Why are Infants Precocious Language Learners? Implications for Adult Second-Language Learning

Carolyn Quam, Andrew Lotto, Kimberly Golisch, Celeste Gallegos, and LouAnn Gerken[1]

1. Introduction

Typically developing infants and young children show better language learning outcomes than typical adults learning a second language (L2; Newport, 1990; Flege, 1995, 1999). It has been suggested that the infant brain might be particularly well suited for language learning (Newport, 1990; Thompson-Schill et al., 2009). We propose that infants' reliance on implicit rather than explicit learning may in large part account for their successful language learning. We argue that infants are particularly good at flexibly learning new linguistic structure because they are less directed processors of information than older learners; they are more likely to let the particular structure of the present language input guide their learning. This contrasts with adults, who tend to identify one salient dimension to attend to, particularly early in learning (Johansen & Palmeri, 2002).

In typical development, the neural structures supporting explicit learning (the Declarative Memory system) develop more slowly than those supporting implicit learning (the Procedural Memory system; Jones & Herbert, 2006; Richmond & Nelson, 2007). Infants' immature explicit learning, and consequent reliance on implicit learning—likely in combination with limited experience—appears to cause their memories for both words (Singh, Morgan, & White, 2004; Singh, White, &

---

[1] University of Arizona; corresponding author: Carolyn Quam, cmquam@email.arizona.edu
Affiliations: Department of Speech, Language, and Hearing Sciences (Quam, Lotto, & Gallegos); Department of Psychology (Quam, Gerken); Undergraduate Program in Neuroscience and Cognitive Science (Golisch); Department of Linguistics (Gerken).

Morgan, 2008; Thiessen & Yee, 2010) and visual information (e.g., Robinson & Pascalis, 2004; Hayne, Boniface, & Barr, 2000) to contain more "irrelevant" contextual detail than the memories of older learners. Younger learners are also more likely to treat non-phonological dimensions as relevant to tasks like word learning (e.g., Singh, Hui, Chan, & Golinkoff, 2014; Hay, Graf Estes, Wang, & Saffran, 2014; Namy & Waxman, 1998; Woodward & Hoyne, 1999).

While young learners' overweighting of non-phonological dimensions has been couched in the literature as immaturity, we propose that it also reflects "open-mindedness" about what is relevant to language learning, which is crucial for learning new linguistic structure. Learners will be more successful at learning new linguistic categories if they are willing to attend to whatever dimensions are important to the particular categories, rather than attending only to dimensions that their prior experience with language indicates are relevant. There is good reason to suspect that adults are less open minded about attending to new dimensions than infants. For adults, reliance on explicit- rather than implicit-learning strategies (Filoteo, Lauritzen, & Maddox, 2010; Ashby & Maddox, 2010), in combination with extensive experience with the native language (Best & Tyler, 2007; Best, McRoberts, & Goodell, 2001; Flege, 1995), may lead to less flexibility in attending to and integrating non-native dimensions when learning new categories. However, adults can be shifted to learn in a more infant-like, open minded manner. For example, adults integrate two dimensions to learn categories more effectively if their access to explicit learning is blocked by taxing working memory (Filoteo et al., 2010). In **Experiment 1**, we ask: can we make adults learn more like infants, by blocking their access to both native-language biases and the explicit-learning mode?

Experiment 2 reports a preliminary study with typically developing preschoolers. The work is motivated by evidence that preschoolers with Specific Language Impairment (SLI) may share commonalities with adults in relying more heavily on explicit (vs. implicit) learning than young typically developing learners. A recent proposal argues that the core deficit in Specific Language Impairment (SLI) is the procedural-memory system, which supports and underlies a form of implicit learning (Lum, Conti-Ramsden, Page, & Ullman, 2012; Lum, Conti-Ramsden, Morgan, & Ullman, 2014). Children with SLI display impairments to language grammar (e.g., Rice & Wexler, 1996) that sometimes co-occur with phonological deficits. The primary source of grammatical deficits in

SLI has recently been proposed to be weak implicit learning. Lum, Conti-Ramsden, Page, and Ullman (2012) found that a large sample of children with SLI showed impaired procedural memory (which is related to implicit learning). The authors found spared declarative memory (related to explicit learning) in the children with SLI. Other studies have also found evidence of impaired implicit learning in children and adults with SLI (Tomblin, Mainela-Arnold, & Zhang, 2011; Evans, Saffran, & Robe-Torres, 2009; Hsu, Tomblin, & Christiansen, 2014; Richardson, Harris, Plante, and Gerken, 2006; though see Plante et al., 2010). Impaired implicit learning would put children with SLI in striking contrast with typically developing infants, for whom explicit memory develops more slowly over the first year than implicit memory (Jones & Herbert, 2006). Impaired implicit learning in SLI also dovetails with our proposal that a primary source of adults' language-learning difficulty is their over-reliance on explicit vs. implicit learning mechanisms, suggesting parallels between the two groups. We thus predict that, in a category-learning task in which integrating multiple dimensions is best done using implicit-learning mechanisms, children with SLI will learn less optimally than their peers.

