The relative benefits of ensembles of words on early word recognition for learners at different levels of reading skill

Matthew J. Cooper Borkenhagen[1,2] & Devin M. Kearns[3,4]

[1]School of Teacher Education, Florida State University [2]The Florida Center for Reading Research
[3]Department of Educational Psychology, The University of Connecticut [4]Department of Teacher Education & Learning Sciences, North Carolina State University

FLORIDA CENTER *for* READING RESEARCH

Background

Early word recognition develops in large part as a function of the experiences learners have early on with printed words.

Measuring the relative impacts of different ensembles of printed words is difficult due to the temporally diffuse nature of early print experiences. They happen gradually and over many learning trials. Instruction accelerates this process (Foorman et al., 1991) as can words selected based on their properties (Apfelbaum et al, 2013), but learning can only be partially attributed to instructional experiences.

Learners of different levels of skill likely benefit from ensembles of words that contain particular properties of print and speech (see Compton et al., 2014 for discussion), though these effects have not been proved at scale.

The scientific community takes as self-evident that differences in words contribute meaningfully to learning outcomes from print exposure, though the nature of these effects is not established.

Questions

Do different ensembles of words learned during early reading development contribute to different word reading outcomes?

Are different ensembles more or less effective for learning at different levels of reading skill?

Do specific words vary in performance as a function of experience and is the effect modulated across levels of reading skill?

Methods

Computational models that map print to speech were used, like those used in other simulation work (Cox et al., 2018; Seidenberg & McClelland, 1989; Harm & Seidenberg, 1999).

The words selected for training and the representational capacity of each model were both manipulated. 10K models in each of four capacity conditions (ranging from 10 to 100 hidden units) were used ($N = 40{,}000$). Models were matched across levels of hidden unit such that the same learning environment appeared in each level.

Models learned 300 words across 50 learning trials (epochs) and were all tested on the same independent set of 300 holdout words in order to examine generalization. Monosyllabic words were collected from popular children's books, and learning was weighted by frequency.

