Developmental Continuity in Neural Semantic Representations

When we learn a word, what exactly have we learned, and how is this information represented and stored in the mind and brain? These questions have been at the forefront of cognitive science for centuries, engaging philosophers, linguists, lexicographers and (more recently) psychologists, computer scientists, and neuroscientists. The advent of powerful machine learning methods for interpreting brain data promises to revolutionize our search for the answers. In principle, we could use these methods to understand how word knowledge is represented and organized in the brain – both in maturity as well as in early stages of development.

However, there are two major limitations on work to date. First, functional neuroimaging with children is challenging, given children's limited attention span and tendency to fidget. Second, most work to date has taken a data-driven approach that neither speaks to -- nor makes use of -- the rich theoretical and computational tradition of models of linguistic semantic knowledge. Thus, we present the first in a series of studies addressing these issues, making use of a promising new method for recording EEG in children ("free listening"; Brennan, et al. 2019; Levari & Snedeker, 2018) and applying existing, theory-driven, quantitative representations of word meaning.

27 children (ages 5-10, mean=7.5) and 21 adults listened to a 1594-word recorded excerpt from *Matilda*, by Roald Dahl (2003). While free listening sacrifices some experimental control, it has distinct advantages: it is more naturalistic, more engaging for the subjects, and results in large amounts of 'trials' (i.e., individual words) in a relatively short amount of time. Critically, studies using free listening replicate well-established effects from the ERP literature (Alday et al., 2017).

EEG was recorded at 500hz using Brainvision's Actichamp System with 32 active electrodes placed at International 10-20 System locations and on the left and right mastoids, and filtered to 0.1-40 Hz. Data were epoched from -200ms to 1000ms relative to onset of each critical word and baseline corrected using the pre-stimulus time window (-200–0ms). Eye artifacts were removed through independent component analysis. Trials were discarded if they contained artifacts greater than 90μV.

Analyses focused on the 107 distinct verbs and 138 distinct nouns in the excerpt. For types with multiple tokens, we averaged ERPs over tokens. Data were analyzed using Representational Similarity Analysis (RSA; Fig. 1; Kriegeskorte et al., 2008). RSA is essentially a correlation of correlations, measuring the degree to which stimuli that are similar under one metric are similar under another. Initial analyses established reliability: separate RSAs for verbs and nouns were significant and well above chance for both children and adults (Tab. 1). Interestingly, RSAs comparing children to adults were nearly as strong as RSAs comparing children to other children or adults to other adults, suggesting little developmental change.

We first replicated prior findings of systematic relationships between ERPs and neural network word embeddings such as fastText, which capture substantial amounts of lexical semantics and morphosyntax (Bojanowski et al., 2017; He, et al. 2022). RSAs were significant for both nouns and verbs for both children and adults, though only between ⅓ and ½ the size of the between-subject RSAs, suggesting word embeddings capture only some of the systematic variability in ERPs (Tab. 1).

It is not yet clear exactly how word embeddings represent semantics or how they map onto more formal theories. Critically, RSAs were nearly as strong when using pairwise distances in WordNet (a hierarchical ontology of meaning) or human pairwise similarity judgments (Small World of Words; De Deyne et al., 2019) (Tab. 2). However, we found no relationship between ERPs for verbs and pairwise similarity in participation in distinct argument structures (as recorded in VerbNet; Kipper et al., 2006), despite prior evidence that the latter is highly correlated with semantics. **Strikingly, we see little evidence of developmental change.** We discuss these results in terms of theories of word meaning and language learning, and prospects for using this method to tie specific neural mechanisms to formal theories of language.

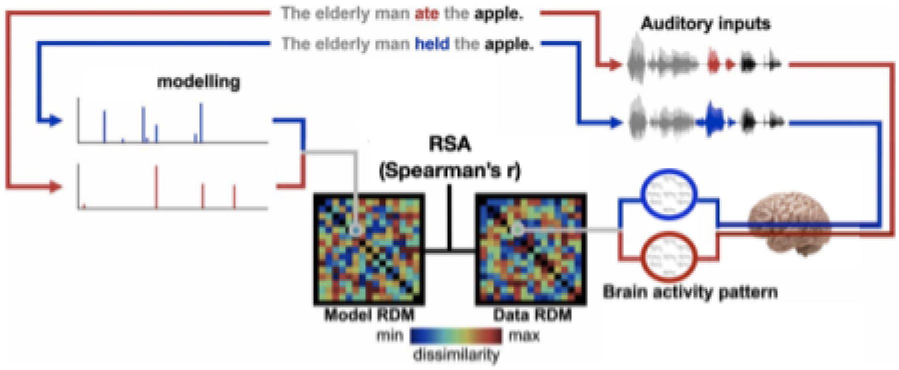

Fig. 1. Schematic of Representational Similarity Analysis (RSA). In this example, for each subject, we calculate the pairwise similarity in ERPs for each stimulus. We also calculate the pairwise similarity for each stimulus in terms of some semantic model (e.g., WordNet). We then correlate these pairwise similarities to determine whether words that evoke similar ERPs are also similarly represented in the model.

Matilda Dataset: Verbs

	adults	children	fastText
adults	.34 (.0004)		
children	.33 (.06)	.37 (.0002)	
fastText	.16 (.04)	.14 (.05)	1 (0)

Matilda Dataset: Nouns

	adults	children	fastText
adults	.26 (0.0002)		
children	.24 (.05)	.28 (.0002)	
fastText	.08 (.03)	.08 (.02)	1 (0)

Table 1. Results of reliability analyses. Adult-Adult and Child-Child correlations represent a noise ceiling on further analyses. All results are significant.

Matilda Dataset: Verbs

	WordNet	VerbNet	SWoW
adults	.14 (.05)	.002 (.036)	.07 (.02)
children	.11 (.06)	-.01 (.04)	.07 (.02)
WordNet	1 (0)	.17 (0)	.12 (0)
VerbNet	.17 (0)	1 (0)	.02 (0)

Matilda Dataset: Nouns

	WordNet	SWoW
adults	.08 (.03)	.06 (.02)
children	.09 (.03)	.06 (.02)
WordNet	1 (0)	.11 (0)

Table 2. RSA results comparing adult and child neural responses with WordNet, VerbNet, and Small World of Words similarities. Note that VerbNet cannot be applied to nouns. All results are significant except RSAs comparing adults or children to VerbNet.

Bibliography

Alday, P. M., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2017). Electrophysiology reveals the neural dynamics of naturalistic auditory language processing: event-related potentials reflect continuous model updates. eneuro, 4(6).
Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. Transactions of the association for computational linguistics, 5, 135-146. **Dahl, R. (2003).** Matilda. Penguin UK. **De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019)**. The “Small World of Words” English word association norms for over 12,000 cue words. Behavior research methods, 51(3), 987-1006. **He, T., Boudewyn, M. A., Kiat, J. E., Sagae, K., & Luck, S. J. (2022).** Neural correlates of word representation vectors in natural language processing models: Evidence from representational similarity analysis of event-related brain potentials. Psychophysiology, 59(3), e13976. **Kipper, K., Korhonen, A., Ryant, N., & Palmer, M. (2006, May)**. Extending VerbNet with novel verb classes. In Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC’06). **Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008).** Representational similarity analysis-connecting the branches of systems neuroscience. Frontiers in systems neuroscience, 2, 4. **Levari, T & Snedeker, J. (2018).** Children’s N400 is sensitive to both predictability and frequency: Evidence from natural listening. BUCLD 43, Boston, MA.