Language Acquisition and Virtual Grammars

On the continuity of cognition

Mike de Kreek

February 17, 2003

Abstract

The ideas that led to the theory of Data Oriented Processing [Scha 1990] consider analogy-making processes as the core of human language cognition. This theory has become an attractive framework for the modeling of adult language performance as it explains more language phenomena than most other theories. The core of the theory is an analogy-making algorithm which assigns an interpretation to new linguistic input, using a corpus of structured past language experiences.

Obviously, an infant does not arrive in this world with a set of past language experiences. The main topic in this work is how the first, more or less structured language experiences could come about. Or, in other words: what kind of language processing mechanisms must be added to the DOP-framework in order to account not only for adult language use, but also for language acquisition?

The power to make generalizations plays a major role in the research tradition that grounds the first linguistic constructions in prelinguistic structures [Chang et al. 2001]. Characteristic features of this tradition and those of the DOP-framework are integrated and extended theoretically to explain the emergence of first multi-word structures. The one-word, two-word and multi-word stages are shown to be logical stages within one continuous language acquisition process. Furthermore, once the two-word stage is reached, the DOP-framework is shown to facilitate the combinatorial explosion characteristic for the multi-word stage. As a side-effect, claims will be made about performance models versus competence models of language use, the innateness of linguistic knowledge, the autonomy of syntax and the continuity of cognition.

Reflectie en dank

"Het duurt effe, maar dan heb je ook wat, zuwe maar zeggen". 15 mb aan backups over de laatste 4 jaar en dan heb ik het alleen over tekst. Vele versies, verslagen en plannen van aanpak zijn deze voorgegaan. Kortom een tijd lang divergerend proces, door sommigen een esthetische ervaring genoemd. En een mooie.... Maar goed, aan een esthetische ervaring mag, of beter, kan geen waardeoordeel toegekend worden. Daarom dit stuk, waarin in ieder geval de leuke inzichten, en hopelijk ook de belangrijke zaken, van het doorgelopen proces te herkennen zijn.

Velen hebben last gehad of meegenoten van de totstandkoming van deze scriptie. Bij deze dank en excuses. Ik noem de belangrijkste spelers. Opa en oma voor financiele steun in de vorige eeuw. Mieke voor het meelezen en bergijpen van de eerdere versies en het eeuwige vertrouwen. Cynthia voor de voorplaat. De OrO-club voor mijn "afwezigheid" en het pushen tot doorgaan. De ZAMM-club voor de inspirerende discussies en glazen wijn. Anouk voor haar ongelofelijke enthousiasme en het feit dat zij ook besmet is met "performance". Remko Scha voor zijn begeleiding en nooit ophoudende stroom aan nieuwe visioenen. Renata Bartsch voor het zich op korte termijn opstellen als tweede lezer.

Mike de Kreek

Voor Opa en Gees.

Contents

List of Figures

Chapter 1

Introduction

This chapter gives a general idea of where this thesis is taking the reader in section 1.1. The coherence between the chapters is offered in the thesis overview in section 1.2.

1.1 The context

Recently, more and more linguistic theories have left the competence-based way of doing linguistics, and instead attempt to model language processing by taking its performance as a basis. Competence grammars focus on (rewrite) rules that supposedly underlie a language, whereas performance models are more interested in how language is used.

One of the performance models that were proposed during the last decades is Data Oriented Processing (DOP). DOP is built on the assumption that people are able to comprehend and produce language by checking their corpus, which contains all past language experiences a person has had. Statistical data from the corpus influence the decision on which analysis is perceived or produced. The analyses that thus emerge are in turn added to that person's corpus, which therefore grows continuously.

When it comes to language acquisition, however, DOP is faced with an interesting problem: given the fact that DOP needs a corpus to analyze input and that the same corpus is built up by analyzing sentences, how does the initial corpus come about?

This is the main question I will try to answer in this paper. Moreover I will show that it is possible that the same mechanisms which account for adult language processing also account for the acquisition of language by an infant. In that way I will meet what performance-oriented theories have described as "continuity of language".

1.2 Thesis overview

Chapter 2 goes into detail with respect to the competence versus the performance approach to language and language acquisition. This is mainly done by introducing a paper of [Seidenberg et al. 1999] who proposes a performance model for language processing and acquisition on a neural basis: a probabilistic constraints approach.

In chapter 3 the attractive and implemented performance-oriented approach to adult language use will be introduced: the DOP-framework, Data Oriented Processing ([Scha 1990]). Many properties of this framework show similarities with the properties of the performance model mentioned by ([Seidenberg et al. 1999]). In this framework the phenomenon of language acquisition still needs to be explained. Readers well acquainted with the DOP-framework may skip this chapter.

Chapter 4 describes the role of prelinguistic structures within the field of language acquisition which also are important for this thesis. The theory behind this approach will be introduced by explaining two papers [Chang et al. 2001] and [Maia et al. 2001]. The way Chang and Maia describe the emergence of the first grammatical constructions by means of an analysis, a hypothesis and a reorganization procedure, will be an important inspiration for the proposal in the remaining chapters. If the reader accepts the role of prelinguistic structures and the way these can be mapped onto acoustic signals, it suffices to scan through the chapter.

The processes in chapter 4 only covered a limited domain explicitly. I broaden this domain by formulating three key-questions in chapter 5. Next, I show why the one-word, two-word and multi-word stages are logical stages, caused by the learning system. Futhermore, I explain the role the DOP-framework has with respect to a multi-word stage instead of a three-word stage. And finally, it is indicated the learning process is a continuous process that applies to language acquisition, but also to prelinguistic concept formation. Obviously, chapter 5 is the core of this thesis.

Finally, chapter 6 wraps up the answers found by following the proposed approach to modeling language acquisition. I also try to indentify the technical details that play a role in simulating the proposal. And last but not least, philosophical issues connected with the proposal are discussed.

Chapter 2

Competence versus performance

The phenomenon of language acquisition can be approached from various frameworks for adult language use. Each of them lead to different assumptions, answers and solutions. After a short description of the most common properties of language, this chapter will describe the preference of performance-oriented models over competence-oriented ones to confront language acquisition.

2.1 Properties of language

Language is a communication system. Somehow an utterance in a certain context can be used to convey information from the speaker to the hearer. Different kinds of context play a role: the discourse context and the situational context. Individual experience of speaker and hearer can be important. Most of the time an utterance obeys the rules (syntax and morphology) of the language, or at least of the speakers of the language, in order to optimally represent a certain meaning (semantics). Intonation has important influence on syntax as well as semantics. Prerequisite to understanding an utterance is the ability to recognize the sounds of the language (phonology) and to segment out the different words in auditory input.

The abilities of writing and reading are learned after the process of language acquisition and thus will not be considered in this thesis, although the cover shows some of the first writings of my niece Cynthia.

2.2 The performing competence

For many years the standard view on language acquisition was based on Chomsky's Generative Grammar. In this approach it was assumed that a child learns the rule system it needs to understand a language by setting parameters in the Universal Grammar. This Universal Grammar is supposed to be innate, which means every child is born with it. Over the past decades, a new approach reared up its head, with interest in the statistical and probabilistic aspects of language: the performance models.

[Allen et al. 1999] wrote a paper in which they, among other things, go into the differences between the two approaches, especially when it comes to language acquisition. I will give a short overview of their argumentation.

The standard, competence-oriented view says that knowing a language is knowing a grammar of that language. Therefore, learning a language must be learning a grammar, e.g. the rule system of a language. The fact that an infant can learn a language, which is characterized as being extremely complex, so rapidly and that the child can produce many more sentences than it could possibly have heard, gives rise to the idea of an innate Universal Grammar. The commonalities between different languages adds to these ideas.

There is, however, a difference between what a grammar allows and what people can produce. An example is for instance the case of deep recursive stuctures. These are grammatical according to the rules, but totally incomprehensible. This is obviously a problem, as when knowing a language is knowing the rules of the language, this should not be happening. The other way around happens as well; many ungrammatical sentences are perfectly understandable.

In order to deal with this, competence grammars allow all these structures to be generated, but performance constraints account for the difficulty to produce or comprehend certain outcomes. Basically, what people know about language (competence) and what they do with this knowledge (performance) is kept seperate.

This distinction disregards data that could be seen as essential to understanding language, such as false starts and errors. Moreover, statistical and probabilistic properties of the input are completely ignored. How often certain structures are used an how they are combined is taken to be of no importance characterizing a language.

Also, it seems that without performance there can be no competence. After all, the rules of the competence grammar have come about by looking at grammaticality judgements, given by people who actually *use* language, and whose judgements are not free from influences, other than their linguistic competence.

> "It is conceivable that competence in this sense of a statistically represented knowledge does not exist. (...) Since such a scenario would demand a major rethinking of the goals of the field of linguistics, I will not deal with it further." (a quote from Schutze used in [Seidenberg et al. 1999])

In the meantime this rethinking has taken place, and has led to the emergence of performance models. These performance models are to handle all and only those strutures that people can. If we adopt this way of thinking about linguistics, this has great effect on the issue of language learning. Instead of assuming that the task of an infant is to learn rules by setting parameters, we view the task of learning a language as learning to *use* a language, which means is must know how to produce and comprehend utterances. The fact that a child can eventually tell whether a sentence is grammatical is no more than a bonus.

In the following overview the two approaches can be compared on the main aspects of any linguististic theory; what is knowledge of a language, how is it acquired an how is it used in production and comprehension.

2.3 The competence-oriented approach

The vast generative tradition has the following answers to what the knowledge of a language is, how the language is acquired and how this knowledge is used in production and comprehension ([Chomsky 1965], [Pinker 1995] and ([Seidenberg et al. 1999]).

1. Knowing a language involves knowing a (generative) grammar. The grammar specifies how language is structured at different levels of representation (syntax, semantics, morphology etc.). It permits the creation of a nearly infinite set of well-formed utterances and it provides a basis for distinguishing well-formed from ill-formed sentences.

2. The grammar is a characterization of the knowledge of an idealized speaker-hearer. This is called "competence": the knowledge of the language that a language user has in theory, abstracting away from his actual language use ("performance").

3. Language acquisition mainly involves the identification of the competence grammar: the child has to converge on the knowledge structures that constitute the generative grammar. Testing whether a grammar rule is applicable to process certain input is part of this process.

4. Children are born with certain liguistic knowledge, a universal grammar, often with parameters to be set. This can only explain, among other phenomena, why children acquire language so fast, why language exhibits structures for which there is no evidence in the input and why there are linguistic universals.

Some of the main problems with this approach are the following. Adult language users often do not agree on whether a sentence is well-formed or ill-formed. Performance often contains "mistakes", e.g. false starts. Generative grammars do not account for the selection humans are able to perform with respect to ambiguous structures. Finally, many parameters are involved in finding the grammar of the mother tongue.

2.4 The performance-oriented approach

Seidenberg and MacDonald advocate a probabilistic constraints approach to deal with language acquisition and processing ([Seidenberg et al. 1999]). They provide alternative perspectives on all of the four issues mentioned in section 2.3, avoiding the problems described.

1. Knowing a language is knowing how to perform the communicative tasks of comprehension and production. The representation of this knowledge is embedded in a neural network, therefore it is not a grammar consisting of deterministic formal rules in line with the generative tradition.

2. The model can account for the performance of what people can comprehend and produce. This performance system ideally incorporates memory capacities, statistical and probabilistic aspects of language, reasoning capacities in comprehending text or discourse and perceptual systems employed in language use.

3. Language acquisition is learning to use the language, the "performance grammar". Accumulating information about statistical and probabilistic aspects of language is part of this process. Being able to make judgements about whether a sentence is grammatical or not is a by-product of language acquisition.

4. With this change in theoretical orientation the arguments that led to the assumption of an innate universal grammar should be reconsidered. For example, languages may exhibit so many common properties (linguistic universals), because otherwise they could not be processed, given the nature of human perceptual and memory capacities.

One of the most important consequences of these alternative perspectives is the essential continuity between language acquisition and processing. The aim of the probabilistic constraints approach is to develop an integrated theory in which the same principles apply to both. In other words: it is attempted to explain both language acquisition and adult language performance within the same model. The main idea of the constraint-based approach is that:

> "(...) comprehending or producing an utterance involves interactions among a large number of probabilistic constraints over different types of linguistic and non-linguistic information." ([Seidenberg et al. 1999], p. 576)

A neural network has proven to be an excellent tool to support this and the above views (see [Allen et al. 1999] for a limited illustration) for the following reasons:

- A neural network can be designed in such a way that it is both a representation of linguistic knowledge and a processing mechanism. This situation contrasts with other approaches in which knowledge of the language is separate from the performance systems that make use of this knowledge.

- The constraints that the model encodes are probabilistic rather than absolute. For example, a noun phrase at the start of a sentence is typically the agent of the action specified by the verb, but not always.

- The interactions between constraints are non-linear. Types of information that are not very constraining in isolation become highly constraining when acting together. For example, the probability that an sentence-initial noun phrase is an agent goes up consideraly if the noun phrase is animate.

- The levels of linguistic representation can be shown to emerge in the course of acquisition. For example, grammatical categories derive from several sources of correlated information, including meanings, phonological structure and syntactic context.

[Allen et al. 1999]'s simulation-model does not simply involve using syntacic information to acquire semantic information ("syntactic bootstrapping") *or* using semantic information (from the environment) to acquire syntactic information ("semantic bootstrapping"). Rather, it simultaneously performs syntactic and semantic bootstrapping to converge on semantic and syntactic representations of verbs. Seidenberg emphasizes that what is called bootstrapping in acquisition literature is equally an example of constraint-based language processing; the distributional information that Allen's model acquired shows similarities with the constraints used in adult performance. This relationship is not accidental, because the task of the model was not language acquisition per se, but rather a simplified version of what human adult comprehenders do: assigning a representation to each input sentences.

In this thesis, however, we will make use of a different performance model, which is built on statistical and probabilistic mechanisms: Data Oriented Processing. In the next chapter I will give a brief explanation of the DOP-approach.

Chapter 3

Performance model: DOP

In chapter 2 a theoretical change in orientation was advocated with respect to language acquisition and adult language use. A shift from a competence-oriented approach to a performance-oriented one sheds a different light on the phenomenon of language acquisition ([Seidenberg et al. 1999]).

A similar change was proposed by Scha with the theory of virtual grammars ([Scha 1990], [Scha 1992]). In order to be able to connect this theory with language acquisition, its most important aspects will be presented in this chapter in sections 3.1 and 3.2, as well as the most important remaining question with respect to language acquisition in 3.3

3.1 Data Oriented Parsing

The DOP-framework is a "corpus based" approach. As the word "corpus-based" implies, corpora ("big bags") of data form the point of departure in this approach to language modeling. From corpora of linguistic data, knowledge can be extracted or learned about the language in question. If the data in the corpus is "raw", e.g. flat, unanalyzed sentences, the approach is often called unsupervised learning. An example can be found in the work of Cartwright and Brent ([Cartwright et al. 1997]) in which distributional cues are used for syntactic categorization. That is, the sum of all environments of a word is calculated and compared with other word's environments in order to cluster them.

If the data is analyzed in some way, e.g. annotated with linguistic information, the approach is called supervised learning. This is the kind of corpus-based approach that functions as the performance model for adult language processing in this thesis, described as "experience-driven interpretation of input" by ([Scha 1990], [Scha 1992]). This proposal has become known by the name "Data Oriented Parsing" (DOP) (e.g. [Bod 1992] and [Bod 1995]). Note that the word "parsing" betrays the original interest in syntactic language analysis. The framework, however, can also be extended with semantics ([Bonnema et al. 1997]) or applied to language production, so do not let the name misguide you. Presently, the abbreviation stands for "Data Oriented

Processing".

Scha's proposal can best be introduced by the following quotation:

> "All lexical elements, syntactic structures and "constructions" that the language user has ever encountered, and their frequency of occurrence, can have an influence on the processing of new input." [Scha 1990]

This implies that a human being (unconsiously) has these data more or less available and uses them to process language. Artificially, a corpus of sentences with their syntactic analyses and semantic interpretations can mimic the past language experiences of a human if it is combined with a suitable analogy-mechanism. This mechanism should then have access to similar data and their frequencies mentioned above. To be on the safe side with respect to what data are meant, I offer the next quote as well:

> "(...) the whole sentence, and all its constituents, but also all patterns that we can extract from these by introducing "free variables" for lexical elements or complex constituents." [Scha 1990].

Based on these fragments the mechanism attempts to find analogies between new input and the corpus. To see how this functions, consider the following, extremely simple, example taken from Scha 1990. Assume the corpus consists of these two trees:

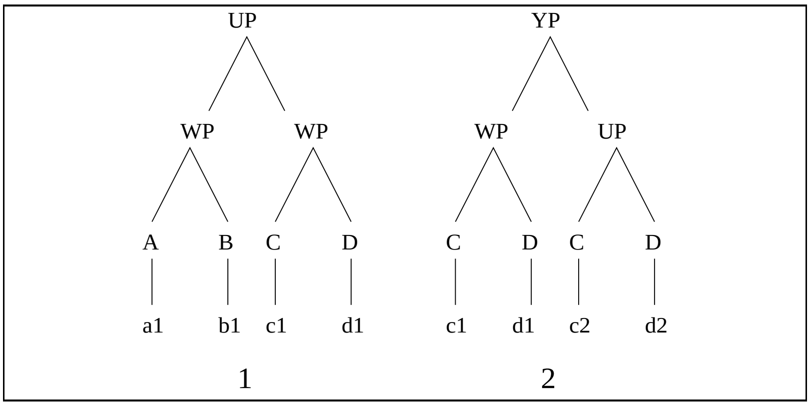

Figure 3.1: Corpus consisting of two trees.

An incomming sequence ("sentence") "c1 d1" will receive two analyses. One can literally be recognized as a WP based on the WP(C(c1), D(d1))-fragment in tree 1 and 2. The other will be recognized as UP in tree 2 through combining UP(C,D) with C(c1) and D(d1). In [Bod et al 1996] it is emphasized that

> "(...) the interesting part of perception theory is not the part which describes the set of possible structures, but the part which describes the preference of one structure above the other (...)".