The primary source of language-learning difficulties in SLI is an active debate in the developmental language disorders field. While some evidence suggests that implicit learning is the core deficit (Lum, Conti-Ramsden, Page, & Ullman, 2012; Lum, Conti-Ramsden, Morgan, & Ullman, 2014; Tomblin, Mainela-Arnold, & Zhang, 2011; Evans, Saffran, & Robe-Torres, 2009; Hsu, Tomblin, & Christiansen, 2014), other evidence suggests that working memory is the core deficit (Weismer, Evans, & Hesketh, 1999; Leonard et al., 2007; Montgomery, 2003; Montgomery & Evans, 2009). It also may be that deficits to multiple memory systems contribute to language-learning difficulties in SLI (Lum & Conti-Ramsden, 2013). Applying the present work to SLI will shed light on this debate.

2. Experiment 1

In a sound-category-learning task, we attempted to shift adults so that they would learn in a more infant-like way: attending more to a non-native sound dimension than they are typically biased to do, and integrating two dimensions in category learning. We manipulated two factors that we proposed would impact their ability to attend to both dimensions: their access to (1) native-language biases and (2) explicit learning.

2.1. Method

2.1.1. Participants

We tested 52 undergraduate participants (36 female; age range of 18-22 years; *M* = 18.6 years; *SD* = 0.8 years), 13 in each of 4 conditions (created by crossing the language-context and the working-memory manipulations). Participants were recruited from the Undergraduate Psychology Subject Pool at the University of Arizona. All included participants were native English speakers as reported on a questionnaire. Seventeen more were excluded for being under 18 (2), more than 3 SDs above the mean age (i.e., over 26; 3), for hearing loss (4), learning disabilities (4), and experimenter error/computer malfunctions (4). (In addition, 42 were excluded for being bilingual or not a native-English speaker.)

2.1.2. Auditory Stimuli

Adults learned two sound categories (synthesized isolated vowels) that varied on both a native-language dimension (second-formant frequency, F2, which is the primary cue for differentiating the vowels /i/ and /u/ in English) and a non-native dimension (fundamental frequency—F0—perceived as pitch; see **Figure 1** for category structure). Ideal learning of the two categories (displayed in the two-dimensional space in **Figure 1** in orange squares and green triangles, respectively) would involve integrating both dimensions and weighting them equally for categorization. Participants' weighting of each dimension (pitch and F2) was assessed in the training phase using responses to a 6x6 grid of stimuli that were distributed across the entire two-dimensional space (see **Figure 1**, blue diamonds).

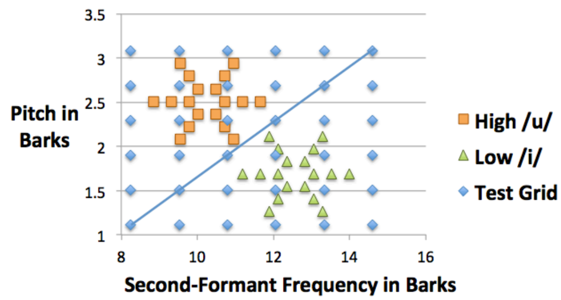

Figure 1. Training categories (orange squares, green triangles) and test grid for computing cue weights (blue diamonds).

Sounds were isolated vowels synthesized using Klatt (Klatt & Klatt, 1990, via the Praat program; Boersma & Weenink, 2008; Weenink, 2009). Each sound was 0.4 seconds in duration, with a final pitch decrement to 75% of the original pitch value, to make it sound less robotic. The last 10 milliseconds of each soundfile were also amplitude-ramped using a Matlab script, so that the soundfile did not end abruptly in a "clicking" sound (as it would otherwise if the soundwave was cut off before or after a zero crossing).

The two-dimensional space in which sounds were distributed spanned the pitch range from 117 Hz to 280 Hz and the F2 range from 1060 Hz to 2195 Hz. These values were based loosely on the first author's pitch and F2 range, so were meant to simulate the range of producible values for one female speaker. However, the endpoints were exaggerated somewhat so that the categories (which did not extend to the endpoints) contrasted sufficiently on each dimension.

Stimuli were spaced according to the Bark scale (which incorporates logarithmic compression meant to simulate the compression of higher frequencies in the human auditory system; Zwicker, 1961), to ensure that distances between the stimuli were perceptually equivalent along the entire span of each dimension.

