Spatial Metaphor Facilitates Word Learning

Ariel Starr,
University of Washington

Alagia J. Cirolia,
University of California, Berkeley

Katharine A. Tillman,
The University of Texas at Austin

Mahesh Srinivasan
University of California, Berkeley

Abstract

Why are spatial metaphors, like the use of "high" to describe a musical pitch, so common? This study tested one hundred and fifty-four 3- to 5-year-old English-learning children on their ability to learn a novel adjective in the domain of space or pitch and to extend this adjective to the untrained dimension. Children were more proficient at learning the word when it described a spatial attribute compared to pitch. However, once children learned the word, they extended it to the untrained dimension without feedback. Thus, children leveraged preexisting associations between space and pitch to spontaneously understand new metaphors. These results suggest that spatial metaphors may be common across languages in part because they scaffold children's acquisition of word meanings that are otherwise difficult to learn.

During the first few years of life, children rapidly acquire thousands of words. One of the many challenges faced by young learners is how to acquire words for abstract concepts. How do children learn words for concepts like time, love, and justice—concepts whose referents cannot be directly perceived by the senses? Notably, across languages and cultures, words that originally had spatial meanings have often developed new, metaphorical meanings that describe more abstract phenomena (Clark, 1973; Lakoff & Johnson, 1980; Sweetser, 1980; Talmy, 1988; Xu, Malt, & Srinivasan, 2017). For example, in English when describing temporal concepts, we might say "the vacation was too *short*" or "the movie release was pushed *back*." Spatial words are also frequently used metaphorically to describe other concepts that are relatively more abstract or fleeting than spatial concepts, such as musical pitch and numerical magnitudes (e.g., "she sang a *high* note"; "1,000,000 is a *big* number"). But why are spatial metaphors so prevalent in natural language? The present study explores the idea that spatial metaphors are common in part because they scaffold children's learning of word meanings that would otherwise be difficult to learn.

Correspondence concerning this article should be addressed to Ariel Starr, Department of Psychology, University of Washington, Guthrie Hall, Seattle, WA 98195. abstarr@uw.edu.

It has previously been suggested that children learn the spatial meanings of words before their more abstract, metaphorical meanings (e.g., learning *high* as it applies to space before learning *high* as it applies to musical pitch; Andrews & Madeira, 1977; Casasanto, Fotakopoulou, & Boroditsky, 2010; Clark, 1973). One reason that children might learn word meanings in this order could be that, unlike words that label spatial concepts, words that label more abstract concepts, like the use of *high* to label pitch, correspond to referents that cannot be seen or touched and tend to be fleeting in experience. Thus, given only situational evidence (e.g., hearing a word used to describe a scene), it may be easier for children to form word-concept mappings for spatial word meanings, like the use of *high* to label spatial position, than for more abstract word meanings, like the use of *high* to label the pitch of a sound. Although pitch is not a purely abstract concept because it can be directly perceived through the senses (Brysbaert, Warriner, & Kuperman, 2014), it can only be perceived through audition, whereas spatial positions or relations can be perceived through a combination of sight, touch, and audition. Thus, in this article, we describe pitch as being "more abstract" than space. The relative abstractness of pitch in comparison to spatial concepts might make it more difficult for children to form a word-concept mapping involving pitch than space.

Broad patterns in early world learning support the hypothesis that learning words for more abstract concepts is more challenging than learning words for more concrete concepts. In the early vocabularies of young children, concrete nouns tend to dominate, with abstract nouns and verbs being relatively less common, and concrete verbs (i.e., action verbs) being more prevalent than more abstract verbs (i.e., mental state verbs; Bloom et al., 1975; Braginsky, Yurovsky, Marchman, & Frank, 2019; Gentner, 1978; Huttenlocher, Smiley, & Charney, 1983). Research on word learning that uses the human simulation paradigm points to an explanation for why this is the case (Gillette, Gleitman, Gleitman, & Lederer, 1999; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005; Snedeker & Gleitman, 2004). In these studies, when adult participants are asked to guess a mystery word that has been used in a scene, they are better able to do so for nouns than for verbs, and better able to guess concrete nouns compared to abstract nouns. This suggests that, given only situational evidence, it is easier to map nouns—particularly concrete nouns—to concepts than it is to map verbs or abstract nouns to their respective, more abstract concepts. Critically, because these results derive from fluent, adult speakers of a language, they demonstrate that even if a learner already possesses all of the relevant concepts to which a new word might refer, it will still be harder for them to map words to some concepts than others.

Although learning verbs is difficult, children do, of course, ultimately succeed. Prior research suggests that one particularly helpful clue for learning verbs is understanding the syntactic structure in which they have been embedded (Fisher, Gertner, Scott, & Yuan, 2010; Gillette et al., 1999; Gleitman, 1990; Gleitman et al., 2005; Snedeker & Gleitman, 2004). Similar to how syntax might help children solve the mapping problem for verbs, we propose that spatial metaphors could help children map words to abstract concepts, and that this may help explain why these metaphors are prevalent in natural language (Srinivasan & Carey, 2010; Xu et al., 2017). In particular, although it might be difficult for children to learn a new word for an abstract concept by observing the scenes in which that word has been used, this process may be simpler if children have already learned a prior, related meaning for the

word. For example, if a child has already learned the spatial meaning of *high*—which should be possible to learn given only situational evidence—this could constrain their hypotheses about what *high* might mean when it is used metaphorically to refer to properties of a sound. By this account, it may be more efficient for children to extend a spatial word that they already know to a new, more abstract meaning (e.g., extending *high* from space to pitch) than to learn an entirely new word for this more abstract meaning (e.g., learning an unambiguous word for pitch).

Notably, this proposal runs contrary to both classical and contemporary models of word learning (Bloom, 2000; Markman, 1990; Trueswell, Medina, Hafri, & Gleitman, 2013; Xu & Tenenbaum, 2007; Yu & Smith, 2007; Yurovsky & Frank, 2015), according to which children assume that a new word will have a single meaning. For example, according to the taxonomic assumption (Markman, 1990), children expect a new word to refer to items from the same taxonomic category rather than other thematically related items. Such assumptions predict that lexical ambiguity should be rare, as it should be difficult for children to learn words that label multiple different categories of meaning, making these words less likely to be transmitted culturally (for discussion, see Dautriche, Chemla, & Christophe, 2016; Srinivasan & Snedeker, 2014). However, lexical ambiguity is common across languages. For example more than 80% of English words carry multiple meanings (Rodd, Gaskell, & Marslen-Wilson, 2004). Recent research suggests that even toddlers can acquire multiple meanings for a new word (Dautriche, Fibla, Fievet, & Christophe, 2018) and, moreover, that in cases of *polysemy* (in which words have multiple, related meanings), children can leverage their knowledge of one meaning of a word to understand and even infer additional, related meanings for that word (Floyd & Goldberg, 2020; Srinivasan, Al-Mughairy, Foushee, & Barner, 2017; Srinivasan, Berner, & Rabagliati, 2018). For example, if preschoolers are taught that *daxing* labels an action involving a tool, they spontaneously anticipate that the tool itself will be called *a dax*, generalizing from their knowledge of how words like *hammer* and *shovel* are used (Srinivasan et al., 2017).

These previous findings suggest that polysemy may be common in part because it is a feature of the lexicon that facilitates word learning: it may be easier for children to learn a lexicon in which words are linked to multiple, related meanings in systematic ways than to learn an unambiguous lexicon in which each meaning has its own unique label (Ramiro, Srinivasan, Malt & Xu, 2018; Srinivasan & Rabagliati, 2015). The present study extends this prior research on polysemy by testing whether children can use their knowledge of one meaning of a word (e.g., *high* in space) to help them learn another, metaphorically related meaning of that word (e.g., *high* in pitch). Children's ability to leverage a prior, learned meaning of a word to acquire a new meaning for that word may be especially valuable in cases in which the new meaning is more abstract or fleeting in experience, as in the case of spatial metaphors for pitch.

