Compositionality in Chinese Characters: Evidence from English-speaking Children

Min Tang (min_tang@berkeley.edu)
Department of Psychology, University of California, Berkeley
Berkeley, CA 94720 USA

Rongzhi Liu (rongzhi@uchicago.edu)
Psychology Department, University of Chicago
Chicago, IL 60657 USA

Fei Xu (fei_xu@berkeley.edu)
Department of Psychology, University of California, Berkeley
Berkeley, CA 94720 USA

Abstract

Compositionality is a core property of language: the meaning of sentences is derived from the meanings of individual words and rules for combining their meanings (Partee, 1984). Human adults have been shown to make compositional generalizations across many domains such as language, visual concept learning, and sequence learning. Few studies have investigated conceptual compositionality in young children. In two experiments with English-speaking 5- to 8-year-old children who have not been exposed to Chinese characters, we found that after a brief training session, they were able to generalize the newly learned radical-meaning pairs to new characters compositionally. Our results suggest that by age 5, children can make meaningfully compositional generalizations.

Keywords: compositionality, cognitive development, Chinese characters

Introduction

One of the fundamental questions in cognitive science is to understand the nature of conceptual representations and linguistic representations. Philosophers, psychologists, linguists, and researchers in machine learning and computational modeling have discussed the issue of *compositionality* extensively over the years (e.g., Fodor & Pylyshyn, 1988; Partee, 1984; Lake & Baroni, 2023; Piantadosi & Jacobs, 2016; Piantadosi & Aslin, 2016; Zhou, Feinman & Lake, 2024). Compositionality is a core property of language, that is, the meaning of sentences is derived from the meanings of individual words and rules for combining their meanings. Furthermore, according to the language of thought hypothesis (Fodor, 1975; Piantadosi & Jacobs, 2016; Quilty-Dunn, Porot & Mandelbaum, 2023), our conceptual representations that underlie our linguistic competence have the same format – concepts can be combined to form more complex conceptual representations. Even more generally, the reuse and recombination of preexisting concepts and knowledge appears to be a benchmark of human intelligence, and we see evidence for this across many domains.

In language, adult learners can understand an infinite number of new combinations of words into novel sentences (e.g., Colorless green ideas sleep furiously; the robotic barista served him a stiff drink). Preschool age children understand the compositional rules of adjective-noun combinations (e.g., Barner & Snedeker, 2008; Hamburger & Crain, 1984). For adults, several recent studies have investigated whether our non-linguistic, conceptual representations also follow the rules of compositionality. For example, Amalric et al. (2017) provided evidence that adult learners can utilize a 'geometric language' to predict sequences of shapes; Zhou, Feinman, and Lake (2024) showed that our visual concepts can be combined to make meaningful compositional generalizations.

What is the developmental origin of conceptual compositionality? Can young children make meaningful compositional generalizations in domains besides language?

To our knowledge very few studies have examined this question. Piantadosi & Aslin (2016) found that 3.5- to 4.5-year-olds can predictively compose two functions. In their task, children had to predict, in a forced choice design, what a car would look like after it went behind a screen (i.e., function). The screen changed the car in different ways according to the images on the screen. For example, if the screen had stars on it, it would put stars on the car, or if the screen was red, the car would turn red. In the critical trials, children had to combine two screens (e.g., a screen with stars and a red screen) to accurately predict what the car would like after it passed both screens (e.g., a car with red stars). 3.5-4.5-year-old children succeeded at predicting the outcome of the car. Using a similar design, Piantadosi et al. (2018) found that 9-month-old infants learned the individual functions but had difficulties composing them in a looking time study.

One limitation of this study is that since the order of applying the individual functions does not matter, it leaves open the question that preschoolers simply added both features to the object (i.e., $f(x)+g(x)$), as opposed to having composed the two functions (i.e., $g(f(x))$). In addition, since the screens (one is red and the other has stars) show the individual functions, children did not have to compose the functions mentally; they could have applied the functions using the screens as prompts. Thus the question of when children demonstrate the capacity for compositional generalization remains open.

