

Abstracts

Natasha Kirkham

The Richness of the Input: Looking at the impact of multiple cues, multiple modalities

Previous studies of infant statistical learning have provided robust evidence that infants across the first year of life are remarkably sensitive to probabilistic relations among visual information, within both temporal and spatial arrays (Fiser & Aslin, 2002; Kirkham et al., 2002; 2007). This work has raised important questions regarding the usefulness of statistical learning in an environment filled with the variability and noise typical of the natural world. This talk will look at how multiple cues across multiple modalities impact on the infant's understanding of her environment.

Padraic Monaghan

Modelling multiple cues in language learning

Multiple cues are prevalent in situations where children are acquiring language. For speech segmentation, word-meaning mappings, and learning grammatical categories of words, phonological and prosodic information from the word itself is supported by distributional information from the co-occurrence of words in phrases, as well as gestural and joint-attention information from the speaker. In this talk, I present corpus analyses and computational models of language learning to highlight the serendipitous manner in which multiple cues are present in the child's language learning environment. I focus on a model of word-meaning mappings, with implementations of gestural, prosodic, and distributional information combining to help learning. The model provides insight into why multiple cues are present, and also why no single cue is perfectly reliable to support learning. The consequence of multiple cues in the environment is enhanced learning, but also a system that is more robust to variability in the environment for future instances of word access and acquisition.

Jessica Horst

The Role of Context in Learning the Meanings of Words

Vocabulary acquisition is an important milestone in early cognitive development. Although much is known about how children guess the meaning of a new word (cf. fast mapping), less is known about how children commit word-meaning associations to memory for later retrieval. This talk will review several empirical studies demonstrating the initial naming context plays a critical role in how well children form robust memory representations of new name-object associations. We will explore word learning contexts in terms of both the to-be-learned targets and the other objects that may be present and competing for children's attention. This series of studies includes both traditional fast mapping tasks as well as teaching children words from reading storybooks. Overall, it's not just what is named that matters—but the context in which the initial naming occurs.

Gert Westermann

From preverbal to language-based categorization

A considerable amount of research has addressed the early abilities of infants to form perceptual categories of objects. Other work has asked what the role of emerging language is in shaping these categories. I will describe empirical work that addresses these questions and then present an artificial neural network model that provides insight into the mechanisms underlying early categorization at the transition from preverbal to language-based categorization. The model accounts for several empirical results from familiarization-novelty-preference studies and predicts that infants will familiarize faster for objects for which a label is known. Furthermore, the model suggests that language-based conceptual representations do not replace earlier perceptually-based category representations but instead modify these earlier representations to strike a balance between perceptual similarity and shared labels.

Julian Pine (for Caroline Rowland)

Predicting individual differences in language development: Going beyond the input

Historically, research on individual differences in language development has tended to focus either on differences in the language-learning environment or on differences in processing factors such as phonological working memory or processing speed. However, there is increasing evidence that differences in processing ability are themselves determined by the child's linguistic experience. In this talk, I will argue for the need to develop computational models that specify how constructs such as phonological working memory capacity and processing speed interact with the input and with children's existing knowledge if we are to understand how children build linguistic knowledge and why some children build such knowledge more quickly than others.

Gary Jones

Predicting individual differences in language development: A computational modelling account

Nonword repetition is a simple task where children repeat aloud spoken nonsense words. The task is highly predictive of language performance, later reading skill, and is a clinical marker of language impairment. This talk presents a computational account of nonword repetition which learns a pared-down version of the language input by applying a straightforward sequence learning mechanism. Surprisingly, this model can account for children's performance across a variety of different repetition tests and between the ages of two and six years purely from the linguistic knowledge that has been acquired as learning progresses. The computational account provides an ideal platform for investigating how changes to model parameters affect performance over time and hence how nonword repetition may change across different individuals.

Michael Ramscar & Cyrus Shaoul
Discriminative learning and language

The development of morphological processing has been the focal topic in a debate over the nature of language, learning and the mind in cognitive science. Particular attention has been paid to the systematic nature of children's morphological errors (for example children tend to go through a phase of saying "mouses" as they learn English nominal morphology). Because these errors are not explicitly corrected, it has been argued that the transition to adult language cannot be explained by learning, and that the acquisition of even relatively simple aspects of grammar must involve innate, language specific mechanisms. We describe the background to this debate, along with some models of nominal learning based on discrimination learning that generate clear and surprising predictions. In particular, that exposure to regular plurals (e.g. rats) can decrease children's tendency to overregularize irregular plurals (e.g. mouses). We review empirical results showing that testing memory for items with regular plural labels does lead to a decrease in irregular plural overregularization in six-year-olds, yet an increase in four-year-olds. These models and results indicate that when the learning problem facing children is properly characterized, overregularization both arises and is resolved as a result of the discriminative nature of human learning systems, and the way learning responds to the distribution of evidence in the linguistic environment. The models and results we review indicate that, far from being evidence for language specific mechanisms, the behaviour manifest in overregularization bears all the hallmarks of basic learning mechanisms that we share with a number of other animals. We discuss the implications of this for our broader understanding of language and learning, as well as the nature of cognitive development across the lifespan.

