

Modeling Impairments in Lexical Development

Michael Vinos (michael.vinos@gmail.com)

Graduate Program in Basic and Applied Cognitive Science, University of Athens, Greece

Angeliki Andrikopoulou (ang.andrikopoulou@gmail.com)

Department of Philosophy and History of Science, University of Athens, Greece

Christina F. Papailiou (papailiou@rhodes.aegean.gr)

Department of Preschool Education and Educational Planning, University of the Aegean, Rhodes, Greece

Athanassios Protopapas (aprotopapas@phs.uoa.gr)

Department of Philosophy and History of Science, University of Athens, Greece

Abstract

We implemented the connectionist model of social-pragmatic word learning (Caza & Knott, 2012) to test the hypothesis that reduced joint attention between infant and mother would increase the difference in acquisition between nouns and verbs as observed in Autistic Spectrum Disorder (ASD). The ratio of objects to actions in the observed event stream was manipulated to create an original noun-verb asymmetry. Ten simulations were run for each of the combinations of three conditions of communicative reliability and two conditions of unfiltered random associative learning, which is regarded by some researchers as the primary mechanism of language learning in ASD. The simulations indicated that the reduction in the reliability of communicative actions does not lead to increased noun-verb asymmetry within the originally planned training epochs. A trend in the predicted direction appeared toward the end of training, suggesting that further simulations may help resolve the issue within the current architecture.

Keywords: connectionist model; lexical development; social-pragmatic approach; joint attention; bootstrapping; ASD; SLI; noun-verb asymmetry; associative learning

Introduction

The pronounced rate of vocabulary acquisition during the second year of life constitutes a major milestone in cognitive development (Ganger & Brent, 2004). Laboratory research as well as cross-linguistic studies of natural vocabulary acquisition demonstrate that verbs often lag behind nouns (Gentner, 2006; Bornstein et al., 2004; Childers & Tomasello, 2006). The differentiated learning rate may be attributed to conceptual differences, as verbs are considered less stable and their meaning depends on nouns (Waxman et al., 2013). Alternatively, input distributions may account for the discrepancy, as nouns typically predominate in infant-directed speech (Sandhofer, Smith, & Luo, 2000). It is also proposed that other factors may be involved, such as shape, individuation, concreteness and imageability (Gentner & Boroditsky, 2001).

This pattern of differentiated acquisition is also observed in children with Autism Spectrum Disorder (ASD) and Specific Language Impairment (SLI). However, the categorical asymmetry is greatly exaggerated in ASD (Blach & Chiat, 2003; Marshall, 2003; Ellis Weismer et al., 2011).

The situation for SLI is less clear and there are relatively few studies comparing this population with children with ASD or typically developing (TD) children. There is some evidence to suggest that the frequency dependence and poor retention noted in children with SLI may particularly affect verb learning (Rice et al., 1994; Windfuhr et al., 2002). Weismer et al. (2011) found no significant differences between ASD and SLI in the number of action words, while other studies found significant differences between the two groups in the use of verbs describing cognitive and mental states (Ziatas et al., 1998).

In an attempt to provide a unified explanation of word learning, Hollich et al. (2000) proposed the Emergentist Coalition Model. This hybrid theory holds that children rely differentially on multiple cues over developmental time to map words onto referents – whether they are nouns or verbs. At first, infants are sensitive to perceptual cues, mapping a word to the referent that is most interesting or salient. They then use the social intent of a speaker, along with linguistic cues, to home in on word reference.

The ability to understand the intentions of others develops at the end of the first year and seems to constitute a crucial prerequisite for language development (Tomasello & Carpenter, 2007; Trevarthen, 1994). Consistent with this approach, over 50% of the variance in language production and comprehension can be attributed to factors such as the amount of time in joint attention between infant and mother (Carpenter et al., 1998). Deficits in joint attention have been implicated in ASD as predictors of “concurrent language ability” (Dawson et al., 2004). Moreover, recent studies suggest that language problems in SLI may be partially explained by observed deficits in joint attention (Papakaloudou & Papailiou, 2016). Nevertheless, different processes are involved in acquiring nouns and verbs in joint attention episodes. For example, unlike what is the case with nouns, children benefit from hearing verbs if a verb is heard either before or after an event, and do not show a similar benefit if a new verb and event occur simultaneously (Tomasello & Kruger, 1992; Childers et al., 2007).

