Effects of Meaning-Preserving Corrections on Language Learning

Dana Angluin *
Department of Computer Science
Yale University, USA
dana.angluin@yale.edu

Leonor Becerra-Bonache
Laboratoire Hubert Curien
Université de Saint-Etienne, France
leonor.becerra@univ-st-etienne.fr

Abstract

We present a computational model of language learning via a sequence of interactions between a teacher and a learner. Experiments learning limited sublanguages of 10 natural languages show that the learner achieves a high level of performance after a reasonable number of interactions, the teacher can produce meaning-preserving corrections of the learner's utterances, and the learner can detect them. The learner does not treat corrections specially; nonetheless in several cases, significantly fewer interactions are needed by a learner interacting with a correcting teacher than with a non-correcting teacher.

1 Introduction

A child learning his or her native language typically does so while interacting with other people who are using the language to communicate in shared situations. The correspondence between situations and utterances seems likely to be a very important source of information for the language learner. Once a child begins to produce his or her own utterances, other people's responses to them (or lack thereof) are another source of information about the language. When the child's utterances fall short of adult-level competence, sometimes the other person in the conversation will repeat the child's utterance in a more correct form. A number of studies have focused on the phenomenon of such corrections and questions of their frequency in child-directed speech and whether children can and do make use of them; some of these studies are discussed in the next section.

*Research supported by the National Science Foundation, Grant CCF-0916389.

In this paper we construct a computational model with a learner and a teacher who interact in a sequence of shared situations. In each situation the teacher and learner interact as follows. First the learner uses what it has learned about the language to (attempt to) generate an utterance appropriate to the situation. The teacher then analyzes the correctness of the learner's utterance and either generates an utterance intended as a correction of the learner's utterance, or generates another utterance of its own appropriate to the situation. Finally, the learner uses information given by its own utterance, the teacher's utterance and the situation to update its knowledge of the language. At the conclusion of this interaction, a new interaction is begun with the next situation in the sequence.

Both the learner and the teacher engage in comprehension and production of utterances which are intended to be appropriate to their shared situation. This setting allows us to study several questions: whether the teacher can offer meaningful corrections to the learner, whether the learner can detect intended corrections by the teacher, and whether the presence of corrections by the teacher has an effect on language acquisition by the learner. For our model, the answer to each of these questions is yes, and while the model is in many respects artificial and simplified, we believe it sheds new light on these issues. Additional details are available (Angluin and Becerra-Bonache, 2010).

Proceedings of the Fifteenth Conference on Computational Natural Language Learning, pages 97–105,
Portland, Oregon, USA, 23–24 June 2011.

2 Meaning-preserving corrections

Formal models of language acquisition have mainly focused on learning from positive data, that is, utterances that are grammatically correct. But a question that remains open is: Do children receive negative data and can they make use of it?

Chomsky's poverty of stimulus argument has been used to support the idea of human innate linguistic capacity. It is claimed that there are principles of grammar that cannot be learned from positive data only, and negative evidence is not available to children. Hence, since children do not have enough evidence to induce the grammar of their native language, the additional knowledge language learners need is provided by some form of innate linguistic capacity.

E. M. Gold's negative results in the framework of formal language learning have also been used to support the innateness of language. Gold proved that superfinite classes of languages are not learnable from positive data only, which implies than none of the language classes defined by Chomsky to model natural language is learnable from positive data only (Gold, 1967).

Brown and Hanlon (Brown and Hanlon, 1970) studied negative evidence understood as explicit approvals or disapprovals of a child's utterance (e.g.,"That's right" or "That's wrong.") They showed that there is no dependence between these kinds of answers and the grammaticality of children's utterances. These results were taken as showing that children do not receive negative data. But do these results really show this? It seems evident that parents rarely address their children in that way. During the first stages of language acquisition children make a lot of errors, and parents are not constantly telling them that their sentences are wrong; rather the important thing is that they can communicate with each other. However, it is worth studying whether other sources of negative evidence are provided to children. Is this the only form of negative data? Do adults correct children in a different way?

