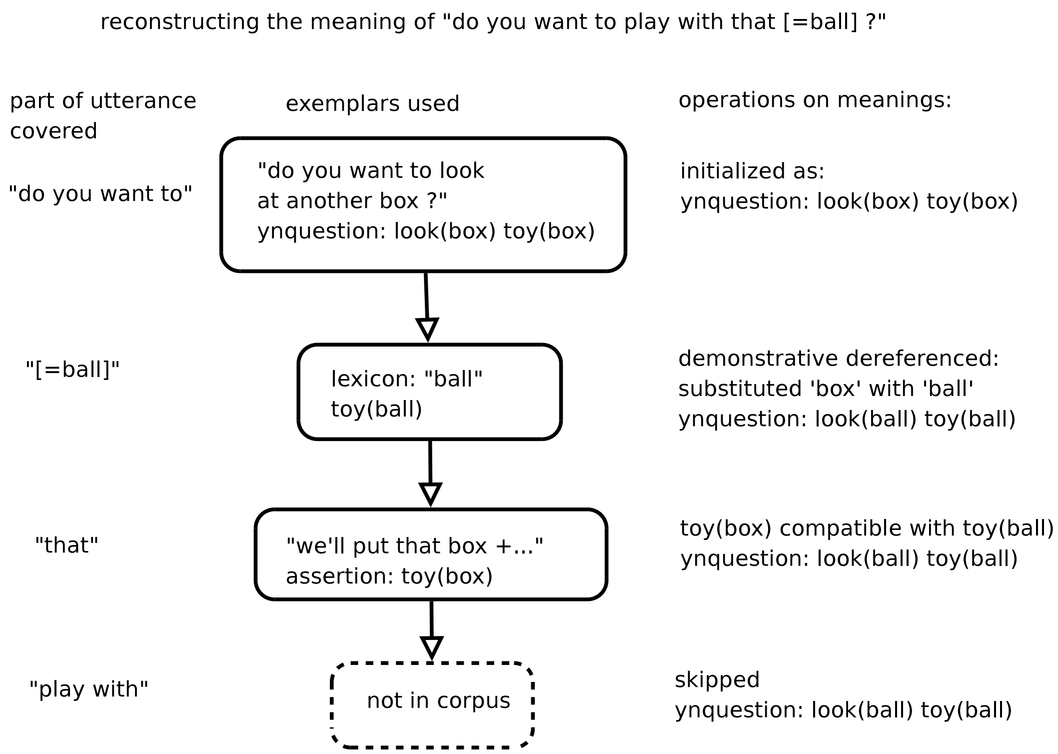

Figure 2: Interpretation process in a step-wise fashion

the longest matching construction is used first. This process performs backtracking on the choice of initial exemplar, because the initial exemplar is crucial for arriving at the best interpretation in that it provides a template to which the other exemplars must conform. Backtracking on all exemplars would be more expensive and suffers from diminishing returns, so it seems to be a more reasonable trade-off to limit backtracking to the initial exemplar. In order to rank the different derivations three heuristics are employed (in this order):

1. minimize number of exemplars used to cover utterance
2. penalize clauses in interpretation whose word form is not present in the utterance according to the lexicon, with an exponential decay according to salience of clause (salient clauses weigh most).
3. maximize length of interpretation (most specific)

These heuristics work well for the current corpus, but a larger corpus might employ more sophisticated use of frequencies and do away with the latter two (somewhat ad-hoc) heuristics.

See figure 2 and 3 for depictions of the steps involved with interpreting an utterance.

For verbalization, the meaning to be expressed is first sought in the corpus of exemplars. If it is present, the utterance associated with it is reduced by filtering out words not in the lexicon. This reduction should not be viewed as merely throwing words away, but as a form of conservatism (lexical knowledge is corroborated more strongly than that of individual exemplars) and limitations in recall of rote memory.