Each of the two training categories was composed of 18 stimuli: 6 equally spaced points along each of 3 concentric circles (see **Figure 1**). The two categories were placed on either side of the diagonal dividing the entire two-dimensional space. Because the concentric circles increased in size as they moved away from the center point, the stimuli were denser toward the center of the distribution and tapered off towards the edges. This Gaussian-like distribution attempted to simulate the structure of phonetic categories in language (Holt & Lotto, 2006).

The training stimuli were mapped onto the grid using trigonometry to create the three radiating circles, each with progressively larger radii (see **Figure 1**). We first used two formulas to locate each of the thirty-six stimuli in an abstract two-dimensional space: (1) x-coordinate = H + r*cos(t); (2) y-coordinate = K + r*sin(t), where H is the x-coordinate and K is the y-coordinate of the center of the circle, and r is the radius for each concentric circle. The radii across the 3 circles increased by a constant. After each of the 36 stimuli was positioned in the abstract two-dimensional space, that space was then transformed into the ranges of pitch and F2 in Barks (see above). The 36 test stimuli were equally spaced (in Barks) along a 6x6 grid spanning the entire range of both dimensions.

2.1.3 Apparatus and Procedure

The experiment was administered on a Mac Mini computer with an attached Dell monitor, Apple keyboard, and headphones. Adults were told they were learning sounds from an alien language, Qixian. In each trial, a sound played and participants pressed one of two buttons to categorize it. In training, they received feedback (smiley/frowny faces) as to whether they had correctly assigned each sound to the category.

Across participants, we manipulated two factors that we predicted would facilitate adults' attention to the non-native dimension and integration of the two dimensions: (1) reducing access to native-language phonetic biases by placing people in a foreign-language context, and (2) reducing access to explicit-learning abilities by taxing working memory. The language-context manipulation was instantiated by placing either native-language letters ("ee" vs. "ooh") or unfamiliar/foreign symbols (ξ and ♉) on the response keys that participants used to assign each sound to a category (see **Figure 2**). The working-memory task, interspersed between category-learning trials, involved remembering digits. To tax working memory, in between each category-learning training trial, an array of 4 numbers flashed on the screen followed by a probe (Filoteo et al., 2010). Participants responded as to whether the probe was in the original array. If their overall accuracy was below 90% after any trial, a message appeared on the screen instructing them to try harder.

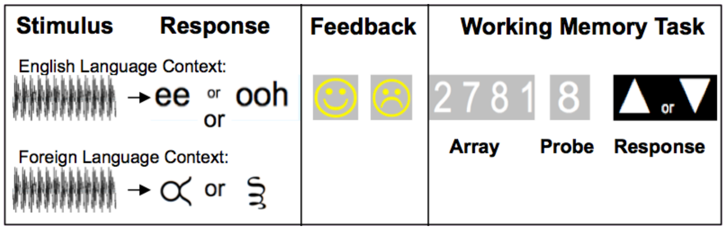

Figure 2. Structure of training trials.

Each of the 36 training stimuli appeared once each during each of 3 training blocks. After each training block, participants completed a test block (making 3 alternating training and test blocks, for a total of 6 blocks of trials). In test blocks, the 36 test stimuli were presented once each. Test trials were similar to training trials except that no feedback was provided and the working-memory task was removed. Test responses allowed us to assess participants' cue weight for each dimension, by correlating responses to test stimuli—which comprised a 6x6 grid spanning the entire perceptual space (see **Figure 1**—blue diamonds indicate test stimuli)—with both the pitch and F2 values of the stimuli (Holt & Lotto, 2006).

2.2. Results

We first assessed adults' accuracy across the three blocks of training trials, to ensure that participants had learned the novel sound categories. Overall accuracy across all 52 participants across all three blocks was 85.2% (*SD*, 7.1%), and accuracy scores ranged from 66.7% to 98.1% (chance performance being 50%). Accuracy was significantly higher for the high-"ooh" category (high pitch, low F2) than for the low-"ee" category (low pitch, high F2; $t(51) = 3.94$, $p < .001$; mean difference, 5.3%), but both were well above chance levels. Overall accuracy increased across the three test blocks, with significantly higher accuracy in block 2 than block 1 ($t(51) = 2.8$, $p < .01$) and in block 3 vs. block 1 ($t(51) = 3.7$, $p < .001$), though blocks 2 and 3 did not differ significantly (means (SDs): block 1, 84.3% (8.8%); block 2, 87.5% (7.2%); block 3, 88.1% (8.0%)). Across participants, accuracy between the three blocks was highly correlated, indicating stable individual differences in accuracy (all three $r > 0.5$; all $p < .001$).

Half of participants (N=28) also completed the working-memory task in between category-learning trials. For these participants, we also assessed overall accuracy on the working-memory task, to ensure that they were attending sufficiently to the task that their working-memory resources were in fact taxed. Overall working-memory accuracy was high (*M,* 93.9%; *SD,* 5.9%), and ranged from 70.4% to 98.1% correct.

