A Collaborative Tool for the Computational Modelling of Child Language Acquisition

Kris Jack

CEA LIST, Laboratoire d'ingénerie de la connaissance multimédia multilingue
18 route du Panorama, BP6
Fontenay-aux-Roses, F-92265 France
mrkrisjack@gmail.com

Abstract

A large number of computational language learners have been proposed for modelling the process of child language acquisition. Comparing them, however, can be difficult due to the different assumptions that they make, the diverse test results presented, and the different linguistic behaviours investigated. This paper introduces a toolkit that allows different language learners to be trained, tested and analysed under standardised conditions. The results can be easily compared with one another and with typical child language development to highlight the relative advantages and disadvantages of learners.

1 Introduction

The computational modelling of language acquisition can help in understanding the acquisition process by estimating the problem faced by children and designing algorithms that solve it in a similar way as they do. Many such models have been produced in recent years, tackling various linguistic behaviours. Like in any relatively new domain of research, however, the treatment of one problem often reveals the presence of several more that in turn require new solutions of their own. This has led to the design and implementation of numerous learners that differ in either subtle or fundamental ways. Given such variety, it is not yet clear which kind of model, or combination of models, can best account for the overall behaviour witnessed during child language development.

When surveying the computational language acquisition literature, the relative advantages and disadvantages of language learners are not always clear. As such, it can be very difficult to compare different learners with one another. The main problem is the lack of standardisation in the field. Language learners are constructed with different underlying assumptions, largely due to the lack of consensus in linguistic theory, are trained using different data (that can vary from miniature languages to full blown natural languages) and are tested using different testing measures (some of which include the '*Looks good to me*' approach).

In this paper, a toolkit for investigating the computational modelling of child language acquisition is proposed. The Language Acquisition Toolkit (LAT) allows researchers to work collaboratively in solving the modelling task, while addressing the problems introduced. It is an attempt to bring the field forward by creating a standardised way for testing and implementing language acquisition learners. The issues addressed in this paper are largely driven by engineering concerns although the choices that are made by the modeller will impact not only on their learner but also on the associated language theory. The driving motivation behind the LAT is that the best way to compare different language learners is to compare the behaviours that they produce. The closer a learner's behaviour is to the behaviour witnessed in children, the better the model.

The LAT is a computational framework that can train, test and analyse the linguistic performance of a computational language learner in order to chart developmental linguistic trajectories. The motivation for the LAT shall first be explored before describing it in detail, discussing its features and considering future directions.

2   Background

The process of modelling child language acquisition is very complex, as many of the first attempts confirmed (Feldman *et al.*, 1990; Suppes, Liang & Bottner, 1991). Rather than modelling the process in entirety, an undoubtedly daunting task, modellers took the simplified approach of focusing upon individual linguistic behaviours, leading to much research into relatively constrained problems such as understanding over- and under-generalisation errors (Plunkett, Sinha, Moller & Strandsby, 1992), single word learning (Regier, 2005), syntactic category acquisition (Redington, Chater & Finch, 1988) and past-tense learning (Rumelhart & Mcclelland, 1986). While such models have led to valuable insights in the domain, it can be difficult to see how each of them is related to one another given the lack of standardised learning, testing and analysis.

Often, the variety found in computational models reflects the divisions between linguistic theories pertaining to child language acquisition (Kaplan, Oudeyer & Bergen, 2008). Given that linguists remain divided about how children learn language, it is not surprising to find a similar division in the computational modelling community as well. One of the fundamental issues that separates modellers is the kind of data that the learner learns. This can range from the use of plain textual data (Elman, 1993), to grounded sensor-based input (Roy, 2008). Standardising the type of learning data would thus be useful for comparing language learners.

Typical computational models are often tested under different circumstances and using different techniques. For example, while some papers offer a general analysis of the model's behaviour, others focus on particular features, while some test language comprehension, others test language production, and while some consider developmental growth, others consider only the start and end points of training. Although this is often justifiable in the context of the research problem, it makes it difficult to directly compare two models. It would be useful to put all models through the same set of rigorous tests in order to find out how they are similar and how they differ from one another. Such standardised testing will often reveal important differences that may have previously been hidden.