Yu and Ballard (2007) modeled social-pragmatic word learning by exploiting social cues such as joint attention and the prosody of the maternal speech. Despite excluding

“theory of mind” skills and prior knowledge of the characteristics of the language, their model outperformed mere statistical associative learning. To model simultaneous learning of individual word meanings and inferring the speaker’s communicative intentions—the “chicken and egg” problem of child’s early vocabulary acquisition—Frank et al. (2009) posited a computational structure of variables representing the task of word-learning linked together by a set of probabilistic dependencies corresponding to task assumptions. This *intentional model* outperformed several cross-situational word learning models (including the one by Yu & Ballard, 2007) in both aspects of the vocabulary acquisition process.

More recently, a connectionist model of the *social-pragmatic approach to word learning* (Tomasello, 2000) was proposed by Caza and Knott (2012; Knott, 2014). In this comprehensive approach, two subsystems bootstrap each other: One subsystem learns to identify favorable learning opportunities, that is, communicative events that involve reliable mappings between concepts and words. The other subsystem learns to associate words with observed concepts. This model implements the intuition that although plain co-occurrence of word forms and objects may be too noisy to permit successful learning, the reliability of co-occurrences is far from uniform across time. Specifically, in a social-pragmatic framework, co-occurrences are likely to be more reliable when an intent to communicate can be identified. However, in this model events do not come initially flagged as conducive to word learning or not. Rather, the reliability of learning itself is used to tune the network to the appropriate types of events. Therefore, as the model gradually learns that certain events, namely communicative actions of the mother, present favorable learning opportunities, communicative actions and intentions are increasingly recognized and vocabulary

acquisition is accelerated, in a mutually beneficial and reinforcing circle. In this way, development of joint attention enables and underlies efficient word learning.

Method

The separation between object and action words in the model of Caza and Knott (2012) makes it suitable for investigating the differential learning of nouns and verbs. In the present study we were interested in the effects of impaired joint attention on the noun-verb asymmetry. We developed our version of the model on the Emergent platform (Aisa, Mingus, & O’Reilly, 2008) using the Leabra algorithm for training the connection weights, which is more biologically plausible than backpropagation (O’Reilly & Munakata, 2000). The architecture of the model, depicted in Figure 1, consists of (a) a “filter” subnetwork, (b) two word learning subnetworks working in parallel to associate pairs of perceived object and action concepts to concurrent phonological representations of a word, and (c) a reward system, which updates the connections of the filter depending on whether a concept-word mapping has been determined by the model to be successful or not. The model assumes that reliable perception and identification of objects, actions, and phonological word forms is in place prior to the modeled process of lexical learning, that is, of the associations between these types of representations.

In this model, an action, an object, and a word form are perceived at each time frame. The word-learning subnetworks predict word forms based on the observed action and object, generating error signals by comparison to the word form that is actually perceived. At the same time, the action-object pair constitute an observed *event*, which is used by the filter subnetwork to gate a learning signal governing learning in the two word learning subnetworks. In this way, the filter subnetwork controls the extent to which each event leads to modification of the connection weights in the word learning subnetworks. In other words, it modulates the extent to which the current action/object-word pairings are allowed to affect current word knowledge.

Conversely, successful word learning rewards the filter, leading to learning of events that reliably predict correct concept-word pairs. That is, when an action/object-word pairing is predicted correctly by the model, the event preceding it is tagged as informative by a modification of the corresponding weight in the filter subnetwork. In this way, events (i.e., action-object pairings) that are followed by correct word predictions gradually become more influential in word learning. The activation threshold of the gate is set to 0.55 for all the simulations described below. The coordinated actions of these subsystems lead the model to recognize maternal communicative actions as ideal candidates for successful concept-word mapping.