Cyrus Shaoul & Michael Ramscar
Discriminative learning models

Researchers interested in language overwhelmingly assume a model of learning in which cognitively, events that co-occur together in experience become associated (much as in neuroscience, the metaphor of neurons firing together and wiring together remains pervasive). This naïve understanding of associative learning has long been shown to be false, and incapable of accounting even for the most basic phenomena associated with Pavlovian conditioning. (Rescorla & Wagner, 1972) In fact, research has revealed the basic processes of learning to be an evolutionary twist on Sherlock Holmes' dictum: "When you have eliminated the impossible, whatever remains, however improbable, must be the truth". If a dog associates a bell with food, it is only because learning has systematically weighed every other potential association in its environment and found it wanting (Rescorla & Wagner, 1972; Ramscar et al, 2013c). The mechanisms and principles of discriminative learning are introduced.

Elena Lieven
From strings to grammar

There are good theoretical and empirical reasons to assume that children build their language not only out of individual words but also out of multiword (and morpheme) strings. Evidence comes from the nature of the input; the ways in which children construct novel utterances; and the computational modelling of children's grammars. These strings are the basis for the development of schemas containing slots. Productivity arises from the increasing semantic and pragmatic breadth of the items filling the slots and from the development of connections between schemas on the basis of both meaning and form. The learning of these schemas accounts both for the error-free learning of constructions, as well as for some types of errors. I will briefly illustrate this with reference to *wh*-questions and long-distance dependencies before turning to two studies of 1st person pronoun case errors in English children's speech.

Morten H. Christiansen
Language Acquisition as Learning to Process: The Importance of Chunking

Language happens in the here-and-now. If the linguistic input is not processed immediately, nothing can be learned from it. To successfully deal with the continual deluge of linguistic information, the brain must compress and recode the input as rapidly as possible. As a consequence, incoming language incrementally gets recoded into chunks of decreasing granularity, from sounds to constructions and beyond. To illustrate, I present a usage-based computational model of language acquisition that learns in a purely incremental fashion, through on-line processing of simple statistics, and offers broad, cross-linguistic coverage while uniting comprehension and production within a single framework. The model achieves strong performance across over 200 single-child corpora representing 29 languages from the CHILDES database. I conclude that the immediacy of language processing provides a fundamental constraint on accounts of language acquisition, implying that acquisition involves learning to process, rather than inducing a grammar.

Laurence B. Leonard
Specific Language Impairment Across Languages: Candidate Symptoms for Computational Models

Salient symptoms of grammatical deficits in children with specific language impairment (SLI) will be presented, with an eye toward their possible inclusion in computational models. Symptoms will include a disproportionate weakness in tense/agreement morphology in children with SLI acquiring English and other Germanic languages, inappropriate sentence position placement of finite verbs in German and negative particles in Swedish, reduction of suffix sequences in agglutinating languages, and "near-miss" errors in languages with a rich verb inflection system. Several of these symptoms can probably be incorporated into current models with only minor adjustments, though near-miss errors may require a model that is sensitive to features such as person, number, and tense.

Daniel Freudenthal

**Incorporating defaulting effects into MOSAIC: Building a two-factor model of the
Optional Infinitive stage**

In several languages, children go through a stage during which they produce Optional Infinitive (OI) errors: non-finite verb forms in contexts in which a finite verb form is required. MOSAIC is a computational model of grammatical development that simulates the cross-linguistic patterning of OI errors by learning reduced compound finite structures from the input. However the current version of MOSAIC is unable to simulate the very high rate of OI errors in early child English.

In this talk, we investigate the possibility that the high rate of OI errors in English reflects the operation of two processes: 1) the learning of reduced compound finite structures and 2) defaulting to the most common form of the verb in the input (for most English verbs, the bare stem, but for most Spanish verbs, the 3sg present tense). We describe a version of MOSAIC in which the original mechanism for producing reduced modal structures is complemented by a mechanism that defaults to the most frequent form of the verb in utterance-final strings of increasing length in the input.

We show that this version of the model provides a better fit to data from both English and Dutch without affecting the model's previously good fit to Spanish. This result reflects the tendency of the model to show frequent and persistent defaulting to the bare form in English, a small amount of defaulting to the 3sg in Spanish, and defaulting to the infinitive at low MLUs, but to the bare stem at higher MLUs in Dutch. The model thus provides a plausible two-factor model of the OI stage, which captures cross-linguistic differences in the patterning of both OI and defaulting errors. It also has the potential to serve as the basis for a model of verb-marking errors in children with SLI.