Title
Impossible to _gnore: Word-Form Inconsistency Slows Preschool Children's Word-Learning

Journal
Language Learning and Development, 10(1)

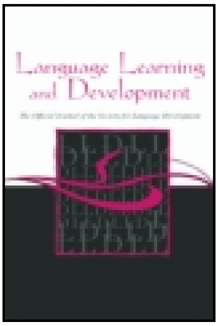
ISSN
1547-5441

Author
Creel, Sarah C

Publication Date
2014-01-02

DOI
10.1080/15475441.2013.803871

Peer reviewed



Language Learning and Development
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hlld20

Impossible to _gnore: Word-Form Inconsistency Slows Preschool Children's Word-Learning

Sarah C. Creel [a]

[a] Department of Cognitive Science, University of California San, Diego

Published online: 01 Nov 2013.

Language Learning and Development, 00: 1–28, 2013

ISSN: 1547-5441 print / 1547-3341 online
DOI: 10.1080/15475441.2013.803871

Routledge
Taylor & Francis Group

Impossible to _gnore: Word-Form Inconsistency Slows Preschool Children's Word-Learning

Sarah C. Creel
Department of Cognitive Science, University of California San Diego

Many studies have examined language acquisition under morphosyntactic or semantic inconsistency, but few have considered *word-form* inconsistency. Many young learners encounter word-form inconsistency due to accent variation in their communities. The current study asked how preschoolers recognize accent-variants of newly learned words. Can preschoolers generalize recognition based on partial match to the learned form? When learning in two accents simultaneously, do children ignore inconsistent elements, or encode two word forms (one per accent)? Three- to five-year-olds learned words in a novel-word learning paradigm but did not generalize to new accent-like pronunciations (Experiment 1) unless familiar-word recognition trials were interspersed (Experiments 3 and 4), which apparently generated a familiar-word-recognition pragmatic context. When exposure included two accent-variants per word, children were less accurate (Experiment 2) and slower to look to referents (Experiments 2, 5) relative to one-accent learning. Implications for language learning and accent processing over development are discussed.

A sizable body of research has explored how learners process syntactic and morphological input variability (e.g., Hudson Kam & Newport, 2005, 2009). However, little work has addressed consequences of *word-form* inconsistency, which is present when language learners experience multiple accents in the same language. This variability may have strong effects on language learning. While numerous studies (e.g., Bradlow & Bent, 2008; Maye, Aslin, & Tanenhaus, 2008) have asked how adult listeners deal with unfamiliar accents, fewer studies address the effects of accent variability on young children. Further, little work connects adult accent-comprehension tasks using overt recognition to infant and toddler studies of accent processing that use looking time measures, which might be considered more implicit. A step toward connecting the infant and adult literatures is to look at effects of accent variability in intermediate age ranges, such as the preschool years (see also Ramon-Casas, Swingley, Sebastián-Gallés, & Bosch, 2009), using both implicit and explicit measures to discern how the different paradigms map onto each other. This is the approach taken in the current study.

Accents are variants of a language that have different realizations of the same word forms. This can occur via nonnative speakers of a language or due to regional or social-group differences within native speaking populations. A speaker from Boston may say "yard" in a way that other

Correspondence should be addressed to Sarah C. Creel, University of California San Diego, Department of Cognitive Science, 9500 Gilman Drive, Mail Code 0515, La Jolla, CA 92093-0515. E-mail: creel@cogsci.ucsd.edu

American listeners hear as “yad”; Russian speakers of English may produce “yard” much like “yart” because Russian devoices consonants at the ends of words. Due to prevalent language and dialect variability throughout the world, many children will hear familiar and novel words from speakers of two or more different accents. For some children, the predominant home accent may mismatch the accent of their community (see Floccia, Delle Luche, Durant, Butler, & Goslin, 2012). Other children may hear multiple accents in the home—their mother may refer to a writing instrument as a /pɪn/, while their father calls the same object a /pɛn/. Hearing multiple accents violates the assumption that words’ forms will be consistent. How do children recognize and learn words under these conditions?