Updates of the weights in the filter subnetwork were gradually disengaged from randomly occurring associative learning. Departing somewhat from the original simulations of Caza and Knott (2012), who let the probability of unfiltered associative learning diminish to zero, we retained

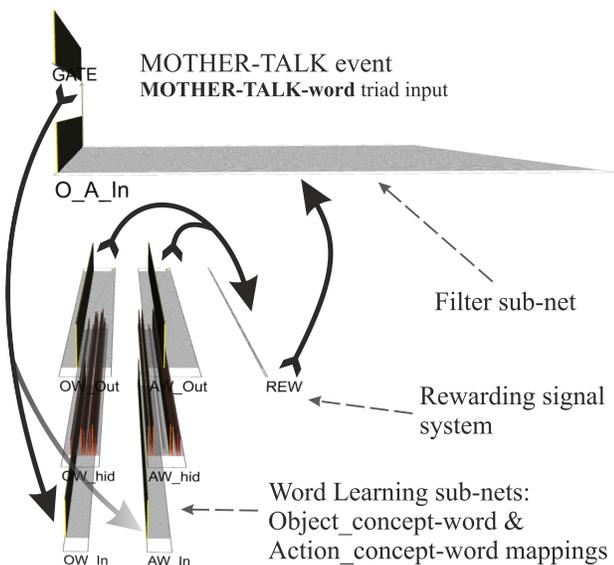


Figure 1: Network architecture of our version of the Caza & Knott (2012) model as implemented in Emergent.

it minimally operative throughout the simulations as a more plausible approach to normal and impaired word acquisition. This is consistent with the literature showing that typically developing children as well as children with ASD do not rely solely on intention monitoring and pragmatic cues for mapping words to objects but also on an associative infrastructure as well (Preissler, 2008; Preissler & Carey, 2005). This particular feature of the model also permitted us to examine the effect of residual plain associative learning. Two settings were used for this minimal *residual associative learning* (RAL), one with probability of 0.001 in each trial (low RAL) and one with probability of 0.005 (high RAL). These values were reached in a designated number of trials from an initial value of 0.5 set at the onset of training. This manipulation goes beyond the demonstrations of Caza and Knott (2012) that bootstrapping is effective in getting the filter subnetwork to efficiently gate relevant training signals to the word learning subnetworks, and was aimed at testing the extent to which unfiltered training, which would likely include a large proportion of incorrect concept-word pairings, might be supportive or disruptive of overall word learning.

In the context of ASD, excessive associative learning is considered by some researchers to constitute the primary foundation of language learning (Luyster & Lord, 2009), not gated by the detection of communicative intent. When combined with an overall reduced rate of communicative event detection, as appropriate for ASD, an abnormally high rate of associative learning—compared to the baseline provided by the simulations of Caza and Knott (2012)—would be expected to track more closely the rate of occurrence of particular word types in the input. Therefore, it might interact with word class frequency in a way that can provide useful clues for understanding lexical development

under such abnormal circumstances.

A mixture of training patterns constitutes the input stream. A baseline difference in noun-verb learning was modeled by setting the ratio of reliable object to action concepts to 4:3 in all simulations. Successful joint attention in this model amounts to correct identification of the action-object pair that is attended by the mother. In a communicative context, it is assumed for present purposes that the mother mostly speaks about the event that is currently in her attention. Following Caza and Knott (2012), a communicative action was represented as an event (object-action pair) composed of MOTHER (object) and TALK (action). Joint attention was modeled by the proportion of maternal communicative actions that were reliable, i.e., followed by a valid word-concept pair. Thus, successful joint attention naturally leads to more successful learning opportunities. Simulations tested the hypothesis that a decrease of communicative reliability would exaggerate the noun-verb asymmetry, approaching an ASD profile, rather than simply delay acquisition, resulting in an SLI profile.