Some researchers, however, have made the opposite claim, suggesting that the metaphorical nature of pitch words may *hinder* children's learning of them (Costa-Giomi & Descombes, 1996; Flowers & Costa-Giomi, 1991). In one study, Costa-Giomi and Descombes (1996) examined 7-year-old French-speaking children's understanding of pitch terms. In a pretraining assessment, children rarely used the correct words to label changes in musical

pitch. Children were then trained to use either the pitch-specific terms *aigu* and *grave* or the spatial metaphorical terms *haut* and *bas* (which translate to *high* and *low*, respectively) and then asked to label pitch changes in a familiar song. Interestingly, children who learned the pitch-specific terms were more accurate in labeling the pitch changes compared to the children who learned the metaphorical terms. From these and other similar findings, Costa-Giomi and colleagues suggested that it is more difficult for children to learn metaphorical pitch terms in comparison to pitch-specific terms because children are confused by the multiple meanings conveyed by the metaphorical terms. By this account, children may not perceive a structural similarity between musical pitch and space, which then makes it difficult for them to metaphorically extend spatial words to additional pitch meanings.

However, more recent studies challenge the assumption that children do not understand structural mappings between pitch and space. Infants, even in the first year of life, spontaneously associate space with other dimensions such as pitch, time, and number (e.g., de Hevia, Izard, Coubart, Spelke, & Streri, 2014; Dolscheid, Hunnius, Casasanto, & Majid, 2014; Lourenco & Longo, 2010; Srinivasan & Carey, 2010; Walker et al., 2010). Infants, for example, prefer to look at objects that move upward in space when paired with a tone that increases in pitch compared to one that decreases in pitch, and prefer to look at smaller objects paired with higher pitched tones compared to lower pitched tones, consistent with the ways in which pitch and space are metaphorically linked in different languages (Dolscheid et al., 2014; Mondloch & Maurer, 2004; Walker et al., 2010). Preschoolers also explicitly match stimuli between pitch and space in similar ways (Starr & Srinivasan, 2018; but see Shayan, Ozturk, Bowerman, & Majid, 2014). These findings suggest that children may be able to leverage preexisting cross-domain associations when extending words to new meanings in other domains. For example, if children already associate high spatial locations with high pitches, they may be able to guess that a word that has labeled high spatial positions is also likely to label high sound frequencies, and vice versa.

To test this hypothesis, we taught adults and 3- to 5-year-old English-learning children a novel adjective in the domain of space or pitch, evaluated their comprehension of the novel adjective in the trained dimension, and then tested their ability to extend the novel adjective to the other, untrained dimension (Figure 1). For example, in one condition, participants were taught that the novel adjective *daxy* describes a high spatial location (using two training exemplars) and were then tested on their understanding of *daxy* with novel exemplars. Next, in the extension phase, these participants heard two sounds—one high in pitch and the other low—and were asked which of these sounds was *daxy*. Critically, in this extension phase, participants were not provided with training exemplars or given any feedback on their choices. This lack of training and feedback ensured that children's responses would reflect their own intuitions about what the word means when used again in a novel domain. Through this approach, we assessed potential asymmetries in word learning (i.e., if it is easier to initially learn words for space than pitch) and children's propensity to extend words metaphorically across domains. This enabled us to probe the relative difficulty of learning novel words that describe a spatial versus auditory attribute, as well as whether children can leverage their knowledge of a word learned in one domain to predict what that same word will mean when it is used in another domain.

We predicted that in the training phase, children would be more successful at learning the novel word in the spatial context compared to the auditory context. In comparison to auditory attributes, which tend to be fleeting, spatial attributes are more concrete and are often available for simultaneous comparison. Therefore, we expected that—even when given equivalent evidence in the training phase—it would be easier for children to learn a novel adjective that describes a spatial attribute compared to one that describes an auditory attribute. Although this prediction is intuitive and has been suggested previously (Casasanto et al., 2010; Clark, 1973), there is little direct evidence that children learn the spatial senses of words more easily than their more abstract, metaphorical senses (but see Shusterman & Wagner, 2005).

We were particularly interested in how children would extend the novel word to the untrained domain in the extension phase and whether they would extend the word in ways that are consistent with how metaphorical language links space and pitch in English and in other languages. For example, if children have learned that *daxy* means high in space, will they take advantage of preexisting associations between space and pitch and spontaneously expect that *daxy* will also mean high in pitch, in absence of any training or feedback? We were also interested in whether there would be an advantage for extending the novel word in one direction compared to the other. In particular, because words for concrete concepts are more likely to acquire additional, more abstract meanings over time than the reverse (Xu et al., 2017), and because it has been hypothesized that metaphorical targets are conceptualized in terms of their source domains (Boroditsky, 2000; Casasanto, 2010; Gibbs Jr, Gibbs, & Gibbs, 1994; Lakoff & Johnson, 1980), it may be easier for children to extend a newly learned spatial word to the domain of pitch than to extend a newly learned word for pitch to the domain of space.

Importantly, our study design also took advantage of cross-linguistic variation in spatial metaphors for pitch to test how children's ability to extend words across domains might depend on their experience with and knowledge of metaphorical language. Although English describes musical pitch in terms of spatial height, other languages, including Turkish and Farsi, describe pitch in terms of thickness (i.e., such that *thicker* sounds are lower in pitch). Therefore, although testing children's understanding of novel height-pitch metaphors may recapitulate the learning of metaphorical polysemy that already exists in children's vocabularies, the thickness-pitch metaphors are entirely novel to English-learning children and provide insight into how metaphorical polysemy is acquired. We hypothesized that if children rely on their knowledge of English-language metaphors for pitch during the task, then they should more systematically extend new words between pitch and height compared to between pitch and thickness. In contrast, if the ability to generalize between space and pitch does not depend on experience with specific metaphorical language, and if children can instead draw on their preexisting knowledge of space-pitch mappings (e.g., de Hevia et al., 2014; Dolscheid et al., 2014; Lourenco & Longo, 2010; Srinivasan & Carey, 2010; Starr & Srinivasan, 2018; Walker et al., 2010), they may be equally systematic when extending new words between pitch and height and between pitch and thickness.

We tested children between the ages of 3 and 5 to assess the effects of age on both word learning and cross-domain extension. We expected that although younger and older children

might be relatively good at learning a new spatial adjective, younger children might struggle with learning a new adjective for pitch. We were also interested in the effect of age on the ability to systematically interpret new metaphors in the extension phase, and whether increased familiarity with English-language spatial metaphors for pitch might make it more (or less) difficult for older children to generalize between space and pitch in the thickness-pitch condition compared to that in the height-pitch condition.

More generally, we reasoned that children's performance in the word learning and extension phases of our study would shed light on why spatial metaphors are so prevalent across languages. Our hypothesis is that spatial metaphor may be prevalent across languages in part because it confers a functional advantage in scaffolding the acquisition of abstract word meanings, thereby making the lexicon more learnable (Srinivasan & Carey, 2010; Xu et al., 2017). In particular, if learners can leverage their preexisting associations between space and more abstract domains, this may enable them to extend easy to learn spatial words to more abstract meanings and circumvent the difficulty of learning dedicated words for these more abstract meanings.