In this case the preferred analysis (optimal analogy) is the first one, because it demanded no combinations (the second one *did*) and because this analysis occurs twice (the second analysis only *once*) in the corpus. So, an analysis of an input becomes more probable if it is constructed with larger and more common fragments. The most probable analysis is considered to be the perceived analysis. Fur a further description of the procedure involved in this process I refer to section 5.6 about the multi-word stage.

The technical details about how all possible subtrees could be generated and how the optimal analogy is calculated can be found in [Bod 1995].

3.2 Virtual grammars

In many linguistics-related books one can encounter the discussion about grammaticality of sentences and utterances. With the DOP-framework the distinction between grammatical and ungrammatical sentences has lost its absolute character and has become a continuum. With respect to the example in the last section, this means that the second analysis was *less* grammatical than the first analysis. This is rather convenient, because this reflects the fact that, without exception, all of us from time to time come across a sentence not being able to decide whether it is grammatical or not. Moreover, since DOP models the language use of an *individual*, it is possible that one person has different grammatical intuitions than another person. Note, though, that these intuitions do not play a significant role in everyday language. We all use ungrammatical sentences all the time, which does not stop us from understanding each other. Within the DOP-framework one could go one step further. Just like different people have different strategies doing mental arithmetic, it might be that different people also have different language-structures for the same utterances in their memory. To perform data oriented processing it does not matter what the structures in the corpus look like, as long as there is a high degree of consistency.

Consider the uncommon sentence “the young thesis was writing a minestrone of arguments”. Some people might give the verdict “ungrammatical” to this sentence. It would be interesting to study the differences in grammatical judgements between (groups of) people. To give a tentative example: it might be that people who have read many poems, judge sentences, like this one, different from people who hardly read at all. Results like this would completely fit in with the DOP-framework with the past language experiences as its basis. Almost trivial, but not mentioned before, is the fact that if the corpus of past experiences is empty or underdeveloped, grammatical judgements, as adult perform them, will not be possible or, more likely, irrelevant. This indicates that grammatical judgements *are* relevant after passing a certain point in development of the corpus. I will return to this aspect in section 6.2.3 discussing the phenomenon of syntax.

Where does syntax start (or end)?

Connected with this observation is the fact that a language user sometimes judges one and the same sentence as grammatical or ungrammatical depending on the information (linguistic or situational, together called “the context") processed just before processing this sentence. For example, a sentence with a certain structure sometimes enables a language user to recognize the structure of the next sentence needed for the right interpretation. The "priming" of the first sentence plays an important role here. This phenomenon can very well be accounted for within the actual DOP-framework by assigning more weight to recent analyses in the corpus.

All this means that grammar as a fixed set of rules that describes the syntax and semantics of a language does not exist; the best case scenario is that a subset of all utterances of a language is regarded grammatical by *all* speakers of that language. Within the DOP-framework, grammar has a **virtual character**; it does not have the appearance which most linguists expect it to have. As Luc Steels puts it in "Volume I - words and meanings" about the Talking Heads experiment [Steels 1999]:

> (...) Grammars can be seen as ecologies, where form-meaning pairs compete in the population. New syntactic and semantic categories, new constructions and new uses of grammatical conventions are continuously created, and existing ones may destabilize and become in disuse. Each language user employs his own idiolect which is as well as possible co-ordinated with that of other language users but there is never complete similarity and never absloute stability. This explains perhaps why linguists have such a hard time to pin down *the* language of a community. (...) (p. 46)

Grammar emerges from the corpus by means of the analogy mechanism. In the remaining text the word "grammar" will denote a grammar with this virtual character.

The robustness of the DOP-framework is also an important point to emphasize. If a sentence mistakenly gets a wrong analysis and, as such, ends up in the corpus, this will be fixed later by a majority of correct analyses. On the other hand, when the structure of a utterance cannot be accounted for by the corpus, but the meaning is clear, then it will be attempted to project a structure onto the sentence. Under the influence of DOP a person adapts to the language community. These utterances only sound "funny" during a certain period of time. This is an example of the notion of "language change"; over a period of time more and more people use the same structure until everybody does it, so that the phenomenon has become common. Even within periods of, say, 10-15 years, people seem to be able to "change" their language, e.g. adapting to another dialect. The suggestion about assigning more weight to more recent analyses certainly plays a role here. The DOP-framework accounts for this phenomena, although implementations that tackle language change have not been tested.

3.3 Remaining question

The designers of the DOP-framework have one important question remaining which concerns the issue of language acquisition. In [Scha 1990] it is stated that within the DOP-framework there is continuity between earlier and later stages of language use. No distinct process of language acquisition needs to be postulated. This means that there is continuity with respect to the development of an earlier corpus into a later corpus of a language learner. Since an infant does not come into this world equipped with a corpus full of basic linguistic data, the central question is:

> **question**: How does an infant build up "the first" corpus which forms the departure point for development into later, more complex, corpora?

This question will be split up in subquestions in chapter 5 where I explain about the linguistic and the non-linguistic component involved in language acquisition. First, I will introduce a model in the next chapter that grounds the first grammatical constructions in prelinguistic representations.

Chapter 4

Grounding language in perception: Chang and Maia's model

In 2001 Chang and Maia published two interesting papers about modeling language acquisition. This chapter deals with the main features of their model.

Section 4.1 first introduces the most important aspects of language acquisition. In the same section it is explained on which of these aspects ([Chang et al. 2001] and [Maia et al. 2001]) *do* and *do not* concentrate. What Chang and Maia offer is a detailed model which explains how infants learn to make mappings from word-orders to role-bindings, which come from observed situations. Furthermore, ideas are introduced about how word meanings could be acquired before the "mapping skill" is in place. And finally, a tentative picture is given about how grammatical categories could be acquired after having acquired the "mapping skills".

The sections 4.2 and 4.3 cover the technical details of the two papers. This is done quite thoroughly because their model offers a basis for some aspects of chapter 5 in which the DOP framework is extended to cover language acquisition.

4.1 Language acquistion

Important for this work to mention are the following, commonly acknowledged stages in language acquistion: the babbling stage, the one-word stage, the two-word stage and the early multi-word stage. The most common research topics within these stages consist of:

- **Segmentation**
  Somehow an infant finds out where a certain word begins and where it ends in the acoustic input.

- **Learning words**
  From all the different sensorical impressions the infant receives it is able to learn the meaning of single words.

- **Semantic bootstrapping**
  The meaning of a word gives cues with respect to what syntactic category it belongs to.

- **Child language**
  An infant goes through stages in which it has its own syntactic categories and linguistic constructions.

- **Distributional categorization**
  Similar distribution of words is a reason to put them in the same syntactic category.

- **Syntactic bootstrapping**
  The distribution of a certain word gives cues with respect to what it might mean.

On all these points [Chang et al. 2001] and [Maia et al. 2001] make implicit or explicit claims:

The **segmentation** of words from acoustic signals is mentioned but not explained in [Chang et al. 2001]. It can be inferred from their text that they at least assume a mechanism which is able to memorize acoustic signals and their accompanying situational observations. Building up this memory a certain sound sequence somehow becomes connected with re-occuring parts of situational observations by a converging process. An early conceptual ontology plays a crucial role in that process. Although segmentation is not my main focus either, the subject will return in the next chapter connected with learning the meaning of words in the one-word stage (chapter 5).

Learning words is closely related to segmentation and not a part of the technical model of [Chang et al. 2001] either. They do mention some important aspects of word-learning, though. The first one is that they assume that there is an early one-word stage in which uttered word-meanings are tightly coupled with specific events, actions or contexts. The second one, related to the former, is that the early one-word utterances are subject to polysemy effects. The same one-word utterance might be encountered in multiple distinct but related situations. Adults might assume that an infant has learnt the fullblown word meaning, but the child might be using the word in these situations as if it has different meanings. The third aspect is that despite their initial context-bound nature, most word meanings eventually converge towards representations that are neutral with respect to their speech act. I will return to these aspects in detail in chapter 5.

[Maia et al. 2001] shows a straightforward interpretation of **semantic bootstrapping**. The prelinguistic conceptual categories that get connected with words start out to function as the (early) “syntactic” word categories. The prelinguistic representations of scenes (which are represented by frames) form the basis of the first linguistic constructions. Semantic features lead to generalizations over the early "syntactic" word categories and the first linguistic constructions. These generalizations are a step towards categories and constructions that resemble those of the adult language community.

The examples in [Chang et al. 2001] include lexical constructions that are formed by mapping linguistic forms on prelinguistic action-specific representations (see also [Maia et al. 2001]). This leads to transitive verb-specific constructions (e.g. **Human**-THROW-**object**). The algorithm that [Chang et al. 2001] offers is limited to the processes behind the emergence of these constructions (details in sections 4.2 and 4.3). Constructions like this are representative for what is called "**Basic child grammar**". They claim that once the child is operating with more abstract prelinguistic representations, mapping linguistic forms can lead to a more general transitive construction. Chang and Maia, however, do not give any example in which it becomes clear what a lexical construction for *throw* looks like in this stage and how its constraints can lead to a more general transitive construction.

In [Chang et al. 2001] a certain level of basic child grammar is achieved processing a few examples. It remains unclear, though, whether without a suitable situation the utterance (e.g. *you throw ball*) can be processed and understood. In connection with this, Maia and Chang do not mention the two-word stage in their work. It is likely that the two-word stage has something to do with the first two-word constructions, so I will return to this in chapter 5.

Distributional cues are tentatively assumed to drive the merging process further than the semantic bootstrapping process could achieve. This way syntactic categories that resemble those of the adult language community come within reach. For example "non-physical action" verbs end up in the same category as "physical action verbs". Chang and Maia refer to other research on this last assumption and do not incorporate it in their model. I will pick this up tentatively in chapter 6, where the consequences of the proposals in chapter 5 are discussed.

Syntactic bootstrapping is not an issue in [Chang et al. 2001]. Still, they probably assume the existence of such a process, because in their view only a small number of abstract relational frames can be learned prelinguistically. The syntactical environment in which an unknown (more complete) verb, like *remember*, occurs might gives cues finding the meaning of it.

The above shows that Chang and Maia concentrate on the emergence of basic child grammar. Prelinguistic structures play a key role in their model. I offer a summary of [Maia et al. 2001] about these structures in the next section 4.2. I decsribe the emergence of the first grammatical constructions according to [Chang et al. 2001] in section 4.3.

4.2 Prelinguistic structures

Before it acquires the first word meanings, the infant has reached a certain level of conceptualizing the world around it. [Maia et al. 2001] advocate the existence of these prelinguistic structures because they are useful "for behaving in the world". The term "prelinguistic" can actually be understood in two ways: before language ever existed and before an individual acquires a language. Chang and Maia base their assumption on an evolutionary line of reasoning (phylogeny):

> "(...) language entered the scene very late in evolutionary time and as such was built on top of older cognitive skills that we share with many of our relatives in the animal kingdom, such as the ability to move, perceive scenes, act on objects, interact with conspecifics, etc." (p. 1 [Maia et al. 2001])

The development of an individual (ontogeny) shows a parallel to the evolution of mankind:

> "(...) ontogeny seems to proceed in the same manner – before acquiring a language, the infant goes through an extended period of physical and social interaction with the environment." (p. 1 [Maia et al. 2001])

The above shows that language acquisition should be explained in terms of the relation of language to cognitive systems that precede it, both ontogenetically and phylogenetically. In other words, language acquisition shows parallels with the emergence of language ([Steels 1999]). Accordingly, the cognitive systems are claimed to be able to produce the beginnings of a stable ontology (corresponding to people, objects, settings and actions) well before the first recognizable words. Also, aspects of the surrounding social and cultural context are firmly in place prelinguistically.

Based on research on humans and animals within the field of neurosciences, [Maia et al. 2001] offer an explanation of the process which eventually leads to the emergence of abstract relational frames to represent an observed scene. Low level features of concepts are involved in this process, which are not within the scope of their work.

The notational conventions are not consistent between [Chang et al. 2001] and [Maia et al. 2001]. I will use a simplified version of their notation, suitable for the purposes of my thesis, as shown in figure 4.1. The examples will be explained in the following sections.

What?	How?	Examples?
frame name	capital, bold	**PUSH, DIRECTED ACTION**
role/relation	small capital	PUSHER, ACTOR
instance/entity	regular, bold	**push, sister**
utterance	italic	*ball*, *throw the ball*
operator	regular, small	construction, before
construction	capital	THROW, PUSH-TRANSITIVE
natural categories	Capital-regular bold	**Human**, **Physical action**

Figure 4.1: Notational conventions.

[Maia et al. 2001] consider the emergence of a representation for a pushing situation involving the brother and the sister of the observer. Based on the unspecified low level features, representations for the persons, **brother** and **sister**, become available. The concept for the action **push** can also be recognized at a certain stage, independent of who is doing the pushing and who is pushed. [Maia et al. 2001] mention tests with monkeys which provided evidence that they had a general representation for **grasp** independently of who was the GRASPER or the GRASPED. This and other evidence leads them to the conclusion that humans also are able to represent the pushing action independently of the actual roles involved, just by **push**. In other words **push** can be experienced as a "stand alone" concept.

Once these representations are in place, the mind may also develop some way of indicating the one who is doing the pushing and the one who is undergoing the pushing, in order to fully represent the pushing actions. These lead to the roles PUSHER and PUSHED, the bindings between actions and entities. Once these concepts (**push**, **brother**, PUSHER etc.) are available, new input can be represented by them in an "action specific relational frame":

PUSH[PUSHER: **sister**. PUSHED: **brother**][1]

It is confusing that [Maia et al. 2001] claim that the infant is able to apply this representation the first time ever it observes a pushing situation. I would expect that the infant has to built up towards this representations by observing several pushing-situations without recognizing them at the time. I will return to this remark in section 5.3. Furthermore, the representation **push** is acknowledged as a concept that could be triggered on its own in a pushing situtation, without roles or instances of **Human** necessary. Interestingly enough it is not incorporated into the action specific relational frame. On the other hand it *does* play a role in the abstract relational frames further on. This is why I assume that **push** must be present in the **PUSH** frame, but is omitted. For simplicity reasons I will assume in chapter 5 that **PUSH** and **push** are the same thing.

Observing more pushing situations in which different humans are doing the pushing leads to the observation that the PUSHERs share an important feature: **Human**. This also goes for the PUSHEDs. The following generalized frame is the result:

PUSH[PUSHER: **Human**. PUSHED: **Human**].

Note that although **sister**, **brother** etc. are known to be **Human**, it takes several examples for the infant to build up the statistical evidence for such a generalized frame to form.

Another process enables the infant to generalize action-specific frames like the ones above towards frames that encompass several actions. Assume

[1] Maia and Chang use symbolic frames for notational convenience. They actually have a neural representation in mind in the spirit of structured connectionism.

that the following frames have been formed (A **physical object** is learned in a similar way as before):

PUSH[PUSHER: **Human**. PUSHED: **Human** or **Physical object**] and **KICK**[KICKER: **Human**. KICKED: **Human** or **Physical object**].

On the one hand the infant uses the similarities[2] between the roles (PUSHER – KICKER and PUSHED – KICKED) in these two frames to generalize across them towards ACTOR- and PATIENT-roles. On the other hand the situation in which a human acting physically on a human or an object is so common that it will invite the child to generalize towards a category **Physical action**. Maia and Chang consider the infant to be able to form the following "abstract relational frame":

DIRECTED ACTION[ACTION: **Physical action**. ACTOR: **Human**. PATIENT: **Physical object** or **Human**.].

In this stage, the infant uses this abstract relational frame processing a scene in which the infant's sister pushes the infant's brother[3]. The representation of this scene will be:

DIRECTED ACTION[ACTION: **push**. ACTOR: **sister**. PATIENT: **brother**.].

Obviously, in the stage when situations are represented by action-specific relational frames, the infant is not able yet to "see" similarities between a kicking situation and a pushing situation. Should the child start operating with action-specific relational frames to map words onto, then possibly acquired words (e.g. *push* and *kick*) will not have anything in common, because their meaning has nothing in common. In other words: *push* is represented by a **PUSH**-frame and *kick* by a **KICK**-frame. Once the child is operating with abstract relational frames such as the **DIRECTED ACTION** frame above, generalization has taken place on the representation level. This means that both *push* and *kick* will be represented by this frame. In this case the words have this aspect in common.

Another generalization example is the mapping of a case marker. When the child maps a case marker to an abstract relational frame[4], that form automatically becomes available for all actions covered by that frame. There exists crosslinguistic evidence that children make linguistic generalizations across prototypical transitive events, like the directed pushing action described above, that do not carry over to non-prototypical events:

> "In [adult] Russian, for example, the direct object is marked by an inflection, regardless of the type of event; when Russian children first start applying this inflection, however, they use it only to mark the

[2]These similarities could be represented by lower level features underlying the concepts. Here, only the relational structures are higlighted. An important factor is that both (PUSHER and KICKER) are represented in the premotor cortex.

[3]Chang and Maia mention that the "old" levels of representation may stay available.

[4]This is coined "making a construction" in the next section.

> direct objects of verbs involving direct, physical action on things."
> (p. 4 [Maia et al. 2001])

Consider a group of verbs that do not involve direct, physical action on things. According to the quote above there is a stage in which the inflection is not known to be suitable for this group. Once the infant learns that the inflection *is* applicable to one verb from this group, the whole group becomes available for the inflection in question.

The above shows that early grammar reflects prelinguistic structures which in turn arise from the architecture of the brain. This leads Maia and Chang to propose the following picture of the organization of language.