Some researchers have studied other kinds of negative data based on reply-types (e.g., Hirsh-Pasek et al. (Hirsh-Pasek et al., 1984), Demetras et al. (Demetras et al., 1986) and Morgan and Travis (Morgan and Travis, 1989).) These studies argue that parents provide negative evidence to their children by using different types of reply to grammatical versus ungrammatical sentences. Marcus analyzed such studies and concluded that there is no evidence that this kind of feedback (he called it noisy feedback) is required for language learning, or even that it exists (Marcus, 1993). He argued for the weakness, inconsistency and inherently artificial nature of this kind of feedback. Moreover, he suggested that even if such feedback exists, a child would learn which forms are erroneous only after complex statistical comparisons. Therefore, he concluded that internal mechanisms are necessary to explain how children recover from errors in language acquisition.

Since the publication of the work of Marcus, the consensus seemed to be that children do not have access to negative data. However, a study carried out by Chouinard and Clark shows that this conclusion may be wrong (Chouinard and Clark, 2003). First, they point out that the reply-type approach does not consider whether the reply itself also contains corrective information, and consequently, replies that are corrective are erroneously grouped with those that are not. Moreover, if we consider only reply-types, they may not help to identify the error made. Hence, Chouinard and Clark propose another view of negative evidence that builds on Clark's principle of contrast (Clark, 1987; Clark, 1993). Parents often check up on a child's erroneous utterances, to make sure they have understood them. They do this by reformulating what they think the child intended to express. Hence, the child's utterance and the adult's reformulation have the same meaning, but different forms. Because children attend to contrasts in form, any change in form that does not mark a different meaning will signal to children that they may have produced an utterance that is not acceptable in the target language. In this way, reformulations identify the locus of any error, and hence the existence of an error. Chouinard and Clark analyze longitudinal data from five children between two and four years old, and show that adults reformulate erroneous child utterances often enough to help learning. Moreover, these results show that children not only detect differences between their own utterance and the adult reformulation, but that they make use of that information.

In this paper we explore this new view of negative data proposed by Chouinard and Clark. Corrections (in form of reformulations) have a semantic component that has not been taken into account in previous studies. Hence, we propose a new computational model of language learning that gives an account of *meaning-preserving corrections*, and in which we can address questions such as: What are the effects of corrections on learning syntax? Can corrections facilitate the language learning process?

3 The Model

We describe the components of our model, and give examples drawn from the primary domain we have used to guide the development of the model.

3.1 Situation, meanings and utterances.

A **situation** is composed of some objects and some of their properties and relations, which pick out some aspects of the world of joint interest to the teacher and learner. A situation is represented as a set of ground atoms over some constants (denoting objects) and predicates (giving properties of the objects and relations between them.) For example, a situation s_1 consisting of a big purple circle to the left of a big red star is represented by the following set of ground atoms: $s_1 = \{\mathit{bi1}(t_1), \mathit{pu1}(t_1), \mathit{ci1}(t_1), \mathit{le2}(t_1, t_2), \mathit{bi1}(t_2), \mathit{re1}(t_2), \mathit{st1}(t_2)\}$.

Formally, we have a finite set P of **predicate symbols**, each of a specific arity. We also have a set of **constant symbols** $t_1, t_2, \ldots$, which are used to represent distinct objects. A **ground atom** is an expression formed by applying a predicate symbol to the correct number of constant symbols as arguments.

We also have a set of of **variables** $x_1, x_2, \ldots$. A **variable atom** is an expression formed by applying a predicate symbol to the correct number of variables as arguments. A **meaning** is a finite sequence of variable atoms. Note that the atoms do not contain constants, and the order in which they appear is significant. A meaning is **supported** in a situation if there exists a **support witness**, that is, a mapping of its variables to *distinct* objects in the situation such that the image under the mapping of each atom in the meaning appears in the situation. If a meaning is supported in a situation by a unique support witness then it is **denoting** in the situation. We assume that both the teacher and learner can determine whether a meaning is denoting in a situation.

We also have a finite alphabet W of words. An **utterance** is a finite sequence of words. The **target language** is the set of utterances the teacher may produce in some situation; in our examples, this includes utterances like *the star* or *the star to the right of the purple circle* but not *star of circle small the green*. We assume each utterance in the target language is assigned a unique meaning. An utterance is **denoting** in a situation if the meaning assigned to utterance is denoting in the situation. Intuitively, an utterance is denoting if it uniquely picks out the objects it refers to in a situation.