There were three levels of joint attention, modeled as the ratio of reliable to nonreliable communicative events: high reliability (500:300), medium reliability (400:400), and low reliability (300:500). The latter condition amounts to inappropriate or unsuccessful detection of communicative intent. The two levels of RAL (low and high) were modeled as the probability of a currently observed action/object-word pairing to affect the word learning subnetworks regardless of the filter subnetwork gate. Ten simulations of 120 epochs were conducted for each combination of joint attention and RAL. Modules were created within Emergent to produce training sequences with the appropriate constraints, and to track the progress of training and network performance. An *endogenous* criterion of learning evaluated word prediction

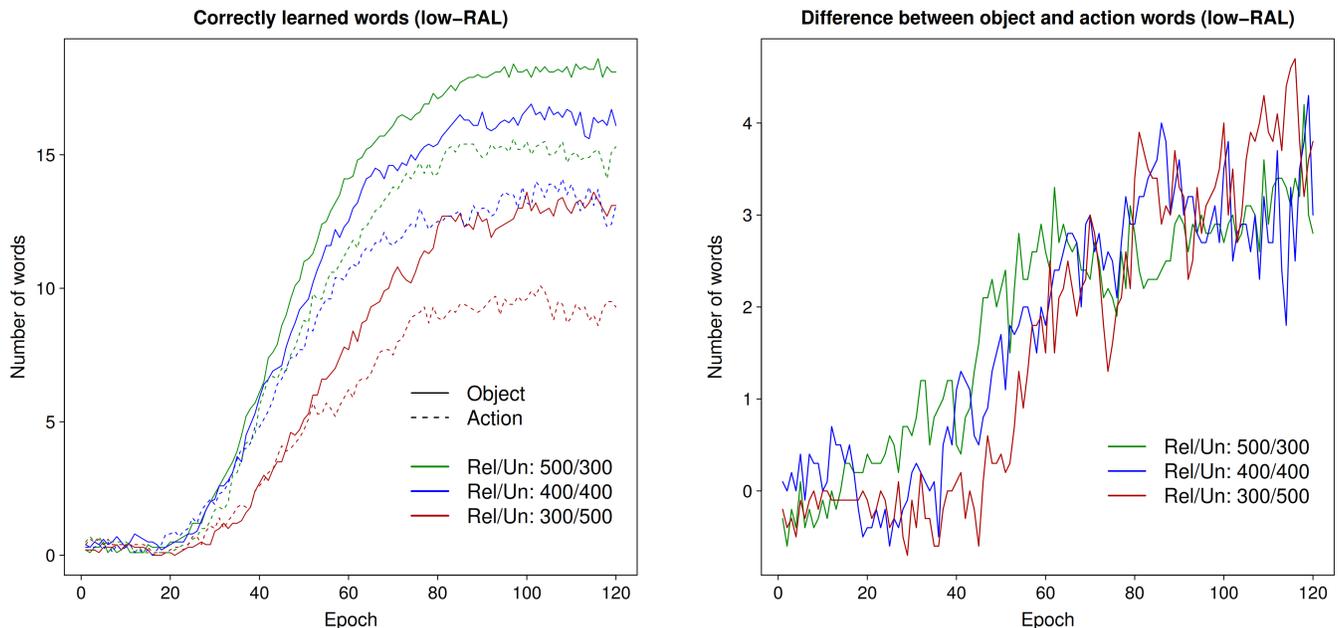


Figure 2: Number of correctly learned nouns and verbs (left) and difference between the number of correctly learned nouns and verbs (right) at the end of each epoch of training in the low-RAL condition.

success in every trial, causing a *reward* or *punishment* signal to the gating component regarding the previous event.

Results and Discussion

Results are displayed in Figures 2 (for low RAL) and 3 (for high RAL). In both RAL conditions the decrease in the reliability of maternal communicative events did not markedly increase the difference between nouns and verbs. Therefore, failure of joint attention did not produce an ASD profile. A tendency toward increased divergence between object and action words at the end of the training was noted but the situation is far from clear within the implemented training epochs. As this pattern of slightly differentiated asymmetry occurred far from ceiling, it invites further study adopting a more protracted training scheme.