The visual and premotor cortex and specialized circuits in the brain tend to give rise to different conceptual categories (e.g. **sister, brother, push, etc.**); these in turn will tend to form the prototypes (e.g. the biologically natural categories **Physical action, Human, Physical object, etc.**) of syntactic categories (e.g. noun, verb, preposition, etc.). Clearly there is not a simple one-to-one correspondence between the prototypes and the syntactic categories. Being confronted with "non-physical action" words like *see* and *remember* in the same syntactic context (distribution) as physical action verbs, the child will eventually assign them to the same syntactic category. This way, the child is able to gradually form a syntactic category, which will exhibit prototypical, radial and graded effects. The words *ball* and *block* would be a prototypical example of a noun in adult language use, originating from the biologically natural category **Physical object**. The word *war*, on the other hand, is put in the same category, because it is encountered in similar environments as the other two words. Obviously it does not originate from the biologically natural category and, moreover, in a set of sentences the word *block* can be replaced by *ball* easier than by *war*. This makes *war* a less prototypical noun.

Chang and Maia describe the algorithmical aspects of the generalization process towards the prototypes of syntactic categories in [Chang et al. 2001] and the first constructions based on them. The claim about the emergence of syntactical categories is not worked out technically.

Because there is only a small number of biologically natural categories (**Physical action, Human** etc.), [Maia et al. 2001] claim that there will also be a small number of abstract relational frames. These frames are claimed to form the basis for grammatical constructions found in child language. How this leads to the emergence of verb-specific grammatical constructions is explained in the next section.

4.3 Grammatical constructions

Once the prelinguistic structures are in place, input consisting of an observed situation (a relational frame) and a short utterance can be considered.

Before being able to learn grammatical constructions a certain level of the one-word stage must be achieved. One-word utterances are assumed by Chang and Maia to be tightly coupled with a specific event, context or even purpose [Chang et al. 2001] with which they have co-occurred. Through cross-situational observation a single word may become representative for certain commonalities in the observed situations. This way the infant is able to acquire a number of primitive speech acts, sufficient for expressing requests, comments, refusals and even queries. The context-boundness can lead to polysemy effects, since the same form may be encountered in multiple distinct (though possibly related) contexts, which may be diverse enough to resist a single generalization. Despite their initial context-bound nature, most word meanings eventually become generalized toward representations that are neutral with respect to the speech act. As the language learner collects an increasing set of relatively stable form-meaning-pairs, the linguistic cues become better correlated with the surrounding environment and accordingly more informative. Chang and Maia concentrate on modeling the emergence of larger grammatical constructions and leave the one-word stage for what it is.

[Chang et al. 2001] propose a model for the acquisition of grammatical constructions on the basis of examples[5], focussing on the earliest multi-word constructions. A grammatical construction is a mapping between form and meaning, typically a set of form relations (e.g. word order) corresponding to a set of meaning relations (relational bindings).

Frames are represented in terms of individual relational bindings. The representation

THROW[THROWER:**Human**. THROWED:**Object**.]

contains two roles and fillers that can be represented by the following relational bindings, in which **throw** is a semantic entity. Like I mentioned in the previous section **throw** suddenly shows up in the abstract relational frame. Here, the same occurs: cutting up the **THROW**-frame in relational bindings leads to a sudden presence of **throw**. This already forced me to conclude that it must have been present in the **THROW**-frame already. The names for roles and relations are kept similar for simplicity reasons (see also figure 4.1 for the notational conventions):

throw.THROWER: **Human** and **throw**.THROWED: **Object**.

A verb-specific example of a construction in which tx_f and tx_m stand for the form of tx and the meaning of tx respectively:

[5]These examples are taken from the CHILDES corpus ([CHILDES]) of child language interaction.

construction THROW-TRANSITIVE
  constituents:
    construction t1 of meaning type **Human**
    construction t2 of type {THROW}
    construction t3 of meaning type **Object**
  formal constraints:
    $t1_f$ before $t2_f$
    $t2_f$ before $t3_f$
  semantic bindings:
    $t2_m$.THROWER:$t1_m$
    $t2_m$.THROWED:$t3_m$
designates $t2_m$

Figure 4.2: THROW-TRANSITIVE construction.

Together with the lexical constructions JOHN (t1), THROW (t2) and THE-BALL (t3) the above construction leads to a **constructional analysis** of the sentence *john throws the ball*. Although not explained in the text, the lexical construction for e.g. THROW will probably look like this (based on figure 1 from [Chang et al. 2001]):

construction THROW
  formal constraints:
    throw
  semantic bindings:
    throw.THROWER:**Human**
    throw.THROWED:**Object**

Figure 4.3: Lexical construction THROW

Consider an infant which only has the lexical construction for THROW available. In processing *throw john*, without a suitable situation, it does not know whether *john* is the THROWED or the THROWER. The THROW-TRANSITIVE construction is a verb specific construction, because it cannot be applied to other transitive verbs, assuming there is only one lexical construction THROW. In a later stage the type of t2 can become **Physical action** and the semantic bindings ACTOR and PATIENT. In this stage the construction would be applicable to all physical action verbs.

The main assumption in learning grammatical constructions is that a learner expects correlations between what is heard and what is perceived. At any point in the learning process, some of these correlations may already have been encoded, thus accounting for some previously learned constructions. The tendency to account for new unexplained data leads to the formation of new constructions.

Three processes are distinguished in the model in order to arrive at new constructions:

1. Some mappings between an utterance and a situation can be predicted by known constructions. This is seen as a precursor to language comprehension, in which the same mappings will evoke meanings not present in the situation. Both require an **analysis** procedure that determines which constructions are potentially relevant, given the utterance. The procedure will also have to find the best fitting subset of those by checking their constraints in context.

2. Once the predictable mappings have been collected, not all of the input can be accounted for. The learner must have a procedure (called **hypothesis** below) for determining which new mappings may best account for the unexplained data. That is, finding form relations corresponding to meaning relations. The primary candidates for these mappings will be relations over lexical elements whose form-meaning mapping has already been established.

3. Another way constructions can arise is by **reorganization** of the set of known constructions. This procedure generalizes similar or co-occurring constructions.

The algorithms belonging to the three described processes are introduced in the following sections. How they function is shown by processing some examples.

4.3.1 Analysis

Consider the following input (input 1) consisting of utterance U and situation S which consists of entities S_e and role bindings as well as attributes of individual entities S_r:

> input 1:
> U = "*you throw a ball*"
> S_e = {**self**, **ball**, **block**, **throw**, **mother**, ...}
> S_r = {**throw**.THROWER:**self**, **throw**.THROWED:**ball**, **ball**.COLOR:**yellow** , ...}

A set of known constructions is assumed which contains lexical entries for BALL, THROW, BLOCK, YOU, SEE, etc. (see figure 4.3 for an example), as well as a two-word THROW-BALL construction associating the before (*throw,ball*) word-order constraint with the binding of **ball** to the THROWED-role of the **THROW** frame:

construction THROW-BALL
constituents:
construction t1 of type {THROW}
construction t2 of type BALL
formal constraints:
$t1_f$ before $t2_f$
semantic bindings:
$t1_m$.THROWED:$t2_m$
designates $t1_m$

Figure 4.4: THROW-BALL construction.

In order to start the analysis, all constructions (lexical or complex) that are associated with these words in the utterance have to be identified in Step 1. The algorithm finds all the constructions (C_{cued}) that contain any of the words in the utterance (F_{known} = {*you, throw, ball*}) within their form constraints. In this case, C_{cued} = {YOU, THROW, BALL, THROW-BALL}.

Analyze utterance. Given an utterance U in situation S and current set of constructions C, produce the best fitting analysis A:

1. Extract the set F_{known} of familiar form units from U and use them to cue the set C_{cued} of constructions.
2. Find the best-fitting subset C_A of C_{cued} for utterance U in situation S. Let F_A be the set of form units and form relations in U used in C_A, and M_A be the set of meaning elements and bindings in S accounted for by C_A. Then $A = <C_A, F_A, M_A>$. A has a associated cost $Cost_A$ providing a quantitive measure of how well A accounts for U in S.
3. Reward constructions in C_A; penalize cued but unused constructions, i.e., those in $C_{cued} \backslash C_A$.

Figure 4.5: Construction analysis

Step 2 is very similar to the analyze algorithm used within the DOP-framework, although its functioning is more transparant (see section 5.6). The contraints and bindings specified by these cued constructions must be matched against the input utterance and situation. More analyses are possible, but the one with the lowest cost, $Cost_A$, will be the best fitting analysis. $Cost_A$ depends on the popularity of the involved constructions. The form constraints (e.g. before (*throw,ball*)) of the found constructions are all met by the utterance, just like the semantic constraints (e.g. **throw**.THROWED:**ball**) are met by the situation. In the eventual best fitting analysis A, the constructions used are C_A = {YOU, THROW,BALL,THROW-BALL}, which covers the forms and form relations in F_A = {*you, throw, ball,* before(*throw, ball*)} and maps the meanings and meaning relations in M_A = {**self**, **throw**, **ball**, **throw**.THROWED:**ball**}.

In step 3 the popularity of constructions is modeled. How constructions are rewarded or penalized and how $Cost_A$ is calculated is omitted, just like some other details:

> "(..) At the stage of interest here (...) we assume that all constraints are simple and few enough that exhaustive search should suffice, so we omit the details about how cueing constructions, checking constraints and finding the best analysis proceed." ([Chang et al. 2001])

Within the DOP-framework the statistics of each fragment determine its popularity; no penalizing is nessecary (section 5.6).

4.3.2 Hypothesis

At this stage in processing the input **throw**.THROWER:**self** and before(*you,throw*) are not yet mapped onto each other.

Hypothesis construction. Given an analysis A of utterance U in situation S, hypothesize a new construction C_U including correlated but unused form and meaning relations:

1. Find the set F_{rel} of form relations in U that hold between the familiar forms F_{known} and the set M_{rel} of meaning relations in S that hold between the mapped meaning elements in M_A.
2. Find the set $F_{rem} = F_{rel} \backslash F_A$ of relevant form relations that remain unused in A, and the set M_{rem} $= M_{rel} \backslash M_A$ of relevant meaning relations that remain unmapped in A. Create the super construction $C_{super} = (F_{rem}, M_{rem})$, replacing terms with references to constructions in C_A where possible.
3. Create a potential construction C_{pot} consisting of pairs of form-meaning relations from C_{super} whose arguments are constructionally related.
4. Reanalyze utterance using $C \cup \{C_U\}$, producing a new analysis $A^{‘}$ with cost $Cost_{A^{‘}}$. Incorporate C_U into C if $Cost_A$ - $Cost_{A^{‘}} \geq$ *MinImprovement*; else put C_U in pool of potential constructions.
5. If U contains any unknown form units, add the utterance-situation pair to the pool of unexplained data.

Figure 4.6: Hypothesis construction

In order to hypothesize mappings covering the unexplained data, all relevant data (form relations and meaning bindings) has to be identified in step 1. All the relevant form relations (F_{rel}) are all relations between the form units from F_{known}, so F_{rel} = {before(*you,throw*), before(*throw,ball*), before(*you,ball)* }. All meaning entities involved in the analysis (found in M_A) are also present in the following meaning relations in situtation S: M_{rel} = {**throw**.THROWER:**self**, **throw**.THROWED:**ball**}.

In step 2 the already mapped items have to be disregarded. The unused data is F_{rem} = {before (*you,throw*), before(*you,ball)*} and M_{rem} = {**throw**.THROWER:**self**,}, because M_A and F_A contained the other meaning relations and form relations already. The super construction C_{super}, derived by replacing terms with constructional references is made up of a form pole {before(YOU$_f$,THROW$_f$), before(YOU$_f$,BALL$_f$)} and a meaning pole {THROW$_m$.THROWER:YOU$_m$}. C_{super} contains a potential construction.

The potential construction C_{U1} is obtained in step 3 by retaining only those relations in C_{pot} that hold over correlated arguments. This excludes before(YOU$_f$,BALL$_f$):

({before(YOU$_f$,THROW$_f$)}, {THROW$_m$.THROWER:YOU$_m$})

Reanalyzing the utterance in step 4, including C_{U1} (in the algorithm called C_U) in existing set of constructions C, should ensure a minimum reduction in cost. If not, then C_{U1} ends up in the pool of potential constructions. This pool contains constructions that may render useful processing future examples.

Unexplained units of form (e.g. *a* in this example) leads to maintaining a pool of utterance-situation pairs in step 5 that are partially unexplained. Further examples involving similar units may together lead to a correct generalization.

4.3.3 Reorganization

Considering known constructions can lead to generalization towards biologically natural categories and constructions covering more than two lexical items. Chang and Maia offer a less detailed description of a data driven bottom-up reorganization process based on similarities among and co-occurences of multiple constructions (figure 4.7). This can lead to a change in the set of existing constructions of what [Chang et al. 2001] call the “constructicon”.

Reorganize constructions. Incorporate a new construction C_n into an existing set of constructions C, reorganizing C to consolidate similar and co-occurring constructions if necessary:

1. Find potential construction pairs to consolidate.
   - **Merge** constructions involving correlated relational mappings over one or more pairs of similar constituents, basing similarity judgements and type generalizations on the conceptual ontology.
   - **Compose** frequently co-occuring constructions with compatible constraints.
2. Evaluate constructions; choose the subset maximizing the posterior probability of C on seen data.

Figure 4.7: Reorganize constructions.

In order to explain the merging of items, consider C_{U1} as known to the infant:

$(\{\text{before}(\text{YOU}_f,\text{THROW}_f)\}, \{\text{THROW}_m.\textsc{thrower}:\text{YOU}_m\})$

Assume also that processing the utterance *she's throwing a frisbee* with the appropriate situation produces C_{U2} (other constructions, e.g. containing *frisbee*, involved in the analysis of this utterance do not play a role here):

$(\{\text{before}(\text{SHE}_f,\text{THROW}_f)\}, \{\text{THROW}_m.\textsc{thrower}:\text{SHE}_m\})$

C_{U1} and C_{U2} bear some obvious similarities. Both contain the same form relations and meaning relations involving the constituent construction THROW. The only thing that differs is that in C_{U1} the other constituent construction is YOU and in C_{U2} it is SHE. This is evidence that YOU and SHE might have something in common. Moreover, the meaning-poles of these two constructions are expected to show a high degree of similarity[6]. The overall similarity between C_{U1} and C_{U2} can lead to a merge of the constructional constituents (SHE and YOU), resulting in a merged potential construction:

$(\{\text{before}(\text{X}^h_f,\text{THROW}_f)\}, \{\text{THROW}_m.\textsc{thrower}:\text{X}^h_m\})$

where X is a variable over a construction constrained to have a **Human** meaning pole (where **Human** is a generalization over the two merged constituents). A similar process including other constructions could produce the following mapping:

$(\{\text{before}(\text{THROW}_f,\text{Y}^o_f)\}, \{\text{THROW}_m.\textsc{throwed}:\text{Y}^o_m\})$

[6] Chang and Maia emphasize that the precise manner by which this is indicated is irrelevant for their work.

in which Y is a variable for a constituent construction constrained to have an **Object** meaning pole.

Besides merging based on similarity, potential constructions may also be composed based on co-occurrence. For example, the generalized **Human**-THROW and THROW-**Object** constructions just described, are likely to occur in many analyses in which they share the same THROW constituent (see figure 4.2 for the functioning of constructions). Since they have compatible constraints in both form and meaning (with respect to the meaning constraints no surprise because they originate from the same conceptual THROW frame) repeated co-occurrence leads to the formation of a larger construction that includes all three constituents:

$(\{{}_{\text{before}}(\mathrm{X}^h_f, \mathrm{THROW}_f),\ {}_{\text{before}}(\mathrm{THROW}_f, \mathrm{Y}^o_f)\},$
$\{\mathrm{THROW}_m.\textsc{thrower}{:}\mathrm{X}^h_m,\ \mathrm{THROW}_m.\textsc{throwee}{:}\mathrm{Y}^o_m\})$

This result is another representation of the THROW-TRANSITIVE construction as shown in figure 4.2.

Both the hypothesis and the reorganization procedure include evaluating these potential constructions using Bayesian criteria which are based on minimum description length.

4.3.4 [Chang et al. 2001] and continuity

The three algorithms produce constructions based on utterance-situation pairs and an existing set of constructions. This makes it possible, according to Chang and Maia, that it can be applied to more advanced stages of language development. The potential continuity between early language acquisition and lifelong constructional reorganization offers hope for the modeling of adaptive language understanding systems, human and otherwise. How exactly the child ends up beyond the stage of verb-specific grammatical constructions is left for the imagination of the reader. Furthermore, [Chang et al. 2001] claim that the skills necessary for the emergence of abstract relational frames are similar to the skills for language acquisition. I will attempt to clarify both issues in chapters 5 and 6.

Chapter 5

The DOP-framework and language acquisition

In chapter 2 I showed how Seidenberg and Allen emphasize the importance of a performance model for approaching language acquisition. Chapter 3 introduced the DOP-framework as an attractive, implemented model for adult language performance. In chapter 4 a model was shown that could explain how the first grammatical constructions are grounded in prelinguistic representations. In this chapter I will try to connect these two models, hoping to find answers to the remaining questions the designers of the DOP-framework still have.

Section 5.1 shows how the non-linguistic component of the DOP-framework is much more mysterious than the linguistic component. Not surprisingly, this non-linguistic component plays a major role in language acquisition, because it is involved in the first mappings between situations and utterances. The remaining question about language acquisition within the scope of the DOP-framework is fine-tuned by splitting it up in subquestions. They will zoom in on the "grey area" where the most important role with respect to language processing is shifting from the non-linguistic component to the linguistic component.

Section 5.3 discusses phenomena with respect to the period of non-linguistic experiences. During this stage the prelinguistic structures are formed which form the basis for the non-linguistic component for language processing. The emergence of these structures involves very important aspects that [Chang et al. 2001] neglect to mention explicitly. I will propose some details in order to make claims about the continuity of the learning processes, not just with respect to language acquisition, but with respect to cognition in general.