Adults' responses on test trials enabled us to assess their weighting of the two dimensions in their category judgments. To assess cue weights, we correlated each participant's category judgments for the set of 36 test stimuli with the pitch vs. F2 values of the stimuli, respectively (see Holt & Lotto, 2006, for a similar procedure). First, we assessed whether any participants had failed to strongly associate either cue with the two categories, by identifying the strongest of the two correlations (in the correct direction) for each participant. The mimimum value for the highest correlation coefficient across participants was $r = 0.34$ ($p < .05$).

We next asked whether cue weights became stronger across the three test blocks. Weighting of the pitch cue increased significantly between blocks 1 and 2 ($t(51) = 2.7$, $p < .01$), but not between blocks 2 and 3 (means (SDs): block 1, 0.23 (0.35); block 2, 0.32 (0.31); block 3, 0.30 (0.32)). Weighting of F2 did not change across blocks (overall mean, 0.58; *SD,* 0.41).

Cue weights were averaged across the 3 test blocks for comparison across the four between-subjects conditions, and normalized so that the absolute-values of the weights of both cues for each participant summed to 1 (though the +/- signs of individual cue weights were retained after normalization in order to differentiate between correctly associated—positive—cue weights and incorrectly associated—negative—cue weights). Analyzing normalized cue weights enabled us to take into account not just raw weights for each dimension, but the *relative* weighting of each dimension, which is more informative for the question of degree of cue integration.

Normalized cue-weights, for pitch and F2 respectively, were entered as dependent variables into analyses of variance (ANOVAs) with between-subjects predictors Language Context (Mandarin or English) and Working Memory (taxed or intact). (Note that data patterns were similar for raw cue weights.) For both ANOVAs, we found a main effect of Language Context (both $F(1,48) > 28$, both $p < .001$). This main effect reflected the fact that participants weighted pitch more heavily and F2 less heavily when foreign symbols were on the response keys (pitch mean (SD): 0.58 (0.43); F2: 0.30 (0.42)) than when "ee" and "ooh" were on the response keys (pitch: 0.11 (0.08); F2: 0.89 (0.08)). Additionally, when F2 weights were the dependent variable, we found a marginal main effect of Working Memory ($F(1,48) = 4.03$, $p = 0.051$) and a significant interaction between Working Memory and Language Context ($F(1,48) = 4.43$, $p < .05$). Follow-up t-tests to investigate the interaction revealed that the working-memory manipulation decreased weighting of F2 in the foreign-language context ($t(23.6) = 2.1$, $p < .05$; working-memory intact: 0.46 (0.42); working-memory taxed: 0.13 (0.37)) but not in the English-language context ($t(22.6) < 1$).

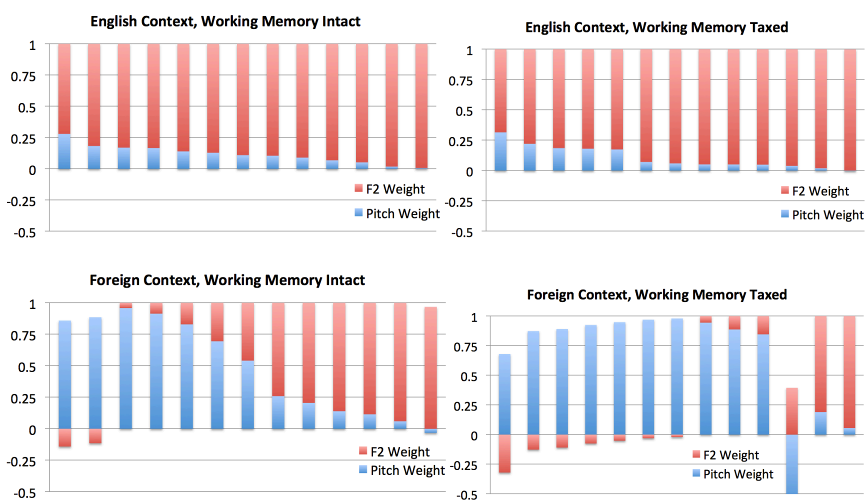

Figure 3. Results in the 4 conditions. Each bar represents a participant's normalized weights of F2 (in red) and pitch (in blue).

2.3. Discussion

Blocking adults' access to both native-language categories and explicit learning reduced reliance on a native dimension, F2. Thus, adults reduced reliance on a native dimension in favor of a non-native dimension when their access was blocked to both native-language biases and explicit-learning strategies. This suggests that both factors interacted to promote more open-minded learning. However, there is an important caveat: adults across the four conditions showed fairly minimal integration of the two cues. Instead, individual adults tended to rely primarily on either pitch or F2. Thus, while we found that adults could be shifted away from native-language biases, we did not find equal weighting of both cues in any condition, which would have represented optimal learning of the two categories.