Method

Participants

Data from 154 children were included in the final sample. This sample included fifty-two 3-year-olds (M_{age}: 3.51 years, range: 3.0–3.99, 28 female), fifty-three 4-year-olds (M_{age}: 4.58 years, range: 4.0–4.99, 28 female), and forty-nine 5-year-olds (M_{age}: 5.31, range: 5.0–5.94, 26 female). All children were native English speakers who were recruited and tested in laboratory, at local preschools, or at local science museums. Data from an additional eight children were excluded for parental interference (one 3-year-old), speaking English as a second language (two 3-year-olds; one 5-year-old), or failure to pay attention during the procedure (two 3-year-olds; two 4-year-olds). The demographics of the child participants were representative of Berkeley, California, and the surrounding area, which is approximately 50% White, 20% Asian, 10% Black, 10% Hispanic or Latino, and 6% from two or more races. Children came from primarily middle- and upper-class families. All children were fluent in English and not regularly exposed to another language. Data from 32 adult participants were also collected to validate the procedure (M_{age}: 20.8 years, range: 18.5–25.7, 17 female). Adult participants and guardians of child participants gave written consent to a protocol approved by the local ethics committee. Adult participants were compensated with university course credit and children received a small gift.

We aimed for a predetermined sample size of 48 children in each age group, equally distributed and randomly assigned across the different conditions. However, a slightly larger number of children was tested so as to allow all eligible children within a given classroom to participate. Our primary contrast of interest was the effect of dimension on word learning and word extension: with 72 participants per dimension condition, collapsed across ages, this gave us 99% power to detect a large-sized effect (Cohen's $d = 0.8$) and 80% power to detect a medium-sized effect (Cohen's $d = 0.5$).

Materials

Spatial stimuli consisted of pairs of familiar objects that differed in their spatial location (height condition) or their horizontal width (thickness condition). Auditory stimuli consisted of pairs of pure tones, each with a duration of 1,000 ms, that were created using Audacity. The frequencies of the lower tones varied between 261 and 494 Hz, and the frequencies of the higher tones were double that of the lower tone in the pair. The auditory stimuli were accompanied by pictures of music notes, and the music notes wiggled for the duration of the tone.

Procedure

Participants were tested individually in a quiet room or corner of their classroom. Stimuli were presented using PowerPoint on a Macbook 15 in. laptop. Participants were randomly assigned to either the height or thickness condition and to either the space-to-sound or sound-to-space condition. Participants in the space-to-sound conditions were taught the novel word in a spatial context, tested on their understanding of this word, and were then asked to extend the word to an auditory context. In contrast, participants in the sound-to-space conditions were taught the word in an auditory context, tested on their understanding of this word, and then asked to extend the word to a spatial context. Within each condition, half of the participants were trained with the novel word corresponding to the unmarked pole of the dimension (high, thick) and the other half were trained with the novel word corresponding to the marked pole of that dimension (low, thin). All participants completed a word training block followed by a word extension block and a posttest block. See Figure 1 for a schematic of the procedure.

Word Training

Participants were first taught and tested on the meaning of the novel word, *daxy*, in the trained dimension. For example, in the space-to-sound condition, participants first saw a screen with two familiar objects on it (e.g., two beach balls in the height condition). While pointing to the objects, the experimenter said, "There are two [name of object] here! This [name of object] is daxy. It is a daxy [name of object]! But look at this [name of object]. It is not daxy. It is not a daxy [name of object]. See this [name of object]? This [name of object] is daxier than this [name of object]." In the sound-to-space condition, participants first saw a screen with two images of music notes on it. The experimenter said, "Let's listen to these sounds! [experimenter plays both sounds] This sound is daxy! [plays first sound] It's a daxy sound! But listen to this sound [plays second sound]. It is not daxy. Listen! This sound [plays first sound] is daxier than this sound [plays second sound]." This demonstration was then repeated for a second pair of objects (in the space-to-sound condition) or sounds (in the sound-to-space condition).

Next, participants were tested on their comprehension of the novel word within the trained dimension. On each of four test trials, participants either saw two objects (in the space-to-sound condition) or heard two sounds (in the sound-to-space condition) that varied along the critical dimension and were asked which object or sound was daxy. Participants responded by pointing to the spatial object or music note and received feedback after each trial. The left

and right position of the target object or music note varied across trials and was counterbalanced across participants.

Word Extension

After participants completed the four test trials within the training dimension, they proceeded to the extension phase. Participants were told "Now let's listen to some sounds!" (space-to-sound phase. Participants were told "Now condition) or, "Now let's look at some pictures!" (sound-to-space condition). Next, on each of four extension trials, participants were asked to interpret the newly learned word in a novel context. In the space-to-sound condition, two sounds were played, one at a time, each accompanied by a wiggling music note. Children were asked "Which sound is daxy? Is it this one [play sound] or this one [play sound]?"; children indicated their response by pointing. In the sound-to-space condition, children saw two pictures and were asked, "Can you point to the daxy [name of object]?" Participants were not given feedback on their responses. Stimuli used in the extension phase were drawn from the stimuli that had been used in the training phase in the other condition. The left and right position of the target object or music note varied across trials and was counterbalanced across participants.

Attention Check

After participants completed the four extension trials, they were tested on their knowledge of the initial, trained meaning of the novel word. In the space-to-sound conditions, participants were tested on the trained spatial meaning of the word. They were told, "Let's look at some pictures again!" and were shown two pictures and asked "Can you point to the daxy [name of object]?" In the sound-to-space conditions, participants were tested on the trained pitch meaning of the novel word. They were told, "Let's listen to some sounds again" and were played two sounds and asked "Which sound is daxy?"

Posttest

At the end of the experiment, participants were prompted to provide their own definition of the novel word: "What does daxy mean?" Summary analyses of these data are available in Supporting Information. In addition, participants were tested on their comprehension of the English metaphorical words that were relevant to the condition they participated in (i.e., *high* and *low* in the height-pitch condition or *thick* and *thin* in the thickness-pitch condition). Participants were presented with one pair of novel spatial stimuli and one pair of novel sounds and were asked which exemplar goes with each word (e.g., "Which picture is high?").

Data Analysis and Sharing

We analyzed performance separately for the word training and word extension phases. All analyses were performed in R. Performance was analyzed using mixed effects models using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017). Significant interactions were followed-up using Tukey contrasts to control for multiple comparisons using the emmeans package (Searle, Speed, & Mil-liken, 1980). All reported analyses and comparisons were planned. However, because we entertained multiple hypotheses about the

outcome of these tests, the overall analyses strategy was more exploratory than confirmatory. Inclusion Bayes factors were calculated using the brms package (Bürkner, 2017). Inclusion Bayes factors ask whether the observed data are more probable under models that include the particular effect (as a main effect and all possible interactions) compared to models that do not include the particular effect. In general, inclusion Bayes factors > 3 are considered strong support for the particular effect, whereas inclusion Bayes factors < 1/3 are considered as strong support in favor of the null. Inclusion Bayes factors between 1/3 and 3 are inconclusive. All materials, raw data, and the RMarkdown analysis code used to produce the results section can be found at https://osf.io/7hztq/.

Results

Adults

Adults performed near ceiling during the word learning phase ($M_{accuracy} = 100\%$) and during the extension phase ($M_{accuracy} = 98.44\%$). Extension performance was analyzed using a mixed effects model that included dimension (space or sound), condition (height or thickness), and their interaction as fixed effects and subject as a random effect (model syntax: glmer(accuracy ~ dimension × condition + (1|subject), family = binomial(link = "logit")). This analysis revealed no significant effects or interactions (dimension: $\beta = -20.86$, $p = 1$; condition: $\beta < .001$, $p = 1$; interaction: $\beta = 20.86$, $p = 1$). This suggests that adults had no difficulty learning or extending the novel word based on the training procedure, and their performance was not influenced by condition or extension direction. Below, we present analyses of the data from our child participants.