The sections 5.4, 5.5 and 5.6 go through the one-word, two-word and multi-word stages respectively, with an explanation of how this all comes about. Each section also contains an analysis, hypothesis and reorganization description as in [Chang et al. 2001], so we can compare these in the last chapter of this thesis. At first sight these processes might seem different, but I will attempt to show that they contain many similarities in chapter 6. In the present chapter it will already become clear that [Chang et al. 2001]'s and [Scha 1990]'s claimed continuity of language acquisition is characterized by the one-word, the two-word and the multi-word stage, because of the natural limitations of the learning mechanism.

The reorganization processes in this chapter, leading to stable prelinguistic and linguistic concepts, involve learning algorithms. Examples of learning algorithms applied to cognitive processes can be found in [Steels 1999], and in the literature about dynamical systems (e.g. [Serra et al. 1990] and [Simon et al. 1995]). The exact details of these algorithms are unimportant for the present goals.

5.1 The linguistic and the non-linguistic component

In chapter 2 I emphasized that language acquisition should be considered as learning how to use the language, rather than identifying its grammar. Therefore, I would like to take the infant, and especially what it might experience and need, as a departure point, postponing the attempt to explain how linguistic structure arises.

In a nutshell this is what happens. From birth (and perhaps before) the infant tries to make sense out of the things that happen around it. Regularities in sensory input become categorized so that new input is recognized based on these categories. Once visual input becomes understood as situations, auditory and situational data are being stored in pairs in memory. Regularities among these pairs of data will be captured by generalization processes, which eventually leads to the one-word stage. In this stage the most rudimentary linguistic corpus becomes available on which a mechanism can operate that finds linguistic analogies. Further generalizations will lead to the two-word and the early multi-word stage. During these stages the infant invents *its own language* ([Scha 1992], [Seidenberg et al. 1999]), with its own categories and structures. This is also called "basic child grammar" in [Chang et al. 2001]. These stages all offer more or less sufficient language skills for the time being. While more and more language experiences (which also gradually become more complex) are processed, the corpus will go through a gradual development and the language that it represents starts to resemble the language of the infant's environment. This process is more or less similar to the language change which a speaker can experience speaking a different dialect or language for a long time (section 3.2). As Scha puts it:

> "Because we do not assume an abstract grammar, there is a complete continuity between the early and the later stages of language use. We do not postulate a separate process of language acquisition." [Scha 1990])

Note that the quote and the text before do not imply that nothing is innate. What *is* assumed to be innate is just different from what is usually assumed: a set of analogy mechanisms[1] or perhaps no more than the psychological principles which lay at the root of these mechanisms ([Scha 1990], [Scha 1992]). These mechanisms have no linguistic features, but are able to operate on different kind of corpora (rudimentary or sophisticated), following the principles of the DOP-framework. If Universal Grammar is defined as the innate tools to acquire a language, then the analogy mechanisms are the DOP-framework's "Universal Grammar":

> "(..) I presume that the innate Universal Grammar is not a grammar, but consists of analogically associating mechanisms. Mechanisms which constitute the basis for matching processes with respect to the corpus, but also for the emergence of new meanings by the development of associations between utterance situations, and for projecting meanings onto utterances. The way in which adult language use can be accounted for by means of matching with respect to a corpus is relatively clear. Compared to that, the question about the beginning of an individual's command of language is much more intriguing: how does our matching algorithm work when there is no corpus yet? ([Scha 1992])

Of course, these mechanisms applies to adults. An adult with a large set of past language experiences uses these together with linguistic context and situational context to reach a decision about what a certain utterance attempts to communicate. By linguistic context we mean the discourse that has taken place sofar. By situational context we mean the situation in which something is said and, possibly, pragmatic information evoked by gestures of the speaker (deixis).

An infant, on the other hand, has to become aware of the fact that something is communicated by the sounds it hears, but in order to discover what this is it can only rely on the situational context. Pragmatic information for an infant is mainly provided by deictic actions of the speaker in a certain situation, whereas for an adult this information is mainly provided by the discourse he or she is engaged in (especially in a conversation which is not about the "here and now"). So, apparently, apart from a linguistic component of language processing (the presently implemented DOP-models), which explains most of adult performance, there must also be a non-linguistic component of language processing, that strongly includes the observed situation. Somewhere in between, the first words get their meanings and next to the non-linguistic component also the linguistic component starts to play a role.

Using this non-linguistic component of language processing the infant tries to map semantic/pragmatic information from the situational context onto the utterances heard. During the language acquisition process the non-linguistic component gets overshadowed by the linguistic component, although the first will keep playing a role. How this works is schematically shown below. Note

[1] Reading Douglas Hofstadter's book "Fluid Concepts and Creative Analogies" ([Hofstadter 1995]) made me consider him as a kind of guru in the field of perception and cognition. The remark "(...) analogy-making lies at the heart of intelligence (...)" and the build up to this remark (p. 63) convinced me that analogy-making also lies at the heart of language processing and acquisition.

that the emergence of prelinguistic structures, as explained in chapter 4, is not covered by the figure.

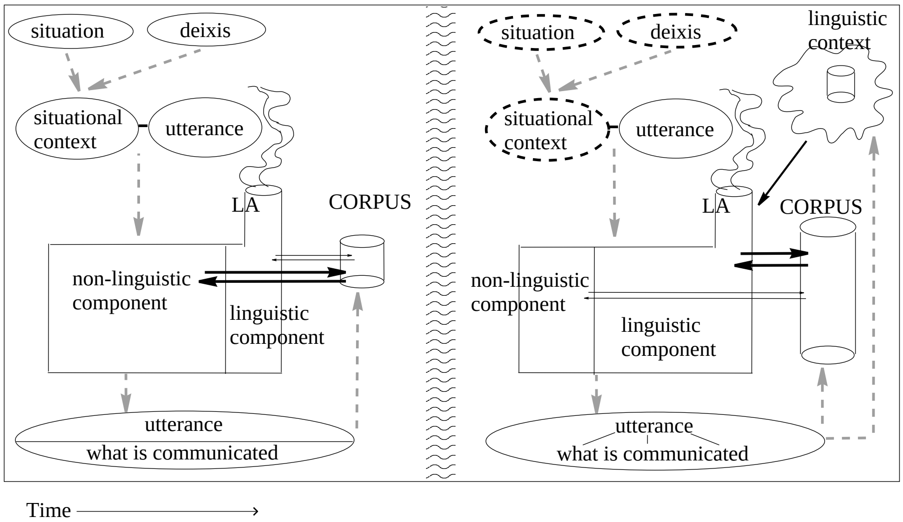

Figure 5.1: Two stages in the language acquisition process.

Consider figure 5.1. The left part shows that situation and deixis at the time of an utterance together form a situational context. In other words, deixis steers the attention of the observer towards a part of the complete situation. The situational context is linked with the utterance[2] and together they form the input for the "language processing factory", which contains a linguistic and a non-linguistic component. At this stage the influence of the linguistic component on the decision on what is communicated is minimal, because the corpus is virtually empty. This is reflected by the small area in the factory taken by the linguistic component in the figure and by the small arrows between the corpus and this component that represent their exchange of information. The output of the factory consists of some kind of representation of the utterance together with a best guess about what is communicated, based on the interpretation of the situational context. This output is added to the corpus. Once the corpus contains data the non-linguistic component will identify regularities which eventually will lead to the situation shown by the right part of the figure. The thick arrows between the non-linguistic component and the corpus indicate this process of identifying regularities. The first regularities found will be approximations of the first word meanings, as we will see further on.

The right part of figure 5.1 shows some changes compared to the left part. First of all it should be noted that the area taken by the linguistic component has become much bigger. This indicates that the corpus contains complex linguistic structures on which the DOP-framework can operate. The

[2] Of course, this is really an acoustic signal, as we will see in the comming sections.

ovals around "situation", "deixis" and "situational context" have become dashed to show that these items have become less (often) important than before for the interpretation process. One could say that "the information that has been communicated" detaches itself from the situational context (real world situations) and attaches itself to sentences and linguistic context (an imaginary world). The linguistic context has partly taken over the role of the situational context. For example, gestures used to identify topic and focus in a situation, play a similar role as definite and indefinite determiners in a discourse. A cloud surrounds the linguistic context to emphasize its short life, because it is built up every time a new discourse starts[3]. In the output of the factory it becomes clear that the connection between what is communicated and the utterance has emerged (see the lines connecting the "utterance"-part with the "what is communicated"-part). Note that the output of the factory is not only added to the corpus but also to the linguistic context. The corpus has grown considerately, because more experiences have been processed. The thick arrows between the linguistic component and the corpus show that there is a lot of traffic between the two. The arrows between the non-linguistic component and the corpus on the other hand show that its activity decreases, but does not disappear. This activity has to do with learning unknown words and constructions at later stages of development, even through the adult age.

The development from the non-linguistic to the linguistic component of language processing is the main concern of this thesis. This is not a trivial development, as becomes clear by the questions in the next section.

5.2 Finetuning the question

> "Our perspective creates the possibility for a plausible model of language acquisition: the gradual development of the linguistic component of language processing, as a result of the gradual growth of the repertoire of linguistic experiences, and the increasing complexity of these experiences. But it will not be simple to describe in detail how language processing takes place in the early, pragmatically and semantically oriented stages, and how the later, more structure-oriented strategies, get bootstrapped out of that." ([Scha 1990])

Based on the information from the previous section and the quote above, the question on page 16: "How does an infant build up 'the first' corpus which forms the departure point for development into later, more complex, corpora?" can be split up in subquestions:

1. What kind of strategies does the non-linguistic component for language processing consist of?
2. How does the continuity with respect to language acquisition come about?
   (a) Do the strategies typical for the non-linguistic component of language processing develop into strategies characteristic for the linguistic component of language processing?

[3] Although a theory covering "past discourse experiences", analogous to the DOP-framework, might be thought of.

(b) If so, how does this development take place?

3. Where does "grammar", as adults know it, begin?

I will attempt to answer these questions implicitly in this chapter and explicitly in chapter 6.

One of my main assumptions is that production of utterances is not lagging behind the interpretation of utterances. That is, the utterances which an infant understands, it is also able to produce. This is obviously not an unproblematic assumption. See e.g. Smolensky's *On the comprehension/production dilemma in Child language* ([Smolensky 1996]), which discusses the claim that infants understand more complex utterances than they produce.

In this thesis an infant's productions are considered to be representative for the different developmental stages of its corpus and offer a helping hand for making certain claims, although an utterance-generating model is not what I am after in this paper. The processing of linguistic and situational input will be the only engine for language acquisition in this work. Moreover, the production of utterances plays no role in finetuning the achieved "grammar level".

The corpus in my interpretation model goes through a zero-word, one-word, two-word and multi-word stage. That is, linguistic units come about consisting of one, two and multiple words, in that order. Going through these stages, the linguistic component starts to get its grip. Following this line of development, I start with the "zero-word" stage in which the prelinguistic structures rule.

5.3 The prelinguistic stage

The non-linguistic component maps the first prelinguistic structures onto acoustic signals in section 5.4, but in this section it will first be explained how the prelinguistic structures will emerge. Past non-linguistic experiences are involved in generalizing towards prelinguistic structures. More specifically, [Maia et al. 2001] describe the development from action-specific relational frames towards abstract relational frames, without language playing a role (section 4.2). They claim that this phenomenon is useful for "behaving in the world" and how it is necessary for the emergence of language. I will investigate this in little more detail in section 5.3.1, before turning to the emergence of a corpus filled with prelinguistic data in section 5.3.2, which in turn forms the basis for further development.

5.3.1 Why prelinguistic structures?

According to my view the abstraction level of the corpus of past experiences makes it possible to mentally simulate never observed situations. Once generalizations have led to categorizations (e.g. **Human**, see section 4.2) and abstract representations containing these categories (e.g. a **push**-frame[4]), the infant can

[4]As I mentioned in section 4.2 I use bold regular (**push**) instead of bold captial (**PUSH**) for frames.

produce different kinds of pushing situations it has never encountered before by imagination. This becomes possible by substituting instances of **Human** in the representation, as indicated by vertical lines in figure 5.2, which shows a situational interpretation.

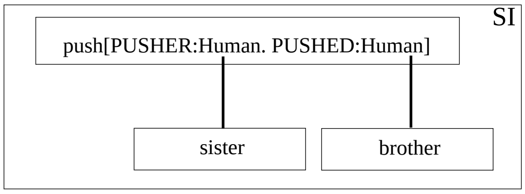

Figure 5.2: Situational interpretation (SI) of a sister-pushing-brother event.

I think that these kinds of generalizations make it possible to imagine and in this way properly predict certain situations, that could possibly happen. These predictions play a role in planning actions by mentally simulating different possible scenarios. This is one of the skills necessary to survive in the world for all primates. Similarly, interpreting events, never exactly encountered before, as being similar to earlier experiences can be of main importance for the next action to take [Hurford 2002].

Furthermore, I think that the creative power of imagining and interpreting by using past experiences is an important prerequisite for language acquisition of an individual and for the emergence of language within a society of organisms. This creativity based on past experiences shows remarkable parallels to the creative power of language, as we will see further on by considering language processing by the linguistic component.

5.3.2 How?

The way prelinguistic structures emerge is not part of the questions (section 5.2) which this thesis explicitly attempts to answer. Nevertheless, I offer a more or less detailed explanation below, because there may be similarities between the processes acounting for acquiring ontology and the processes for acquiring language. This means that the questions about continuity of language acquisition might be extended to the prelinguistic stage. Accordingly, an account of prelinguistic conceptual processes offers a good basis for the details of the one-word stage in the next section (5.4).

As mentioned earlier, past experiences play a major role in conceptual processes leading to prelinguistic structures. Somehow a history of experiences in memory will lead to certain concepts. A data-oriented approach can be found within the theory of Dynamical Conceptual Semantics, in which [Bartsch:1998] describes concept formation processes as follows:

> "In the DCS-model of concept formation and understanding, concept formation on the experiential level consists in creating structures on

growing sets of data. (...) Two types of structures that are established, are on the one hand similarity structures or classifications, i.e. similarity sets of situations formed to begin with under certain basic perspectives such as color, form, taste, touch, sound, form, motoric behaviour, and on the other hand they are contiguity relationships. These are factual relationships in space and time. (...). Thus a banana is classified together with apples and pears as fruit, because of its similarity to these due to identities in the qualities of their contiguity relationships to eating, tasting, touching these things, and, with further experience about these things, to growing on trees and other plants. In this way we become to see situations and objects as belonging to certain classes and to the life histories of individuals or other historical entities.

(...) our set of data does not only grow, but the data also changes in the manner they are understood, and in this way old data can become new data. In principle this spiral like process of understanding a situation can go on, and will practically terminate differently in different settings and different states of development, i.e. according to different structurings available under the perspectives created by different practical contexts, and available at a certain moment in the history of an understanding individual." ([Bartsch:1998])

As we can see low level features like taste, color, form, sound, motoric behavior etc. play a role in concept formation, e.g. leading to the concept of a banana. In addition to these, there are features or contexts to a banana that are close to those of an apple, which eventually leads to the formation of the concept fruit.

In the following, I will attempt to identify the processes involved in concept formation. This leads to three descriptions, which, applying the terminology of [Chang et al. 2001], I will call: REORGANIZATION, ANALYSIS and HYPOTHESIS. Leaving one of these processes out would impair or delay the process leading to concept formation. Consider figure 5.3 .

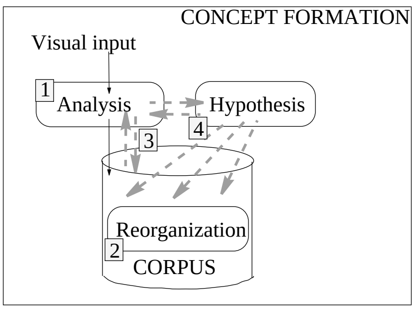

Figure 5.3: A schematical view on concept formation.

The development of the corpus can be described as cycle. The cycle starts with an empty corpus. Accordingly, the input, consisting of visual input[5] is represented in low level features (taste, form, spatial aspects, temporal aspects etc.) by ANALYSIS, without manipulation based on earlier experiences[6] [1]. These representations are collected in the corpus, until certain regularities emerge and are identified as abstractions by REORGANIZATION [2]. From now on visual input can be analyzed by assigning representations to it, based on these abstractions present in the corpus [3]. This implies that REORGANIZATION has reprocessed old data on the achieved abstraction level, otherwise ANALYSIS would not have any data on this abstraction level to make analogies with. Whether the old data disappears is an interesting question, which will be picked up further on. HYPOTHESIS monitors the reorganization process in the corpus and projects its findings [4] onto data ANALYSIS is unable to represent on the new level of abstraction. This way HYPOTHESIS helps ANALYSIS to fulfill the "expectations" evoked by REORGANIZATION, without needing as much statistical evidence. The corpus now gets filled with structures on another abstraction level than we started out with and a new development cycle can start. Furthermore, HYPOTHESIS can speed up the learning process this way. The functioning of these processes can be explained in more detail as follows.

Reorganization

Below I describe how visual input, first represented in low level features, becomes represented on higher level of abstraction.

First of all, I want to emphasize that I do not agree with the way action specific relational frames emerge according to [Maia et al. 2001]. Consider the first time an infant observes a scene in which, according to adults, the infant's sister pushes the infant's brother. According to [Maia et al. 2001] the infant represents this situation with help of concepts that are in fact typical for a stable overall concept of pushing: a frame including the roles PUSHER, PUSHED and **push** (section 4.2).

In my view, these concepts, and thus the relational frame are not available this first time, but they will emerge as follows. The first-time-pushing situation is represented with low level features like color, form, motoric behavior, etc. Let us assume that the sister Anna and brother Bill are conceptualized already. The low level interpretation could then be paraphrased as (for more examples see [Hurford 2002]):

> **anna moves** [pause] **anna touches bill** [pause] **bill moves**

A conceptualization like this is not present in [Maia et al. 2001]. I would like to call it a "situation specific" representation. Although the paraphrase is probably incomplete, it is meant to show that the concept of pushing is unknown. The pauses indicate that the parts are experienced as autonomous units with

[5]Auditory input is not considered in this stage, to be able to focus on the emergence of prelinguistic structures. Obviously, it will play a major role in the processes leading to the one-word stage described in section 5.4.