3. Experiment 2

Experiment 2 is a preliminary study with typically developing 4- to 5-year-olds using a paradigm that we plan to use with children with SLI. In the preliminary study, we ask: can typically developing preschoolers learn the same sound categories that adults learned in Experiment 1? How do their cue-weighting strategies compare to adults'? Will preschoolers' category-learning success change across age?

3.1. Method

3.1.1. Participants

Four- and five-year-olds were recruited from a database maintained by the Tweety Language Development Lab at the University of Arizona. Twenty-three children (19 females, 4 males) between the ages of 49 months, 26 days and 73 months, 8 days (*M:* 60 months, 12 days; *SD*: 7 months, 28 days) participated in a child-friendly version of the adult experiment. One additional child was excluded because of a sensory-input disorder reported by the parent.

3.1.2. Auditory Stimuli

Auditory stimuli were very similar to those used in Experiment 1. Twelve children were trained with the exact same sound categories used in Experiment 1 (in the "Original" condition). After conducting Experiment 1, we realized that the original categories were structured so that one could achieve slightly higher accuracy using only F2 to categorize the training stimuli than using only pitch. Thus, the next 11 children were trained with categories that were rotated slightly from the originals so that they were perfectly symmetrical along the diagonal axis (in the "Symmetrical condition"; see **Figure 4**), so that either unidimensional strategy would lead to the same level of categorization accuracy. The two groups of children (tested in the "Original" vs. "Symmetrical" conditions) did not differ significantly in training accuracy or test cue weights, so we collapsed across groups for all analyses.

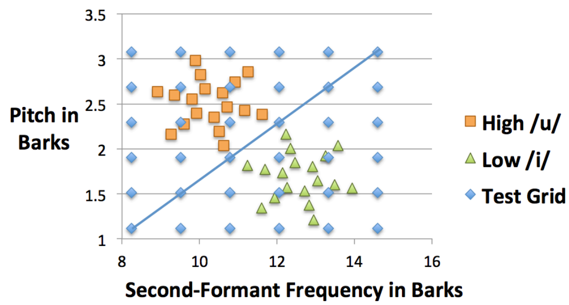

Figure 4. Training categories (orange squares, green triangles) for the Symmetrical version were perfectly symmetrical around the diagonal separating the two categories. The test grid (blue diamonds) was identical to the Original version.

3.1.3 Apparatus and Procedure

Children participated in a child-adapted version of the category-learning task from the adult preliminary study. The experiment was administered on a Mac Mini computer with an attached Dell monitor, Apple keyboard, and child headphones. We made several modifications to make the task more appropriate for 4- to 5-year-olds. Children were introduced to a monster who was sometimes hungry and sometimes thirsty. If he was hungry, he wanted an apple. If he was thirsty, he wanted milk (see **Figure 5**). Children were told that the monster, Leonard, talked in a funny way, so they had to really listen and try to figure out what he was asking for. They responded on each trial by pressing a key depicting either the milk or the apple, and they received feedback: either a smiley face or a frowny face (as in Experiment 1). Each training phase consisted of 36 stimuli, randomly sampled from the two training categories. In the test phase, the feedback was removed and the 36 stimuli came from the test grid. **Figure 6** is a diagram of the within-trial structure for the training and test phases. Whereas adults completed 3 [training + test] phases, children completed only 2, because of the need to keep the experiment under 30 minutes (length was child-dependent but typically about 25 minutes).


Figure 5. Visual stimuli in Experiment 2.

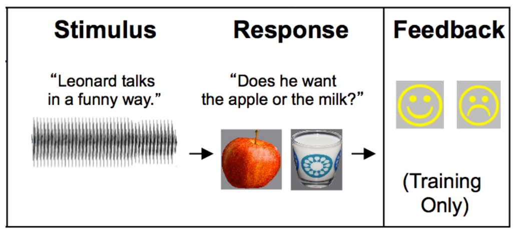

Figure 6. Trial Structure in Experiment 2.

3.2. Results

We first assessed children's accuracy across the two blocks of training trials, to ensure that participants had learned the novel sound categories. Overall accuracy across all 23 participants across all three blocks was 66.4% (*SD*, 16.5%), and accuracy scores ranged from 36.1% to 94.4% (chance performance being 50%), indicating substantially more variability in training accuracy than for adults. Accuracy did not differ for the two categories, nor did it change significantly across the two training blocks. Across participants, accuracy between the two blocks was highly correlated, indicating stable individual differences in accuracy (r = 0.76, $p < .001$). Overall accuracy also increased significantly with age (r = 0.53, $p < .01$); see **Figure 7**.