Practically, however, not all models that are described in the literature are made available for download. As a result, researchers often have to spend time recreating models. This assumes, of course, that the model has been described in enough detail that it can be faithfully recreated. Much time could be saved if such models were available for download, from a common repository, such as the Weka makes machine learning algorithms freely available in a software workbench (http://www.cs.waikato.ac.nz/~ml/).

A good language learner should not just solve language learning problems, but should do so in a similar way as is witnessed in children. Based on psycho-linguistic evidence, several linguistic timetables have been derived containing important linguistic milestones (Brown, 1973; Ingram, 1989; Pinker, 1994; Tomasello, 2005). The character of language development is a significant feature in child language acquisition and modellers should be encouraged to model it to better understand the process. A language learner that demonstrates a good use of syntax at the same time as producing its first words is not very realistic. Instead, there should be a prolonged period in which words are learned followed by the emergence of syntax. Unfortunately, a language model can often produce behaviours at unexpected times, signalling a problem with the linguistic theory that it embodies. A standardised approach to analysing the linguistic development of a language learner would be an advantage.

3   The Language Acquisition Toolkit

3.1   Introduction

The Language Acquisition Toolkit (LAT) is a piece of software that allows researchers to develop and test computational language learners within a standardised environment. The LAT's target users are researchers who have basic skills in software development and are comfortable using the programming language Java. It assumes that the language learner operates under the restrictions imposed in the miniature language paradigm (Feldman *et al.*, 1990). The LAT can be obtained from www.langac.com and is available under a GNU public license meaning that the code can be reproduced and modified without obtaining permission.

The LAT is an attempt to standardise the training, testing and analysing of language learners within an open an accessible environment (Figure 1). In training, the language learner observes a

simulated world in which action-based events occur. Both simulated descriptions and visual data are sent to the language learner for analysis. The LAT then tests both the language learner's comprehension and production capacities. Comprehension is tested by sending a description to the language learner and scoring the visual data that are produced. Similarly, production is tested by sending visual data to language learner and scoring the descriptions produced. The LAT then analyses the results obtained from testing and develops data describing the learner's development.
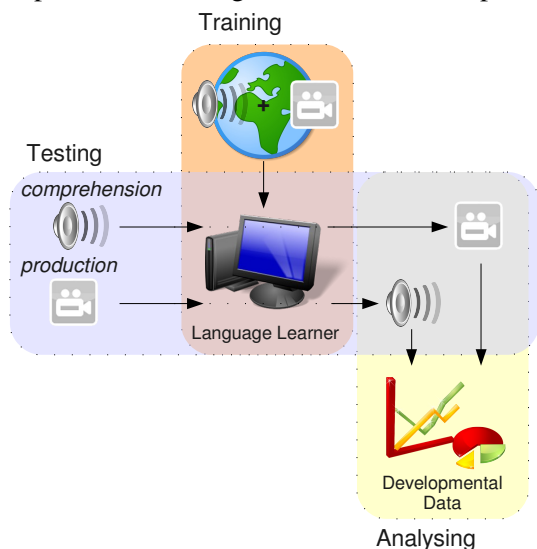

Figure 1: LAT Overview. A language learner is placed in the LAT's simulated world where it learns from simulated audio and visual data. The LAT tests the learner and the results are used to produce data describing its development.

3.2 Training

The LAT can be configured to train different language learners by generating a simulated environment in which action-based events occur. The simulated environment operates within the miniature language acquisition paradigm (Feldman *et al.*, 1990), a simplified simulation of the real-world. A simulation is employed rather than grounding the model in the real-world in order to better control the number and type of problems that are being investigated in a single experiment. While the miniature language paradigm imposes a number of constraints, the proposed simulation contains enough complexity to justify its use.