In the present study, we borrow insights from research that focused on Chinese children's ability to implicitly track the statistical regularities while learning to read Chinese characters. Chinese characters are written with a configuration of strokes (e.g., 狗 for dog). Among the roughly 2500 common characters, about 80-90% are compound characters with two parts, also known as 'radicals.' Radicals are recurring patterns of strokes that appear in multiple characters. A semantic radical provides a hint for the meaning of the character, whereas a phonetic radical provides a hint for the pronunciation of the character (Shu, Chen, Anderson, Wu,& Xuan, 2003). However, radicals only provide clues to meaning or sound probabilistically. A previous study by Yin and McBride (2015) investigated whether Chinese kindergarteners who have not received formal instructions in reading and writing were sensitive to the structure and phonetic regularities in Chinese characters. Five-year-old children were taught unfamiliar characters that either conformed to the statistical regularities of real characters or not. Their results suggest that by age 5, children have noticed the regularities among Chinese characters – both the position and the sound of the radicals (parts that are correlated with either meaning or pronunciation) – and used them for learning unfamiliar characters.

Here we use Chinese characters as a test domain to examine whether form-meaning correlations (semantic radicals) are readily used by English-speaking children who have not been exposed to Chinese writing, and whether they can use semantic radicals to make compositional generalizations, i.e., to guess the meaning of new characters they have not been exposed to during training. Compositionality in language usually refers to how spoken language is understood based on rules of combining the meanings of individual words. Here we use the Chinese writing system and test children with no prior knowledge of it. We tap into meaning via a visual task, so in effect we are testing conceptual compositionality.

In the present studies, we tested 5- to 8-year-old non-Chinese-speaking children in the United States. We sought to determine if these children demonstrate sensitivity to the regularities in Chinese characters, and can succeed in using these regularities to guess the meaning of new characters. In particular, in training trials, we taught children a set of characters with the same radical and similar meanings (e.g., 河 for river, 海 for sea, and 湖 for lake). Then in test trials, we showed children two new characters, one of which contained the same radical and the other did not. We asked children to guess which of the new characters has a similar meaning (e.g., which character means puddle).

In Experiment 1, all the characters were left-right structured and the radicals were always on the left side of the character in both training trials and test trials. In Experiment 2, the radicals were switched from the left side to the right side in the test trials. We hypothesized that even without any Chinese literacy background, children who observed the characters in the training trials would notice the pattern of each set of characters and find the same radical that was common across the set of characters. Subsequently children would be able to apply this knowledge to make guesses about the meanings of new characters. We further hypothesized that children would recognize the radicals in new characters even when the radicals are in a different position in the test trials.

Experiment 1

Methods

Participants Twenty-four children between the ages of 5 and 8 years (mean age = 6.88; range = 5.13 to 8.99; 11 females) participated in the experiment. Participants were tested via Zoom, a video conferencing software. Parents of the participants provided written informed consent prior to the experimental session.

Design Participants were shown 1 familiarization trial and 4 blocks. Each block tested one of the 4 radicals: the water radical (氵), the fire radical (火), the fruit radical (木), and the body part radical (月). Each block consisted of 1 training trial and 4 test trials.

Procedure Families joined the Zoom meeting on their personal devices. The experimenter displayed the stimuli by screen-sharing the PowerPoint slides on Zoom. We instructed parents to set up their screens such that the Zoom software was in full-screen mode, and the video feeds of the participant and the experimenter were either beside the stimuli (i.e., side-by-side mode), or in a corner of the screen not blocking the stimuli. Before each session started, the parent and the child were told by the experimenter that the child would play a game by looking at some Chinese characters and guessing the meanings of some new characters. The experimenter also reminded the parent not to provide hints or suggestions while the child was playing the game. When the study began, children sat in front of the device and their faces were fully captured by the camera. The stimuli, the child's video, and the experimenter's video were recorded throughout the experimental sessions.

Children were asked to listen to the experimenter and look at the characters in the familiarization trial and training trials, and then they were asked to predict the meaning of new characters in the test trials by choosing from two options. After each test trial, the experimenter said, "Good job!" or "Great!" and moved on to the next trial. The experimenter never provided feedback about whether children's choices were correct or not.

All the Chinese characters in the training and test trials were selected according to two criteria: (1) the radical is on the left side and (2) the non-radical part has no more than 5 strokes. Some Chinese characters were not the exact translations of the English words. We chose characters based on simplicity and sometimes we made up English meanings in order to keep the stimuli easily understandable by young children.