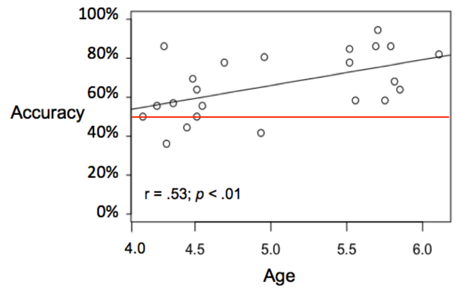

Figure 7. Children's training accuracy was significantly correlated with age (red line indicates chance, 50%; black line plots a linear-regression model).

Children's responses on test trials enabled us to assess their weighting of the two dimensions in their category judgments. As in Experiment 1, to assess cue weights, we correlated each participant's category judgments for the set of 36 test stimuli with the pitch vs. F2 values of the stimuli, respectively. Weights of each cue across all 23 participants increased between block 1 and block 2, but not significantly. Thus, we collapsed across blocks for all subsequent analyses. Overall weighting of F2 across participants was 0.26 (*SD*, 0.36); overall weighting of pitch was 0.11 (*SD,* 0.24), but these weights did not differ significantly ($t(22) < 1.6$, $p = 0.13$). Unlike adults, all of whom reliably correlated at least one dimension with category judgments, just under half of children (11/23) showed significant (10 children) or marginal (1 child) correlations for at least one dimension (see **Figure 8, top;** the highest cue-weight for these 11 children ranged from $r = 0.30$ to $r = 0.87$). Like adults in Experiment 1, most of these 11 children relied primarily on one cue (usually F2) or the other, rather than integrating both dimensions equally into their category judgments.

Because children were much more variable than adults in whether they learned categories (see training accuracy above) and in how strongly they weighted dimensions, we analyzed *raw* cue weights rather than normalizing weights so that they summed to 1. Normalizing very weak correlations could lead to the false appearance of cue integration. Raw weights of both F2 ($r = 0.44$, $p < .05$) and pitch ($r = 0.51$, $p < .05$) increased significantly with age (see **Figure 9**).

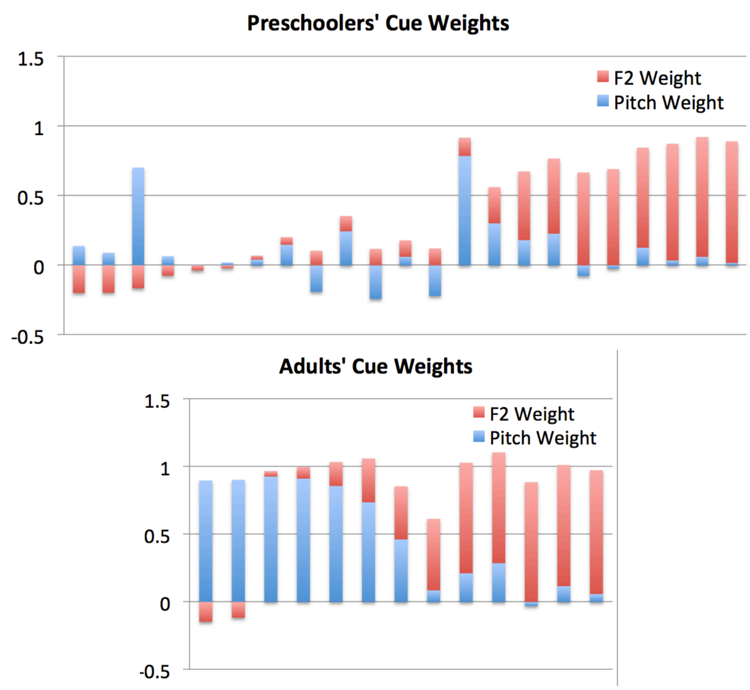

Figure 8. Children's (top) and adults' (bottom) raw cue weights for F2 (in red) and pitch (in blue); adults are from the foreign-language, intact-working-memory condition of Experiment 1.

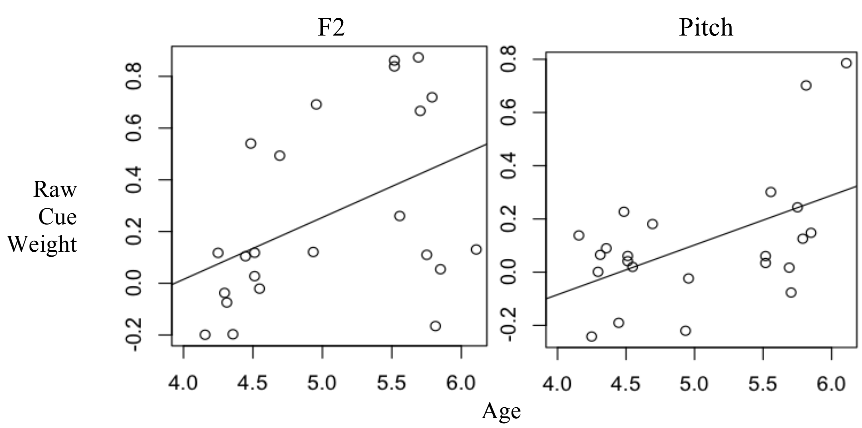

Figure 9. Children's weights for both F2 (left) and pitch (right) were significantly correlated with age.