The learner is trained by watching an event that is simulated in the blocks world in which a number of geometric objects can be found. When an event occurs, a symbolic representation of the description and visual data are generated.

More concretely, an event is the pairing of a simulated description and a action, $e = \langle d, a \rangle$. Events are represented following evidence from child studies. First, it is assumed that the learner can establish a triadic relationship between an object, a speaker and themselves in order to associate a description with an action. This kind of relationship is typically called joint-attention and does not appear in children until around 12 months-old (Tomasello, 1995). As such, the symbolic content present in descriptions and actions are limited to those found in child literature during the first year of life.

An infant's acoustic sensitivity is so attuned that from four-days-old she demonstrates the ability to differentiate between native and non-native speech (Mehler *et al.*, 1988). Such discrimination lies in rhythmic properties that differ over language groups (Dehaene-Lambertz & Houston, 1998; Mehler, Dupoux, Nazzi & Dehaene-Lambertz, 1996) and is likely to be syllable-based since infants detect change in syllable quantity, but not in phoneme quantity over samples of speech (Bijeljac-Babic, Bertoncini & Mehler, 1993). Infants also detect vowel change, a syllable covariant, more readily than consonant change (Bertoncini *et al.*, 1988), further supporting a syllabic base. A description is thus represented as a non-zero length ordered list of syllables in the LAT. Word segmentations are not included as there is no acoustic equivalent of the blank space in written language.

In terms of visual sensitivity, infants can identify objects through retinal and object displacement during motion from four months-old (Kellman, Gleitman & Spelke, 1987), and make relative spacial distinctions between left and right, and above and below, from three to ten months old (Quinn & Schyns, 2003). Infants can also make use of shape and colour to differentiate between objects in the first year of life (Landau, Smith & Jones, 1988). The LAT thus describes the physical properties of objects that inhabit the blocks world (e.g. shape, colour, size and position), referred to as features. An action is defined as a non-zero length ordered list of feature sets, where each feature set is associated with a unique time interval. A set of features describes all objects that can be seen in an event. Note that actions in this terminology do not relate to actions in terms of verbs in natural language, but to a list of descriptions of scenes. Properties such as

push and pull are thus not explicitly represented as symbolic features.

Two types of events can occur in the blocks world: action-based; and descriptive. In the case of an action-based event, an object performs an action while in the case of a descriptive event, objects do not change. As a result, action-based events contain different feature sets, giving the impression of change, while descriptive events contain the same feature sets, indicating no change. The description in an action-based event describes the action while the description in a descriptive event describes an object in the static scene. Objects can perform several actions including moving, flashing, growing, shrinking, appearing, disappearing, destroying another object, hitting another object, pushing another object and pulling another object.

The LAT randomly generate events that can be used as training data. It can create objects, make them perform actions, and describe the events by instantiating appropriate grammar fragments. To encourage the use of standardised sets of training data, a number of sets of data have been randomly generated that each contain 10,000 events. These data have been generated from different parameters (e.g. amount of noise, probability that an object will perform an action in an event, probabilities for each action to occur, number of time intervals for an event, number of colours/shapes/sizes/actions possible) with different language properties (e.g. recursion present/ not present, number of rules, language in use).

To provide concrete examples of typical LAT training data, one data set, called the Appearance data set will be presented in detail. The appearance data set is inspired from a study with real participants. Participants sat in front of a computer screen that initially showed a blank white screen. They were asked to describe all changes that were made to the screen in enough detail that a stranger could recreate the scene using only their descriptions. By pressing a key on the keyboard, a new geometric object appeared on the screen and the change was described by the participant. While the addition of an object to a scene appears to be a trivial change, participants produced complex linguistic descriptions that revealed a deep knowledge of their language. For example, descriptions such as "a blue circle appeared to the upper right of the green square at the bottom" and "a red circle appeared between

the four squares making the shape of a cross" (Jack, 2005).