In our model the goal of the learner is to be able to produce every denoting utterance in any given situation. Our model is probabilistic, and what we require is that the probability of learner errors be reduced to very low levels.

3.2 The target language and meaning transducers.

We represent the linguistic competence of the teacher by a finite state transducer that both recognizes the utterances in the target language and translates each correct utterance to its meaning. Let A denote the set of all variable atoms over P. We define a **meaning transducer** M with input symbols W and output symbols A as follows. M has a finite set Q of states, an initial state $q_0 \in Q$, a finite set $F \subseteq Q$ of final states, a deterministic transition function δ mapping $Q \times W$ to Q, and an output function γ mapping $Q \times W$ to $A \cup \{\varepsilon\}$, where ε denotes the empty sequence.

The transition function δ is extended in the usual way to $\delta(q, u)$. The **language** of M, denoted $L(M)$ is the set of all utterances $u \in W^*$ such that $\delta(q_0, u) \in F$. For each utterance u, we define $M(u)$ to be the **meaning** of u, that is, the finite sequence of non-empty outputs produced by M in processing u. Fig. 1 shows a meaning transducer M_1 for a limited sublanguage of Spanish. M_1 assigns the utterance *el triangulo rojo* the meaning $(\mathit{tr1}(x_1), \mathit{re1}(x_1))$.

3.3 The learning task.

Initially the teacher and learner know the predicates P and are able to determine whether a meaning is

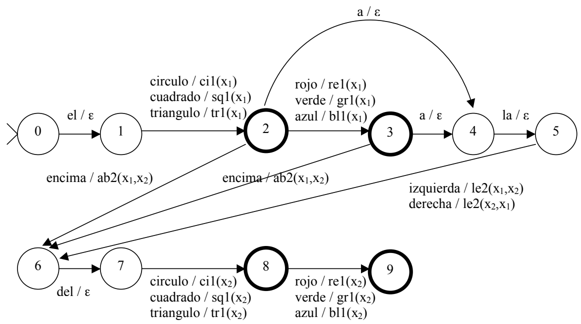

Figure 1: Meaning transducer M_1.

denoting in a situation. The learner and teacher both also know a shared set of **categories** that classify a subset of the predicates into similarity groups. The categories facilitate generalization by the learner and analysis of incorrect learner utterances by the teacher. In our geometric shape domain the categories are shape, size, and color; there is no category for the positional relations. Initially the teacher also has the meaning transducer for the target language, but the learner has no language-specific knowledge.

4 The Interaction of Learner and Teacher

In one interaction of the learner and teacher, a new situation is generated and presented to both of them. The learner attempts to produce a denoting utterance for the situation, and the teacher analyzes the learner's utterance and decides whether to produce a correction of the learner's utterance or a new denoting utterance of its own. Finally, the learner uses the situation and the teacher's utterance to update its current grammar for the language.

In this section we describe the algorithms used by the learner and teacher to carry out the steps of this process.

4.1 Comprehension and the co-occurrence graph.

To process the teacher's utterance, the learner records the words in the utterance and the predicates in the situation in an undirected **co-occurrence graph**. Each node is a word or predicate symbol and there is an edge for each pair of nodes. Each node u has an occurrence count, $c(u)$, recording the number of utterances or situations it has occurred in. Each edge (u, v) also has an occurrence count, $c(u, v)$, recording the number of utterance/situation pairs in which the endpoints of the edge have occurred together. From the co-occurrence graph the learner derives a directed graph with the same nodes, the **implication graph**, parameterized by a noise threshold θ (set at 0.95 in the experiments.) For each ordered pair of nodes u and v, the directed edge (u, v) is included in the implication graph if $c(u, v)/c(u) \geq \theta$. The learner then deletes edges from predicates to words and computes the **transitively reduced implication graph**.

The learner uses the transitively reduced implication graph to try to find the meaning of the teacher's utterance by translating the words of the utterance into a set of sequences of predicates, and determining if there is a unique denoting meaning corresponding to one of the predicate sequences. If so, the unique meaning is generalized into a **general form** by replacing each predicate by its category generalization. For example, if the learner detects the unique meaning $(\mathit{tr1}(x_1), \mathit{re1}(x_1))$, it is generalized to the general form $(\mathit{shape1}(x_1), \mathit{color1}(x_1))$. The learner's set of general forms is the basis for its production.