Results

Figure 1

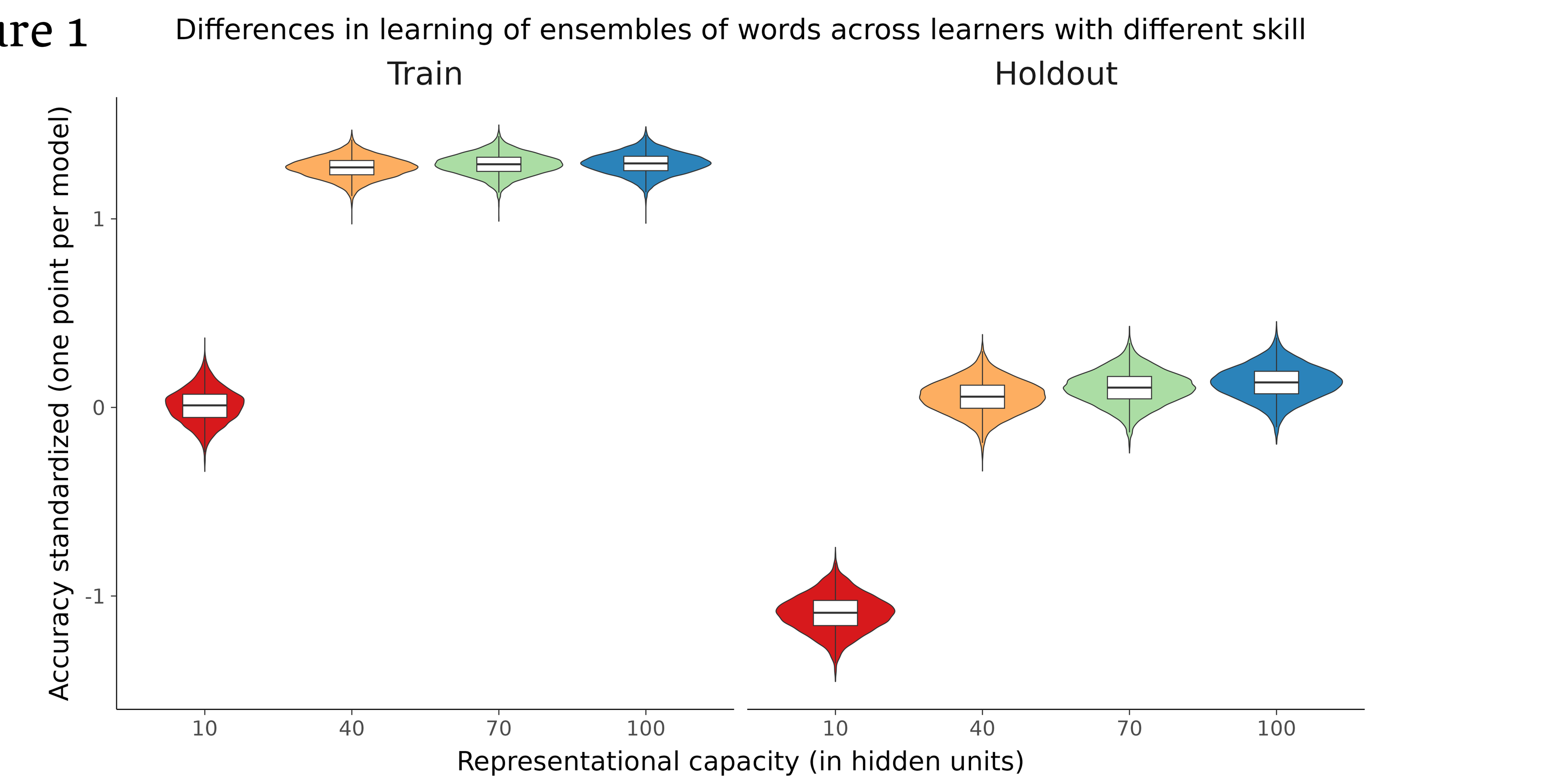

Figure 1: Differences in word reading ability (average accuracy for all words) in the training and holdout sets are shown for models of 10, 40, 70, and 100 hidden units at the end of learning. Models with fewer hidden units are less capable of learning relationships between letters and sounds. Average accuracy and its spread vary across levels of hidden units in both train and holdout conditions. Greater variation in learning the training ensemble exists for models with lower representational capacity (10 hidden unit condition).