The present study explored how preschool-aged children deal with word-form inconsistency in language input. Two situations were considered: *generalizing* newly learned words to a novel accent and *learning words in two accents* simultaneously. For newly learned words, children may be able to recognize accented variants via a partial match to the learned representation or may need greater familiarity with the accent. For learning words with accent variability, however, lexical knowledge is not accessible. Two major possibilities were considered. First, children might regard accent-variants of a word-form as the same word. Second, they might regard accent-variants—and learn them—as two different word-forms. A third possibility, discussed but not strongly tested here, is that children perceive differences between accent variants, but treat them as contextually-determined forms of the same word, analogous to allophones. Each learning situation, and the constraints upon it, is discussed in turn.

RECOGNIZING LEARNED WORDS IN AN ACCENT

Both adults (Bradlow & Bent, 2008; Clarke & Garrett, 2004; Weil, 2001) and children (Best, Tyler, Gooding, Orlando, & Quann, 2009; Nathan, Wells, & Donlan, 1998; Schmale, Cristiá, Seidl, & Johnson, 2010; Schmale & Seidl, 2009) have a harder time recognizing familiar words in accented speech than in unaccented (familiarly accented) speech. Two perceptual factors appear to ease this difficulty: the similarity of the accented form to the canonical (familiarly accented) form and the listener’s prior experience with hearing the accent.

Adults can recognize familiar words in unfamiliar accents based on partial match to a known word form (Bradlow & Bent, 2008; Maye et al., 2008), and there is some evidence that infants and young children do as well. Previous studies have shown that infants as young as 14 months look more to a picture of a dog when they hear *doggy* than when they hear the mispronunciation **toggy* (Swingley & Aslin, 2000, 2002). This has been taken as evidence of infants’ fine-grained phonetic knowledge of words. However, considering accented-speech processing, we need to reverse the question: how well do children recognize the *mispronunciations*? In the same studies, children show delayed, but nonzero, recognition of the mispronounced words, looking more at a dog than a shoe when they hear *toggy* (Swingley & Aslin, 2000, 2002). White and Morgan (2008) found that 19-month-olds also look more to a dog than a novel object upon hearing a single-feature mispronunciation such as *toggy*, with more novel-object looks as phonological distance from the familiar form increased. Preschool-aged children are also sensitive to partial match to a familiar word form (Creel, 2012; Gerken, Murphy, & Aslin, 1995; Merriman & Schuster, 1991; Storkel, 2002). For instance, Creel (2012) showed that preschoolers most often select the familiar referent—not a novel picture—when hearing a mispronounced familiar word (e.g., **buzzle* for *puzzle*), especially if it differs by a single phonological feature. These studies suggest that infants and young children, like adults in accented-speech recognition studies (e.g. Bradlow & Bent, 2008; Maye et al., 2008), may have some capacity to recognize accent-like variants of highly-familiar words. How similar the word must be to the original form is not as certain, though a good guess would be that single-feature (and single-phoneme) changes are the easiest to recognize. This may be modulated by the degree of contextual support for a familiar-word interpretation (e.g., Fallon, Trehub, & Schneider, 2002).

In short, listeners of all ages recognize familiar words in unfamiliar accents from partial matches. Moreover, recent research suggests that recognition by partial match to a known form may even hold for newly learned words, which one might expect to have more fragile representations (see Stager & Werker, 1997). Schmale, Hollich, and Seidl (2011) found that 30-month-olds generalized a single *newly learned* word (neech) learned in a familiar (American) accent to a novel (Spanish) accent; that is, they looked more to the paired object when hearing the accented version of the trained word (neech$_S$) but looked more to a novel object when hearing an untrained accented word (moof$_S$). This suggests that 30-month-olds recognized the similarity between the never-before-heard Spanish form and the learned American form.