Another effect seen mainly in the high-RAL condition (Figure 3) was an increase of the noun-verb asymmetry in the high reliability condition around the middle of training. It unclear whether this is amenable to a developmental interpretation or due to a ceiling effect. Additionally, in both RAL conditions a symmetry in learning action and object words was noted in the initial one-third of training epochs in the low and medium reliability conditions, despite their differential frequency of encounter in the training trials, suggesting perhaps that the associative learning mechanism, mostly activated in the beginning of training, made up for the frequency difference. Note that in the entire set of 60 simulations a low baseline gating threshold was adopted, corresponding to the minimum gate unit activation required for gating the word learning subnetworks. This led to a large number of events other than MOTHER-TALK activating the word learning system throughout training, essentially constituting a noise-adding mechanism, possibly obscuring

the model behavior. This choice (and other parameters) was found necessary for the bootstrapping process to kick in.

Under low-RAL, gating filter efficiency reached 100% at a mean number of 147k trials ($SD = 23k$) in the 300:500 condition, at 113k trials ($SD = 19k$) in the 400:400 condition, and at 114k trials ($SD = 20k$) in the 500:300 condition. Under high-RAL, 100% filter efficiency was reached at 130k ($SD = 21k$), 121k ($SD = 26k$), and 120k ($SD = 18k$) trials in the three conditions, respectively. Despite the high variability, the model seems to struggle more to overcome the low reliability burden (300:500 ratio) in the low-RAL condition compared to high-RAL, suggesting a contribution from the residual associative mechanism.

To further study the full extent of the model's range of behaviors, in a second set of simulations a (less plausible) *exogenous* criterion of learning was implemented, in which the reward signal was hardcoded into the training sequences, to closely track actual communicative reliability rather than correct word prediction. The resulting difference between object and action word learning in the two RAL conditions is displayed in Figure 4. Again, failure of joint attention did not produce an ASD profile, leading instead to a protracted rate of acquisition, with decreased difference between nouns and verbs, consistent with SLI. A possible exception may be noted under high RAL past the initial one-third of training.

Future Work

Following up on these observations, further simulations (currently in progress) can help us shed light on the role of social-pragmatic factors in learning word meanings. In addition to thorough exploration of the dynamic interplay between the ratio of communicative event types, the residual amount of associative learning, and the gating activation