Given the complexity of the language produced, a simplified version the task was constructed in which only the appearance of one object next to another object was considered. By restricting the context, there is less demand for a computational language learner to have a rich semantic representation of scenes. This served as a reasonable starting point from which to conduct the investigation. The actions in the Appearance data set were constructed by randomly generating one object and placing it in the middle of a 3x3 grid scene and then adding a second object, which was also randomly generated, in a different position. Eight colours and shapes were used. Each action was also accompanied by an appropriate description that was generated using a grammar fragment (Figure 2).

$E \rightarrow NP_1$ PAR $_2$	
PAR1 $\rightarrow NP_1$ PART	PAR$_2 \rightarrow$ REL NP$_2$
RELT $\rightarrow$ REL Det$_2$	REL $\rightarrow$ REL$_1$ | REL$_2$
REL$_1 \rightarrow$ a bove | be low | to the REL$_4$	REL$_2 \rightarrow$ REL$_3$ REL$_4$
REL$_3 \rightarrow$ to the low er | to the u pper	REL$_4 \rightarrow$ left of | right of
NP$_1 \rightarrow$ Det$_1$ N$_{bar}$	NP$_2 \rightarrow$ Det$_2$ N$_{bar}$
N$_{bar} \rightarrow$ SHP COL	
Det$_1 \rightarrow$ a	Det$_2 \rightarrow$ the
COL $\rightarrow$ black | blue | grey | green | pink | black | red | ye low	SHP $\rightarrow$ cir cle | cross | dia mond | heart | rec tang le | star |square | tri ang gle

Figure 2: Miniature Language from Appearance Data Set. All strings are syllable segmented rather than word segmented.

Events from this data set have actions that are described using a 2-frame time interval, where the first set of features describes the state of the scene before the action occurs and the second set of features describes the scene after the action occurs (Figure 3). Note that it is assumed that the learner can identify concepts such as colour, shape and position and that such symbolic information is associated with a particular object. The notion of object-hood, where the first object in the scene is O_1 and the second object is O_2, is carried across time intervals with O_1 being recog-

nised as the same object before and after an action occurs.

Before action (t=1)	After action (t=2)
O_1: square O_1: blue O_1: x2 O_1: y2	O_1: square O_2: circle O_1: blue O_2: yellow O_1: x2 O_2: x3 O_1: y2 O_2: y3
a ye low cir cle to the u pper right of the blue square	

Figure 3: Sample Event from Appearance Data Set. Two time frames represented graphically and as feature sets. The accompanying syllable segmented description of the event is also shown.

The remaining data sets contain more complex events in which more actions and richer miniature languages are employed. Actions are randomly generated, with respect to the constraints imposed on the data set (e.g. number of colours, shapes, and actions) and appropriate descriptions are generated. These descriptions are produced by following a heuristic that minimises the number of syllables that can appear in a single description. This reduces the production of unnatural sentences. For example, take the case where an object appears in a scene amongst 10 other objects. A description could be generated to describe the action with respect to one other object, two other objects or as many as 10 other objects. While such descriptions are all valid, many of them would sound unnatural if employed. The algorithm selects descriptions by favouring those that have fewer syllables. A parser is then employed that eliminates invalid descriptions that can be misinterpreted. By making a parsimonious use of syllables, more natural descriptions tend to be produced. More abstract language can also be found such as the use of the word 'bullying' to describe pushing, pulling and hitting.

3.3 Testing

The LAT monitors the linguistic development of a language leaner by testing its comprehension and production capacities. The learner's comprehension and production are tested at every round of training.

For each set of training data, there is an associated set of testing data, ensuring a standardised test procedure for language learners. Test data is produced using grammar rules for producing descriptions and heuristics for producing actions. The tests are constructed to reflect the properties found in the training data's miniature language. As such, the learner is only tested on the kind of descriptions and actions that it has the opportunity to learn through observing events. Concretely, a testing set is a set of events where each event relates one or more descriptions to one or more actions. The set of testing data associated with the Appearance training data set can be used to test the learner's vocabulary, certain multi-word combinations and full sentences. Using the terminology found in Appearance's grammar fragment (Figure 2), the LAT tests for the comprehension of shapes (SHP), colours (COL), objects (N_{bar}), indefinite objects (NP_1), definite objects (NP_2) and events (E).

