



Developmental Continuity in Neural Semantic Representations

Tanya Levari¹, Stefano Anzellotti², Aidas Aglinskas², Albert Kim³, Emily Prud'hommeaux², Joshua Hartshorne²

¹Harvard University, ²Boston College, ³University of Colorado Boulder



BACKGROUND

When we learn a word, what exactly have we learned, and how is this information represented and stored in the mind and brain?

Powerful machine learning methods for interpreting brain data may allow us to understand how word knowledge is represented and organized in the brain across development.

Previous Limitations:

- 1) Functional neuroimaging with children is challenging.
- 2) Data-driven methods overlook rich theoretical models of semantic knowledge.

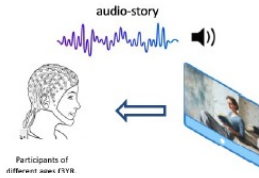
We present a study addressing these issues using a new method for recording continuous EEG in children ("free listening"; Brennan, et al. 2019; Levari & Snedeker, 2018) and applying existing, theory-driven, quantitative representations of word meaning.

METHODS

ERP Recordings

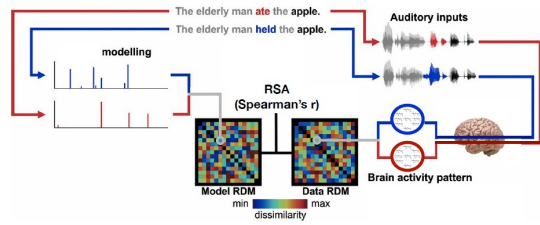
27 children (ages 5-10, mean=7.5), 21 adults

Naturalistic Listening Paradigm: Participants listen to a 1594-word excerpt from *Matilda*, by Roald Dahl (2003) while continuous EEG is recorded



Analyses include ERP responses to **107 distinct verbs** and **138 distinct nouns**. For types with multiple tokens, we averaged ERPs over tokens.

Data were analyzed using Representational Similarity Analysis

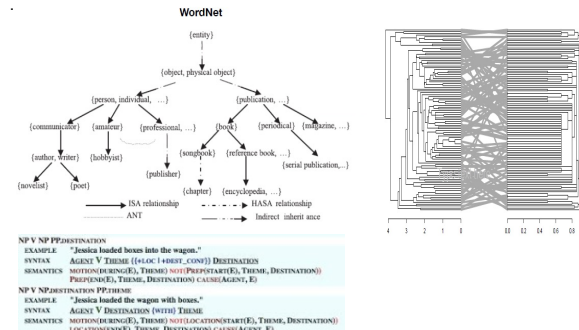


Models of Words Meaning

WordNet: An ontological model of word meaning which represents words hierarchically by their superordinate and subordinate categories. RSA calculated over path similarity (Miller, 1995)

VerbNet: Database of verb argument structure behavior. RSA calculated over feature vectors for each verb (Schuler, 2005)

Small World of Words (SWoW): a crowd-sourced dataset of the top 300 word associations for each of 12,292 English words (De Deyne et al, 2019).



RESULTS

Analyses of Reliability

We first evaluated the relationships between adult and child ERPs and replicated systematic relationships between ERPs and neural network word embeddings such as fastText (Bojanowski et al., 2017; He, et al. 2022)

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)

RSAs were significant for both nouns and verbs for both children and adults, though only between 1/4 and 1/2 the size of the between-subject RSAs, suggesting word embeddings capture only some of the systematic variability in ERPs.

Analyses of Semantic Models

RSA analyses comparing ERP responses in children and adults to corresponding relationships between verbs in WordNet, VerbNet, and SWoW, and nouns in WordNet and SWoW.

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)

DISCUSSION

ERPs capture information about semantic relationships across words.

We replicated prior findings of systematic relationships between ERPs and neural network word embeddings, as captured by fastText. Critically, RSAs were nearly as strong when using pairwise distances in WordNet and human pairwise similarity judgments, like Small World of Words.

No relationship between ERPs and verb argument structure participation.

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. Note: could reflect limitations in EEG analyses or VerbNet.

Little evidence of developmental change.

No differences in patterns between children and adults. RSAs comparing children to adults were nearly as strong as RSAs comparing children to other children or adults to other adults

References

Fellbaum, C. (2010). WordNet. In *Theory and applications of ontology: computer applications* (pp. 231-243). Springer, Dordrecht.

He, T., Boudewyn, M. A., Klat, 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.

Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM*, 38(11), 39-41.

Schuler, K. K. (2005). *VerbNet: A broad-coverage, comprehensive verb lexicon*. University of Pennsylvania.

Levari, T. & Snedeker, J. (2019). Children's N400 is sensitive to both predictability and frequency: Evidence from natural listening. *BUCLD 43*, Boston, MA.

Brennan, J. R., & Hale, J. T. (2019). Hierarchical structure guides rapid linguistic predictions during naturalistic listening. *PLoS one*, 14(1), e0207741.

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.

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.

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)*.