4.2 Production by the learner.

Each general form denotes the set of possible meanings obtained by substituting appropriate symbols from P for the category symbols. To produce a denoting utterance for a situation, the learner finds all the meanings generated by its general forms using predicates from the situation and tests each meaning to see if it is denoting, producing a set of possible denoting meanings. If the set is empty, the learner produces no utterance. Otherwise, it attempts to translate each denoting meaning into an utterance.

The learner selects one of these utterances with a probability depending on a number stored with the corresponding general form recording the last time a teacher utterance matched it. This ensures that repeatedly matched general forms are selected with asymptotically uniform probability, while general forms that are only matched a few times are selected with probability tending to zero.

4.3 From meaning to utterance.

The process the learner uses to produce an utterance from a denoting meaning is as follows. For a meaning that is a sequence of k atoms, there are two related sequences of positions: the **atom positions** $1, 2, \ldots, k$ and the **gap positions** $0, 1, \ldots, k$. The atom positions refer to the corresponding atoms, and gap position i refers to the position to the right of atom i, (where gap position 0 is to the left of atom a_1.) The learner generates a sequence of zero or more words for each position in left to right order: gap position 0, atom position 1, gap position 1, atom position 2, and so on, until gap position k. The resulting sequences of words are concatenated to form the final utterance.

The choice of what sequence of words to produce for each position is represented by a decision tree. For each variable atom the learner has encountered, there is a decision tree that determines what sequence of words to produce for that atom in the context of the whole meaning. For example, in a sublanguage of Spanish in which there are both masculine and feminine nouns for shapes, the atom $\mathit{re1}(x_1)$ has a decision tree that branches on the value of the shape predicate applied to x_1 to select either *rojo* or *roja* as appropriate. For the gap positions, there are decision trees indexed by the generalizations of all the variable atoms that have occurred; the variable atom at position i is generalized, and the corresponding decision tree is used to generate a sequence of words for gap position i. Gap position 0 does not follow any atom position and has a separate decision tree.

If there is no decision tree associated with a given atom or gap position in a meaning, the learner falls back on a "telegraphic speech" strategy. For a gap position with no decision tree, no words are produced. For an atom position whose atom has no associated decision tree, the learner searches the transitively reduced implication graph for words that approximately imply the predicate of the atom and chooses one of maximum observed frequency.

4.4 The teacher's response.

If the learner produces an utterance, the teacher analyzes it and then chooses its own utterance for the situation. The teacher may find the learner's utterance correct, incorrect but correctable, or incorrect and uncorrectable. If the learner's utterance is incorrect but correctable, the teacher chooses a possible correction for it. The teacher randomly decides whether or not to use the correction as its utterance according to the **correction probability**. If the teacher does not use the correction, then its own utterance is chosen uniformly at random from the denoting utterances for the situation.

If the learner's utterance is one of the correct denoting utterances for the situation, the teacher classifies it as **correct**. If the learner's utterance is not correct, the teacher "translates" the learner's utterance into a sequence of predicates by using the meaning transducer for the language. If the resulting sequence of predicates corresponds to a denoting meaning, the learner's utterance is classified as having an **error in form**. The correction is chosen by considering the denoting utterances with the same sequence of predicates as the learner's utterance, and choosing one that is "most similar" to the learner's utterance. For example, if the learner's utterance was *el elipse pequeno* and $(\mathit{el1}, \mathit{sm1})$ corresponds to a denoting utterance for the situation, the teacher chooses *la elipse pequena* as the correction. If the learner's utterance is neither correct nor an error in form, the teacher uses a measure of similarity between the learner's sequence of predicates and those of denoting utterances to determine whether there is a "close enough" match. If so, the teacher classifies the learner's utterance as having an **error in meaning** and chooses as the possible correction a denoting utterance whose predicate sequence is "most similar" to the learner's predicate sequence. If the learner produces an utterance and none of these cases apply, then the teacher classifies the learner's utterance as **uninterpretable** and does not offer a correction.