Figure 2

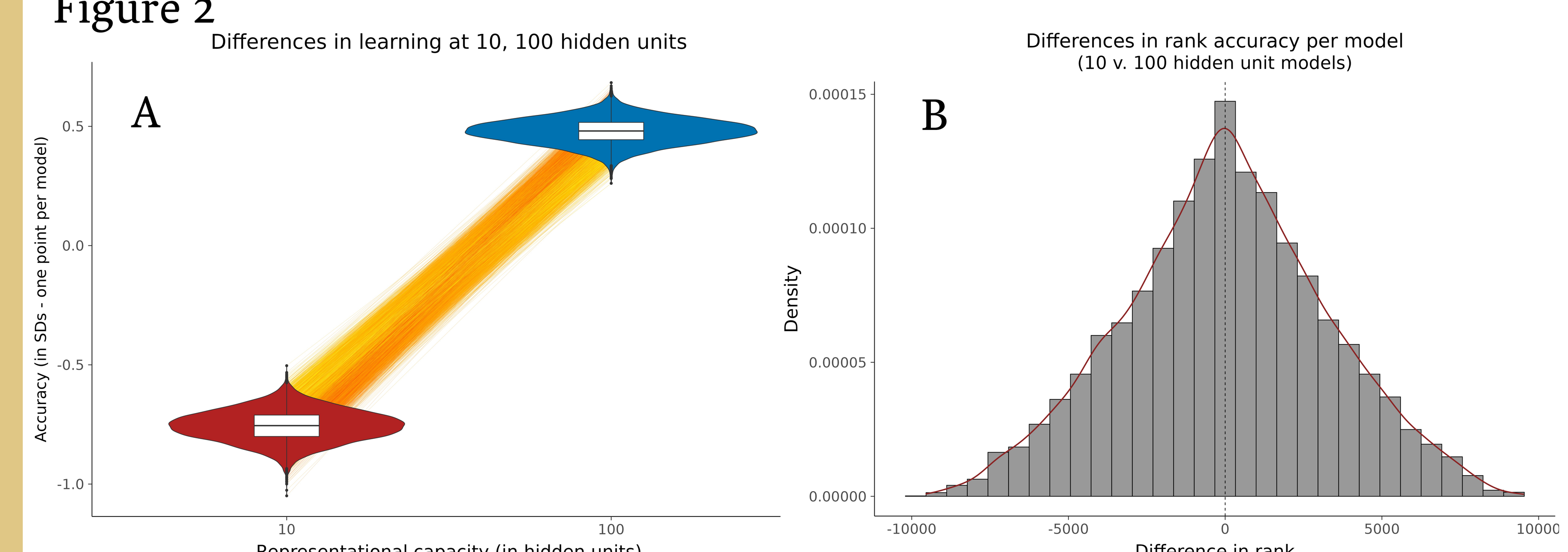

Figure 2: (A) Differences in word reading accuracy in 10 and 100 hidden unit groups are shown, with accuracy aggregated across training and test items. Lines connect models that are paired in terms of their learning environment (i.e., only the number of hidden units varies). Ranked achievement in one distribution does not necessarily carry over to the model in the other. (B) The distribution of differences in ranking among all models in terms of word reading accuracy. The difference value is calculated as the rank for a given model in its 10 hidden unit version minus the same model's rank in 100 units. The extreme values (which is limited to ±10K) represent the greatest possible change in rank performance for a model across the 10 and 100 hidden unit conditions.