An alternative explanation for recognizing accented forms is *exposure*—to particular accented word forms, and to general accent properties. Children may regard *toggy* as more likely to mean *doggy* than *novel thing* because they have already heard *toggy* prior to the experiment, via speech errors or accented talkers. Children also benefit from previous exposure *to an accent itself* (like adults: Bradlow & Bent, 2008; Clarke & Garrett, 2004; Weil, 2001) due to consistency in how a particular accent modifies sounds, presumably driven by feedback from existing lexical knowledge (e.g., Eisner & McQueen, 2005; Kraljic & Samuel, 2005; Norris, McQueen, & Cutler, 2003). White and Aslin (2011) showed that children as young as 19 months may adjust phoneme categories via lexical feedback: *dog* becoming *dag* implies that *sock* will become *sack*. Schmale, Cristia, and Seidl (2012), in a paradigm similar to Schmale et al. (2011), showed that 24-month-olds succeeded in cross-accent recognition of a newly learned word when they received exposure to the accent prior to word learning. Thus, young children benefit from even brief accent exposure in comprehending accented words.

Several questions remain regarding recognition of newly learned words. First, when task difficulty is scaled up—for instance, learning more words, using more similar words, or requiring overt responses rather than just looking time—does accent difficulty persist into later age groups? Second, how far can listeners get in recognition by partial match: must they have heard the accent or the specific accented word form previously, or is partial match sufficient?

LEARNING NEW WORDS IN MULTIPLE ACCENTS

While familiar or even recently learned words may be recognizable based on partial match to the known form, lexical knowledge cannot be brought to bear if a word is completely unknown. How does word learning take place when learners hear the same words but in different accents? Previous work on category formation (e.g., Posner & Keele, 1968) and studies of language acquisition and change (Hudson Kam & Newport, 2005, 2009; see Smith & Wonnacott, 2010, for related adult evidence) suggest that human learners pick up on consistent or regular input patterns. For word forms, learners might ignore the less-prevalent (see Hudson Kam & Newport, 2005, 2009) or less-salient variant, or might extract consistent elements while ignoring the inconsistent ones. Recent work by Floccia et al. (2012) suggests that young word learners may select a single form as the "correct" one: British English-learning 20-month-old children only recognized r-containing words (like *bird*) when they were spoken in the r-containing local dialect and not in an r-less dialect—even if it was their parents' dialect. This implies that young children given accent inconsistency may downweight one of the variants in their input. Further, studies of phonological-rule learning, though they do not examine word-meaning mapping, suggest that learners can *extract consistent phonological patterns*. For instance, Richtsmeier, Gerken, and Ohala (2011) found that 4-year-olds generalized phonotactic patterns to new environments when provided with both word-form variability and talker variability (see also Chambers, Onishi, & Fisher, 2011; Newport & Aslin, 2004). To do this, listeners must be ignoring or collapsing across variable segments. This might entail a broadening of the variable category (see Brunellière, Dufour, Nguyen, & Frauenfelder, 2009, for evidence that exposure to regional accents can decrease listener sensitivity to native phoneme contrasts), or even ignoring its value completely. If listeners can ignore inconsistent segments to extract a consistent segment pattern, it is possible that they can map that pattern to a referent.

A second approach to multiple-accent input is simply storing two separate word forms, one per accent. Research on adults exposed to multiple accents (Sebastián-Gallés, Echeverria, & Bosch, 2005; Sebastián-Gallés, Vera-Constán, Larsson, Costa, & Deco, 2009; Sumner & Samuel, 2009) supports the notion of dual word-form representations. For instance, Catalan natives often fail, in a lexical decision task, to reject nonwords that were created by mispronouncing Catalan /ɛ/ words with the more Spanish-like /e/ (Sebastián-Gallés et al., 2005, 2009). This occurred despite listeners easily distinguishing the two Catalan sounds in both behavioral and ERP paradigms (Sebastián-Gallés et al., 2009). The authors argued that these Catalan speakers have encoded Spanish-accent word variants in addition to native-Catalan word forms. If this holds for child learners, then they must encode dual forms (one for each accent-variant) of each accent-variable word they encounter. This as much as *doubles* how many forms the child must encode relative to a child exposed to a single accent, potentially slowing lexical learning. Further, word-form variation might impair formation of object categories. For instance, while infants are aided in category formation by hearing different labels for each category (e.g., Xu, 2002), infants hearing different labels applied to the same category are impaired in category formation (Waxman & Braun, 2005). These studies, taken together, suggest that if learners in multiple-accent situations are acquiring duplicate word-forms, their learning will be slowed.