When the teacher has produced an utterance, the learner analyzes it and updates its grammar of the language as reflected in the co-occurrence graph, the general forms, and the decision trees for word choice. The decision trees are updated by computing an alignment between the teacher's utterance and the learner's understanding of the teacher's meaning, which assigns a subsequence of words from the utterance to each atom or gap position in the meaning. Each subsequence of words is then added to the data for the decision tree corresponding to the position of that subsequence.

If the learner has produced an utterance and finds that the teacher's utterance has the same meaning, but is expressed differently, then the learner classifies the teacher's utterance as a **correction**. In the current model, the learner reports this classification, but does not use it in any way.

5 Empirical Results

We have implemented and tested our learning and teaching procedures in order to explore questions about the roles of corrections in language learning. We have used a simplified version of the *Miniature Language Acquisition* task proposed by Feldman et al. (Feldman et al., 1990). Although this task is not as complex as those faced by children, it involves enough complexity to be compared to many real-word tasks.

The questions that we address in this section are the following. (1) Can the learner accomplish the learning task to a high level of correctness and coverage from a "reasonable" number of interactions (that is, well short of the number needed to memorize every legal situation/sentence pair)? (2) What are the effects of correction or non-correction by the teacher on the learner's accomplishment of the learning tasks?

5.1 The specific learning tasks.

Each situation has two objects, each with three attributes (shape, color and size), and one binary relation between the two objects (above or to the left of.) The attribute of shape has six possible values (circle, square, triangle, star, ellipse, and hexagon), that of color has six possible values (red, orange, yellow, green, blue, and purple), and that of size three possible values (big, medium, and small.) There are 108 distinct objects and 23,328 distinct situations. Situations are generated uniformly at random.

For several natural languages we construct a limited sublanguage of utterances related to these situations. A typical utterance in English is *the medium purple star below the small hexagon*. There are 168 meanings referring to a single object and 112,896 meanings referring to two objects, for a total of 113,064 possible meanings. The 113,064 possible meanings are instances of 68 general forms: 4 referring to a single object and 64 referring to two objects. These languages are the **68-form languages**.

We consulted at least one speaker of each language to help us construct a meaning transducer to translate appropriate phrases in the language to all 113,064 possible meanings. Each transducer was constructed to have exactly one accepted phrase for each possible meaning. We also constructed transducers for reduced sublanguages, consisting of the subset of utterances that refer to a single object (168 utterances) and those that refer to two objects, but include all three attributes of both (46,656 utterances.) Each meaning in the reduced sublanguage is an instance of one of 8 general forms, while most of the lexical and syntactic complexity of the 68-form language is preserved. We refer to these reduced sublanguages as the **8-form languages**.

5.2 How many interactions are needed to learn?

The level of performance of a learner is measured using two quantities: the correctness and completeness of the learner's utterances in a given situation. The learning procedure has a test mode in which the learner receives a situation and responds with the set of U utterances it could produce in that situation, with their corresponding production probabilities. The **correctness** of the learner is the sum of the production probabilities of the elements of U that are in the correct denoting set. The **completeness** of the learner is the fraction of all correct denoting utterances that are in U. The averages of correctness and completeness of the learner for 200 randomly generated situations are used to estimate the overall correctness and completeness of the learner. A learner reaches a **level** p **of performance** if both correctness and completeness are at least p.

In the first set of trials the target level of performance is 0.99 and the learner and teacher engage in a sequence of interactions until the learner first reaches this level of performance. The performance of the learner is tested at intervals of 100 interactions. Fig. 2 shows the number of interactions needed to reach the 0.99 level of performance for each 68-form language with correction probabilities of 0.0 (i.e., the teacher never corrects the learner) and 1.0 (i.e., the teacher offers a correction to the learner every time it classifies the learner's utterance as an error in form or an error in meaning.) For correction probability 1.0, it also shows the number of incorrect utterances by the learner, the number of corrections offered by the teacher, and the percentage of teacher utterances that were corrections. Each entry is the median value of 10 trials except those in the last column. It is worth noting that the learner does not treat corrections specially.

	0.0	1.0	incorrect	corrections	c/u%
English	700	750	25.0	11.5	1.5%
German	800	750	71.5	52.5	7.0%
Greek	3400	2600	344.0	319.0	12.3%
Hebrew	900	900	89.5	62.5	6.9%
Hungarian	750	800	76.5	58.5	7.3%
Mandarin	700	800	50.0	31.5	3.9%
Russian	3700	2900	380.0	357.0	12.3%
Spanish	1000	850	86.0	68.0	8.0%
Swedish	1000	900	54.0	43.5	4.8%
Turkish	800	900	59.0	37.0	4.1%

Figure 2: Interactions, incorrect learner utterances and corrections by the teacher to reach the 0.99 level of performance for 68-form languages.