If the meaning to be expressed does not occur in the exemplars the corpus of constructions is checked. If it is found neither in the exemplars nor in the constructions, the model must fall back to the inferred lexicon. First it is attempted

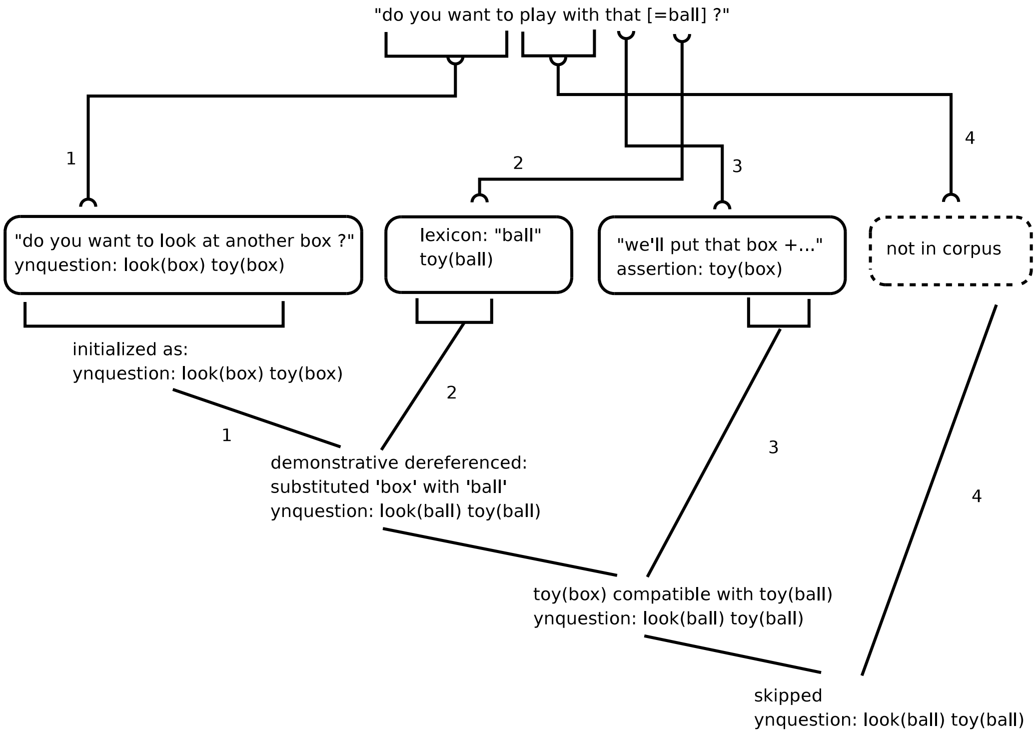

Figure 3: Interpretation depicted as resolution process

to express the first clause, if it is in the lexicon. If that fails the clause is unpacked in a predicate and argument, to form a topic-comment structure. The first predicate (comment or operator) is always expressed, and sometimes the second predicate (topic) is expressed as well. This is currently decided at random (for lack of a better method)

It is possible to supply the referent for demonstratives in sentences with square brackets, for example when we are pointing to a cow in a book, and we ask this question:

```
Parent:  what does this animal [=cow] do ?
initial exemplar:
          ('what does a bunny do ?', 'whquestion: do(X) animal(bunny)')
substituted (cow) for (bunny)
demonstrative dereferenced: whquestion: do(X) animal(cow)
          'this' in 'this is a lamb .'
        and 'assertion: point(lamb) animal(lamb)'
        matches 'whquestion: do(X) animal(cow)'
skipping 'animal'
        interpretation: whquestion: do(X) animal(cow)
```

Interpreting this exemplar caused the meaning of “cow” to be inserted, in this case by substituting the meaning of “cow” for `bunny`. The response is generated by substituting `assertion` for `whquestion`, and unifying the resulting representation with an exemplar. In the process, a discourse topic is established by noting the occurrence of `animal(cow)` in both the question and the response:

```
instantiated (X) with (moo)
        reaction: assertion: do(moo) animal(cow)
reduced: +^ moo@o .
```

```
        topic: animal(cow)
Child:  moo@o
```

The discourse topic (as contrasted with sentence topic) makes it possible to prime the interpretation of an utterance with the topic from the previous utterance. This is implemented by attempting to unify the topic with the interpretation of an utterance; failing silently if it is incompatible, but merging the topic if the utterance contained demonstratives, for example. What follows is a demonstration of this. First we establish 'ball' as the discourse topic:

```
Parent:  that's a ball
initial exemplar:
         ("that's a swing .",
          'assertion: point(swing) toy(swing)')
substituted (ball) for (swing)
         'ball' in 'ball .'
        and 'assertion: toy(ball)'
        matches 'assertion: point(swing) toy(swing)'
        interpretation: assertion: point(ball) toy(ball)
        reaction: assertion: point(ball) toy(ball)
trying to express: point(ball)
        topic: toy(ball)
Child:  ball
```