Word Training

We first analyzed the data from the word training phase of the study to assess how well children learned the novel word in the trained dimension (Figure 2). We included age (3, 4, or 5 years), training condition (spatial height, spatial thickness, or pitch), and their interaction as fixed effects and included subject as a random effect (model syntax: glmer(accuracy ~ training condition × age group + (1|subject), family = binomial(link = "logit")). This analysis revealed that there was a main effect of trained dimension ($BF_{inclusion} = 3.22$). Performance in the pitch condition was significantly lower than that in the two spatial conditions ($\beta = -1.71$, $p = .001$). However, performance in the two spatial conditions did not differ ($\beta = 4.69$, $p = .988$). The effect of age was also not significant ($\beta = 1.11$, $p = .181$, $BF_{inclusion} = 0.89$), nor was there a significant interaction between training condition and age ($\beta = -1.76$, $p = .083$, $BF_{inclusion} = 3.48$; note that the inclusion Bayes factor suggests that these factors do interact). Follow-up t-tests indicated that training performance was above chance for children in each age group in each training condition (3-year-olds height: $m = 90\%$ [81, 98], $t = 8.95$, $p < .001$, Cohen's $d = 2.48$; 3-year-olds thickness: $m = 95\%$ [88, 100], $t = 11.54$, $p < .001$, $d = 3.08$; 3-year-olds pitch: $m = 66\%$ [55, 77], $t = 2.87$, $p = .008$, $d = 0.57$; 4-year-olds height: $m = 96\%$ [88, 100], $t = 8.95$, $p < .001$, $d = 2.48$; 4-year-olds thickness: $m = 100\%$ [100, 100]; 4-year-olds pitch: $m = 81\%$ [73, 88], $t = 2.87$, $p = .008$, $d = 0.57$; 5-year-olds height: $m = 98\%$ [94, 100], $t = 23$, $p = 0$, $d = 6.64$; 5-year-olds thickness: $m = 87\%$ [67, 100], $t = 4.4$, $p = .001$, $d = 1.22$; 5-year-olds pitch: $m = 89\%$ [82, 95], $t = 11.48$, $p < .001$, $d = 2.34$). These results demonstrate that children were

more successful at learning the novel word in the two spatial conditions compared to that in the pitch condition and that this pattern may be stronger for younger children. However, on average, children successfully learned the novel word in all conditions.

Word Extension

The next series of analyses focused on children's ability to extend the word to the novel, untrained dimension (Figure 3). For these analyses, we used a logistic mixed effects model with age (3, 4, or 5 years), extension direction (space-to-sound or sound-to-space), and condition (height or thickness), as well as interaction terms for all combinations of these variables, as fixed effects and included subject as a random effect (model syntax: glmer(accuracy ~ extension direction × training condition × age group + (1|subject), family = binomial(link = "logit")). This analysis revealed a significant interaction between age and condition ($\beta = 1.302$, $p = .047$, $BF_{inclusion} = 1.89$). No main effects (all βs < -0.173, ps $> .639$, $BF_{inclusion}$ 0.46–1.2) or other interactions reached significance (all βs < -0.787, ps $> .395$, $BF_{inclusion}$ 0.49–0.85). Follow-up pair-wise Tukey comparisons indicated that 3-year-olds performed worse than 5-year-olds in the thickness condition ($z = -3.16$, $p = .02$), regardless of extension direction. No other pairwise comparisons reached significance (zs < -2.47, ps $> .132$).

Follow-up t-tests comparing performance to chance indicated that children's extension performance was above chance in all conditions, with the exception of 3-year-olds in the thickness condition (3-year-olds height: $m = 71\%$ [62, 80], $t = 4.45$, $p < .001$, $d = 0.91$; 3-year-olds thickness: $m = 50\%$ [38, 62], $t = 0$, $p = 1$, $d = 0$; 4-year-olds height: $m = 66\%$ [55, 77], $t = 2.93$, $p = .007$, $d = 0.54$; 4-year-olds thickness: $m = 65\%$ [51, 76], $t = 2.17$, $p = .04$, $d = 0.44$; 5-year-olds height: $m = 76\%$ [66, 88], $t = 4.4$, $p < .001$, $d = 0.9$; 5-year-olds thickness: $m = 82\%$ [70, 91], $t = 5.74$, $p < .001$, $d = 1.2$). These results demonstrate that in general, children succeeded in extending the novel word from the trained dimension to the new dimension, reflecting their ability to spontaneously understand new metaphorical word meanings. In addition, the lack of a significant main effect or interaction involving extension direction suggests that although children were more easily able to learn words in a spatial context than in an auditory context, as demonstrated by the word training data, there was no evidence that they were better at extending the novel word in either direction. We also explored how accuracy in the word training phase related to word extension accuracy. These analyses, which are reported in full in Supporting Information, indicated that training accuracy improved the model fit overall, but no individual predictors or interaction terms significantly explained unique variance.

It is worth emphasizing that 4- and 5-year-old children's extension performance was not significantly different as a function of whether the novel word was metaphorically extended along a familiar space-pitch relation in English (height-pitch) or along an unfamiliar relation that exists in other languages like Turkish and Farsi (thickness-pitch). This finding is striking, as it suggests that children's success in understanding the extensions of the novel word cannot be explained by mapping the novel word onto a familiar metaphorical English word that describes both space and pitch. Although this strategy was available to children in the height-pitch condition (e.g., who could have mapped *daxy* to English *high* or *low*), it was

not available to children in the thickness-pitch condition, and yet children performed equally well across these conditions.

Posttest Word Comprehension

Children of all ages understood the height words in the spatial context (t-tests comparing accuracy to chance; all $ts > 6.17$, all $ps < .001$; Figure 4). Four- and five-year old children also understood the height words in the auditory context ($ts > 4.04$, all $ps < .001$), though 3-year-olds' comprehension was not significantly different from chance ($t = 1.58$, $p = .13$). This pattern of results provides direct evidence that children acquire the spatial meanings of *high* and *low* at an earlier age than their pitch meanings. By contrast, children were at chance in understanding the thickness terms in both the spatial and auditory contexts (all $ts < 1.3$, all $ps < .001$), with the exception of 3-year-olds who were significantly below chance at understanding thickness terms in the spatial context ($t = -3.53$, $p = .002$).

Discussion

A critical challenge faced by word learners is how to learn words whose referents cannot be seen or touched. Here, we explored how spatial metaphor may help children overcome this challenge. We found that preschool-aged children were more proficient at learning a novel word that described a spatial attribute compared to one that described musical pitch. However, once children had learned the novel word, they were able to extend it to the other dimension, regardless of the dimension in which the word was initially learned. These results suggest that children are readily able to understand new metaphorical meanings for words and that this ability may support word learning in abstract contexts.

As predicted, children learned the meaning of the novel word more easily when it was taught in a spatial context compared to than when it was taught in an auditory context. This may be because when children learned the word in the spatial context, they were able to directly compare the two possible referents because they were presented simultaneously. For example, as soon as the experimenter labeled one of the objects as *daxy*, that object could be compared to the other object on the screen. By contrast, in the auditory context, the tones were played sequentially and could not be simultaneously compared. Therefore, when one of the tones was labeled as *daxy*, participants needed to compare the *daxy* tone to the other tone based on their memory of each tone. This design was meant to mimic the types of situations in which children would likely naturally encounter these types of words—it is more common to be able to directly compare two spatial properties than it is to be able to compare two auditory properties presented simultaneously. More generally, spatial properties tend to persist in time, whereas auditory properties are fleeting and last only as long as the duration of the sound, and spatial properties can be experienced though multiple senses, whereas auditory properties can only be heard. Because spatial words can be learned by directly observing the scenes in which those words have been used, it may be easier for young children to learn words that describe concrete attributes, such as spatial extent, height, or thickness, than to learn words for attributes of sound, time, and other more abstract properties. Our finding from the posttest that children understand the spatial meanings of

high and *low* at earlier ages than they understand their pitch meanings also supports this hypothesis.