In testing the learner's comprehension, the LAT sends a description as input and receives a set of actions as output. The output is automatically scored by comparing it with the expected output that is associated with the description. Actions are compared based on the feature values that are relevant to the given description. Given the description "a ye low cir cle to the u pper right of the blue square" (Figure 3), the colours, shapes and relative positions of the objects are relevant whereas their exact positions are not. The LAT equally accepts a yellow circle that appears higher or further right than its idealised position with respect to the blue square, as long as the relative positions remain correct.

Borrowing from research in child language acquisition studies, four kinds of incorrect responses are identified: over-extended; under-extended; mismatched; or incorrect. For example, the meaning of the description "square" is under-extended if the learner only uses it to refer to red squares, blue squares and green squares, but not to squares of other colours. Similarly, the meaning of the description "red square" is over-extended if it refers to red squares, blue squares and red circles. A mismatch is found if the description "square" is used to refer to objects other than squares, for examples circles and triangles, but never to squares themselves. Results that deviate from these cases are simply considered incorrect.

The LAT can score both single words and phrases based on these categories.

In addition, the output produced by the learner can also be described using the standard information retrieval measures of precision, recall, and the e-measure which is a weighted combination of the two former values (van Rijsbergen, 1979).

The process of testing the learner's production is similar to that of testing comprehension. Rather than the LAT sending a description as input, however, it sends an action. The learner then produces a set of descriptions as output. Results from production are scored using the same principles as applied during comprehension. That is, the learner's output is compared to the expected output and it is scored as either correct, over-extended, under-extended, mismatched or incorrect.

3.4 Analysing

Both the comprehension and production results that are produced from testing are used to evaluate the learner's linguistic stage of development. Several types of analysis have been designed to ease learner comparisons: round-based; trial-based; and learner-based. Round-based analyses analyse the results produced from a single round of testing. Trial-based analyses take round-based statistics and compare them with previous rounds in order to find behavioural trends in the data. Finally, learner-based analyses compare trial-based data for several trials in order to extract general behavioural trends. By performing analyses at all three levels of detail, a more complete account of the learner's behaviour is produced.

The LAT is currently able to perform a number of round-based analyses that are often found in the literature: summary of test results in terms of correct results and errors; chart the linguistic generativity of the learner; and present evidence of syntactic activity.

Round-based analyses produce results that are then used to determine the model's stage of linguistic development using data from child language studies: pre-linguistic; holophrastic; early multi-word; late multi-word; and abstract stages.

A number of trial-based analyses are performed using these data, in order to identify particular linguistic behaviours: linguistic development; vocabulary acquisition; comprehension/production imbalance. With the creation of a linguistic development timetable, all data can also be presented in terms of stages. For example, the number of words that are correctly comprehended and the rate of vocabulary acquisition can be shown by stage.

Model-based analyses can performed when the results from several trials are available. Each of the results, such as the rate of vocabulary acquisition during a stage, are compared across trials to identify general behavioural trends.

The LAT thus offers a standardised platform for training, testing and analysing language models. The results from all analyses can be automatically compared to determine the differences between learners and which learner best fits child language data.

4 Discussion

The LAT is a freely available tool that offers a standardised environment from which language modellers can develop their language learners. It is an attempt to advance the domain by offering a platform where common goals can be focussed upon in a collaborative environment. It aims to standardise the training, testing and analysing of language learners by understanding the needs of language modellers through collaboration.