The child-friendly experiment, in which children assigned sounds to the “milk” or “apple” categories, was most similar to the Experiment 1 condition in which adults categorized sounds using foreign response-key symbols (because of their arbitrary relationship with sound properties) in the absence of the working-memory task. Thus, we compared children’s raw weights of both F2 and pitch to adults’ raw cue weights in the foreign-language, no-working-memory condition (see **Figure 8**). Weights of both dimensions increased numerically between preschool and adulthood, but the difference was only statistically reliable for the non-native dimension, pitch ($t(17.5) = 3.3$, $p < .005$; children’s mean (SD): 0.11 (0.24), adults: 0.49 (0.39)); the two age-groups did not differ significantly in F2 weights ($t(22.8) < 1.5$, $p > .2$; children: 0.26 (0.36), adults: 0.43 (0.40)).

3.3. Discussion

In a child-friendly adaptation of the category-learning task with 4- and 5-year-olds, we found that both training accuracy and cue weights of F2 and pitch increased significantly with age. Just under half of children (11/23) reliably associated at least one cue with the categories (in contrast to adults, all of whom did so). These 11 children, like adults in Experiment 1, showed fairly minimal integration of the two cues. Instead, individual children tended to rely primarily on either pitch or F2. When compared to adults from Experiment 1 who learned categories under the most analogous conditions (with foreign symbols on the response keys and intact working memory), children weighted both dimensions more weakly, but this difference was significant only for the pitch dimension.

4. General Discussion

In Experiment 1, we found that blocking access to both native-language biases and explicit-learning abilities reduced adults’ reliance on the native acoustic dimension, second-formant frequency (F2). Adults thus reduced reliance on a native dimension in favor of a non-native dimension when their access was blocked to both native-language biases and explicit-learning strategies. This suggests that both factors interacted to promote more open-mindedness in interpretation of a non-native acoustic dimension. In Experiment 2, children showed evidence of learning the same categories as adults. Children's category-learning accuracy and the strength of their cue weights was correlated with age, and their accuracy and cue weights were also somewhat weaker than adults'.

Neither adults nor children strongly *integrated* the two dimensions. Instead, most participants relied primarily one dimension or the other to make category judgments. Integration of both dimensions would be the strongest evidence of infant-like, open-minded learning, since the categories were designed to vary equally on both dimensions. We might see more cue integration if categories were more overlapping along the diagonal boundary. The current category structure put minimal pressure on participants to integrate dimensions, since fairly high accuracy could be achieved by using either unidimensional strategy (relying solely on F2 or on pitch; see **Figure 1**). In ongoing work with both adults and children, we are using a category structure for which high accuracy *requires* integrating both dimensions equally, because categories are more spread along the diagonal (two-dimensional) boundary.

We anticipate that investigating the conditions under which different groups of learners can shift their attention to a non-native dimension and integrate it with a native dimension in sound-category learning will ultimately lead to insights that can improve language instruction for adult L2 learners, as well as to shed light on mechanisms of typical development. In addition, we plan to explore implications of the present work for developmental Specific Language Impairment (SLI). Cue integration in category learning has been argued to require implicit learning (e.g., Filoteo et al., 2010), upon which both typical adult learners and children with SLI may rely less heavily than typically developing children.

Acknowledgements

We owe tremendous thanks to the families who participated in Experiment 2 and the undergraduates who participated in Experiment 1. Funding was provided by NIH-NIDCD K99DC013795 and NIH-NICHD F32 HD065382 grants to CQ, and NSF grant 0950601 to LAG. We also thank Elena Plante and Todd Maddox for providing advice on the project as part of their roles as mentors/advisors on CQ's K99 award.

References

Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy of Sciences, 1224,* 147-161.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13-34). Amsterdam: John Benjamins.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America, 109,* 775-794.

Boersma, P., & Weenink, D. (2008). Praat: doing phonetics by computer (Version 5.0.30) [Computer program]. Retrieved from http://www.praat.org/.

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research, 52,* 321-335.

Filoteo, J. V., Lauritzen, S., & Maddox, W. T. (2010). Removing the frontal lobes: The effects of engaging executive functions on perceptual category learning. *Psychological Science, 21,* 415-423.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research (pp. 233-276). Timonium, MD: York Press.