A third possibility strikes a balance between the first two: learners might learn two forms, their contexts, and their relationship, representing the forms as allophonic variants of a single underlying form. This "phonological translation" capacity (Oller, Cobo-Lewis, & Eilers, 1998) would allow learners to use a single representation while keeping track of variation and expecting the correct form in the correct context. However, listeners might need substantial exposure to calculate relations between accents, making it difficult to detect in the lab. This possibility is not tested deeply in the current study, but is revisited in the General Discussion.

Studies of naturalistic word-form variability are somewhat consistent with either merging representations or knowing phonological translation equivalents. Mattock, Polka, Rvachew, and Krehm (2010) found that 17-month-old French-English bilingual infants learned accent-variable words more readily than did monolingual infants of either language. This implies that those infants either did not distinguish differences between the two variants or understood their equivalence (interestingly, Sundara, Polka, & Molnar, 2008, show that younger French-English bilingual infants can distinguish *sounds* between languages). Ramon-Casas et al. (2009), in a looking-while-listening paradigm, found that bilingual toddlers' recognition of mispronounced Catalan words was not impaired when the mispronunciation involved a vowel pair only present in Catalan (/e/ vs. /ɛ/). Among preschoolers, only Catalan-dominant (but not Spanish-dominant) bilingual children's word recognition was impaired by such mispronunciations. Their results suggest that, unless children experience a high proportion of Catalan input, they are indifferent to words mispronounced with the Catalan /e/ - /ɛ/ vowel difference—again, consistent either with merging representations or understanding the representations' equivalence. Ramon-Casas et al. (2009), as well as Albareda-Castellot, Pons, and Sebastián-Gallés (2011), suggest that indifference to Catalan vowel changes in word recognition may be driven by the relatively high cognate overlap between Catalan and Spanish. These cognates largely differ in their vowels, which, the argument goes, leads children to attend less to subtle differences in vowel sounds (see Ramon-Casas & Bosch, 2010, for evidence that Catalan vowels are preserved in noncognate words). As pointed out by Albareda-Castellot et al., accent-variable environments might be regarded as an extreme case of cognate overlap. If so, then accent variability might lead learners to disregard inconsistencies between differently accented word-forms. This hypothesis has not been tested experimentally.

In sum, many questions remain regarding learning under word-form variability. Will learners ignore differences between words that map to the same referent? Or instead, will they be slowed in learning because they must encode multiple word-forms (one per accent)? While these questions have received some attention in infant populations, little is known about the older age group (3-5-year-olds) examined here.

THE CURRENT STUDY

In summary, current evidence hints that word-form inconsistency may affect young children's abilities to map novel-accented forms onto existing representations and their ability to encode words. Yet several questions remain. First, how readily do children recognize newly learned words in never-before-heard accented forms—must their input contain accent variants before they can recognize such variants, or does partial match to the newly-learned word form suffice? Second, how do children contend with accent variability *as they are learning* words? Do children readily tune out variable elements, or does exposure to word-form variability slow learning because they must learn two words (both accented forms) instead of one?