In the column for correction probability 0.0 there are two clear groups: Greek and Russian, each with at least 3400 interactions and the rest of the languages, each with at most 1000 interactions. The first observation is that the learner achieves correctness and completeness of 0.99 for each of these languages after being exposed to a small fraction of all possible situations and utterances. Even 3700 interactions involve at most 16.5% of all possible situations and at most 3.5% of all possible utterances by the teacher, while 1000 interactions involve fewer than 4.3% of all situations and fewer than 1% of all possible utterances.

5.3 How do corrections affect learning?

In the column for correction probability 1.0 we see the same two groups of languages. For Greek, the number of interactions falls from 3400 to 2600, a decrease of about 24%. For Russian, the number of interactions falls from 3700 to 2900, a decrease of about 21%. Corrections have a clear positive effect in these trials for Greek and Russian, but not for the rest of the languages.

Comparing the numbers of incorrect learner utterances and the number of corrections offered by the teacher, we see that the teacher finds corrections for a substantial fraction of incorrect learner utterances. The last column of Fig. 2 shows the percentage of the total number of teacher utterances that were corrections, from a low of 1.5% to a high of 12.3%.

There are several processes at work in the improvement of the learner's performance. Comprehension improves as more information accumulates about words and predicates. New correct general forms are acquired, and unmatched incorrect general forms decrease in probability. More data improves the decision tree rules for choosing phrases. Attainment of the 0.99 level of performance may be limited by the need to acquire all the correct general forms or by the need to improve the correctness of the phrase choices.

In the case of Greek and Russian, most of the trials had acquired their last general form by the time the 0.90 level of performance was reached, but for the other languages correct general forms were still being acquired between the 0.95 and the 0.99 levels of performance. Thus the acquisition of general forms was not a bottleneck for Greek and Russian, but was for the other languages. Because the teacher's corrections generally do not help with the acquisition of new general forms (the general form in a correction is often the same one the learner just used), but do tend to improve the correctness of phrase choice, we do not expect correction to reduce the number of interactions to attain the 0.99 level of performance when the bottleneck is the acquisition of general forms. This observation led us to construct reduced sublanguages with just 8 general forms to see if correction would have more of an effect when the bottleneck of acquiring general forms was removed.

The reduced sublanguages have just 8 general forms, which are acquired relatively early. Fig. 3 gives the numbers of interactions to reach the 0.99 level of performance (except for Turkish, where the level is 0.95) for the 8-form sublanguages with correction probability 0.0 and 1.0. These numbers are the means of 100 trials (except for Greek and Russian, which each had 20 trials); the performance of the learner was tested every 50 interactions.

Comparing the results for 8-form sublanguages with corresponding 68-form languages, we see that some require notably fewer interactions for 8-form

	0.0	1.0	% reduction
English	247.0	202.0	18.2 %
German	920.0	683.5	25.7 %
Greek	6630.0	4102.5	38.1 %
Hebrew	1052.0	771.5	26.7 %
Hungarian	1632.5	1060.5	35.0 %
Mandarin	340.5	297.5	12.6 %
Russian	6962.5	4640.0	33.4 %
Spanish	908.0	630.5	30.6 %
Swedish	214.0	189.0	11.7 %
Turkish	1112.0*	772.0*	30.6 %

Figure 3: Interactions to reach the 0.99 level of performance for 8-form languages. (For Turkish: the 0.95 level.)

sublanguages (English, Mandarin, and Swedish) while others require notably more (Greek, Hungarian and Russian.) In the case of Turkish, the learner cannot attain the 0.99 level of performance for the 8-form sublanguage at all, though it does so for the 68-form language; this is caused by limitations in learner comprehension as well as the differing frequencies of forms. Thus, the 8-form languages are neither uniformly easier nor uniformly harder than their 68-form counterparts. Arguably, the restrictions that produce the 8-form languages make them "more artificial" than the 68-form languages; however, the artificiality helps us understand more about the possible roles of correction in language learning.