Figure 3

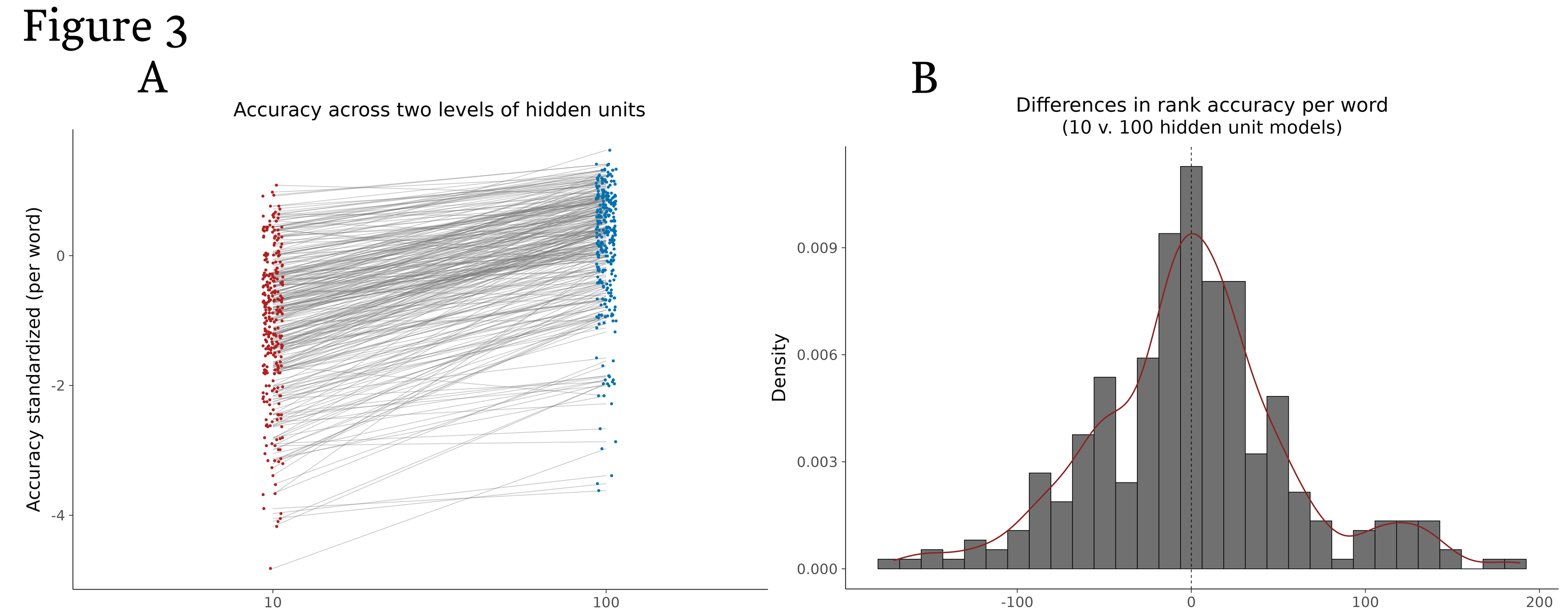

Figure 3: Analogous data to Fig. 2 except averaged with respect to words and done for a subsample of 40 models across all training epochs in 10 and 100 hidden unit conditions (300 holdout words for 80 models across 50 epochs = 1,200,000 observations). (A) Difference in aggregate accuracy of each word is shown. Lines connect identical words in each distribution. (B) The distribution of differences in ranking for each word, similar to Fig. 2b. For example, a ranking of -50 indicates that a word dropped in its ranking on accuracy across models in the 100 hidden unit condition relative to the 10 hidden unit condition.

Discussion

Dramatic differences in performance exist in learning within and across levels of representational capacity in the models studied (Fig. 1). For example, the difference between the poorest performer and highest performer among the models with 10 hidden units (left facet) in Figure 1 is .71 of an SD, and the difference between the overall poorest performing and best performing model (across all conditions) is 2.98 SDs in average word reading accuracy.

Models with impoverished phonology are associated with greater differences in the effect of word ensemble (100 hidden unit group has an absolute difference of .51 of an SD as compared to .71 for the 10 hidden unit condition).

The effects of word ensemble vary as a function of the representational capacity of the learner; a set of words that are good for one learner are not necessarily good for that same learner if they were not impaired in their ability to represent connections between print and speech (Fig. 2a-b).

Individual words vary in their learning impact across different levels of learner ability. A word that achieves a high level of overall accuracy may decrease substantially for a learner with more cognitive resources (and vice versa; Fig. 3a-b).

Conclusions

These simulations represent prima facie evidence that ensembles of words have dramatic impact in learning outcomes, and that what is good for a learner of one level of skill (e.g., a child with poor phonology) may be quite different for a learner at a different level of skill.

This suggests that care should be put into identifying words for learning in early reading development.

References

Compton, D. L., Miller, A. C., Elleman, A. M., & Steacy, L. M. (2014). Have we forsaken reading theory in the name of "quick fix" interventions for children with reading disability?. *Scientific Studies of Reading, 18*(1), 55-73.

Cox, C. R., Borkenhagen, M. C., & Seidenberg, M. S. (2019, July). Efficiency of Learning in Experience-Limited Domains: Generalization Beyond the Wug Test. In *CogSci* (Vol. 2019, pp. 1566-1571).

Foorman, B. R., Francis, D. J., Novy, D. M., & Liberman, D. (1991). How letter-sound instruction mediates progress in first-grade reading and spelling. *Journal of Educational Psychology, 83*(4), 456-469.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychological Review, 106*(3), 491-528.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological review, 96*(4), 523-568.

Apfelbaum, K. S., Hazeltine, E., & McMurray, B. (2013). Statistical learning in reading: variability in irrelevant letters helps children learn phonics skills. *Developmental Psychology, 49*(7), 1348.

Acknowledgements

MCB was supported by the College of Education, Health, and Human Sciences at Florida State University and the Florida Center for Reading Research, as well as an award from the Council for Research and Creativity at FSU.

FLORIDA STATE UNIVERSITY