By using the LAT, the language modeller accepts the need to work with standardised training data. Such standardisation is widespread in computational linguistics. For example, in the field of automatic text classification, there are several databases of pre-classified documents (e.g. Reuters-21578, Reuters Corpus Volume 1 and 20 Newsgroups) that researchers can use to evaluate different algorithms and to compare their results. The LAT offers different sets of training data that are constrained by principles of the miniature language paradigm. In using such data, the modelling task differs from the task that a child faces in a number of ways. In particular, the learning problem is simplified in that the real-world contains many more objects and that natural language has far more linguistic structures and words than the language fragments. It is for these reasons, however, that such a paradigm is attractive. Many language learning problems can be effectively investigated by first simplifying the problem and then developing solutions. When such problems in the miniature language paradigm have been adequately solved, it is envisaged that the LAT can be grounded in a real

environment where vast volumes of data are available for processing.

The results from learning can then be tested using a standardised set of tests. The learner is treated as a black box, meaning that the LAT evaluates its output alone without entering into its inner workings. This helps to keep the LAT's functionality independent from the learner by focussing on the way in which it behaves rather than how it produces particular behaviours, similar to the relationship found between the linguist and child in the real world. By testing both comprehension and production on a large set of descriptions and actions, a complete picture of the learner's linguistic state can be derived. The LAT also checks for language errors such as over-extensions, under-extensions and mismatches. Individual results are made available to the researcher in a tabular format as well as providing overall recall, precision and e-measure scores.

By standardising the test results, different language learners can be easily compared with one another. The LAT can analyse these results to discover behavioural trends in the data with can be used in further comparisons. It is also interesting to note that the LAT makes an attempt to compare the behaviour produced by a language learner with that of children. Inspired by child language development timetables, a set of milestones has been derived that are used to characterise the learner's behaviour in terms of stages. The LAT attempts to encourage researchers to consider the developmental behaviour of their language learners over time.

It is important to note that the LAT is a work in progress. This disclaimer is likely to remain true for many years. Developing a gold standard is a difficult task and one that risks to evolve over time. The LAT should be regarded as a proposal for standardisation. Being a collaborative project, any contributor can challenge this proposal by offering their own solutions. Contributors are encouraged to create their own data and algorithms and to upload them to the LAT. A gold standard can only emerge from the selections that are made by other modellers, who vote by using certain data and algorithms in their own modelling tasks. In this sense, the proposed instantiation of the LAT described in this article is less important than the idea behind the LAT itself.

5   Future Considerations

In designing the LAT, it quickly became clear that the task was not straight-forward. Designing a tool that can make useful and standardised comparisons between language learners is a complex task. A balancing act between not excluding certain types of learners and creating a constrained, manageable environment is not without its difficulties. As such, it is worth considering future developments for the LAT. While still in a preliminary state of development, it is hoped that a collaborative approach to the task will allow it to be steered in the directions that are best adapted to its potential users. A number of these directions are now considered.

The miniature language paradigm is at the heart of the LAT. This language can be extended to include more complex linguistic constructions and a larger vocabulary. It is suggested that a systematic approach is followed in which the learning task is made progressively complex by adding linguistic features that tend to be witnessed in children during development. It seems reasonable to follow a longitudinal approach to development. Contributors are also encouraged to create and submit new training data sets in order to explore how complex a miniature language can become.

The type of information that is available to the learner could also be changed. At present, the descriptions lack acoustic information such as tone. Such data is indispensable in investigating certain languages such as Mandarin and Swahili. Similarly, the symbolic representations of visual objects can be refined to better represent reality. Colours can be represented by RGB values rather than linguistically-related symbols, as it is unlikely that children start with such pre-defined semantic categories from the outset of learning.

It is also worthwhile considering more complex testing and analysis algorithms. It is likely that they will be developed in step with new linguistic phenomena that are investigated, building a useful catalogue of tools. In addition, it may be useful to develop learner-dependant analysis tools in order to demonstrate how the inner workings are related to the outward behaviour.