Even though in the case of the 8-form languages there are only 8 correct general forms to acquire, the distribution on utterances with one object versus utterances with two objects is quite different from the case of the 68-form languages. For a situation with two objects of different shapes, there are 40 denoting utterances in the case of 68-form languages, of which 8 refer to one object and 32 refer to two objects. In the case of the 8-form languagues, there are 10 denoting utterances, of which 8 refer to one object and 2 refer to two objects. Thus, in situations of this kind (which are $5/6$ of the total), utterances referring to two objects are 4 times more likely in the case of 68-form languages than in the case of 8-form languages. This means that if the learner needs to see utterances involving two objects in order to master certain aspects of syntax (for example, cases of articles, adjectives and nouns), the waiting time is noticeably longer in the case of 8-form languages.

This longer waiting time emphasizes the effects of correction, because the initial phase of learning is a smaller fraction of the whole. In the third column of Fig. 3 we show the percentage reduction in the number of interactions to reach the 0.99 level of performance (except: 0.95 for Turkish) from correction probability 0.0 to correction probability 1.0 for the 8-form languages. For each language, corrections produce a reduction, ranging from a low of 11.7% for Swedish to a high of 38.1% for Greek. This confirms our hypothesis that corrections can substantially help the learner when the problem of acquiring all the general forms is not the bottleneck.

6 Discussion and Future Work

We show that a simple model of a teacher can offer meaning-preserving corrections to the learner and such corrections can significantly reduce the number of interactions for the learner to reach a high level of performance. This improvement does not depend on the learner's ability to detect corrections: the effect depends on the change in the distribution of teacher utterances in the correcting versus non-correcting conditions. This suggests re-visiting discussions in linguistics that assume that the learner must identify teacher corrections in order for them to have an influence on the learning process.

Our model of language is very simplified, and would have to be modified to deal with issues such as multi-word phrases bearing meaning, morphological relations between words, phonological rules for word choice, words with more than one meaning and meanings that can be expressed in more than one way, languages with freer word-orders and meaning components expressed by non-contiguous sequences of words. Other desirable directions to explore include more sophisticated use of co-occurrence information, more powerful methods of learning the grammars of meanings, feedback to allow the learning of production to improve comprehension, better methods of alignment between utterances and meanings, methods to allow the learner's semantic categories to evolve in response to language learning, and methods allowing the learner to make use of its ability to detect corrections.

References

D. Angluin and L. Becerra-Bonache. 2010. A Model of Semantics and Corrections in Language Learning. Technical Report, Yale University Department of Computer Science, YALE/DCS/TR-1425.

R. Brown and C. Hanlon. 1970. Derivational complexity and the order of acquisition in child speech. In J.R. Hayes (ed.): *Cognition and the Development of Language*. Wiley, New York, NY.

M.M. Chouinard and E.V. Clark. 2003. Adult Reformulations of Child Errors as Negative Evidence. *Journal of Child Language*, 30:637–669.

E.V. Clark 1987. The principle of contrast: a constraint on language acquisition. In B. MacWhinney (ed.): *Mechanisms of language acquisition*. Erlbaum, Hillsdale, NJ.

E.V. Clark 1993. *The Lexicon in Acquisition*. Cambridge University Press, Cambridge, UK.

M. J. Demetras, K. N. Post and C.E. Snow. 1986. Brown and Hanlon revisited: mothers' sensitivity to ungrammatical forms. *Journal of Child Language*, 2:81–88.

J.A. Feldman, G. Lakoff, A. Stolcke and S. Weber 1990. Miniature Language Acquisition: A Touchstone for Cognitive Science. *Annual Conference of the Cognitive Science Society*, 686–693.

E.M. Gold. 1967. Language identification in the limit. *Information and Control*, 10:447–474.

K. Hirsh-Pasek, R.A. Treiman M. and Schneiderman. 1984. Brown and Hanlon revisited: mothers' sensitivity to ungrammatical forms. *Journal of Child Language*, 2:81–88.

G.F. Marcus 1993. Negative evidence in language acquisition. *Cognition*, 46:53–95.

J.L. Morgan and L.L. Travis. 1989. Limits on negative information in language input. *Journal of Child Language*, 16:531–552.