[6]This makes the recognition of low level features an innate skill.

a temporal aspect connecting them. Spatial features underlie the moving actions and motoric behavior underlies the touching. Assume that several of these events are observed including different humans and differences in the force of the moving and the touching. Abstracting over the low level features and the referents involved leads the three units to become merged into a coherent whole, an action specific relational frame:

push[PUSHER:**Human**. PUSHED:**Human**.]

The process described above is a development from situation-specific low level representations towards action specific relational frames. A similar process leads to the emergence of abstract relational frames (see also section 4.2). These reorganization processes can be seen as induction processes; conclusions, based on re-occurring phenomena. Summarizing the above processes leads to the following:

REORGANIZATION Given a current set A of low level situational interpretations, create higher level concepts by identifying co-occurring regularities among its members and add these concepts to A.

The description above does not necessarily imply that the old data disappear. Instead, structures are created on sets of old data. In other words, the "reason" that data is clustered is represented by a new, more abstract, structure. When these abstract structures, now also representing the old data, are stable enough, they will be involved in representing future input instead of the lower level representations. This is why the set of existing analyses is said to have been reorganized. I will show how this works below by describing the interpretation process.

Analysis

Assume that REORGANIZATION has done its work in the first development cycle. The input still consists of visual stimuli, but they are not consciously experienced as low level features anymore (more on this observation in chapter 6).

Interpretation is taken care of by ANALYSIS which attempts to assign a representation to new input. Consider ANALYSIS processing a never encountered john-push-anna situation at a stage where the corpus already contains the situational interpretation **push**[PUSHER:**john**. PUSHED:**bill**.] for a john-push-bill situation and **push**[PUSHER:**bill**. PUSHED:**anna**.] for a bill-push-anna situation. A representation can now be achieved as follows. First of all I assume that parts of the input are recognized and represented on the new abstraction level. I assume furthermore that an analogy mechanism is able to derive representations for the new input by combining these parts of known analyses. In this case the mechanism would lead to two derivations by combining different parts of old data based on their shared "**Human**-category":

1. **Human[john]** combined with **push**[PUSHER:**Human**. PUSHED:**anna**.]

2. **push**[PUSHER:**john**. PUSHED:**Human**.] combined with **Human[anna]**

The two derivations lead to the same structure (see also the example in fig. 4.2). In general, the analogy mechanism can lead to different possible structures. The probabilities of the structures are calculated based on the probabilities of their derivations. The most probable structure will be the perceived structure. More schematically, this leads to:

ANALYSIS Given a visual input and a current set of situational interpretations A, produce the best fitting situational interpretation based on earlier experiences and add it to A.

Hypothesis

When ANALYSIS produces an analysis on the previous level of abstraction, HYPOTHESIS might be able to make a proposal on the current level of abstraction. Consider analyzing a john-push-ball situation at a stage where **push**[PUSHER:**John**.PUSHED:**Human**.] is available and **push**[PUSHER:**John**. PUSHED:**Object**.] is not. Assume also that **ball** has been categorized as **Object**, which in turn is present in other analyses than those covering a pushing situation. HYPOTHESIS might "bluntly" propose the analysis **push**[PUSHER:**John**. PUSHED:**ball**.] instead of REORGANIZATION which would need more evidence. Knowing that **ball** belongs to the category **Object** makes the abstraction **push**[PUSHER:**Human**. PUSHED:**Object**.] available. If the hypothesis happens to be wrong it will not be reinforced later and, thus, play an continually decreasing role in future interpretations (see also section 3.2 about virtual grammars). If it happens to be applicable to later interpretations it will "survive", because its statistics will increase. In a nutshell, HYPOTHESIS can be described as follows:

HYPOTHESIS Assist ANALYSIS in assigning interpretations to new input on the present level of abstraction instead of using low level concepts, based on knowledge of the changes that REORGANIZATION has accomplished in the corpus.

5.3.3 Remarks

The processes of REORGANIZATION, ANALYSIS and HYPOTHESIS show a remarkable overlap with the experience-driven interpretation of input mentioned in chapter 3 which introduced the DOP-framework, and with the algorithms of [Chang et al. 2001] treated in chapter 4. Together they seem to cover a re-occurring cycle that represents input in the present level of abstraction until the next level of representation is achieved. This seems to be an ongoing process. I will return to these observations in the next sections, where making analogies and generalizations within the corpus is easier to explain, because of the decreased importance of the lower level features of concepts.

I assume that in the prelinguistic stage the infant is able to build up a corpus that contains structures representing situational interpretations in the way [Chang et al. 2001] suggest. These can be either abstract relational frames, like the ones represented in figure 5.2, or action specific relational frames, depending on the stage the infant is in (section 4.2). Bod and Scha are interested in the development from a corpus with non-linguistic experiences towards a corpus with linguistic experiences ([Bod et al 1996]). I will show in the next sections that in between these stages there is a period in which non-linguistic experiences are enriched with linguistic ones.

5.4 One-word stage

As stated before, the one-word stage is assumed to be reflected by the corpus. An earlier corpus containing prelinguistic structures together with acoustic signals will lead to this new corpus. In this section I describe how the corpus becomes structured in such a way that it starts to contain words associated with instances of **Human**, **Physical objects**, **Physical actions**, **Spatial locations** etc., the biologically natural categories mentioned by [Maia et al. 2001]. Instead of a lexical construction as in figure 4.3, I would like to represent the eventual lexical entry for *push* and *brother* as an abstraction over accompanying situational interpretations as in figure 5.4. The difference in representation is mainly visual. Here the abstract relational frame is used as the meaning for *push* and the semantic category of *brother* is **Human**:

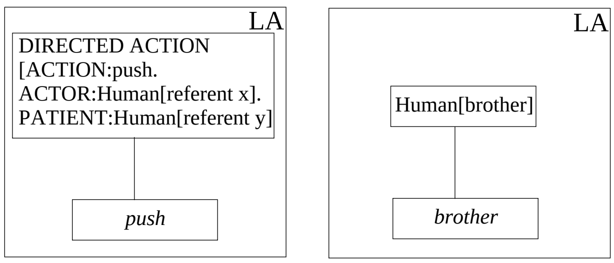

Figure 5.4: Linguistic abstraction (LA) *push* and *brother*.

But how is this situation achieved? First of all, I would like to show what the one-word stage is, looking at some real world data. Next, I will show my view on how the one-word stage could come about by a REORGANIZATION process. Then, aspects of interpretation, ANALYSIS and HYPOTHESIS, in the one-word stage will be discussed. In section 5.4.3 I conclude by making some important remarks.

5.4.1 What is it?

I mentioned earlier that an infant invents its own language (see section 5.1). It will do so based on the linguistic and situational input it gets from its immediate environment. At an early age that will be in and around the house. The parents account for a large part of the linguistic input. The fact that the infant is inventing its own language is the reason I scrutinize the linguistic data involving only one child. This way it becomes possible to catch some of the specific regularities (e.g. a frequently occurring coffee-drinking event accompanied by appropriate utterances) provided in a specific environment (a family at home with a certain stable number of entities and events etc.).

The databases of the Childes-project contain suitable linguistic data for this purpose. This project was started in 1981 [CHILDES]. The abbreviation stands for Child Language Data Exchange System. The system provides tools for studying conversational interactions. These tools include a database of transcripts, programs for computer analysis of transcripts, methods for linguistic coding and systems for linking transcripts to digitized audio and video. Moreover, the databases contain both child language and child directed speech.

The data I will look at consists of Jacqueline Sachs' longitudinal study of her daughter, Naomi, who was born June 8, 1968. The transcripts cover the time from age 1;2.29 to 4;9.3 [7]. In this period every infant goes through the one-word stage and the two-word stage to eventually reach a mature stage of multi-word language. The recordings took place with an average of two times a month. I will consider the period leading to the early multi-word stage, until around age 2;2.0.

The first words uttered by Naomi are words like:

> *mommy*, *daddy*, *boy*, *girl*, *doggie*, *snow*, *book*, *eggies*, *cheese*, *money*, etc..

These examples do not show that, for example, the concept of a mother is acquired. Many of these words are used by Naomi as proper names, although the use of *girl* and *boy* seems to indicate knowledge of the concept of gender.

Also some other words like:

> *yes*, *no* and *again*

are used for basic interaction. The word

> *down*

is the first word which might be an instance of the biologically natural category **Spatial location** or **Path**. Although this word's frequency is high in the child directed speech, the real reason that Naomi acquires it before other tokens of this category, might be the repeated use with short intervals in the same situation (changing the diaper and telling Naomi to lay down all the time).

Around the age of 1;06.0, the first actions labeled by Naomi with words are:

> *look*, *see*, *eat*, *hit* etc.

[7] The age of "1;2.29" means 1 year, 2 months and 29 days.

The most frequent verb through the whole recording period is

> *want.*

Its frequency is twice as high as that of the next verb.

From the conversations in the database it can be inferred that uttering any of the mentioned words covers more than just the meanings as adults know them. On many occasions the utterances mean:

> *that is a [book],*
> *this is [cheese],*
> *do that [again],*
> *I am going to [eat] it,*
> *I [want] that one.*

All of these utterances contain implicit arguments (*that, it,* etc.). But there is also a different kind of meaning not demanding an implicit argument:

> *there is a [book],*
> *look, a [doggie]*

The first examples demanding implicit arguments can be considered as predicates, the latter as propositions. Why these observations are interesting and more peculiarities of the one-word stage will be explained in the next subsection.

5.4.2 How?

The main assumption that is crucial for [Chang et al. 2001] and many others is that the infant expects correlations and mappings between what is heard and what is perceived. I would like to rephrase this as: the infant *finds out* that it can expect correlations and mappings between what is heard and what is perceived. To be more specific, I think that the infant discovers the possibility of the mapping of frames onto words while it develops towards the one-word stage, not earlier. Similarly, expecting other correlations (e.g. mapping word-order onto role-bindings) does not happen until the first correlations have emerged (section 5.5.2). I also assume that the information conveyed by an utterance is often salient to the infant in the situational context:

> "In interacting with live human speakers, who tend to talk about the here and now in the presence of children, the child can be more of a mind-reader, guessing what the speaker might have meant (...). That is, before children have learned syntax, they know the meaning of many words, and they might be able to make good guesses as to what their parents are saying based on their knowledge of how the referents of these words typically act (for example, people tend to eat apples, but not vice-versa). In fact, parental speech to young children is so redundant with its context that a person with no knowledge of the order in which parents' words are spoken, only the words themselves, can infer from transcripts, with high accuracy, what was being said (...)." ([Pinker 1995])

Figure 5.5 below zooms in on the non-linguistic component[8] of figure 5.1. A new development cycle is about to start where the previous one ended in section 5.3.2. There, it resulted in a corpus with prelinguistic structures and an ANALYSIS which was able to represent new visual input based on these structures instead on low level features. The current cycle starts out with this state of affairs, with one important difference. Before, the auditory input remained unconsidered, but now the prelinguistic structures in the corpus are connected to their accompanying acoustic signals. So from now on the input consists of visual and auditory input. The acoustic input will conveniently be represented as words in an utterance, e.g. *john pushes bill*.

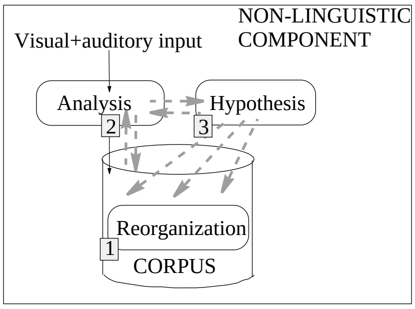

Figure 5.5: The non-linguistic component.

Obviously, figure 5.5 has much in common with figure 5.3. In this cycle of the learning process towards linguistic structures REORGANIZATION operates on situational interpretations together with acoustic signals to converge on abstractions over situational interpretations [1]: approximations of word meanings. Once approximations of word meanings become available ANALYSIS get its first linguistic grip, because it will be able to associate new acoustic input with these approximations [2]. So, next to analyzing visual input it will also analyze auditory input. HYPOTHESIS will actively search for more "word meanings" [3], mimicking the results achieved more passively by REORGANIZATION. Mapping uncovered parts from acoustic signals onto parts from the situational interpretation, and vice versa, is part of its strategy.

First, I will show below how I think the first "word-to-meaning" mappings emerge in more detail by describing the REORGANIZATION process. Then the processes of ANALYSIS and HYPOTHESIS follow.

Reorganization

In the previous section (5.3.2) I showed that regularities in the prelinguistic input lead to the emergence of relational frames. At a certain point this process goes hand in hand with a reorganization with respect to auditory data. In other words, regularities within the situational interpretations become linked

[8] The non-linguistic component maps prelinguistic structures onto acoustic signals, whereas the linguistic component only operates on complex linguistic structures.

with regularities within their accompanying acoustic input. The two forces steering this process can be described as follows:

1. Regularities within the set of situational interpretations trigger a search for co-occurring regularities in the acoustic signals.

2. Regularities in the acoustic signals trigger a search for regularities in the accompanying situational interpretations.

The two described processes 1 and 2 will attempt to find a balance by checking eachother's findings. This implies that they function parallel and exchange data until a stable compromise is achieved. Imagine that in process 1 *push* emerges as a co-occurring acoustic signal with [PUSHER:**Human**] (section 5.3.2). In process 2 it might become clear that all of the tokens of *push* have a broader overlap than just [PUSHER:**Human**] in their accompanying situational interpretations. This result is then given back to process 1 to check its viability, and so on. Note that the process can also start with a regularity in the acoustic signals. Summarizing the following could describe REORGANIZATION typical for the one-word stage:

REORGANIZATION Given a set of analyses A, consisting of situational interpretations and acoustic signals, let co-occurring regularities among both identify "meaning-word" pairs. Reprocess old data to incorporate the findings in A.

The results are added to the corpus of past experiences as a linguistic interpretation. This means that up to now the following analyses populate the corpus:

- Situational interpretations from the stage when acoustic signals are regarded as noise.

- Situational interpretations together with acoustic signals consisting of unknown words.

- Situational interpretations together with linguistic interpretations containing an approximation of a word meaning unified with the situational interpretation.

The way the first word meanings emerge by REORGANIZATION leaves room for some special cases. For example, it makes it possible that a certain sequence of acoustic signals, known to adults as distinct words, gets glued together as one word with a certain meaning. The infant often does not know where one word stops and another starts. This way *boot off* can become a single word according to its corpus. This indicates that the words of child language do not have to be the same as those of adult language. In a later phase of language learning these "words" will be split up into the "right" parts, based on encountering more tokens of *boot* and *off* in other contexts.

Here is another special example. Imagine a family in which the father is the only one smoking cigarettes. The infant only has father-smoking-cigarettes situations available. This means that the prelinguistic representation of these situations does not get the chance to reach an abstract level, because the input does not contain enough variation (section 5.3.2). It is likely then that a word like *smoking* gets mapped onto a situational representation which does not contain slots for different arguments. Instead the father and the cigarettes also belong to the meaning of smoking. Once the infant's world is not limited anymore by the family home, it will have the chance to encounter other people smoking and other things being smoked. The meaning of *smoking* will then get slots for **Humans** and **Objects** (or possibly the subcategory "smokeable objects").

These two examples clearly show that the regularities in the input influence the characteristics of the "word meanings" that emerge. The meanings are clearly different than how adults know them. For example, the last example can be considered as a word-to-meaning mapping at an early stage of language acquisition, which covers a more specific situation than *smoking* does for adults. It also indicates that more than one meaning can emerge, which in turn can merge in a later stage, leading to a more abstract meaning. Of course, the meanings, in their turn, influence the process of ANALYSIS and HYPOTHESIS discussed below.

Analysis

The input consists of visual and auditory input. The visual input leads ANALYSIS to achieve situational interpretations as explained by the previous development cycle. Once the first approximations of word meanings have emerged, ANALYSIS can also use these data for interpreting new input containing these words. Below, *push* is considered to have a meaning according to the corpus. Furthermore, it is assumed that a part of the situation is identified by deixis. If this is not the case, I assume that the infant is aware of the fact that the situation does not contain informative information. Note that it does not always have to be the speaker that identifies part of the situation. It might very well be that the infant looks at something, which triggers the speaker to talk about it.

Consider analyzing input in which the situational interpretation contains informative information and the acoustic input contains *push*. The meaning of *push* is retrieved from the corpus. If the meaning is a more general case of the observed situation containing a pushing action, the meaning gets reinforced, by adding the meaning to the corpus. So if the meaning is **push**[PUSHER:**Human**. PUSHED:**Human**.] and the situation contains **push**[PUSHER:**anna**. PUSHED:**bill**.], then these two are unifiable which indicates that the meaning is appropriate. The linguistic interpretation including the unified data is added to the corpus.

On the other hand, if the retrieved meaning is different, but equally or more specific than the situational representation, this could trigger REORGANIZATION. Imagine that all observed pushing situations involved **john** pushing a **ball**, then the meaning of *push* also contains these items. This is what I called a situation specific representation (5.3.2), which can, as we see here, become a (temporary) word meaning. If the situation contains a never encountered situation specific representation of Anna pushing Bill, then the meaning of *push* is equally specific as the current situational representation. This observation can start a reorganization process in order to achieve a more general meaning.

It is possible that it is clear to the observer that the situational interpretation contains informative details, but that these are absolutely not compatible with the retrieved word meaning. In that case HYPOTHESIS might be triggered to introduce a new word meaning.

In case it is obvious to the observer that what is said is not about the "here and now", the situational representation plays no role in the analysis. Instead, ANALYSIS has to rely on the word meanings in the corpus to make something out of the input. There might be word meanings that cover a complete situation, like the specific meaning of *push* above, containing **john** and **ball**. In this case ANALYSIS will retrieve this meaning and the observer will mentally simulate this situation, whether it was meant by the speaker or not.