Then it can be referred to (implicitly) by making an utterance that is compatible with this discourse topic:

```
Parent:  can you throw it ?
initial exemplar:
         ('can you throw it [= ball] to Mommy ?',
          'imperative: throw(X) toy(X) family(mommy)')
instantiated (X) with (ball)
        interpretation: imperative: throw(ball) toy(ball) family(mommy)
        reaction: acknowledgement  throw(ball) toy(ball) family(mommy)
        topic: family(mommy)
Child:  0
```

One of the advantages of the algorithm just described is its graceful degradation. Given sufficient redundancy, words can be misperceived or left out, and the remaining words might still enable correct interpretation. This feature enables natural interpretation of ellipses without specialized mechanisms:

```
Parent:  kitty do ?
initial exemplar:
        ("what's a kitty say ?", 'whquestion: do(X) animal(cat)')
        'do' in '+^ what does a bunny do ?'
        and 'whquestion: do(X) animal(bunny)'
        matches 'whquestion: do(X) animal(cat)'
        interpretation: whquestion: do(X) animal(cat)
instantiated (X) with (meow)
        reaction: assertion: do(meow) animal(cat)
reduced: meow@o .
        topic: animal(cat)
Child:  meow@o
```

(a) Childes fragment		(b) Model output	
*MOT:	what shall we do ?	*MOT:	what shall we do ?
*CHI:	eat.	*CHI:	eat
*MOT:	shall we eat cookies ?	*MOT:	shall we eat cookies ?
*CHI:	ah	*CHI:	0
*MOT:	shall we ?	*MOT:	shall we ?
*CHI:	mmhmm.	*CHI:	0
*MOT:	where are the cookies ?	*MOT:	where are the cookies ?
*CHI:	in bag.	*CHI:	cookie bag
*MOT:	cookie in the bag ?	*MOT:	cookie in the bag ?
*CHI:	baby eat.	*CHI:	cookie bag
*MOT:	..	*MOT:	..
*CHI:	baby eat.	*CHI:	
*MOT:	..	*MOT:	..
*CHI:	cookies.	*CHI:	
*MOT:	baby eat cookies ?	*MOT:	baby eat cookies ?
*CHI:	eat cookies.	*CHI:	eat

(a) Childes fragment, as used in van Turnhout (2007)

(b) Model output

Figure 4: Comparison of Childes data and the responses generated by the model under discussion

3.5 Results

After testing the model interactively during development, a more systematic form of evaluation was introduced. By re-enacting fragments of dialogues from Childes data, it becomes possible to juxtapose the responses of a real child to those of the model developed in this thesis. As a first soundness check I have tested how well a dialogue can be re-enacted using its own exemplars. See figure 3.4 and 3.5 for two such comparisons.

However, this method is limited because it does not require the combination of multiple exemplars to generalize or produce new utterances (the model may still do so, however).

Another soundness check is to try to combine words in a non-sensical manner (colorless green ideas...). This should cause interpretation to fail and should not elicit a response (no production without comprehension):

```
Parent:  what does a ball say ?
        interpretation:
        reaction:
reduced: 0 [=! grunt] . [+ trn]
Child:
```

This result seems to be supported by the observation that cildren's utterances are always meaningful combinations of words[17], and a reaction to their most meaningful interpretation of the parent's utterance.

Another method is to take a sentence from the corpus of another child, and generalize from the current set of exemplars. Here is a question interpreted with a different topic from the exemplar:

[17]Except in the babbling phase, or when there is clearly a game of intentionally producing polysyllabic nonsense going on.

```
*MOT:  that's the cow .

*MOT:  what's this ?
*CHI:  yyy .
*MOT:  is that a donkey ?
*CHI:  donkey .
*MOT:  right .

*MOT:  that's a donkey .
*CHI:  0 .
*MOT:  what's this ?
*CHI:  duck .
*MOT:  what does a duckie say ?
*CHI:  0 [<] .
*CHI:  quack@o .
```

(a) Childes fragment, New England corpus, Christopher, July 19th, 1984

```
*MOT:  that's the cow .
*CHI:  cow
*MOT:  what's this ?
*CHI:
*MOT:  is that a donkey ?
*CHI:  donkey
*MOT:  right .
*CHI:
*MOT:  that's a donkey .
*CHI:  donkey
*MOT:  what's this [=duckie] ?
*CHI:  duckie
*MOT:  what does a duckie say ?

*CHI:  quack@o
```