The familiariz ation trial In the familiarization trial, the experimenter introduced basic ideas of Chinese characters to the non-Chinese-speaking participants. Children were shown a set of Chinese characters on the first slide and the experimenter said, “Chinese characters are different from the English letters. Each Chinese character has a meaning, just like a word in English.” Then, the experimenter displayed a picture of a golden retriever, the English word “Dog” and the Chinese character “狗” (meaning “dog”) on the screen and told the child this is what the word “dog” looks like in Chinese characters. After this, the experimenter displayed the pictures, the English words and the Chinese characters of pig (猪), cat (猫) and fox (狐) in sequence on the same slide. A red square with a dividing line in the middle overlaid on top of each Chinese character to help the child notice the two parts of the character - left part and right part. At the end of the familiarization trial, the experimenter asked the child if he/she found three patterns from these characters, which were: 1) they were all names of animals; 2) they all had two parts - left part and right part; and 3) the left parts of the characters were all the same.

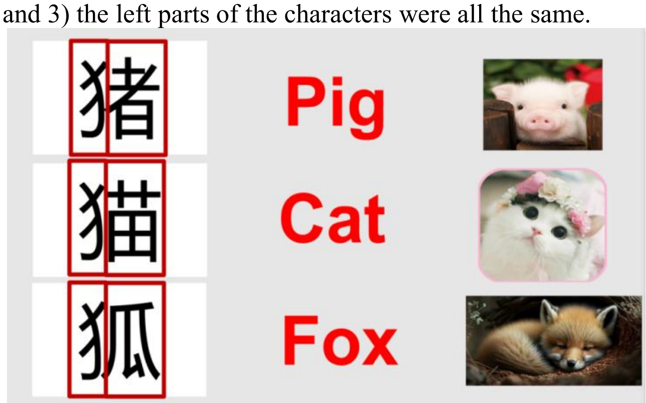

Figure 1: Characters shown in the familiarization trial.

The training trials In each training trial, the experimenter introduced a set of 6 Chinese characters all containing one of the radicals (the water radical, the fire radical, the fruit radical, or the body part radical) on the left side. Before the training trial, the experimenter told the child the category of the set of characters that they are about to learn (e.g., fruit), and reminded the child to find patterns if they can. The experimenter displayed the characters and the English meanings one by one, along with red squares. The experimenter told participants the meaning of each character and drew their attention to the structure of each character. For example, the experimenter said, “See this character? It means river, like the Nile River. This character means river. (Then the experimenter clicked the mouse and the red box showed up) Look, it has two parts, the left part and the right part. This character means river.”

Figure 2: Characters shown in the training trials.

The test trials After each training trial, participants completed 4 test trials testing the radical they just learned. In each test trial, the child was asked to choose which of the 2 new Chinese characters corresponded to a similar meaning within the category. On the left side of the slide, the experimenter first showed 3 characters the child just learned in the training trial, and said, “Remember these 3 characters? This means sea, this means rain, and this means tear. They are all about water.” Then 2 new characters appeared on the screen, one of which contained the target radical and the other did not. Then an English meaning within the learned category appeared on the slide in red. The child was asked to look at the 2 new characters and guess which one corresponded to the English meaning. For example, in the water radical set, the child was asked to point out which characters meant “lake”, “puddle”, “juice” or “waterfall” in Chinese. The character containing the target radical is the correct response.

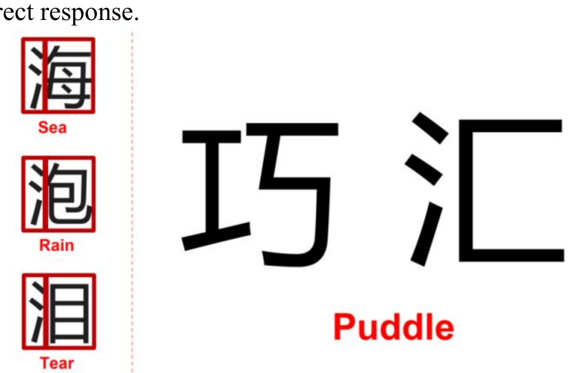

Figure 3: Characters shown in the test trials.

In one of the 4 test trials, children chose from 2 new characters that were a minimal pair (i.e., they only differed in the radicals on the left side). For example, in Figure 4, the two characters had the same part on the right side and different radicals on the left side.

Figure 4: Minimum pair in the test trials.

Results

The average number of correct responses was 11.6 (out of 16 test trials). Figure 5 shows the proportion of trials that children chose the correct character by radical type and age. We used mixed-effects logistic regression to predict children's answers (correct = 1, incorrect = 0) from radical type, trial type (whether the choices were a minimal pair), block order, trial order, age (z-scored), gender, and their interactions. We did not find any significant effects.