An alternate explanation for why children might learn the spatial meaning of words before their pitch meanings is that caregivers may more frequently label or draw children's attention to spatial relations compared to auditory relations. For example, it is possible that children hear the words *high* and *low* in reference to space more frequently than to pitch in their language environments. By this account, because of differences in input frequency, children may form explicit categories for spatial positions or size earlier than they form explicit categories for auditory pitch and are therefore better at learning novel words in spatial contexts. However, it is unlikely that this explanation can fully account for our results. In particular, we found that 4- and 5-year-olds—who already understood the words *high* and *low* in the context of both space and pitch—were more successful at learning novel words with these meanings in a spatial context, suggesting that learning words for pitch is more difficult than learning words for space, even when children already possess all of the necessary concepts. This finding suggests that part of the difficulty children have with learning pitch words stems from the challenge of mapping known concepts to words, not just learning the concepts themselves. This echoes prior findings from the human simulation paradigm, which demonstrate that, given only situational evidence, it is easier to map known concepts to words for concrete concepts than for more abstract concepts (Gillette et al., 1999; Gleitman, 1990; Gleitman et al., 2005; Snedeker & Gleitman, 2004). Prior work that has explicitly focused on children's ability to verbally describe relative differences and changes in pitch also underscores the difficulty of learning words for pitch (Costa-Giomi & Descombes, 1996; Creel, 2016; Flowers & Costa-Giomi, 1991; Hair, 1977). Our findings are also consistent with the findings of a recent, large-scale cross-cultural analysis of the contents of children's early vocabularies that found that concreteness is a strong predictor of the age at which words are acquired (Braginsky et al., 2019).

Our second key finding is that children were generally successful at extending the novel word from the trained dimension to a new dimension, in essence comprehending new metaphors in real time. Once children had learned an initial meaning of the word, they were able to flexibly comprehend it when it was used with an entirely new sense, even though they were not provided with any feedback or training on this new meaning. Thus, although children struggled when they had to directly learn a new word for pitch in training (with feedback), they were able to extend a newly learned word for space to the domain of pitch (without feedback). This was true both when children were extending between height and pitch, which reflects a familiar metaphorical mapping in English, as well as when they were extending between thickness and pitch (with the notable exception of 3-year-olds, as discussed below), which reflects a metaphorical mapping found only in other languages like Turkish and Farsi. This demonstrates that children were not simply mapping the novel word to English words already in their vocabulary and relying on English-language metaphors to guide their extension of the novel word. Instead, children may have drawn on preexisting associations between pitch and space, which have been documented in both preverbal infants and preschool-aged children (Dolscheid et al., 2014; Mondloch & Maurer, 2004; Starr & Srinivasan, 2018; Walker et al., 2010). Because these mappings are already present in early childhood, children could utilize them to predict how a word that has been learned in one of

these domains may also be used to describe another domain. Together, our findings suggest that spatial metaphor could be a tool that scaffolds children's acquisition of words for more abstract concepts, which might otherwise be difficult to learn.

Although 3-year-olds did not reliably extend the novel word in the thickness condition, they did reliably extend the novel word in the height condition. Because our posttest measure also found that 3-year-olds did not yet understand the meaning of the words *high* and *low* as they apply to pitch, it is unlikely that these children were using their knowledge of English height-pitch metaphors to guide extension performance. Still, it was surprising that 3-year-olds failed to understand new metaphorical meanings in the thickness condition, especially given prior work with infants and preschoolers, which has demonstrated that young children map pitch to both spatial height and thickness (Dolscheid et al., 2014; Mondloch & Maurer, 2004; Starr & Srinivasan, 2018; Walker et al., 2010). One possible explanation for why 3-year-olds succeeded at extending the meaning of the word in the height condition but not in the thickness condition is that while children are learning the words *high* and *low* as they apply to pitch, attention to the height-pitch mapping may be strengthened at the expense of attention to the thickness-pitch mapping. This could also help explain why previous studies with children learning languages with height-pitch metaphors have reported conflicting findings with regard to 3-year-olds' ability to map pitch to thickness (Shayan et al., 2014; Starr & Srinivasan, 2018). A similar example of how word learning may shift children's attentional biases has been found in the domain of relational reasoning. For 3-year-old children learning English, the prevalence of nouns in the language input may increase children's attention toward object features and decrease their attention toward relational features, resulting in a temporary decline in performance on some relational reasoning tasks relative to younger children (Carstensen et al., 2019; Hoyos, Shao, & Gentner, 2016; Walker, Bridgers, & Gopnik, 2016).

Interestingly, although children successfully learned the novel word when it referred to thickness, our posttest indicated that children did not understand the words *thick* and *thin* in their spatial context. Children's difficulty understanding these words may be related to underexposure rather than to difficulty with the concepts themselves. In our posttest (summarized in Supporting Information), we found that when we asked children in the spatial thickness condition what they thought the novel word meant, they frequently provided spatial words like *big* and *small*, which demonstrates that they were picking up on the spatial features that were relevant to the novel word, even if they did not specifically know the words *thick* and *thin*. In comparison to the words *thick* and *thin, big* and *small* are much more frequently used in American English (raw frequency counts in the Corpus for Contemporary American English for big: 600,364 and small: 382,563; raw frequency counts for thick: 41,420 and thin: 44,274; Davies, 2009). Therefore, children may be slow to learn the words *thick* and *thin* because they do not encounter them very often in their environment, even though they understand the spatial concepts that these words reference.

We chose to focus on pitch as a test case for understanding how spatial metaphors contribute to word learning. Although pitch is not a purely abstract concept because it does have a direct perceptual correlate (i.e., sound frequency; Brysbaert et al., 2014), the perceptual experience of sounds tends to be more fleeting than the experience of spatial height or

thickness because tones have a limited temporal duration and two tones cannot be simultaneously compared. This relative abstractness may explain why learning labels for pitch, as well as acquiring a conceptual understanding of pitch, appears to be a difficult task for children (Andrews & Madeira, 1977; Costa-Giomi & Descombes, 1996; Flowers & Costa-Giomi, 1991; Hair, 1977, 1981). Pitch also shares important features with more quintessential abstract concepts, like time and number: spatial language is frequently used to describe each of these domains, and infants and young children have intuitions about how these domains align with space (Dolscheid et al., 2014; Mondloch & Maurer, 2004; Starr & Srinivasan, 2018; Walker et al., 2010). Our findings thus open the possibility that spatial metaphors may facilitate word learning in other abstract domains, like time and number.

The present findings provide insight into why metaphorical language is so prevalent across languages, and why metaphorical language is often asymmetric, such that spatial language is more often co-opted to describe more abstract concepts than the reverse (Lakoff & Johnson, 1980; Xu et al., 2017). If dedicated, unambiguous words for abstract concepts are difficult to learn, as our findings suggest, then these types of words will be less to enter the lexicon and be transmitted across generations. By contrast, if spatial words are relatively easy to learn and if speakers share intuitions on how these words can be metaphorically extended to more abstract meanings (e.g., based on existing cross-domain associations available in infancy), then spatial metaphors may develop in a lexicon because this will make the lexicon more learnable by scaffolding the acquisition of words for abstract concepts. Indeed, our findings suggest that a child who has learned that *high* labels space will have an advantage in understanding *high* in the context of pitch, an advantage that would not exist if English had separate, unambiguous words for space and pitch. This account predicts that, within a lexicon, it should be relatively common for words for space to accrue additional, more abstract meanings over time, and relatively rare for words for abstract concepts to arise *de novo* and accrue more concrete meanings (Gerrig & Gibbs Jr, 1988; Srinivasan & Carey, 2010; Thibodeau & Durgin, 2008). Supporting this prediction, recent work suggests that over the past millennium, English words with more external and concrete meanings have been metaphorically extended to more internal and abstract meanings more often than the reverse (Xu et al., 2017). Notably, this explanation for asymmetry in metaphorical language —which appeals to the pressures of word learning on historical semantic change—contrasts with views in which metaphorical asymmetry reflects a conceptual asymmetry in which abstract concepts are understood in terms of, and grounded in more concrete concepts (Boroditsky, 2000; Casasanto, 2010; Gibbs Jr et al., 1994; Lakoff & Johnson, 1980).