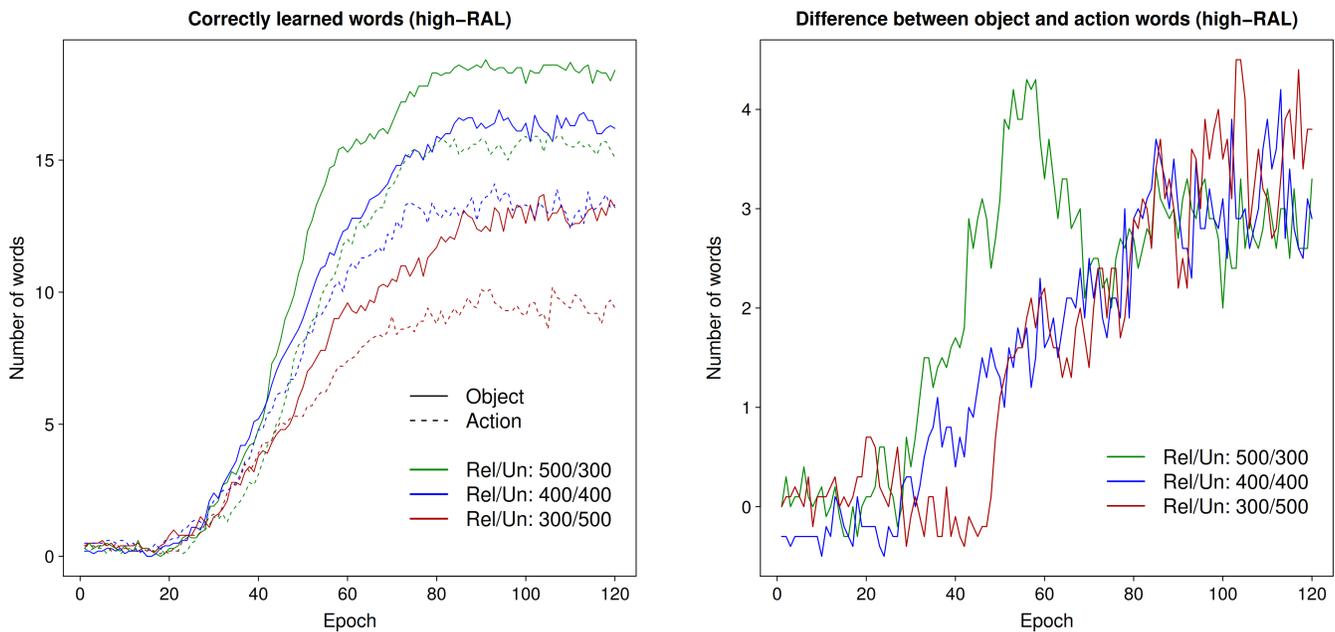


Figure 3: Number of correctly learned nouns and verbs (left) and difference between the number of correctly learned nouns and verbs (right) at the end of each epoch of training in the high-RAL condition.

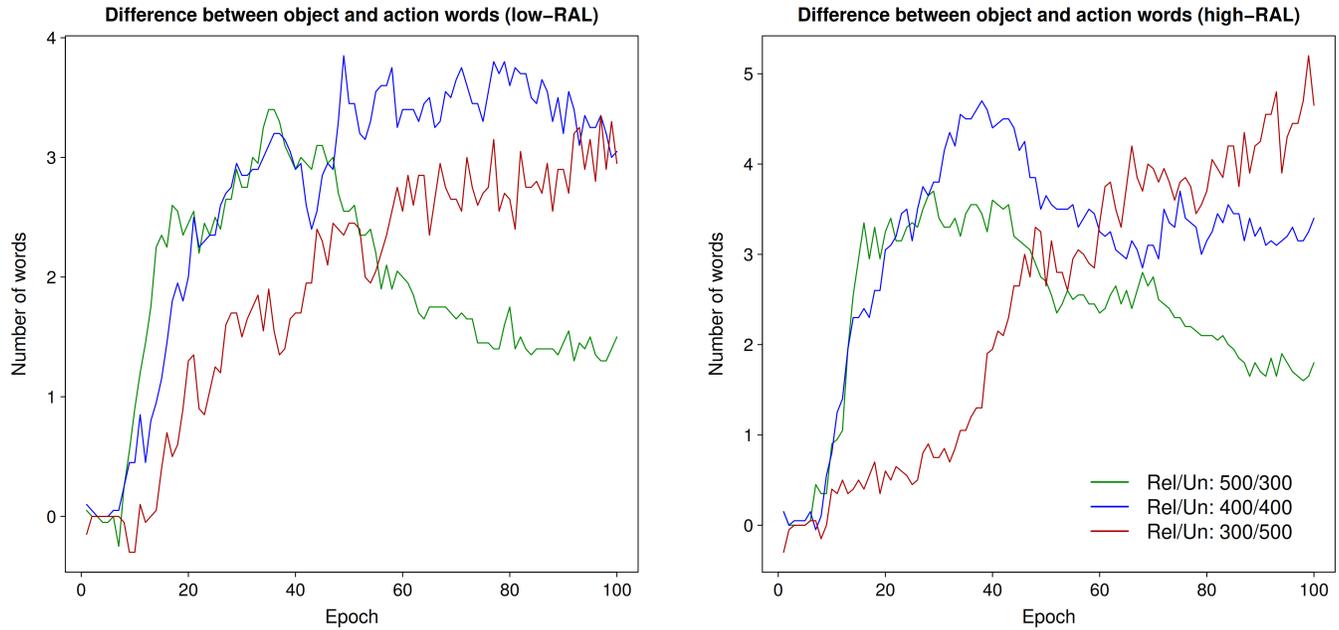


Figure 4: Difference between the number of correctly learned nouns and verbs at the end of each epoch of training in the low-RAL condition (left) and in the high-RAL condition (right) for the *exogenous* rewarding approach.