The current study assessed how readily children recognized accent-like variants of newly learned words, using both an implicit measure (eye tracking) comparable to measures used with infants and toddlers, and an explicit measure (pointing). Convergence between explicit and implicit measures will imply that looking time infant/toddler measures are continuous with more explicit measures in older populations. In Experiment 1, children learned words in one accent and were then tested on the original accent and an altered accent. This simulated learning in one accent, followed by generalization to never-before-heard accent-variants (see Schmale et al., 2011, for a related approach). (Note that the term "accent variant" is used here to mean a word form produced in a particular accent, rather than a particular accent such as southeastern U.S. English.) In Experiment 2, children learned two forms of each word from the outset, with each talker using one "accent." This simulated word learning with two-accent input. The last three experiments addressed whether providing a familiar-referent pragmatic context during recognition—by interspersing test trials with familiar words, with pictured referents—affected generalization to an altered accent (Experiments 3 and 4) and integration of two-accent input (Experiment 5).

EXPERIMENT 1

The first experiment tested a new word-learning paradigm with 3-5-year-old children. An artificial accent was employed to control for declines in intelligibility in naturalistic accented speech (though note that it is an open question as to whether naturalistic accents themselves might serve as a cue to interpret following words more leniently). The "accent" change was an alternation between the vowels /i/ as in *bean* and /ɪ/ as in *bin*. This vowel pair was chosen because it mimics a vowel-merger found in Spanish-accented English (Fox, Flege, & Munro, 1995). Though participants were monolingual, they were residents of a city with a large Hispanic population (29% in 2010; www.quickfacts.census.gov), and thus these children had likely had incidental exposure to Spanish accents. A vowel change was used because cross-linguistically, dialects tend to vary in vowels more than consonants, particularly in English. These factors should predispose children in the current experiment to accept alternate forms as equivalent. Children saw cartoon creatures labeled with the novel words. Test trials presented both original pronunciations (OPs) and accented pronunciations (APs). This experiment speaks to the question of whether children can recognize an accent variant without having experienced it previously—whether they can *generalize* based on partial matching. If so, children should show above-chance accuracy and visual fixations to the APs. If children cannot recognize an accent variant without explicit *accent exposure*, they should not show above-chance accuracy or visual fixations to APs.

Method

Participants

$N = 24$ monolingual English-speaking preschoolers (mean age = 4.5, range: 3.3–5.5; 14 female) took part. Consent forms asked parents to indicate all languages that the child experienced in the home, and which ones were understood and spoken by the child. Children were placed in this study only if parents reported that the child heard only English. That is, any indication from the parent that the child was exposed to or understood another language led us not to include that child. Ambient cultural exposure to Spanish or Spanish-accented English may still have occurred, which one might think would *decrease* accent-comprehension difficulty—thus, this population serves as a conservative test of difficulty comprehending accents.

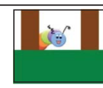
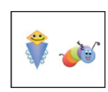
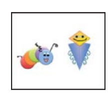
Stimuli

One female talker (HP) and one male talker (GW) recorded sentences describing the cartoon creatures (Table 1). Sentence frames omitted the vowel /ɪ/ (as in *pin*), so that carrier phrases in Experiment 2 did not need to contain accented forms—accented carrier sentences might add difficulty to word *segmentation*, a potentially different problem from word *recognition*. The novel words (see Tables 1 and 2) were *gif, geef, kib,* and *keeb* (with “hard” g’s, as in “gum”). Equal numbers of children learned *gif* and *kib, geef* and *kib, gif* and *keeb,* and *geef* and *keeb.* Words with similar onset consonants (/g/ and /k/) were used so that the words were mildly similar to one another. This meant that the task was not trivially easy, but APs were still more similar to their OP than to the other word. Words in test sentences averaged 459 ms ($SD = 48$) for the female talker, and 665 ms ($SD = 64$) for the male, with mean duration of 562 ms ($SD = 118$).