Although we have argued that spatial word meanings may generally be easier to learn and subsequently extend to other domains, it is important to acknowledge that the types of concepts that are most salient are not constant across all learning environments. For example, in some situations, children more readily learn words for action-based roles rather than perceptual features (Yin & Csibra, 2015). In addition, not all spatial concepts are necessarily easy to learn. For example, the concepts of *left* and *right* are notoriously difficult for children to learn, and these words are typically not mastered until the age of 5–7 years (Cole & Cole, 1996; Hermer-Vazquez, Moffet, & Munkholm, 2001). Spatial concepts vary in their relative concreteness, and this may influence which types of spatial words are most easy to learn and thus most likely to serve as source domains for metaphorical extension.

One limitation of the present work is that we did not explicitly test whether it is more difficult for children to learn an unambiguous word for pitch than it is to extend a known spatial word to the domain of pitch. This would be an interesting extension for future work, as it would more directly test the idea that spatial metaphors make the lexicon more learnable. In addition, we lacked the statistical power needed to examine the effect of markedness on children's word learning and extension performance. Thus, another open question is whether there is an advantage for learning or extending the word referring to the unmarked pole of a dimension (e.g., high, thick) relative to the marked pole (low, thin). Finally, although our frequentist statistical analyses suggested that most of our predictors of interest (extension direction and age) did not have significant effects on children's ability to extend the newly learned word, our Bayesian analyses did not provide strong evidence in favor of the null hypotheses. Therefore, although children were largely successful at extending the word to new contexts, additional work is needed to further understand how extension direction and other factors influence children's comprehension of new metaphorical word meanings.

In conclusion, the present results suggest that spatial metaphors contribute to word learning by enabling children to reuse already-learned words in new contexts. Although learning novel words for relatively abstract concepts is difficult, children are surprisingly flexible in their ability to extend previously learned spatial words to nonspatial contexts. This asymmetry in word learning, in combination with children's shared intuitions about how spatial dimensions map onto nonspatial dimensions like pitch, may help explain why spatial metaphors are widespread across natural languages.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Andrews ML, & Madeira SS (1977). The assessment of pitch discrimination ability in young children. Journal of Speech and Hearing Disorders, 42, 279–286. 10.1044/jshd.4202.279

Bloom P (2000). How children learn the meanings of words. Cambridge, MA: MIT Press.

Bloom L, Lightbown P, Hood L, Bowerman M, Maratsos M, & Maratsos MP (1975). Structure and variation in child language. Monographs of the Society for Research in Child Development, 40, 1–97. [PubMed: 1102959]

Boroditsky L (2000). Metaphoric structuring: Understanding time through spatial metaphors. Cognition, 75 (1), 1–28. [PubMed: 10815775]

Braginsky M, Yurovsky D, Marchman VA, & Frank MC (2019). Consistency and variability in children's word learning across languages. Open Mind, 16. 10.1162/opmi_a_00026

Brysbaert M, Warriner AB, & Kuperman V (2014). Concreteness ratings for 40 thousand generally known English word lemmas. Behavior Research Methods, 46, 904–911. 10.3758/s13428-013-0403-5 [PubMed: 24142837]

Bürkner PC (2017). brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software, 80. 10.18637/jss.v080.i01

Carstensen A, Zhang J, Heyman GD, Fu G, Lee K, & Walker CM (2019). Context shapes early diversity in abstract thought. Proceedings of the National Academy of Sciences of the United States of America, 116, 13891–13896. 10.1073/pnas.1818365116 [PubMed: 31235570]

Casasanto D (2010). Space for thinking. In Evans V & Chilton P (Eds.), Language, cognition, and space: State of the art and new directions (pp. 453–478). London: Equinox Publishing.

Casasanto D, Fotakopoulou O, & Boroditsky L (2010). Space and time in the child's mind: Evidence for a cross-dimensional asymmetry. Cognitive science, 34(3), 387–405. [PubMed: 21564218]

Clark HH (1973). Space, time, semantics, and the child. In Cognitive development and acquisition of language (pp. 27–63). Academic Press.

Cole M, & Cole S (1996). The development of children. New York, NY: W.H. Freeman.

Costa-Giomi E, & Descombes V (1996). Pitch labels with single and multiple meanings: A study with French-speaking children. Journal of Research in Music Education, 44, 204–214. 10.2307/3345594

Creel SC (2016). Ups and downs in auditory development: Preschoolers' sensitivity to pitch contour and timbre. Cognitive Science, 40, 373–403. 10.1111/cogs.12237 [PubMed: 25846115]

Dautriche I, Chemla E, & Christophe A (2016). Word learning: Homophony and the distribution of learning exemplars. Language Learning and Development, 12, 231–251. 10.1080/15475441.2015.1127163

Dautriche I, Fibla L, Fievet A-C, & Christophe A (2018). Learning homophones in context: Easy cases are favored in the lexicon of natural languages. Cognitive Psychology, 104, 83–105. 10.1016/j.cogpsych.2018.04.001 [PubMed: 29778004]

Davies M (2009). The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, architecture, and linguistic insights. International Journal of Corpus Linguistics, 14, 159–190. 10.1075/ijcl.14.2.02dav

Dolscheid S, Hunnius S, Casasanto D, & Majid A (2014). Prelinguistic infants are sensitive to space-pitch associations found across cultures. Psychological Science, 25(6), 1256–1261. [PubMed: 24899170]

Fisher C, Gertner Y, Scott RM, & Yuan S (2010). Syntactic bootstrapping. Wiley Interdisciplinary Reviews: Cognitive Science, 1, 143–149. 10.1002/wcs.17 [PubMed: 26271229]

Flowers PJ, & Costa-Giomi E (1991). Verbal and nonverbal identification of pitch changes in a familiar song by English- and Spanish-speaking preschool children. Bulletin of the Council for Research in Music Education, 107, 1–12. Retrieved from https://www.jstor.org/stable/40318416

Floyd S, & Goldberg AE (2020). Children make use of relationships across meanings in word learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10.1037/xlm0000821

Gentner D (1978). On relational meaning: The acquisition of verb meaning. Child development, 49, 988–998.

Gerrig RJ, & Gibbs Jr RW (1988). Beyond the lexicon: Creativity in language production. Metaphor and Symbol, 3(3), 1–19.

Gibbs RW Jr, Gibbs RW, & Gibbs J (1994). The poetics of mind: Figurative thought, language, and understanding. Cambridge University Press.

Gillette J, Gleitman H, Gleitman L, & Lederer A (1999). Human simulations of vocabulary learning. Cognition, 73(2), 135–176. [PubMed: 10580161]

Gleitman L (1990). The structural sources of verb meanings. Language Acquisition, 1(1), 3–55.

Gleitman LR, Cassidy K, Nappa R, Papafragou A, & Trueswell JC (2005). Hard words. Language Learning and development, 1(1), 23–64.

Hair HI (1977). Discrimination of tonal direction on verbal and nonverbal tasks by first grade children. Journal of Research in Music Education, 25, 197–210. 10.2307/3345304

Hair HI (1981). Verbal identification of music concepts. Journal of Research in Music Education, 29(1), 11–21. 10.2307/3344675

Hermer-Vazquez L, Moffet A, & Munkholm P (2001). Language, space, and the development of cognitive flexibility in humans: The case of two spatial memory tasks. Cognition, 79, 263–299. 10.1016/S0010-0277(00)00120-7 [PubMed: 11165214]

de Hevia MD, Izard V, Coubart A, Spelke ES, & Streri A (2014). Representations of space, time, and number in neonates. Proceedings of the National Academy of Sciences, 111(13), 4809–4813.