On the other hand, the meaning of *push* might have reached a more abstract meaning. If ANALYSIS retrieves this meaning the empty slots are not informative enough to simulate a situation. Instead, ANALYSIS might return to the corpus and retrieve the most frequent linguistic interpretation for *push*. This might be the situation that the mother pushes the infant, assuming that it has been frequently observed.

ANALYSIS in the one-word stage can be summarized as follows:

> ANALYSIS
> Given a current set of analyses A and an input containing known words, find the best fitting interpretation based on A and add this to A.

Hypothesis

As soon as the first word meanings become available HYPOTHESIS will actively attempt to identify more word meanings. Consider an input consisting of *push car* and a situational representation containing a car pushed by John. Assume that *push* and *john* are known, but *car* is not. ANALYSIS will be able to account for everything except for *car* and **car**, by unification. HYPOTHESIS maps these parts onto eachother, which leads to a potential word meaning. An implemented example can be found in the work of [Siskind 1994]. This process reflects the fact that the infant already in the one-word stage expects words to have meanings (5.4.2). Of course this expectation speeds up the learning process. Depending on the future applicability of this hypothesis it will either survive when it is used frequently or disappear when it is never used again (see also section 3.2 about virtual grammars). In a nutshell, HYPOTHESIS can be described as follows:

HYPOTHESIS
Assist ANALYSIS in assigning interpretations to new auditory input on the present level of abstraction instead of the previous level, based on knowledge of the changes that REORGANIZATION has accomplished in the corpus.

5.4.3 Remarks

As shown above, once stable linguistic interpretations start to populate the corpus, the linguistic component starts to get its first, albeit soft, grip. During the interpretation of a situation and an utterance every familiar word will be represented with its meaning, which is an abstraction over linguistic interpretations. Compatibility with the interpretation of the situation will be checked. Errors here might for example indicate an ambiguous word or the wrong focus of attention. Words that are familiar but not compatible with the situational interpretation do not end up in the analysis. On the other hand, familiar words that are compatible with the situation do become part of the linguistic interpretation. In figure 5.6 an example is shown[9]. The complete analysis is added to the corpus.

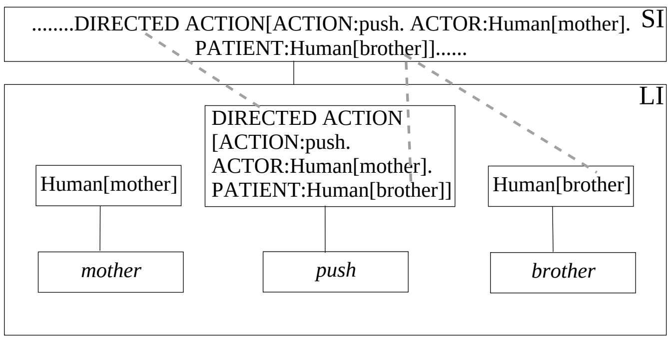

Figure 5.6: An analysis consisting of a situational (SI) and linguistic interpretation (LI) in the one-word stage.

The dotted lines in the figure visualize some traces of unification processes between the situational and the linguistic interpretation. The dots within the situational interpretation show that it might contain more than just the **DIRECTED ACTION**.

Important is the assumption (nr 1) that the infant is trying to make sense of the world and with observing a situation and hearing some compatible words this is going just fine. This also implies that word order is not an issue in the one-word stage; hearing and understanding each word in *john eats a sandwich* is just as good as *a sandwich eats john* while observing a situation in which John eats a sandwich. Processing the utterance *mother push brother* and

[9] I have incorporated **Human** into the analysis in order to preserve the constraints (compare figure 5.2).

the situational interpretation in figure 5.6 means experiencing each word on its own and checking its compatibility with the situation. Experiencing these words close to eachother keeps them together, so in the linguistic representation all words are present. This is analogous to the low level interpretations of a pushing situation kept together by temporal aspects in section 5.3.2. To be more specific, the infant does not "know" the phenomenon of combining words, because the corpus does not contain structures like this. This is why within the linguistic interpretation in figure 5.6 the items are not connected with each other.

With the approximations of word meanings available claims can be made about production. In a pushing situation an infant might utter *push*, whether the word meaning is specific or more abstract. In the first case the meaning of *push* might contain **john** and **ball**. It may be uttered, even in a situation in which Anna pushes Bill, because of its similarities with this situation. In the other case, the more abstract meaning of *push* plays the following role. Uttering *push*, the infant knows that the roles of the meaning of *push* are filled in by the observed ACTOR and PATIENT. This means that *push* takes implicit arguments, which is comparable with the unification mentioned earlier. Obviously, it is not easy, if not impossible, to identify whether the infant has a specific or more abstract word meaning available.

[Maia et al. 2001]'s examples cover a limited range of biologically natural categories (section 4.2). As I noticed in section 5.4.1 the frequency of *want* in Naomi's utterances is very high and she starts to use it almost as early as the **Physical action** verbs. It seems to me that she must have a stable representation available at an early stage to map *want* onto. This is probably also the case for some other early acquired words (*know*). What these representations look like and how they come about remains uncertain for now.

On the other hand, many function-words, e.g. *the* and *a*, that become more valuable participating in discourses are not mapped in this stage, because there is not yet enough discourse experience to have achieved representations to map them onto. That is why in this stage they do not end up in representations like figure 5.6.

The similarity of the learning process, typical for the one-word stage, to the one in section 5.3.2 about the prelinguistic structures is striking. This observation will be connected to the continuity of the language acquisition process in chapter 6. The presence of data in the corpus like depicted in figure 5.6 is the departure point for the emergence of the two-word stage, which is described in the next section.

5.5 Two-word stage

The two-word stage constitutes an important and logical stage in the development of the corpus. Explaining it in the same way as the preceding sections gives interesting insights.

During the one-word stage analyses consisting of abstract relational frames (situational interpretations) together with familiar words as linguistic interpretations started to populate the corpus (figure 5.6, section 5.4.2). In the linguistic interpretations the words were not combined yet.

Subsequently, I will show that REORGANIZATION among these analyses in the corpus leads to the first two-word structures. This in turn results in the expectation of such structures in new input, which means that the ANALYSIS and the HYPOTHESIS procedure have these two-word structures to operate on. The most important aspect of the two-word stage is that the phenomenon of unary predicates emerges. Before going into detail of how this comes about in section 5.5.2, I will show how we can recognize whether the infant has arrived, or is close to arriving, in the two-word stage.

5.5.1 What is it?

Two-word utterances are regarded as the first "grammatical constructions" ([Chang et al. 2001]). The infant in this stage is considered to have acquired a basic syntactical knowledge ([Pinker 1995]). As explained earlier, a two-word utterance in adult language can in fact be a single word according to the infant's corpus. Consider some of Naomi's supposedly two-word utterances:

get down	*get up*
put down	*put back*
fall down	*fell out*
sit down	*sit up.*

It seems reasonable to assume that she knows how to combine *down* with several action-words and that the utterances are not single words, because of the variation in the combinations. Also present in the corpus are:

eat toast
eat it
eat them
push it
push me
push that.

These utterances also seem to be two-word utterances, but we cannot be sure, because of the following ones:

eat it cheerios
eat it icecream.

In these utterances *eat it* might very well be one item according to the corpus. This might, hypothetically, also be the case for *eat them* and *eat toast*. Of course, if we could look into Naomi's corpus of past language experiences, it would be possible to make better claims about "real" two-word structures. If Naomi's utterances include combinations not literally present in the corpus she can certainly be claimed to have arrived in the two-word stage.

Furthermore, it is interesting that two-word "object-less" constructions like:

piggy eat
kangaroo eat
I eat

are far less frequently uttered than utterances like the "subject-less" sentences above. Not so surprisingly, intransitive verbs occur in Naomi's early utterances in an expected way:

nomi talk
daddy sleeping

Finally, I offer some examples including prepositions:

dress on
jacket on
shoes on
on floor
in there
on table.

In the following section I will deal with the details involved in the emergence of the first two-word structures.

5.5.2 How?

After having reached the one-word stage, more and more stable word-to-representation mappings become available. I showed that these stable constructions can be used by an analogy mechanism that assigns representations to new input, which results in analyses as shown by figure 5.6 (in section 5.4.2, page 50) and figure 5.7 below. These are the kinds of analyses populating the corpus. Remember that the words are not yet combined. From now on, I will assume that the infant represents input (linguistic as well as situational) by abstract relational frames only.

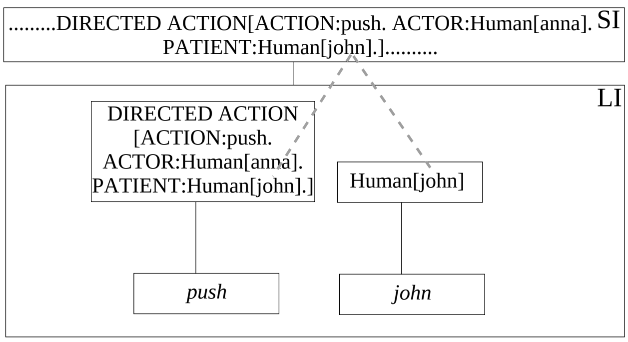

Figure 5.7: A situational and lingustic interpretation close to the two-word stage.

Note that in figure 5.7 the argument slots of meaning of *push* have been unified (the dotted lines) by the referents from the situational interpretation. This was explained as the phenomenon of implicit arguments on page 50 in section 5.4.2. Futhermore, **anna** is the topic of the communicative act and thus omitted in the utterance. The dots in the figure indicate that more things might be going on in the situational interpretation.

A new development cycle can start. The processes REORGANIZATION, ANALYSIS and HYPOTHESIS steering learning and using the two-word constructions can be visualized schematically in a similar way with figure 5.8 below as for the one-word stage.

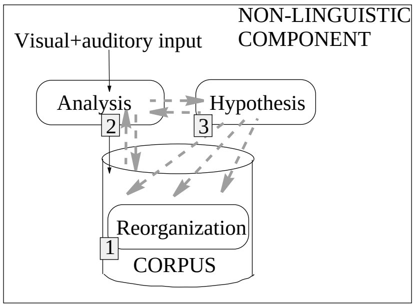

Figure 5.8: The non-linguistic component.

REORGANIZATION now operates on analyses in the corpus built up out of situational and linguistic interpretations. Frequently co-occurring regularities with respect to the semantics and the "word order" within these analyses will lead the REORGANIZATION process to combine linguistic units from the linguistic interpretation preserving these regularities [1]. Once the linguistic combinations populate the corpus ANALYSIS can use them to process new acoustic input, by making analogies [2] it has not made before: two-word structures. If analogous material is absent, HYPOTHESIS might propose a two-word structure [3], mimicking the results of REORGANIZATION. This way, ANALYSIS might be prevented of only applying the former level of abstraction. The three processes of REORGANIZATION, ANALYSIS and HYPOTHESIS are described in more detail below.

Reorganization

From an adult point of view, the input Naomi receives shows certain frequent characteristics. An example of these characteristics is shown in the following fragment; MOT means mother, CHI means child and FAT means father:

CHI: try again
MOT: push with your feet
MOT: push with both feet
MOT: push
MOT: push

MOT: whoopsie Goldie fell down again
MOT: put Goldie through the chimney honey
MOT: put Goldie through the chimney
MOT: there's Goldie again
MOT: put Goldie through the chimney honey put Goldie through the chimney
MOT: Goldie fell down
MOT: poor Goldie
CHI: fell down
MOT: fell down again
MOT: here let's put it like this so she won't fall again.

Even some questions have the following form:

FAT: put baby in the boat ?

Statistically, many utterances from the parents have the characteristics of imperatives, especially when the parents' questions are not taken into consideration. Others have the features of what adults know as a verb phrase, e.g. *fell down again*. In both cases the parents leave out the topic, or given information, of an utterance, so that the focus, or new information, remains. This is especially the case when something is repeated more than once. In the examples above the subject is the topic and is therefore often omitted. These characteristics in the input result in high frequencies of data as shown in figure 5.7 on page 53. To show how REORGANIZATION makes generalizations, I would like to take data like this as an example.

REORGANIZATION identifies two kinds of regularities within a set of analyses as shown in figure 5.7:

1. The frequency of the order in which the lexical mappings labeled with **DIRECTED ACTION** and **Human** are encountered within the linguistic representations is high ([1] in figure 5.9 below).
2. Both meanings of the linguistic interpretations involve the same referent and it always has the same PATIENT-role ([2] in figure 5.9 below).

These re-occurring regularities are indicated by the dotted lines in the following figure, representing the mentioned set:

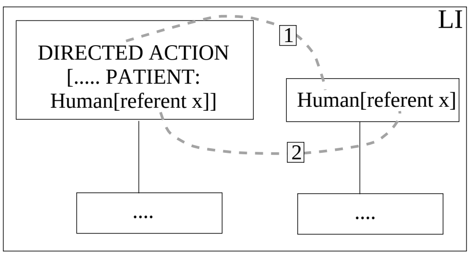

Figure 5.9: An example of co-occurring phenomena leading to linguistic combinations.

By means of unification with the situational representations (figure 5.7) it is this way discovered that the **PATIENT**-referent in the meaning of the **DIRECTED ACTION** is the same as the referent in the meaning of its following neighbour with the label **Human**. This makes the situational representations redundant. I propose figure 5.10 as the result of this discovery: the linguistic abstraction[10] leading to the first two-word structures.

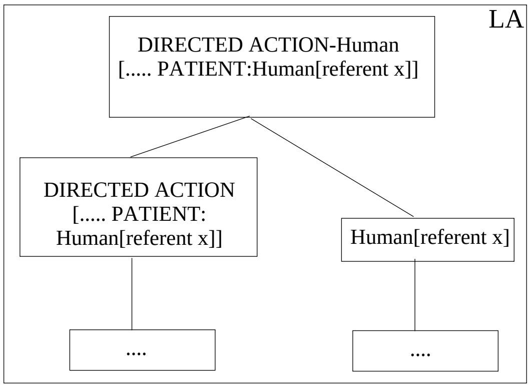

Figure 5.10: A linguistic abstraction leading to two-word structures.

Note that I introduce the term "DIRECTED ACTION-Human" as a mnemonic name for this phenomenon. It should be read as one thing, so from now on I will abbreviate it to DAH (or DAO, involving a **Object**). The resulting meaning is compatible with the original situational interpretation. This new representation applies to the complete set. Thus, all tokens (*push*, *kick*, *hit*, etc.) from the cluster are from now on explained by it.

This means that the infant expects **DIRECTED ACTIONS** to be able to incorporate an entity of type **Human**. Specifically, in the language of Naomi's environment, the referent of the "**Human**-unit" following the "**DIRECTED ACTION**-unit" ends up in its PATIENT slot. This shows that the category order (**DIRECTED ACTION** before **Human**) has been mapped onto the role binding [PATIENT:**Human**[referent x]]. Furthermore, this makes **DIRECTED ACTION** words unary predicates, which take an explicit argument as opposed to the situation in the one-word stage in which they took an implicit argument from the situational interpretation.

I have made the implicit assumption that words become clustered in a group because their meanings are clustered in a group. For example the words *john*, *anna*, *bill*, etc. are clustered because their meanings are clustered into the category **Human**. That is why I use **Human** as a category label for words belonging to this group. One could say that at this level of abstraction the linguistic and semantic categories coincide. Later on this will not be the case

[10]The combination of the two meanings can be seen as a lambda-formula: λreferent-x(meaning left daughter)(meaning right daughter). For readability reasons the present notation is choosen.

anymore. Although the semantic category **Human** will remain, the linguistic generalization will go on (chapter 6).

More abstract representations like this, containing a unary predicate, will emerge by similar generalizations, e.g. **DIRECTED ACTION-Path** (DAP), **LOCATION-Object** (LO), **Object-Path** (OP) or **Human-Path** (HP). Some of Naomi's utterances show this as follows:

> *put down* (DAP)
> *take off* (DAP)
> *on table* (LO)
> *in mouth* (LO)
> *me up* (HP)
> *it up* (HP)

Summarizing REORGANIZATION for the two-word stage leads to:

REORGANIZATION
Given a set of analyses A, consisting of linguistic interpretations containing more known words, let co-occurring regularities among them identify two-word structures. Reprocess old data to incorporate the findings in A.

Having reached this level of abstract representations the interpretation process can treat more complex linguistic structures than in the one-word stage. The next section shows how this works.

Analysis

The input still consists of visual and auditory input. Characteristic for ANALYSIS in the two-word stage is that once the corpus contains two-word structures as depicted in figure 5.11, it is able to make two-word-analogies. Note that the situational interpretation is absent in figure 5.11.

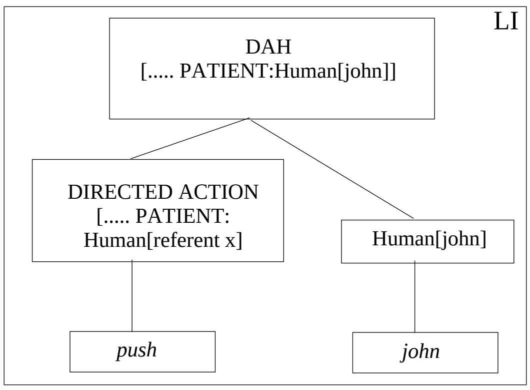

Figure 5.11: A two-word linguistic interpretation for *push john*.

Consider analyzing input with a pushing situation in which John pushes Anna together with the utterance *push anna*. Assume furthermore that this specific utterance has never been processed before. First the meaning of *push* and *anna* are retrieved from the corpus. Since their semantic categories are **DIRECTED ACTION** and **Human** respectively, they are recognized as a DAH structure (figure 5.11). This structure leads to a meaning which can be compatible with the situational interpretation. If it is compatible, then everything is fine and the linguistic interpretation is added to the corpus, which strengthens the statistics of the linguistic structure and the parts involved. It is also possible that the situational interpretation is not compatible with the assigned interpretation. In that case the introduction of a new word meaning might be triggered. This involves HYPOTHESIS.