Finally, it is hoped that the LAT will become a useful resource not just for modellers who are comfortable with coding but also non-programmers. They should be able to implement and experiment with different kinds of models with the

flexibility of looking at different aspects of acquisition under different settings and with different types of data. They can then inform language modellers directly about how particular language models perform well and poorly in certain cases. The collaborative aspect of the LAT encourages not just programmers to share their code, but for everyone to share their ideas.

6 Conclusion

This article proposes a tool that facilitates the consolidation of research into the computational modelling of child language acquisition under the miniature language paradigm. The workshop is being used to launch a first version of the LAT, that is hoped to help language modellers and child language experts to communicate and share their knowledge.

7 Acknowledgements

References

Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P., Kennedy, L. & Mehler, J. (1988). An investigation of Young infants' perceptual representations of speech sounds. *Journal of experimental psychology*, *117*, pp. 21-33.

Bijeljac-Babic, R., Bertoncini, J. & Mehler, J. (1993). How Do 4-Day-Old Infants Categorize Multisyllabic Utterances?. *Developmental Psychology*, *29*, pp. 711-721.

Brown, R. (1973). A First Language: The Early Stages. Harvard University Press.

Dehaene-Lambertz, G. & Houston, D. (1998). Faster Orientation Latencies Toward Native Language in Two-Month-Old Infants. *Language and Speech*, *41*, pp. 21-43.

Elman, J. (1993). Learning and development in neural networks: the importance of starting small. *Cognition*, *48*, pp. 71-99.

Feldman, J., Lakoff, G., Stolcke, A. & Weber, S. (1990,). *Miniature Language Acquisition: A Touchstone for Cognitive Science*.

Ingram, D. (1989). First Language Acquisition: Method, Description and Explanation. Cambridge University Press.

Jack, K. (2005,). *Introducing a Scene Building Game to Model Early First Language Acquisition*, CLUK, Manchester, England.

Kaplan, F., Oudeyer, P. & Bergen, B. (2008). Computational models in the debate over language learnability. *Infant and Child Development*, *17*, pp. 55-80.

Kellman, P., Gleitman, H. & Spelke, E. (1987). Object and observer motion in the perception of objects by infants. *Journal of experimental psychology. Human perception and performance*, *13*, pp. 586-593.

Landau, B., Smith, L. & Jones, S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*, pp. 299-321.

Mehler, J., Dupoux, T., Nazzi, T. & Dehaene-Lambertz, G. (1996). Coping with linguistic diversity: The infant's viewpoint. Lawrence Erlbaum.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J. & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*, pp. 143-178.

Pinker, S. (1994). The Language Instinct. William Morrow.

Plunkett, K., Sinha, C., Moller, M. & Strandsby, O. (1992). Symbol grounding or the emergence of symbols? Vocabulary growth in children and a connectionist net. *Connection Science*, *4*, pp. 293-312.

Quinn, P. & Schyns, P. (2003). What goes up may come down: perceptual process and knowledge access in the organization of complex visual patterns by young infants. *Cognitive Science*, *27*, pp. 923-935.

Redington, M., Chater, N. & Finch, S. (1988). Distributional Information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, *22*, pp. 425-469.

Regier, T. (2005). The emergence of words : Attentional learning in form and meaning. *Cognitive Science*, *29*, pp. 819-865.

Roy, D. (2008). A mechanistic model of three facets of meaning. In M. D. Vega, G. Glennberg & G. Graesser (Eds.), *Symbols and Embodiment*. Oxford University Press.

Rumelhart, D. & Mcclelland, J. (1986). On learning the past tenses of English verbs. In D. Rumelhart & J. Mcclelland (Eds.), *Parallel distributed processing: explorations in the microstructure of cognition*. MIT Press. pp. 216-271.

Suppes, P., Liang, L. & Bottner, M. (1991). Complexity Issues in Robotic Machine Learning of Natural Language. In L. Lam & V. Naroditsky (Eds.), *Modeling Complex Phenomena*. Springer Verlag.

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development*. Erlbaum.

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

van Rijsbergen, C. J. (1979). Information Retrieval. London, Butterworths.