We next compared children's choices against chance. Overall, children chose the correct characters above chance (Exact binomial test: $P_{correct}$ = .73 [.68, .78], p < .001). Children chose the correct characters above chance for all four types of radicals (Body: $P_{correct}$ = .77 [.67, .85], p < .001; Fire: $P_{correct}$ = .77 [.67, .85], p < .001; Fruit: $P_{correct}$ = .69 [.58, .78], p < .001; Water: $P_{correct}$ = .70 [.60, .79], p < .001). In addition, children of all ages chose the correct characters above chance (5-year-olds: $P_{correct}$ = .63 [.54, .72], p = .003; 6-year-olds: $P_{correct}$ = .76 [.65, .85], p < .001; 7-year-olds: $P_{correct}$ = .77 [.68, .84], p < .001; 8-year-olds: $P_{correct}$ = .82 [.71, .91], p < .001).

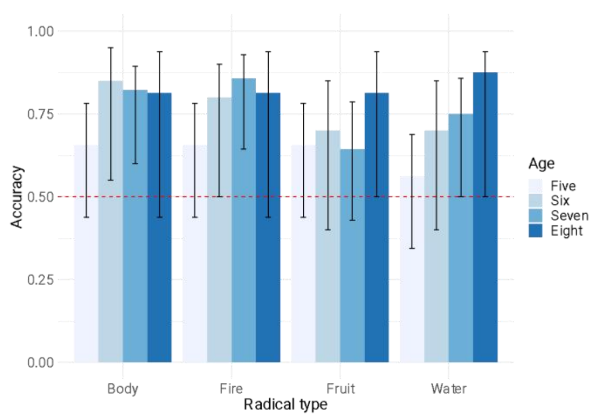

Figure 5: Proportion of trials that children chose the correct character, by radical type and age group, in Experiment 1. The dashed line indicates chance selection (.5), and the error bars indicate bootstrapped 95% CIs.

Discussion

Results from Experiment 1 suggest that after a brief training session, 5- to 8-year-old English-speaking children were able to find the common radicals of Chinese characters, and use them to make guesses about new characters they have never seen before. However, the radicals all appeared on the left side of each character, both in the training and the test trials. Therefore it is possible that children simply picked up this statistical regularity in making generalizations in the test trials. The idea of compositionality requires a stronger test: that a recurring radical could appear in a different position of a character and still preserve its meaning, just like words in English may appear in different positions of sentences (e.g., "the dog" as the subject or object of a sentence) with the same meaning.

In Experiment 2, in order to provide a more stringent test for compositionality, we presented the same training trials to children, with all radicals appearing on the left side of each character. On the test trials, however, the same radicals appeared on the right side of each character.

Experiment 2

Methods

Participants Another group of twenty-four children between the age of 5 and 8 years (mean age = 7.03; range = 5.12 to 8.96; 13 females) participated in the experiment. None participated in Experiment 1.

Stimuli and Procedure The procedure of Experiment 2 was the same as that of Experiment 1, except that the position of the radicals in the test trials was switched from left to right (Figure 6).

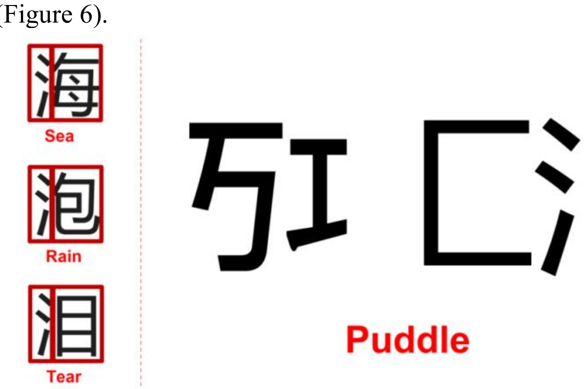

Figure 6: Switched position of radicals in the test trials of Experiment 2.

Results

Figure 7 shows the proportion of trials in which children chose the correct character by radical type and age. We used mixed-effects logistic regression to predict children's answers (correct = 1, incorrect = 0) from radical type, trial type (whether the choices were a minimal pair), block order, trial order, age (z-scored), gender, and their interactions. The best-fitting model included radical type, age, and gender as predictors. Children were less likely to choose the correct characters for the water radical, compared to the body radical (*β* = -1.48, *SE* = 0.43, *p* < .001), the fruit radical (*β* = -1.00, *SE* = 0.40, *p* = .01), and the fire radical (*β* = -0.75, *SE* = 0.39, *p* = .056). Children were more likely to choose the correct characters with increasing age (*β* = 0.96, *SE* = 0.31, *p* = .002). Boys were more likely to choose the correct characters than girls (*β* = 2.30, *SE* = 0.77, *p* = .003).