(b) Model output

Figure 5: Further comparison of Childes data and the responses generated by the model under discussion

```
Parent:  what does a duck say
initial exemplar: ('what does a lion say ?',
                   'whquestion: do(X) animal(lion)')
substituted (duck) for (lion)
        'duck' in 'duck .'
        and 'assertion: animal(duck)'
        matches 'whquestion: do(X) animal(lion)'
        interpretation: whquestion: do(X) animal(duck)
instantiated (X) with (quack)
        reaction: assertion: do(quack) animal(duck)
reduced: quack@o .
        topic: animal(duck)
Child:  quack@o
```

That was a straightforward case, with only one substitution. Here is how a completely novel question (from the perspective of the corpus of exemplars) is interpreted and answered:

```
Parent:  where lives birdie ?
initial exemplar: ("that's where a birdie lives is in a nest .",
                   'assertion: point(nest) animal(bird)')
        '?' in "who's this ?"
        and 'whquestion: point(X) person(X)'
        matches 'assertion: point(nest) animal(bird)'
        interpretation: assertion: point(nest) animal(bird)
        reaction: assertion: point(nest) animal(bird)
reduced: that's where a birdie lives is in a nest .
        topic: animal(bird)
Child:  nest
```

It is also possible to let the model talk to itself. The model successively plays the role of mother and child. The 'dialogue' begins with a random utterance by the mother, to which the model replies. When the reply of the mother to the child would be the same a new random utterance is taken from the set of exemplars which have not already been used (simulating initiative on the part of the parent). A fragment of an example dialogue looks like this:

```
*MOT: this is a gate .
*CHI: gate
*MOT: okay well Mommy will color too .
*CHI: Mommy color
*MOT: what does a cow say ?
*CHI: moo@o
*MOT: oh isn't that [= CHI's paper] nice .
*CHI: nice
```

4 Discussion

The possible contribution of this work to the field of language *acquisition* is small, because the model does little in the way of learning. [18] Instead the focus has been on a psychologically motivated implementation of semantics and language use.

In the model there has been no use of context, except for the limited, solidified context which has become part of exemplars. This is because internalization of exemplars is most useful when there is a certain generality of context. Language should not be purely signal-bound (as behaviorists would have it), although there is nothing wrong with being signal-informed. The focus of the model is mostly on the interpretive power of context-free generalization mechanisms. Including context properly would require making use of perceptual information, which is beyond the scope of this project, which focuses on linguistic performance. The model is concerned with narrow content, which is semantic content without worldly interpretation (as contrasted with wide content in philosophy of mind). The exemplars contain some wide content from the situations in which they occurred, which is re-used by the model. A counter-argument to this is that children's language use is situational rather than conceptual, as with adults (cf. Vygotsky 1962). Representing meaning conceptually might be premature for the two word stage. But perhaps there is no clean break between perception and concepts (cf., observation is theory-laden). Pre-linguistic concepts are definitely perceptual, but consolidating that knowledge in conceptual-semantic representations would have the benefit of greatly compressing information and can not be ruled out as far as I am aware.

To answer the research question, it does indeed seem that an exemplar-based model with the right mechanisms can show difference in performance of comprehension and productivity using the same corpus of exemplars. Aside from being able to re-enact specific dialogues from their exemplars, the model can also generalize and respond to novel utterances. It is not evident whether such a performance model has cognitive plausibility, but parsimoniously the need for innate competence is weakened. Turing's quote about achieving machine intelligence by simulating a child using rote memorization and simple mechanisms appears to have proved applicable in the case of the model presented.

Future work should address language games with more than two moves, more contact with (representation of) situation, learning new representations (which

[18] I have tried adding the feature of recording the interpretation of novel utterances in the corpus of exemplars, making it available to produce utterances based on this exemplar. However, this would require a more sophisticated heuristic to estimate whether the model's interpretation is correct enough to use productively. It would also be possible to detect similarity in meaning between the model's own utterances and the parent's reaction, and assume the latter is a correction if the similarity is high.

requires a notion of relevance, as well as the ability to detect the speech act, predicates and arguments in an utterance). Constraints are currently categorical, but should be probabilistic, as well as influencing other constraints in the current context (spreading activation). It is also possible to move on to the multi-word stage, when the usage of constructions can probably proceed without reduction (i.e., without filtering them according to the knowledge of the lexicon). However, stronger (combinatorial) algorithms should be added for expressing novel meaning representations from multiple exemplars. Even though the focus of this thesis has been on modeling the two word stage, it contains only two instances where the model produces a two word utterance. This shortcoming is solely due to time constraints preventing the development of appropriate mechanisms. I am confident that the corpus of exemplars is rich enough to make two word extrapolations. This can possibly be implemented by reversing the procedure of interpretation.