Hoyos C, Shao R, & Gentner D (2016). The paradox of relational development: Could language learning be (temporarily) harmful? Proceedings of the 38th annual conference of the Cognitive Science Society, Philadelphia, PA.

Huttenlocher J, Smiley P, & Charney R (1983). Emergence of action categories in the child: Evidence from verb meanings. Psychological Review, 90, 72. 10.1037/0033-295X.90.1.72

Kuznetsova A, Brockhoff PB, & Christensen RHB (2017). lmerTest package: Tests in linear mixed effects models. Journal of Statistical Software, 82. 10.18637/jss.v082.i13

Lakoff G, & Johnson M (1980). Metaphors we live by. Chicago. IL: The University of Chicago Press.

Lourenco SF, & Longo MR (2010). General magnitude representation in human infants. Psychological Science, 21(6), 873–881. [PubMed: 20431048]

Markman EM (1990). Constraints children place on word meanings. Cognitive Science, 14(1), 57–77. 10.1207/s15516709cog1401_4

Mondloch CJ, & Maurer D (2004). Do small white balls squeak? Pitch-object correspondences in young children. Cognitive, Affective, & Behavioral Neuroscience, 4 (2), 133–136.

Ramiro C, Srinivasan M, Malt BC, & Xu Y (2018). Algorithms in the historical emergence of word senses. Proceedings of the National Academy of Sciences, 115(10), 2323–2328.

Rodd JM, Gaskell MG, & Marslen-Wilson WD (2004). Modelling the effects of semantic ambiguity in word recognition. Cognitive Science, 28(1), 89–104. 10.1207/s15516709cog2801_4

Searle SR, Speed FM, & Milliken GA (1980). Population marginal means in the linear model: An alternative to least squares means. The American Statistician, 34, 216–221. 10.1080/00031305.1980.10483031

Shayan S, Ozturk O, Bowerman M, & Majid A (2014). Spatial metaphor in language can promote the development of cross-modal mappings in children. Developmental Science, 17, 636–643. 10.1111/desc.12157 [PubMed: 24636133]

Shusterman A, & Wagner L (2005). The acquisition of the space-time metaphor. San Diego, CA: Cognitive Development Society.

Snedeker J, & Gleitman L (2004). Why it is hard to label our concepts. Weaving a Lexicon, 257294.

Srinivasan M, Al-Mughairy S, Foushee R, & Barner D (2017). Learning language from within: Children use semantic generalizations to infer word meanings. Cognition, 159, 11–24. [PubMed: 27880881]

Srinivasan M, Berner C, & Rabagliati H (2019). Children use polysemy to structure new word meanings. Journal of Experimental Psychology: General, 148(5), 926. [PubMed: 30010371]

Srinivasan M, & Carey S (2010). The long and the short of it: On the nature and origin of functional overlap between representations of space and time. Cognition, 116(2), 217–241. [PubMed: 20537324]

Srinivasan M, & Rabagliati H (2015). How concepts and conventions structure the lexicon: Cross-linguistic evidence from polysemy. Lingua, 157, 124–152.

Srinivasan M, & Snedeker J (2014). Polysemy and the taxonomic constraint: Children's representation of words that label multiple kinds. Language Learning and Development, 10(2), 97–128.

Starr A, & Srinivasan M (2018). Spatial metaphor and the development of cross-domain mappings in early childhood. Developmental Psychology, 54, 1822–1832. 10.1037/dev0000573 [PubMed: 30234336]

Sweetser E (1990). From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure (Vol. 54). Cambridge University Press.

Talmy L (1988). Force dynamics in language and cognition. Cognitive science, 12(1), 49–100.

Thibodeau P, & Durgin FH (2008). Productive figurative communication: Conventional metaphors facilitate the comprehension of related novel metaphors. Journal of Memory and Language, 58(2), 521–540.

Trueswell JC, Medina TN, Hafri A, & Gleitman LR (2013). Propose but verify: Fast mapping meets cross-situational word learning. Cognitive Psychology, 66, 126–156. 10.1016/j.cogpsych.2012.10.001 [PubMed: 23142693]

Walker P, Bremner JG, Mason U, Spring J, Mattock K, Slater A, & Johnson SP (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. Psychological Science, 21(1), 21–25. [PubMed: 20424017]

Walker CM, Bridgers S, & Gopnik A (2016). The early emergence and puzzling decline of relational reasoning: Effects of knowledge and search on inferring abstract concepts. Cognition, 156, 30–40. 10.1016/j.cognition.2016.07.008 [PubMed: 27472036]

Xu F, & Tenenbaum JB (2007). Word learning as Bayesian inference. Psychological Review, 114, 245–272. 10.1037/0033-295X.114.2.245 [PubMed: 17500627]

Xu Y, Malt BC, & Srinivasan M (2017). Evolution of word meanings through metaphorical mapping: Systematicity over the past millennium. Cognitive Psychology, 96, 41–53. 10.1016/j.cogpsych.2017.05.005 [PubMed: 28601710]

Yin J, & Csibra G (2015). Concept-based word learning in human infants. Psychological Science, 26, 1316–1324. 10.1177/0956797615588753 [PubMed: 26195636]

Yu C, & Smith LB (2007). Rapid word learning under uncertainty via cross-situational statistics. Psychological Science, 18, 414–420. 10.1111/j.1467-9280.2007.01915.x [PubMed: 17576281]