Flege, J. E. (1999). Age of learning and second language speech. In D. Birdsong (Ed.), Second language acquisition and the Critical Period Hypothesis. (pp. 101-131). Mahwah, NJ: Lawrence Erlbaum Associates.

Hay, J. F., Graf Estes, K., Wang, T., & Saffran, J. R. (2014). From flexibility to constraint: The contrastive use of lexical tone in early word learning. *Child Development,* DOI: 10.1111/cdev.12269.

Hayne, H., Boniface, J., & Barr, R. (2000). The development of declarative memory in human infants: Age-related changes in deferred imitation. *Behavioral Neuroscience, 114,* 77-83.

Holt, L. L., & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *Journal of the Acoustical Society of America, 119,* 3059-3071.

Hsu, H. J., Tomblin, J. B., & Christiansen, M. H. (2014). Impaired statistical learning of non-adjacent dependencies in adolescents with specific language impairment. *Frontiers in Psychology, 5,* doi: 10.3389/fpsyg.2014.00175.

Johansen, M.K., & Palmeri, T.J. (2002). Are there representational shifts during category learning? *Cognitive Psychology, 45*, 482–553.

Jones, E. J. H., & Herbert, J. S. (2006). Exploring memory in infancy: Deferred imitation and the development of declarative memory. *Infant and Child Development, 15,* 195-205.

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America, 87,* 820-857.

Leonard, L. B., Weismer, S. E., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research, 50,* 408-428.

Lum, J. A. G., & Conti-Ramsden, G. (2013). Long-term memory: A review and meta-analysis of studies of declarative and procedural memory in specific language impairment. *Top Lang Disord., 33,* 282-297.

Lum, J. A. G., Conti-Ramsden, G., Morgan, A. T., & Ullman, M. T. (2014). Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance. *Cortex, 51,* 1-10.

Lum, J. A. G., Conti-Ramsden, G., Page, D., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex, 48,* 1138-1154.

Montgomery, J. W. (2003). Working memory and comprehension in children with specific language impairment: What we know so far. *Journal of Communication Disorders, 36,* 221-231.

Montgomery, J. W., & Evans, J. L. (2009). Complex sentence comprehension and working memory in children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research, 52,* 269-288.

Namy, L. L., & Waxman, S. R. (1998). Words and gestures: Infants' interpretations of different forms of symbolic reference. *Child Development, 69,* 295-308.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14,* 11-28.

Plante, E., Bahl, M., Vance, R., & Gerken, L. A. (2010). Children with specific language impairment show rapid, implicit learning of stress assignment rules. *Journal of Communication Disorders, 43,* 397-406.

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of Specific Language Impairment in English-speaking children. *Journal of Speech and Hearing Research, 39,* 1239-1257.

Richmond, J., & Nelson, C. A. (2007). Accounting for change in declarative memory: A cognitive neuroscience perspective. *Developmental Review, 27,* 349-373.

Robinson, A. J., & Pascalis, O. (2004). Development of flexible visual recognition memory in human infants. *Developmental Science, 7,* 527-533.

Singh, L., Hui, T. J., Chan, C., & Golinkoff, R. M. (2014). Influences of vowel and tone variation on emergent word knowledge: A cross-linguistic investigation. *Developmental Science, 17,* 94-109.

Singh, L., Morgan, J. L., & White, K. S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language, 51,* 173-189.

Singh, L., White, K. S., & Morgan, J. L. (2008). Building a word-form lexicon in the fact of variable input: Influences of pitch and amplitude on early spoken word recognition. *Language Learning and Development, 4,* 157-178.

Thiessen, E. D., & Yee, M. N. (2010). Dogs, bogs, labs, and lads: What phonemic generalizations indicate about the nature of children's early word-form representations. *Child Development, 81,* 1287-1303.

Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition without control: When a little frontal lobe goes a long way. *Current Directions in Psychological Science, 18,* 259-263.

Tomblin, J. B., Mainela-Arnold, E., & Zhang, X. (2011). Procedural learning in adolescents with and without Specific Language Impairment. *Language Learning and Development, 3,* 269-293.

Weenick, D. (2009). The KlattGrid speech synthesizer. In Proceedings of the 10th Annual Conference of the International Speech Communication Association (INTERSPEECH), Brighton, United Kingdom.

Weismer, S. E., Evans, J., & Hesketh, L. J. (1999). An examination of verbal working memory capacity in children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research, 42,* 1249-1260.

Woodward, A. L., & Hoyne, K. L. (1999). Infants' learning about words and sounds in relation to objects. *Child Development, 70,* 65-77.

Zwicker, E. (1961). Subdivision of the audible frequency range into critical bands. *Journal of the Acoustical Society of America*, 33, 248-248.