References

Austin, J. L. (1962). *How to Do Things with Words.* Harvard University Press.

Bloomfield, L. (1914). *An Introduction to the Study of Language.* London: G. Bell and Sons/New York: Henry Holt. Reprinted: Amsterdam/Philadelphia: John Benjamins Publishing Company, 1983.

Bod, R. and Scha, R. (1996). Data-oriented language processing: An overview. Technical Report LP-96-13, University of Amsterdam.

Chang, N. and Gurevich, O. (2004). Context-driven construction learning. In *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*, Chicago.

Chomsky, N. (1975). *Reflections on Language.* Pantheon Books, New York.

Chomsky, N. (1999). On the nature, use and acquisition of language. In Ritchie, W. C. and Bhatia, T. K., editors, *Handbook of Child Language Acquisition.* Academic Press.

Cook, V. J. (1988). *Chomsky's Universal Grammar: An Introduction.* Blackwell Publishers, Oxford.

de Kreek, M. (2003). Language acquisition and virtual grammars: On the continuity of cognition. Master's thesis, Universiteit van Amsterdam.

Dreyfus, H. L. (1972). *What Computers can't do: A Critique of Artificial Reason.* Harper and Row.

Fodor, J. A. (1975). *The Language of Thought.* Thomas A. Crowell Co., New York.

Fodor, J. A. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *The Behavioral and Brain Sciences*, 3:63–73.

Griffiths, T. and Steyvers, M. (2002). A probabilistic approach to semantic representation. In *Proceedings of the 24th annual conference of the cognitive science society*, pages 381–386.

Jackendoff, R. (1983). *Semantics and Cognition.* MIT Press, Cambridge/MA.

Janssen, T. (1996). Compositionality. *Handbook of Logic and Language*, pages 417–473.

Krifka, M. (1999). Compositionality. In Wilson, R. A. and Keil, F., editors, *The MIT Encyclopedia of the Cognitive Sciences*. MIT Press.

Lakoff, G. and Johnson, M. (1999). *Philosophy In The Flesh: The Embodied Mind and Its Challenge to Western Thought*. Basic Books, New York.

MacWhinney, B. (1982). Basic syntactic processes. *Language Development (Vol 1): Syntax and semantics.*, 1:73–136.

MacWhinney, B. and Snow, C. (1995). The CHILDES project: Tools for analyzing talk. `http://childes.psy.cmu.edu`.

Odolphi, E. (2008). Formal grammars for early child language. Bachelor's thesis, University of Amsterdam.

Peirce, C. S. (1902). Logic as semiotic: the theory of signs. In Buchler, J., editor, *Philosophical Writings of Peirce*, pages 98–119. Dover, New York (NY).

Rorty, R. (1979). *Philosophy and the Mirror of Nature*. Princeton University Press, Princeton.

Rosch, E., Mervis, C. B., Gray, W., Johnson, D., and Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8:382–439.

Steels, L. (2004). Constructivist development of grounded construction grammar. In *Proceedings of the 42nd Annual Meeting of the ACL*, pages 9–16.

Steyvers, M. and Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, 29(1):41–78.

Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in cognitive sciences*, 4(4):156–163.

Tomasello, M. (2005). *Constructing a Language : A Usage-Based Theory of Language Acquisition*. Harvard University Press.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236):433–460. `http://www.jstor.org/stable/2251299`.

van Cranenburgh, A., Nusselder, A., Peek, N., and van Weelden, C. (2007). Towards a computational model for early language acquisition. 2nd year Bachelor of AI project, University of Amsterdam.

van Kampen, J. (2003). The learnability of syntactic categories. In *Proceedings of GALA*, pages 245–256.

van Kampen, J. and Scha, R. (2007). Modelling the steps of early syntax acquisition. In Cochran, D. and Bod, R., editors, *Proceedings of the Workshop Exemplar-Based Models of Language*, Dublin.

van Turnhout, J. (2007). Learning in dialog: a semantic grammar approach. Bachelor's thesis, University of Amsterdam.

Vygotsky, L. S. (1962). *Thought and language*. MIT Press, Cambridge.

Wittgenstein, L. (1953). *Philosophical Investigations*. Macmillan, New York. G. Anscombe, translator.