Consider analyzing *push anna* when the situation does not explicitly contain Anna being pushed. The same interpretation is achieved as described above. I assume that the observer will attempt to embed this interpretation into its past or future experiences. In other words, Anna might have been pushed in the past, she might become pushed soon or she might be pushed on another location. This attempt is likely to involve recalling what was said before or extending the current conversation: the first signs of a more complex discourse structure.

Summarizing ANALYSIS leads to:

ANALYSIS
Given a current set of analyses A and an input containing known words, find the best fitting interpretation based on A and add this to A.

Hypothesis

Typical for HYPOTHESIS in the two-word stage is the following. Assume that *eat* does not belong to the category **DIRECTED ACTION** and has thus been categorized differently than *push*. Assume furthermore that the meanings of *eat* and *sandwich* are known, but not in combination with eachother. Consider analyzing a situation in which John eats a sandwich together with the utterance *eat sandwich*. ANALYSIS would produce a linguistic interpretation containing both meanings on their own, unified with the situational interpretation. HYPOTHESIS, on the other hand, can propose a combination based on the fact that both meanings contain the same referent, based on the changes REORGANIZATION has brought about on statistical grounds. If this hypothesis is not useful processing input in the future it will disappear from the corpus.

HYPOTHESIS
Assist ANALYSIS in assigning interpretations to new auditory input on the present level of abstraction (two-word structures) instead of the previous level, based on knowledge of the changes that REORGANIZATION has accomplished in the corpus.

5.5.3 Remarks

I stated that the situational representation became redundant in the examples above. This applies to proper names like *john*, because there is only one referent. In the case of common nouns (e.g. *ball*) the situation does not become redundant, because ball is a predicate ("this is a ball"). Either the situational representation or the discourse up to the point where *ball* is heard, should introduce the referent.

As I mentioned earlier, unfamiliar words (including function words) are not represented in the analyses. This does not necessarily mean, though, that the words in the linguistic representations are experienced next to each other. Just like [Chang et al. 2001] assume, there might be a correlation between two words with a word (or noise) in between them. In later work Chang also uses the meets-operator in addition to the before-operator. The former is more strict than the latter, because it has the constraint that two words must be next to eachother ([Bergen et al. 2002]).

The unexplained data (roles and word orders) that [Chang et al. 2001] identify in their algorithm (section 4.3.2) plays a similar role as the data which is about to be interpreted on the previous level of abstraction in my description of HYPOTHESIS. What I have tried to prove is that before information can be labeled as "unexplained" it must be clear that explaining similar data is useful or fun. More variations in the input will reach the infant as it will meet more people and be part of more complex conversations, that will not always be about the "here and now". In order to make its wishes, remarks etc. clear, it will happily make use of the generalizations it has made and probably make a lucky guess once in a while.

The first "linguistically supported" two-word utterances will be built on these abstractions. Note that combinations can be made that are never literally encountered before. I consider these productions (just like interpretations of never encountered input) as the first signs of the creative character of language. Producing a complete new two-word utterance can be triggered by the actual events in the situation, the past, future plans or pure imagination. Obviously, but less extraordinary, this also applies to already processed two-word utterance.

I hope to have shown with the previous sections that in the two-word stage the linguistic component gets a firmer grip on the processing of language and on life in general. This grip will even become bigger in the multi-word stage.

5.6 Multi-word stage

The label "multi-word stage" implies that the infant does not arrive in a three-word stage, after the two-word stage. Instead, combinations of three or more words are being produced from now on. The reason for this is the discovery that the combination of a unary predicate with the appropriate argument can become a unary predicate again. In this section I will explain how this comes about by interpreting new input based on past language experiences. The linguistic component of DOP-framework, which I will call ANALYSIS, plays a major role in this process. REORGANIZATION and HYPOTHESIS, on the other hand, play minor roles explaning the combinatorial explosion characteristic for the multi-word stage. This is why details for these processes are left out of this section. I will pick up on their functioning again in chapter 6.

5.6.1 What is it?

In the multi-word stage many typical short "here and now" utterances occur:

I fall down	*lie me down*	*put on couch*	*you push it*
I get off	*push it down*	*put in mouth*	*I see birdie*
I get down	*lie me over*	*toys in room*	*I drop it*
you sit down	*push it on*	*go under there*	*piggy eating beef*

Also some longer utterances are used by Naomi almost right from the start in the multi-word stage:

daddy will be home tonight
take it out mouth
I want get down
lay down for a nap
is he going to walk

These examples show that some combinatorial, and thus creative, power of the language has become clear to Naomi. They do not necessarily show, however, that Naomi has achieved the similar abstraction level of an adult language user. For example, there might very well be no evidence yet for Naomi that e.g. *sleep* has something in common linguistically with *push*.

5.6.2 Analysis

ANALYSIS has been shown in the previous sections to be able to interpret visual and auditory input on different levels of abstraction depending on the corpus. Here, I will concentrate on the linguistic interpretation of auditory input and thus the linguistic component of language processing, especially on the sudden increase of combinatorial possibilities.

Assume that the following two linguistic interpretations (figure 5.13 and 5.12) are present in the corpus and that *john push ball* is about to be processed, without an informative situational interpretation explicitly being present. Note that the notation is abbreviated compared to earlier sections.

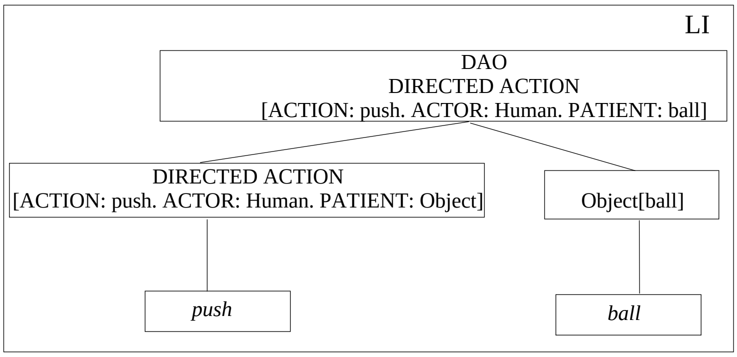

Figure 5.12: Linguistic interpretation *push ball* from the two-word stage.

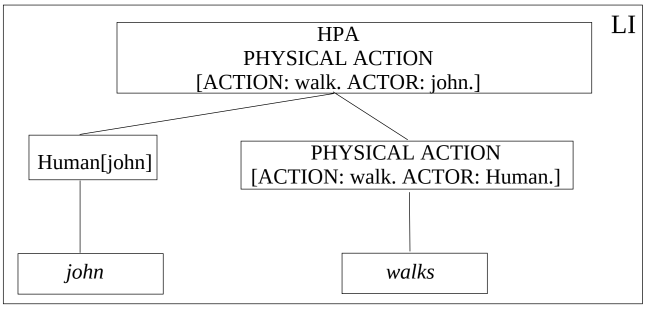

Figure 5.13: Linguistic interpretation *john walks* from the two-word stage.

The meaning of *walk* is known to be able to incorporate the meaning of a **Human**-word, because of the constraints on the ACTOR-slot, with the same constraint. The combination *push ball* also has a ACTOR-slot available. This similarity, achieved by a generalization over linguistic interpretations, makes the units *walk* and *push ball* predicates of the same type. The meaning of *john* is combined with the meaning of *push ball* analogous to the manner in which **john** is combined with **PHYSICAL ACTION**[ACTION:**walk**.ACTOR:**Human**.]. This leads to the following interpretation, which is added to the corpus.

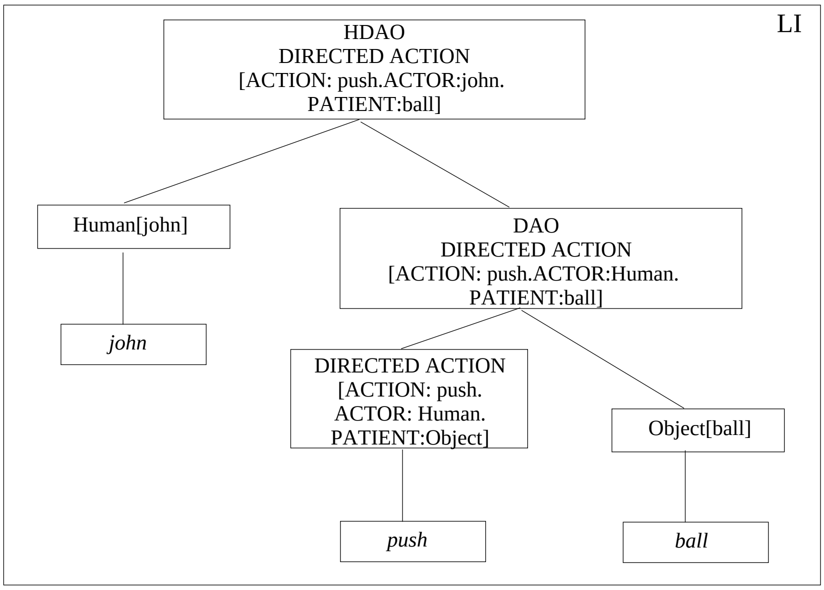

Figure 5.14: Linguistic interpretation of *john push ball.*

It might be questioned what the psychological plausibility is of the binary tree in figure 5.14 compared to a ternary tree with all lexical mappings on the same level. As I said earlier the frequency of the DAO analyses in the corpus is high. This is evidence for the fact that these chunks represent a "stand alone" perception with a certain autonomy. This autonomy is preserved connecting **Human** one level higher in the representation, leading to a binary tree instead of a ternary tree. The **Human** part of the tree covers the topic (given information) the DAO-part covers the focus (new information).

In a similar way the emergence of a linguistic representation for *push me down* comes about. According to the two-word stage the infant has linguistic representations available for *push me* (DAH), *me down* (HP) and *push down* (DAP). In principle they might all play a role in the analyses of this sentences[11]. Assume that *walk down* is available in the corpus as a result of the two-word stage. Accepting the previous example one has accepted that the meanings of *walk* and *push ball* are predicates of the same type. Obviously, this also applies to the meanings of *walk* and *push me.* The meaning of *down* is therefore combined with the meaning of *push me* analogous to the way in which the meanings of *walk* and *down* are combined. This results in the interpretation shown in figure 5.15:

[11] Chang and Maia let before(you,ball) also play a role in the form pole with respect to *you push ball* (chapter 4).

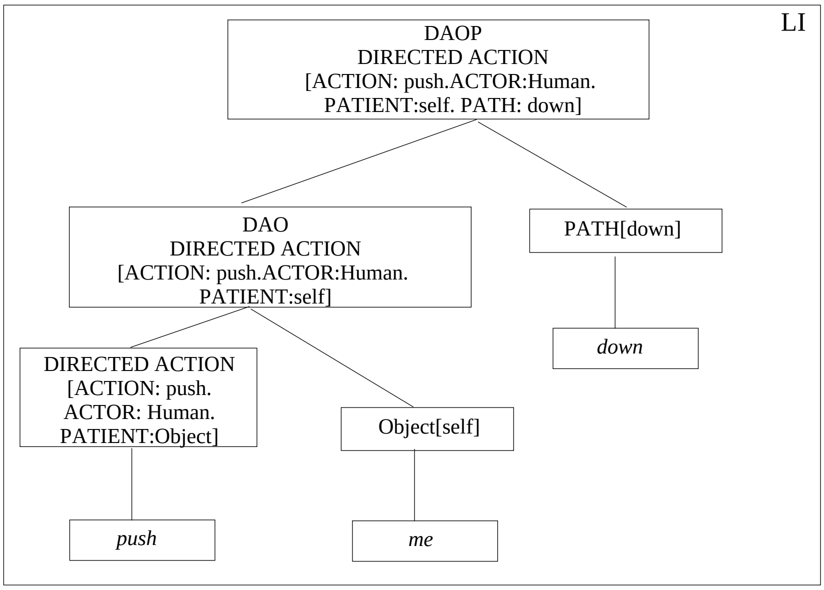

Figure 5.15: Linguistic interpretation of *push me down*.

Processing the sentence *you push me down* has also come within reach in a similar way as described in the previous example. In this case, though, it is hard to predict what the new representation would look like, because the linguistic representations for *push me down* (figure 5.15) and *john push ball* (figure 5.14) could have the similar frequency in the corpus. It might be that the scope of the focus (new information) keeps *push me down* together and that the topic (given information) *you* is connected.

As we have shown, the categories, to which different parts of old structures belong to, play a major role in analyzing new input. This process is steered by an ANALYSIS procedure similar to the one first introduced by [Scha 1990] and described in detail by [Bod 1995]. The role of the situational input decreases (see section 5.5 about the two-word stage), so the best analysis below is calculated purely based on linguistic data.

Analysis
Given a current set of analyses A containing a set of linguistic interpretation LI and an input consisting of a sequence of words, W_{in}, find the best fitting linguistic interpretation $LI_{mostprobable}$ and add it to the set of analyses A.

1. Associate the members of W_{in} with past linguistic interpretations LI and use them to cue a set of potential linguistic fragments LF involved in the interpretation of the sequence of words.

2. Determine the set of all potential linguistic interpretations $LI_{potential}$ by combining the fragments LF. Incorporate fragment statistics to calculate the probabilities of each linguistic interpretation.

3. Calculate the most probable interpretation $LI_{mostprobable}$.

The set of all possible linguistic interpretations, including their individual probabilities, in step 1 and 2 can, for example, be collected by using the parsing algorithms from the theory of Stochastic Tree Subsitution Grammars. This set can contain similar linguistic structures derived in a different way. The probabilities of similar structures are added to eachother in order to be able to calculate the most probable structure in step 3. The way described above the frequencies of "all lexical elements, syntactic structures and 'constructions' that the language user has ever encountered, and their frequency of occurrence, can have an influence on the processing of new input" ([Scha 1990]). For a proper and detailed introduction of the theory behind the DOP-framework I refer to [Bod 1995].

5.6.3 Remarks

I have shown a kind of chain reaction with the previous examples. Once the two-word stage is in place a DOP-mechanism is able to assign an interpretaion to never encountered and rather large sentences. The mechanism takes the categories of structures, lexical or larger, into consideration. Once the larger constructions are added to the corpus, more generalizations can be made over linguistic interpretations exposing similar characteristics. This way the corpus will gradually become populated by even larger and more complex structures.

Although I claimed that the best analysis can be calculated purely based on linguistic representations for the sake of the argument, I have to emphasize that this is not the case. With predicates like **ball** either the situational interpretation or the discourse structure has to introduce a referent.

The infant does not need statistical evidence anymore in order to expect correlations between what is heard and what is seen. More and more utterances will cover other "worlds" than the "here and now". The shown mechanisms make it possible to assign meanings to utterances without explicitly informative situational interpretations. The role of REORGANIZATION in the multi-word stage will be picked up in the next chapter.

Chapter 6

Consequences

The foregoing chapters lead to certain consequences with respect to thinking about and designing models for the phenomenon of language acquisition. I will wrap up the processes described in chapter 5 in section 6.1 towards the design of an overall mechanism. Section 6.2 will take the results of this wrap up and attempt to find answers for the questions formulated in chapter 5. In the last section (6.3) I will draw some conclusions, introduce discussion points and suggest further research.

6.1 Towards an overall mechanism

In this section I take the first steps making a generalization over the development cycles described in chapter 5 towards *one* overall mechanism for language acquisition. In a neural network-model the processes described in chapter 5 would all be gradual. Certain connections in the network will become stronger and others weaker depending on the input. [Chang et al. 2001] and [Allen et al. 1999] have these kind of models in mind mentioning structured connectionism and Hebbian learning. Explaining how statistical information and generalizations are stored in neural networks is not the aim of this work.

Certain is though, that representations achieved by every REORGANIZATION take the upper hand in representing new input. In the present framework this feature somehow needs to be present as well. This is not a trivial issue, but I will try to show how it can be mimicked. First, I will compress the information from chapter 5. Repeating the cycles described in the sections about the prelinguistic stage, the one-word, the two-word and the multi-word stage produces the following step-wise overview:

- Innate ANALYSIS
  Result: a corpus with situational interpretations consisting of situation-specific low level concepts (5.3.2).
- Cycle 1
  Result: a corpus with situational interpretations consisting of action-specific relational frames (section 5.3.2).
- Cycle 2
  Result: a corpus with situational interpretations consisting of abstract relational frames together with accompanying acoustic signals (section 5.3.2).
- Cycle 3
  Result: a corpus with linguistic interpretations containing acoustic signals together with abstractions over situational interpretations (the approximations of word meanings in section 5.4.2).
- Cycle 4
  Result: a corpus with linguistic interpretations consisting of two-word structures (section 5.5.2).
- Cycle 5
  Result: a corpus with linguistic interpretations consisting of multi-word structures (section 5.6.2).
- Cycle 6
  Result: ......

Figure 6.1: Stepwise overview of the processes towards the multi-word stage.

Although a gradual REORGANIZATION, that is active after each input, could be designed within the present framework, I choose not to. What I am after is the emergence of the first linguistic structures in the corpus. So, I propose a once-in-a-while REORGANIZATION engine going through the corpus which is responsible for reaching the "next" abstraction level, without worrying too much about the psychologically more plausible gradualness.

Once one has such a batch-like REORGANIZATION available, it comes within reach to design a mechanism active at any time, even triggered by a (certain kind of) input. In between two REORGANIZATION runs, the new input is represented according to the present abstractions. So, with this and figure 6.1 in mind the "top" (designing top down) algorithm covering all stages would look like figure 6.2 assuming that the first corpus is filled by the innate ANALYSIS.

1. Make a REORGANIZATION run through the corpus.
2. For a set of input I let ANALYSIS interpret every element i based on the current corpus, if necessary, assisted by HYPOTHESIS.
3. Return to step 1.

Figure 6.2: The top algorithm.