Yurovsky D, & Frank MC (2015). An integrative account of constraints on cross-situational learning. Cognition, 145, 53–62. 10.1016/j.cognition.2015.07.013 [PubMed: 26302052]
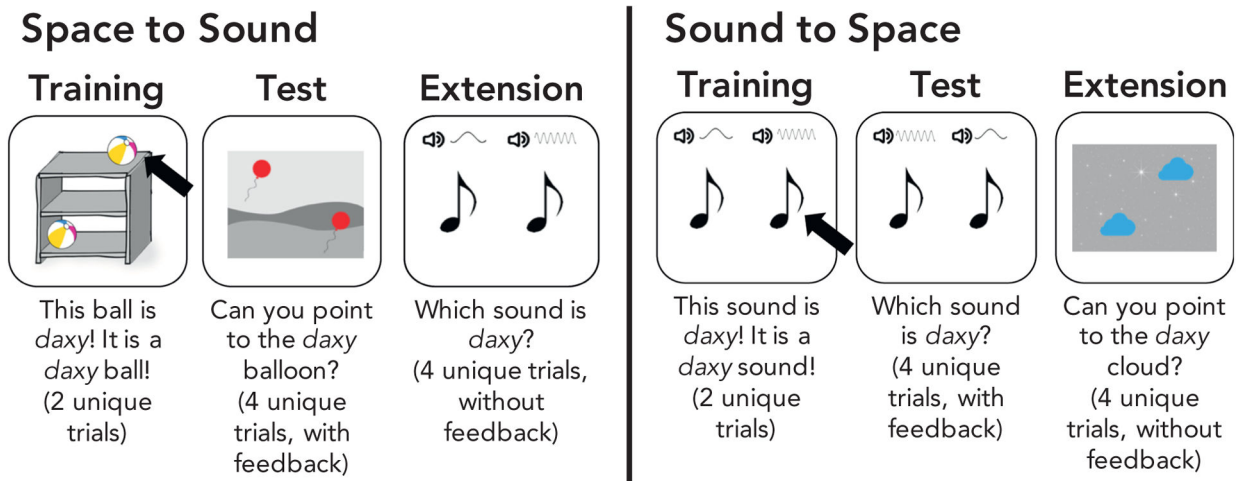
Figure 1.
Schematic of the paradigm in the space-to-sound condition (left) and sound-to-space condition (right). Stimuli are shown for the height condition with the novel word trained on the unmarked pole (*high*). Arrows are shown here for illustration purposes only, to indicate pointing by the experimenter, and the icons indicating audio frequency with squiggles were also not present during test.
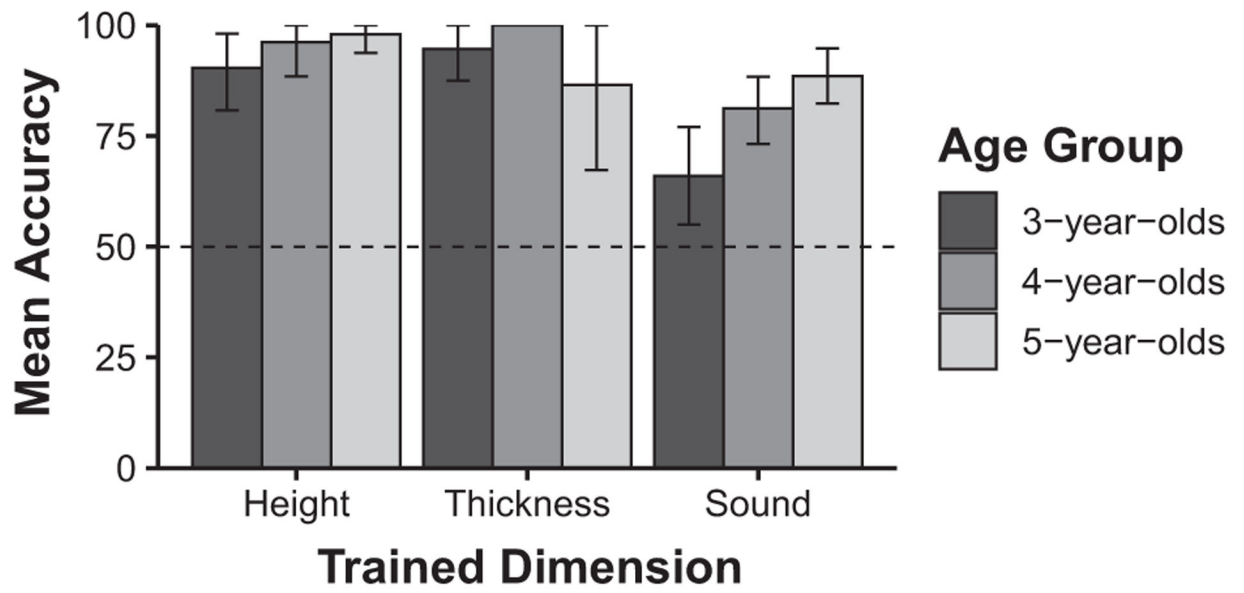
Figure 2.
Children's performance in the word training phase by age group and trained dimension.
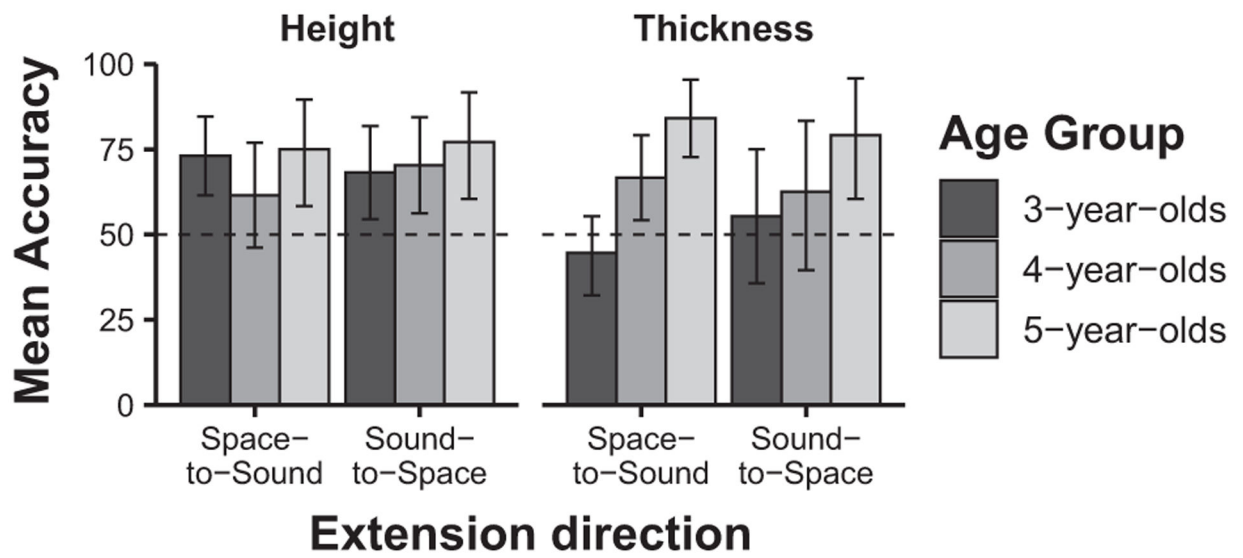Error bars indicate 95% CIs.

Figure 3.
Children's performance in the word extension phase by age group, extension direction, and condition. Note that the age group by direction interaction is not significant but is plotted for comparison with the training data. Error bars indicate 95% CIs.
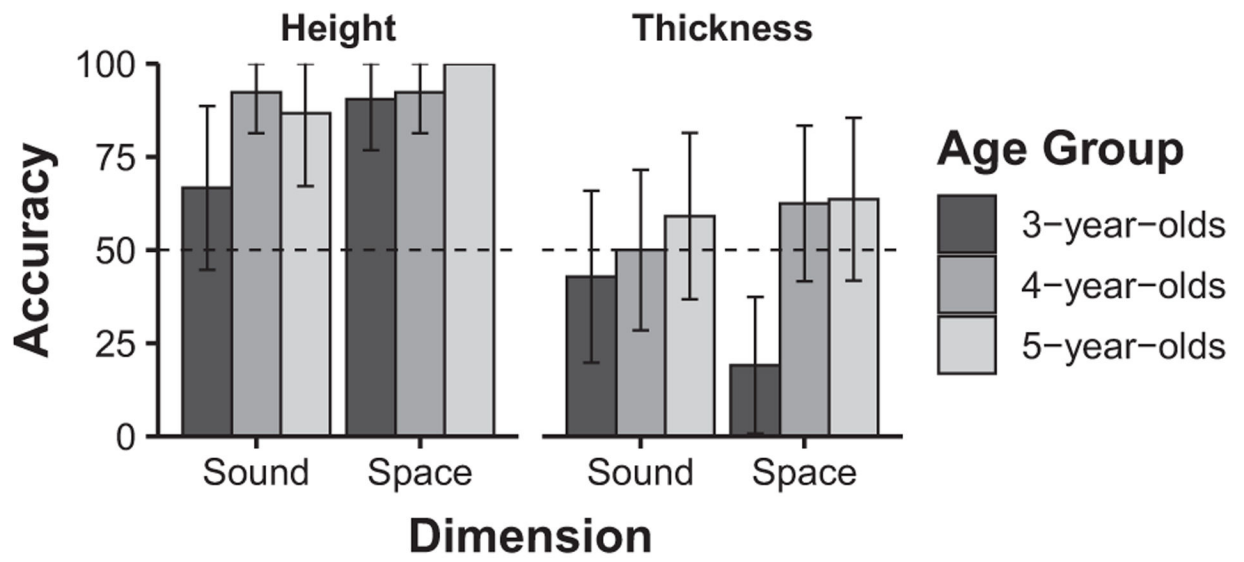
Figure 4.
Children's accuracy in the word comprehension task. Error bars indicate 95% CIs.