We next compared children's choices against chance. Overall, children chose the correct characters above chance (Exact binomial test: $P_{correct}$ = .77 [.73, .81], *p* < .001). Children chose the correct characters above chance for all four types of radicals (Body: $P_{correct}$ = .85 [.77, .92], *p* < .001; Fire: $P_{correct}$ = .77 [.67, .85], *p* < .001; Fruit: $P_{correct}$ = .80 [.71, .88], *p* < .001; Water: $P_{correct}$ = .67 [.56, .76], *p* = .001). In addition, children of all ages chose the correct characters above chance (5-year-olds: $P_{correct}$ = .65 [.56, .74], *p* = .002; 6-year-olds: $P_{correct}$ = .69 [.54, .81], *p* = .01; 7-year-olds: $P_{correct}$ = .80 [.72, .86], *p* < .001; 8-year-olds: $P_{correct}$ = .95 [.88, .99], *p* < .001).

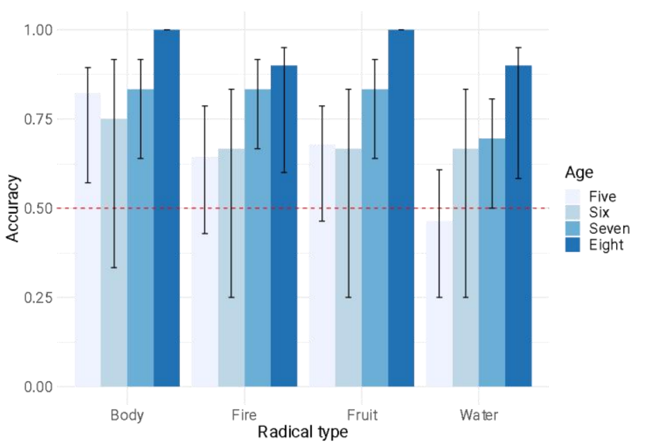

Figure 7: Proportion of trials in which children chose the correct character, by radical and age group, in Experiment 2. The dashed line indicates chance selection (.5), and the error bars indicate bootstrapped 95% CIs.

Discussion

Results from Experiment 2 showed that after a brief training session, 5- to 8-year-old English-speaking children were able to learn the common radicals of Chinese characters and use them to make compositional generalizations on the test trials. They did so when the radicals in the training trials appeared on the left side of each character whereas the same radicals appeared on the right side of each new character in the test trials. These findings provide stronger evidence for compositionality since children understood that the same radicals preserved their meaning even when they were used in a different position in new characters.

We also found some differences across radical type, age, and gender. Children were least likely to choose the correct character for the water radical. One possibility is that the meanings of the characters with the water radical are less coherent compared to the other categories. We also found that children's accuracy increased with age, which could be due to increased pattern learning ability or increased executive function. Lastly, boys performed better than girls in this task. Nevertheless, children of all ages and all genders reliably chose the correct characters above chance.

Comparison Between Experiments 1 and 2

Figure 8 shows the proportion of trials that children chose the correct character across Experiments 1 and 2. We used mixed-effects logistic regression to predict children's answers (correct = 1, incorrect = 0) from radical type, trial type (whether the choices were a minimal pair), Experiment, block order, trial order, age (z-scored), gender, and their interactions. The best-fitting model included radical type and age as predictors. Children were less likely to choose the correct characters for the water radical, compared to the body radical (*β* = -0.96, *SE* = 0.28, *p* < .001) and the fruit radical (*β* = -0.63, *SE* = 0.28, *p* = .02). Children were more likely to choose the correct characters with increasing age (*β* = 0.84, *SE* = 0.28, *p* = .003). There were no significant differences across experiments.

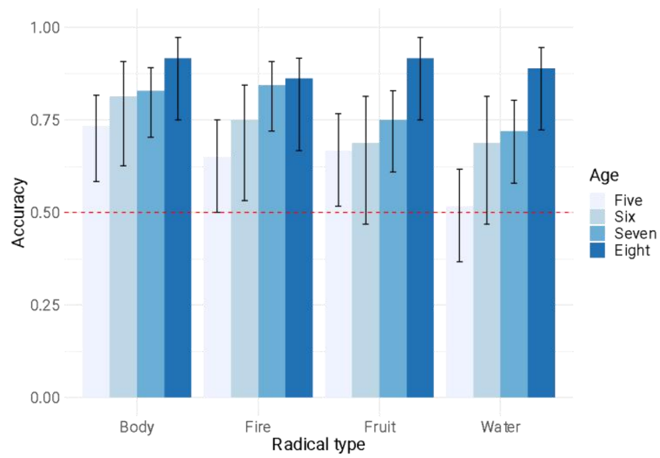

Figure 8: Proportion of trials that children chose the correct character, by radical and age group, in Experiments 1 and 2. The dashed line indicates chance selection (.5), and the error bars indicate bootstrapped 95% CIs.