In the next two subsections I will go into some details behind step 1 and step 2 from figure 6.2.

6.1.1 Step 1: Reorganization

With respect to REORGANIZATION I imagine a process which takes a frequent phenomenon and searches through the analyses in the corpus for a correlating phenomenon. The representation **DIRECTED ACTION**[ACTION:**push**. ACTOR:**Human**. PATIENT:**Human**.] and the **DIRECTED ACTION-Object** order are examples of frequent phenomena; the word *push* and [PATIENT:**Object**] (see sections 5.4 and 5.5), respectively, as their co-occurring phenomena.

The reorganizations take place within the domain of situational interpretations, within the domain of linguistic interpretations and within the combination of the two (see chapter 4 and 5). Although I claim that REORGANIZATION would also apply to the domain of situational interpretations, I would propose to start in cycle 2 in figure 6.1 when implementing a simulation model. In that stage the corpus is filled with abstract relational representations together with sequences of words. This way one does not have to worry about "the obscure lower-level features of the underlying concepts" ([Chang et al. 2001]) and concentrate on the emergence of the first linguistic structures. Furthermore, this is also the initial state that Bod and Scha propose as an interesting starting point in [Bod et al 1996].

Making sure that new input can be represented by ANALYSIS with the abstractions made, I assume that the data, based on which REORGANIZATION achieves an abstraction, are reprocessed to incorporate the abstraction in question. It has to be kept in mind that interpreting new input is based on a mechanism that attempts to find analogies between input and the corpus of past experiences (see chapter 3) which are represented according to the abstractions made. So, these abstractions are not some abstract entities "living somewhere" on their own, like the rewrite rules of a competence grammar are often expected to do. With each REORGANIZATION a set of analyses from the corpus is involved. A logical step would be to reprocess this set and represent every member with the generalization that the set caused.

Note that this does not mean that the old data is disregarded completely. Some reorganizations might lead to abstractions that are not used in future interpretations. Having the old data still available makes a new organization attempt possible. On the other hand if abstractions happen to be useful for future interpretations, then the data they were built on might slowly disappear to the background. We obviously do not experience an orange consciously according to its low level features. Instead, we experience it as an **orange**, as a whole. Another example is the fact that adults with Chinese as their mother tongue, are not able to distinguish an "r" form an "l" in language use. This "malfunction" is not innate but acquired. The past experiences which underlie the ability to experience an "r" have disappeared from the corpus.

Step 1 from figure 6.2 can be divided in the following sub-steps:

Step 1

1. Find all analyses containing a frequent phenomenon.
2. Search for co-ocurring phenomena within these analyses.
3. Integrate these co-ocurring phenomena.
4. Reprocess the involved analyses.

Figure 6.3: Step 1: reorganization.

Some examples of co-occurring regularities:

x y	x a
x y	x b
x y	x c

The phenomena on the left lead to different generalizations than the phenomena on the right. That is, on the left x and y are likely to be clustered, whereas on the right a, b and c. These examples show that there might be many different kinds of co-occurring regularities to be discovered by REORGANIZATION.

I think that the set of input I can be kept stable during a number of cycles through step 1 and step 2 of figure 6.2. The corpus can be checked in between the cycles for the emergence of structures.

6.1.2 Step 2: Analysis and hypothesis

Scha considers Universal Grammar as a set of analogically associating mechanisms (section 5.1). A part of that set has been accounted for by the invention of the performance model for adult language use, described by the present DOP-framework. Step 2 in figure 6.2 is accounting for this process, not only for a corpus with complex linguistic data, but also for a corpus only with situational experiences. The algorithm for this process will have many similar features as the algorithm designed for the present DOP-framework. It will be able to find analogies between input (situational or linguistic) and the corpus of past experiences (situational or linguistic), although the emphasis is on the latter in the present work.

Starting at cycle 2 in figure 6.1, the ANALYSIS step 2 from figure 6.2 will only involve representing familiar linguistic forms. In the beginning this will come down to finding analogies between single words and analyses in the corpus only containing lexical mappings. Later, once REORGANIZATION has led to more complex linguistic interpretations, step 2 will also involve finding analogies between these representations and sequences of words.

The most important feature of HYPOTHESIS is that it is able to identify the difference between a new level of abstraction and a previous level of abstraction. The changes that REORGANIZE evokes in the corpus are therefore closely monitored. Futhermore, it is able to project this knowledge onto interpretations that ANALYSIS is about to make on a previous level of abstraction. In this sense HYPOTHESIS can be seen as a kind of datamining procedure. If it is designed properly it is able to operate on any kind of data, as long as it has access to "the state of affairs before" and "the state of affairs after".

6.2 Answers

In chapter 5.2 I have split up the question the designers of the DOP-framework still have with respect to language acquisition (chapter 3). In this section answers will be given.

6.2.1 New strategies

The strategies not covered by the linguistic component of the DOP-framework are REORGANIZATION and HYPOTHESIS. The reorganizations shown so far take place within three domains: that of situational interpretations, that of situational interpretations together with linguistic interpretations and that of linguistic interpretations.

Important to pay attention to is that these strategies cannot be said to be characteristic only for a non-linguistic component of the DOP-framework. The reason is obvious: not only situational interpretations are affected by these strategies, but linguistic interpretations and their combinations are also affected by them.

This does not come as a surprise if one acknowledges the phenomenon of language change (chapter 3): "New syntactic and semantic categories, new constructions and new uses of grammatical conventions are continuously created (...)" ([Steels 1999]). This means there has to be a mechanism available that accounts for these phenomena in the prelinguistic stage as well as in the linguistic stages. Or, in other words, the mechanisms REORGANIZATION and HYPOTHESIS, taking care of the emergence of new meanings and of projecting meanings onto utterances, stay active during a whole life time.

The mentioned REORGANIZATION probably goes through a certain development, learning from what it has encountered in the past, towards more HYPOTHESIS-like behavior. For an infant just starting to acquire a language, statistical evidence leads to the emergence of a certain linguistic phenomenon. This is taken care of by REORGANIZATION. For an adult the statistical evidence might be less important. An adult knows about the existence of word meanings, grammatical constructions etc. so new phenomenona do not have to be encountered many times in order to converge to the clarity of their use. An adult will be much faster discovering the use of new phenomena, because the hypothesis space is much smaller, building on experience. Obviously, HYPOTHESIS plays the main role with respect to these processes.

6.2.2 Continuity

I have claimed that ANALYSIS is able to find analogies within two domains: (1) between situational interpretations in the corpus and visual input and (2) between linguistic interpretations and acoustic input. I would like to define the DOP-framework in such a way that the claim about the continuity of language proccessing mentioned by [Seidenberg et al. 1999] (chapter 2), [Scha 1992] (chapter 3) and [Chang et al. 2001] (chapter 4) becomes clearer. If the notions of a linguistic and non-linguistic component are abandoned, then the DOP-framework can be defined as:

1. A system which is able to make abstractions within its corpus.
2. A system which is able to interpret input by calculating the optimal analogy between the input and the corpus.
3. A system which is able to make hypotheses if an interpretation is based on the previous level instead of the current level of abstraction.

Whether REORGANIZATION takes place once in a while or while processing every input does not change the picture that I would like to put forward. The DOP-framework defined as above will be active continuously from birth until death, accounting for cognition, including language acquisition.

6.2.3 Where does syntax start (or end)?

The level of abstraction achieved in the beginning of the multi-word stage leads to the emergence of the first basic virtual grammar. But obviously generalizing does not stop at this point. More (conceptually complex) words are acquired, function words come in, recursion emerges etc. Summarizing, language is assumed to contain all kinds of regularities that are waiting to be discovered. I will give one important example of what generalizing might lead to after the emergence of the early multi-word stage.

Cartwright and Brent ([Cartwright et al. 1997]) describe the possible role of distributional analysis in early language acquisition. They "(...) define the distribution of a word as the sum of all its environments and define the environment of a word as its position relative to other words in all utterances in which it occurred. To simplify the description of environments, classes of words could stand in place of individual words (...)" (p. 123). For technical details of their simulations I refer to their work.

Important for the present context is their remark that "(...) it seems to us that a relatively limited form of syntax is required to assist category acquisition, but this is an area for future work (...)" (p. 162). This implies that the proposal worked out in this paper might connect nicely to the work of Cartwright and Brent, because our model can reach a limited form of syntax. Furthermore, categories like DA and H (chapter 5) lead to generalized minimal pairs making the occurrence for true minimal pairs[1] unnecessary. This is convenient because "ver few [true] minimal pairs occur in natural speech (...)" (p. 130).

The simulations Cartwright and Brent run lead to discrete categories that show reasonable resemblance with linguistic categories. They note that "since the internal representation of the grammar and the lexicon are not well understood yet, we can not be certain that discrete categories are the correct form of output (...)" (p. 162). I agree with the assumption that words will be clustered according to their distribution. I furthermore assume that there will be typical examples (e.g. *chair* as a noun) within a certain cluster and that there will be examples at the "outskirts" (e.g. *war* as a noun) of a cluster, being less typical.

I would like to claim that these distributional cues (among others) could make a child realize that something might be right about the sentence *green dreams sleep furiously.* At the same time I would like to claim that an adult, having acquired writing and reading at primary school and highschool, can actually explain why he or she considers this sentence grammatical. This would make school an important environmental factor with respect to the acquisition of competence-like behaviour.

Note that the interpretation process of the given sentence does not converge towards a stable meaning very easily, or maybe not at all. This might have to do with *dream* not being a very typical noun in a typical noun environment. This might lead to what Scha calls an esthetic experience:

> "Saussurean codes then emerge when the interpretive process converges quickly and yields particularly unequivocal results. And esthetic experiences, on the other hand, occur when the process is complex, but in its complexity sufficiently coherent to reach consciousness and to be judged as meaningful. Often, though not necessarily, the interpretive process has a divergent character in this case." ([Scha 1992])

What I have been trying to show tentatively in this section is that some features of what is claimed to be syntax emerge by fargoing generalizations. Where the emergence starts is reasonably clear: once a first virtual grammar has emerged (in the two-word stage). Where it ends, is not so clear: according to my opinion

[1] True minimal pair:
I saw the cat
I saw the dog
Generalized minimal pair:
DA H *cat*
DA H *dog.*

not around the age of five, when a child is usually assumed to have mastered the grammar of its mother tongue.

6.3 Conclusions

Some aspects of implementing a simulation-framework for the proposed theory will be treated in the next section. The big picture of this work is formulated in the final section of this thesis.

6.3.1 Aspects of simulation

Implementing a simulation environment for the proposed model requires at least the following steps:

- **Defining a representation framework that covers the domain of prelinguistic structures**.
  Challenges arise deciding on the limits of prelinguistic structures. Like [Maia et al. 2001] I assume that the major natural categories are present, just like a large set of subcategorizations (**cup**, **toast**, etc.) is in place prelinguistically. But also some sense of more pragmatic aspects, e.g. **topic** and **focus**, must be available in the representation framework.

- **Defining a set of child directed utterances**.
  A subset of the utterances directed to only one child would have my preference, because this way regularities in the input (e.g. repetitions) are preserved. Some variability can be introduced here, e.g. with respect to the length of the utterances.

- **Assigning a situational representation to each utterance**.
  The subset of child directed utterances should be small enough to be able to annotate it with a situational representation, but big enough to be representative. The degree to which the "right" meaning is assigned to each utterance can be varied in order to check the robustness of the system.

- **Designing the analysis algorithm.**
  I mentioned earlier that the algorithms used in the present DOP-implementations will be re-usable to a large extent to designing an analysis procedure. The hypothesis algorithm needs more attention.

- **Designing the reorganization algorithm.**
  This step is the most complex one. Further research is needed to find psychologically plausible, but also technically feasable algorithms. The algorithm should at least be able to find the co-occurrences described in this work, but possibly also some distributional analysis mechanism [Cartwright et al. 1997] could be incorporated.

- **Designing the hypothesis algorithm.**
  As mentioned this algorithm monitors the abstractions made. It could even be considered as making abstractions over the abstractions that have

been made in the previous step. Examples of algorithms applied to cognitive processes like these can be found in the literature about dynamical systems (e.g. [Serra et al. 1990] and [Simon et al. 1995]).

- **Running experiments**
  Certain variable parameters should be defined in order to be able to run different experiments with different, and hopefully explainable, results.

6.3.2 Nativism *and* empiricism

The need for linguistic universals to explain the phenomena of language acquisition is absent in this work. This does not mean, though, that "it all" can be explained by empiricism. Just like [Chang et al. 2001] and [Scha 1990], I assume that certain mechanisms are innate. An example:

> (...) it is probable that the brain circuitry to recognize objects is at least partly innate; on the other hand it is hard to imagine that particular object categories (e.g. **Telephone**) could be. What seems reasonable to assume that there is specialized circuitry for object recognition, and that there is a (possibly specialized) mechanism to categorize objects by interaction with the environment; the particular categories formed will depend on such interaction." ([Maia et al. 2001])

The major prelinguistic conceptual structures lead to the first "syntactical" categories, being part of the infant's individual invention of language. REORGANIZATION plays a major role right through all cognitive processes, and must be innate in some form, just like ANALYSIS and HYPOTHESIS. The first corpus they start to operate on comes about by the innate skill to interpret low level features like taste, form, color, temporal aspects, spatial aspects etc. If the proposed model in this thesis can be proven to be on the right track, then many questions about language acquisition really are questions about cognition, much of which is developed prelinguistically.

Bibliography

[Allen et al. 1999] Joseph Allan and Mark S. Seidenberg. The emergence of grammaticality in connectionist networks. In B. MacWhinney (Ed.) *The emergence of language.*. Mahwah, NJ: Erlbaum.

[Bartsch:1998] Renate Bartsch. *The Formal Relationship between Dynamic Conceptual Semantics and Connectionist Neural Network Modelling.* ILLC Publication Series, LP-1998-06, University of Amsterdam.

[Bergen et al. 2002] B.K. Bergen and N.C. Chang. Embodied Construction Grammar in Simulation-Based Language Understanding. In J.O. Ostman and M. Fried (eds.), *Construction Grammar(s): Cognitive and Cross-Language Dimensions.* Johns Benjamins. Forthcoming, 2002. DRAFT.

[Bod et al 1996] Rens Bod, Remko Scha. *Data Oriented language processing: an overview.* Thechnical report LP-69-13, Institute for Logic Language and Compuation, University of Amsterdam.

[Bod 1995] Rens Bod. *Enriching linguistics with statistics: Performance models of natural language.* Ph. D. thesis. ILLC Dissertation Series 1995-14, University of Amsterdam.

[Bod 1992] Rens Bod. A computational Model of language performance: Data Oriented Parsing. Proceedings Coling 1992, Nantes.

[Bonnema et al. 1997] Remko Bonnema, Rens Bod and Remko Scha. A DOP model for semantic interpretation. Proceedings 35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics (July 7-12, 1997, Madrid, Spain), pp. 159-167.

[Cartwright et al. 1997] T. Cartwright and M. Brent. Syntactic categorization in early language acquisition: formalizing the role of distributional analysis. *Cognition*, vol 63, pp. 121-170.

[Chang et al. 2001] Nancy C. Chang and Tiago V. Maia. Grounded learning of grammatical constructions. In *2001 AAAI Spring symposium on learning grounded representations.*

[CHILDES] http://childes.psy.cmu.edu/ .

[Chomsky 1965] Noam Chomsky. *Aspects of the theory of syntax.* Cambridge MA: MIT Press.

[Hofstadter 1995] Douglas R. Hofstadter and the Fluid Analogies Research Group. *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought.* New York: Basic Books, 518pp.

[Hurford 2002] James R. Hurford. The neural basis of predicate-argument structure. To be published in *Behavioral and Brain Science* (in press). Cambridge University Press.

[Maia et al. 2001] Tiago V. Maia and Nancy C. Chang. Grounding the acquisition of grammar in sensiromotor representations. In *2001 AAAI Spring symposium on learning grounded representations.*

[Pinker 1995] Steven Pinker. Language acquisition. In L. R. Gleitman and M. Liberman (Eds.), *An Invitation to Cognitive Science, 2nd edition: Language*, (pp. 135-182). Cambridge, MA: MIT Press.

[Scha 1992] Remko Scha. Virtuele Grammatica's en Creatieve Algoritmes. Gramma/TTT 1, 1 (1992), pp. 57-77. Originally published in Dutch. Translated into English as: Virtual Grammars and Creative Algorithms.

[Scha 1990] Remko Scha. Taaltheorie en taaltechnologie; competence en performance. In R. de Kort and G.L.J. Leerdam (Eds.) *Computertoepassingen in de Neerlandistiek*. Almere: LVVN. Originally published in Dutch. Translated into English as: Language theory and language technology; comptetence and performance.

[Seidenberg et al. 1999] Mark S. Seidenberg and Maryellen C. MacDonald. A probabilistic constraints approach to language acquisition and processing. *Cognitive Science*, vol 23, pp. 569-588.

[Serra et al. 1990] Roberto Serra and Gianni Zanarini. *Complex systems and coginitive processes.* Springer Verlag, 1990.

[Simon et al. 1995] Tony J. Simon and Graeme S. Halford (Eds). *Developing Cognitive Competence: New Approaches to Process Modeling.* Hillsdale: NJ. Lawrence Erlbaum Associates, 1995.

[Siskind 1994] J. Siskind. *A computational study of cross-situational techniques for learning word-to-meaning mappings. Cognition*, 61, pp. 39-92, 1994.

[Siskind 2001] Jeffrey Mark Siskind. Grounding lexical semantics of verbs in visual perception using force dynamics and event logic. *Journal of Artificial Intelligence Research*, vol 15, pp. 31-90.

[Smolensky 1996] Paul Smolensky. On the comprehension/production dilemma in child language. *Linguistic Inquiry*, vol 23, pp 720-731.

[Steels 1999] L. Steels (VUB Artificial Intelligence Laboratory). *The talking heads experiment - Volume 1 Words an meanings.* Best of publishing, Brussels, 1999.