threshold, future simulations will address the role of specific network parameters on the function of the gating network and the resulting word learning efficiency. A parameter of interest concerns the amount of time for the associative learning mechanism to reach its designated residual value. In the simulations described above this was set to 5000 trials. However, a slower rate of decrease—accompanied by a higher baseline gating threshold—seems promising in the context of ASD (see Figure 5), consistent with the presumed abnormally increased importance of associative learning and reduced communicative reliability in this population.

Moreover, in a future revision of the model, the number of nouns and verbs could be increased substantially, to permit observation of their relative learning rate free from limiting ceiling effects. The small number in the current model may obscure patterns of differential acquisition, as the possibility for further improvement is rapidly curtailed.

In future work we also plan to incorporate additional types of communicative events into the training pattern sequences, as well as carefully studied chains of events abiding to specific constraints, in order to investigate the prospects of certain methods of interventions for impaired language learning, such as parental “synchronous undemanding talking” about the child’s focus of attention (Yoder & McDuffie, 2006) and their interaction with the other modeled language acquisition parameters. Further follow-up work can enrich the lexical learning network with conceptual learning components in order to examine the combined effects of category learning and social interaction impairments typical of ASD.

Acknowledgments

This research has been co-financed by the European Union (European Social Fund) and Greek national funds through the Operational Program “Education and Lifelong

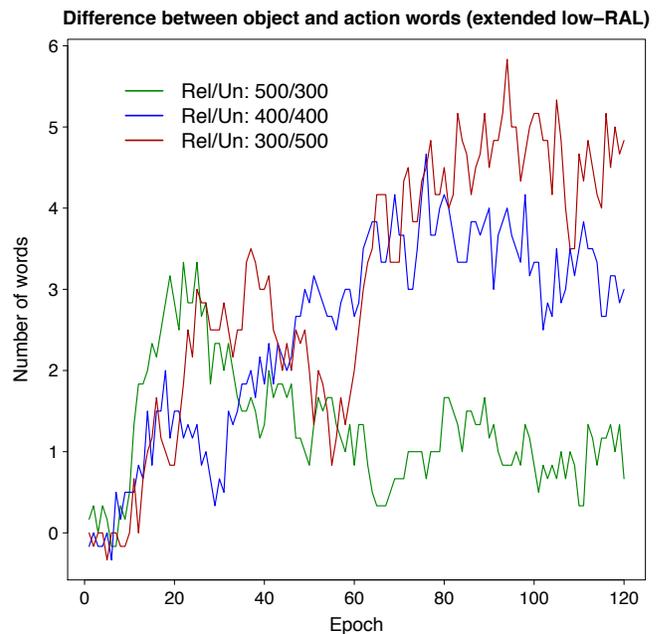


Figure 5: Difference between the number of correctly learned nouns and verbs at the end of each epoch of training in the low-RAL condition with a higher gating threshold and the time for RAL to reach its final value extended to 50000 trials.

Learning” of the National Strategic Reference Framework – Research Funding Programs: THALIS–UOA–COGMEK–375737 and ARISTEIA–BabyAffect–GSRT–3610.