Procedure

There were three cycles of training and testing (Table 1), with distractor sequences interspersed to maintain child attention. Training trials (16 in block 1, 8 in each of blocks 2 and 3) were designed to be engaging. On each training trial, a schematic grass-and-trees background appeared, and then a creature moved from off-screen. Eight different animations were counterbalanced

TABLE 1
Sentence Frames and Trial Sequences in Experiment 1

Phase	*Sample list 1 (2 vowels)*	*Sample list 2 (1 vowel)*	*Display*
Train 1a (8 trials)	Look, a **geef**! That’s a **geef**! Wow, a **geef**! I see a **geef**!	Look, a **keeb**! That’s a **keeb**! Wow, a **keeb**! I see a **keeb**!	
	A **kib**! Do you see the **kib**? A **kib**! Where’s the **kib**?	A **geef**! Do you see the **geef**? A **geef**! Where’s the **geef**?	
Two distractor trials (no words)			
Train 1b (8 trials)	(Same as Train 1a)	(Same as Train 1a)	(Same as Train 1a)
Test 1 (8 trials)	Find the **geef**! Show me the **geef**!	Find the **keeb**! Show me the **keeb**!	 *or*
	Can you find the **kib**? What’s the **kib**?	Can you find the **geef**? What’s the **geef**?	
Two distractor trials (no words)			
Train 2 (8 trials)	(Same as Train 1a)	(Same as Train 1a)	(Same as Train 1a)
Test 2 (8 trials)	Find the **geef**! Show me the ***gif**!	Find the **keeb**! Show me the ***kib**!	(Same as Test 1)
	Can you find the **kib**? What’s the ***keeb**?	Can you find the **geef**? What’s the ***gif**?	
Train 3 (8 trials)	(Same as Train 1a)	(Same as Train 1a)	(Same as Train 1a)
Test 3 (8 trials)	(Same as Test 2)	(Same as Test 2)	(Same as Test 1)

*Note.**indicates changed pronunciations. For greater visibility, displays are not to scale.

TABLE 2
Stimuli Used in All Experiments

Experiment	*Words trained and tested*	*Additional words tested*
Exp. 1: Learn 2 words, generalize to changed versions	giff, kib giff, keeb geef, kib geef, keeb	geef, keeb geef, kib giff, keeb giff, kib
Exp. 2: Learn from talkers with two accents; test with original and "wrong" accents	$giff_F$, kib_F, $geef_M$, $keeb_M$ $giff_M$, kib_M, $geef_F$, $keeb_F$	$giff_M$, kib_M, $geef_F$, $keeb_F$ $giff_F$, kib_F, $geef_M$, $keeb_M$
Exp. 3: Learn 2 words, generalize to changed versions while hearing variable real words	$giff_{F/M}$, $kib_{F/M}$ $geef_{F/M}$, $keeb_{F/M}$ giff, kib geef, keeb	$geef_{F/M}$, $keeb_{F/M}$, $feet_F$, $keys_F$, fit_F, $kizz_F$ giff, kib, $bridge_F$, $fish_F$, $breedge_F$, $feesh_F$ geef, keeb, $feet_F$, $keys_F$, fit_F, $kizz_F$ giff, kib, $bridge_F$, $fish_F$, $breedge_F$, $feesh_F$
Exp. 4: Learn 2 words, generalize to changed versions while hearing OP real words	giff, kib giff, kib geef, keeb geef, keeb	geef, keeb, $pizza_F$, $zebra_F$, pig_F, $milk_F$ geef, keeb, $pizza_M$, $zebra_M$, pig_M, $milk_M$ giff, kib, $pizza_F$, $zebra_F$, pig_F, $milk_F$ giff, kib, $pizza_M$, $zebra_M$, pig_M, $milk_M$
Exp. 5: Learn from talkers with two "accents"; test with original accent plus accent-consistent real words	$giff_F$, kib_F, $geef_M$, $keeb_M$ $giff_M$, kib_M, $geef_F$, $keeb_F$	pizza, zebra, pig_F, $milk_F$, $peeg_M$, $meelk_M$ pizza, zebra, pig_M, $milk_M$, $peeg_F$, $meelk_F$

Note. Each line represents a single version of each experiment. Subscripts refer to the gender of the speaker. If there is no subscript, then both speakers produced a word.

across creatures. For all animations, the creature paused at the center, and after one second, a recorded passage named the creature twice (e.g. *A geef! That's a geef!*). Next, there was a 2-second pause, and then the creature moved off-screen. In each training block, creatures' labels were spoken equally often by each talker.