General Discussion

The present study investigated whether children could make compositional generalizations in an unfamiliar domain. In two experiments with English-speaking children who have not been exposed to Chinese characters, we found that after a brief training session, they were able to generalize the newly learned radical-meaning correlations to new characters. Even more impressive, they generalized to new characters with the radicals appearing in a different position.

This suggests that children did not simply pick up the statistical regularities during training. Instead they understood the basic principle of compositionality, that the same radical will preserve its meaning regardless of where it appears in a character.

As we discussed in the introduction, previous studies have only investigated function learning in children as a test for conceptual compositionality. The findings are hard to interpret given both conceptual and methodological issues. Chinese characters represent an interesting and different domain for investigating conceptual compositionality. On the one hand, learning the meanings of Chinese characters is related to language, as it demonstrates form-meaning pairings; on the other hand, this is also a visual domain, since the learner needs to figure out that the characters may be parsed into multiple parts -- some parts are indicative of meaning but other parts are not. Despite these complexities, children as young as 5 years of age readily learned the radical-meaning pairings and generalized them to new characters.

One limitation of the current study is that children may have chosen the character with the correct radical based on visual similarity. We need to test this alternative hypothesis by providing the same set of test characters with irrelevant meanings. So far at best we have provided some preliminary evidence for conceptual compositionality.

The present study also suggests new ways to examine compositionality in development. In future studies, for instance, we may provide training on Chinese characters with radicals in a left-right arrangement, and test children with new characters where the same radicals appear in a top-bottom arrangement. This would test further the idea that compositional generalization differs from statistical generalization. We can also use visual shapes (Zhou et al. 2024) as another domain for testing the general principle of compositionality. To probe the developmental origin of compositionality, we can also test younger children and infants who have not acquired much language.

References

Amalric, M., Wang, L., Pica, P., Figueira, S., Sigman, M., & Dehaene, S. (2017). The language of geometry: Fast comprehension of geometrical primitives and rules in human adults and preschoolers. *PLoS Computational Biology, 75*, 1–1.

Barner, D., & Snedeker, J. (2008). Compositionality and statistics in adjective acquisition: 4-year-olds interpret tall and short based on the size distributions of novel noun referents. *Child Development, 79*(3), 594–608.

Fodor, J. A. (1975). *The language of thought.* Harvard university press.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*, 3-71.

Hamburger, H., & Crain, S. (1984). Acquisition of cognitive compiling. *Cognition, 17*(2), 85-136.

Lake, B. M., & Baroni, M. (2023). Human-like systematic generalization through a meta-learning neural network. *Nature, 623*, 115–121.

Matthei, E. H. (1981). The acquisition of prenominal modifier sequences. *Cognition*, *11*(3), 301-332.

Partee, B. H. (1984). Compositionality. In *Varieties of Formal Semantics* (pp. 281-). Foris.

Piantadosi, S., & Aslin, R. (2016). Compositional reasoning in early childhood. *PloS One, 11*(9), e0147734–e0147734.

Piantadosi, S. T., Palmeri, H., & Aslin, R. (2018). Limits on composition of conceptual operations in 9-month-olds. *Infancy, 23*(3), 310 – 324

Piantadosi, S. T., & Jacobs, R. A. (2016). Four problems solved by the probabilistic language of thought. *Current Directions in Psychological Science, 25*(1), 54-59.

Quilty-Dunn, J., Porot, N., & Mandelbaum, E. (2023). The best game in town: The reemergence of the language-of-thought hypothesis across the cognitive sciences. *Behavioral and Brain Sciences, 46*, e261.

Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003).Properties of school Chinese: Implications for learning to read. *Child Development, 74*, 27 – 47.

Yin, L., & McBride, C. (2015). Chinese kindergartners learn to read characters analytically. *Psychological Science, 26*(4), 424–432.

Zhou, Y., Feinman, R., & Lake, B. M. (2024). Compositional diversity in visual concept learning. *Cognition, 244*, 105711.