References

- Aisa, B., Mingus, B., & O’Reilly, R. (2008). The emergent neural modeling system. *Neural Networks*, *21*, 1146–1152
- Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S.-Y., Pascual, L., ... Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: spanish, dutch, French, hebrew, italian, korean, and american english. *Child Development*, *75*(4), 1115–1139.
- Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *i*–174.
- Caza, G. A., & Knott, A. (2012). Pragmatic bootstrapping: A neural network model of vocabulary acquisition. *Language Learning and Development*, *8*(2), 113–135.
- Childers, J. B., & Tomasello, M. (2006). Are Nouns Easier to Learn Than Verbs? Three Experimental Studies. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 311–335). New York, NY, US: Oxford University Press.
- Childers, J. B., Vaughan, J., & Burquest, D. A. (2007). Joint attention and word learning in Ngas-speaking toddlers in Nigeria. *Journal of Child Language*, *34*(2), 199–225.
- Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., ... & Liaw, J. (2004). Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Developmental Psychology*, *40*(2), 271.
- Frank, M. C., Goodman, N. D., & Tenenbaum, J. B. (2009). Using speakers’ referential intentions to model early cross-situational word learning. *Psychological Science*, *20*(5), 578–585.
- Ganger, J., & Brent, M. R. (2004). Reexamining the vocabulary spurt. *Developmental Psychology*, *40*(4), 621.
- Gentner, D. (2006). Why Verbs Are Hard to Learn. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 544–564). New York, NY, US: Oxford University Press.
- Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development*. Cambridge University Press.
- Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., ... Rocroi, C. (2000). Breaking the language barrier: an emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, *65*, 1–123
- Knott, A. (2014). How infants learn word meanings and propositional attitudes: a neural network model. In *Communicative Action* (pp. 107–124). Springer.
- Luyster, R., & Lord, C. (2009). Word learning in children with autism spectrum disorders. *Developmental Psychology*, *45*(6), 1774.
- O’Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. MIT press.
- Papakalodouka, P., & Papaeliou, C. F. (2016). The Theory of Intersubjectivity may explain aspects of Specific Language Impairment. *Eleftherna*, *7*. University of Crete.
- Preissler, M. A. (2008). Associative learning of pictures and words by low-functioning children with autism. *Autism: The International Journal of Research and Practice*, *12*(3), 231–248.
- Preissler, M. A., & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition*, *97*(1), B13–B23.
- Rice, M. L., Oetting, J. B., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech and Hearing Research*, *37*(1), 106–122.
- Sandhofer, C. M., Smith, L. B., & Luo, J. (2000). Counting nouns and verbs in the input: differential frequencies, different kinds of learning? *Journal of Child Language*, *27*(03), 561–585.
- Tomasello, M. (2000). The social-pragmatic theory of word learning. *Pragmatics*, *10*(4), 401–414.
- Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science*, *10*(1), 121–125.
- Tomasello, M., & Kruger, A. C. (1992). Joint attention on actions: acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language*, *19*(2), 311–333.
- Trevarthen, C. (1994). Infant semiosis. In W. Nöth (Ed.), *Origins of Semiosis: Sign Evolution in Nature and Culture*. Berlin: Walter de Gruyter.
- Waxman, S., Fu, X., Arunachalam, S., Leddon, E., Geraghty, K., & Song, H. (2013). Are nouns learned before verbs? Infants provide insight into a long-standing debate. *Child Development Perspectives*, *7*(3), 155–159.
- Weismer, S. E., Gernsbacher, M. A., Stronach, S., Karasinski, C., Eernisse, E. R., Venker, C. E., & Sindberg, H. (2011). Lexical and grammatical skills in toddlers on the autism spectrum compared to late talking toddlers. *Journal of Autism and Developmental Disorders*, *41*(8), 1065–1075.
- Windfuhr, K., Faragher, B., & Conti-Ramsden, G. (2002). Lexical learning skills in young children with specific language impairment (SLI). *International Journal of Language & Communication Disorders*, *37*(4), 415–432.
- Yoder, P., & McDuffie, A. (2006). Teaching young children with autism to talk. In *Seminars in speech and language* (Vol. 27, p. 161). Thieme Medical Publishers, Inc.
- Yu, C., & Ballard, D. H. (2007). A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing*, *70*(13), 2149–2165.
- Ziatas, K., Durkin, K., & Pratt, C. (1998). Belief term development in children with autism, Asperger syndrome, specific language impairment, and normal development: Links to theory of mind development. *Journal of Child Psychology and Psychiatry*, *39*(05), 755–763.