On each test trial, the two creatures appeared stationary on a white background, and after 500 milliseconds (ms) a spoken instruction asked the child to, for example, "Point to the X." On test block 1 (8 trials), X was always the original pronunciation (OP). This block made certain that children had learned the OPs of words. On test blocks 2 and 3 (8 trials each), X was either the exact word heard during training, or a vowel-changed AP of the word—/i/ to /ɪ/ or /ɪ/ to /i/. This meant that some children heard both vowels change to /i/ or both to /ɪ/, while for other children, one word changed from /i/ to /ɪ/ and the other changed in the opposite direction. This was included as a factor in the analyses. Side of presentation, sentence frame, and pronunciation were counterbalanced within each block.

Equipment

A Mac Mini (OS 10.4.1) running Matlab (7.6.0) and PsychToolbox 3 (Brainard, 1997) presented the experiment. Children sat in an unbuckled car seat (maintaining consistent viewing distance) approximately 70 cm away from the screen, listening to stimuli presented over child-sized KidzGear headphones (www.gearforkidz.com). Experimenters reported hearing sounds faintly through the headphones but did not detect experimental manipulations (e.g., that the caterpillar-like creature was named different things for different children), suggesting little awareness of word identity. An Eyelink 1000 eye tracker (www.sr-research.com) running in remote mode sat just below the experimental display monitor. The eye tracker was linked to a Dell tower running DOS, which collected gaze position information every 4 ms and integrated experimental timing messages from the Mac. Raw eye tracking data were reprocessed offline by automated Python scripts to map exact gaze coordinates and times into looks to areas of interest relative to the temporal onset of the target words in the speech stream.

Procedure

Experiments were conducted in a quiet area in each child's preschool or day care facility. Children went through an eye tracker calibration sequence as programmed in the Eyelink Toolbox (Cornelissen, Peters, & Palmer, 2002), which was described as a "follow-the-dot game." They then took part in the main experiment, where they were told, "you will see some funny creatures and hear people talking about them." They were prompted to point to the cartoons, and the experimenter clicked the child's responses with a mouse. Generally, children did not need to be prompted to point, though sometimes their points were unclear and the experimenter requested neutrally that the child point again (e.g. "Can you show me again?"). In rare instances, children refused to point. If this happened, they were questioned about both alternatives and typically only verified one. If they refused to answer, data were excluded.

Results

Accuracy

Children reached moderate accuracy on trained pronunciations, but did not generalize to novel pronunciations (Figure 1a; all data summarized in Table 3). Data were analyzed in a logistic regression model, which better accounts for the binomial distribution of binary-choice data than traditional ANOVA (Jaeger, 2008). In all analyses reported, participant intercepts and slopes were included as random factors; items intercepts and slopes were not included because there were so few items.

Accuracy is reported as the percentage of trials in which children chose the creature whose label was most similar to the spoken word. Accuracy in the first test block demonstrated good learning of OPs ($M = 75\%$ correct, $SD = 6.6\%$; estimate = 2.35, $z = 3.65$, $p = .0003$). Overall accuracy in the second and third blocks was lower (65.1% ± 16.9%), driven by the AP trials. A logistic regression model on accuracy was computed with mean-centered factors Block (second, third) and Pronunciation (original [OP], accented [AP]), and Learned Vowels (learned two words with the same vowel, learned two words with two different vowels). Only Pronunciation was significant (estimate = 1.56; $z = 5.32$; $p < .0001$). This resulted from high accuracy on OP trials (78.1% ± 20.6%; estimate = 1.5; $z = 5.44$; $p < .0001$), but chance performance on AP trials (52.1% ± 18.3%, estimate = 0.08; $z = 0.57$; $p = .57$). Results were unchanged when considering only children who scored perfectly in test block 1 ($n = 11$): AP accuracy was only 55.7% ($SD =$ 19.7%).

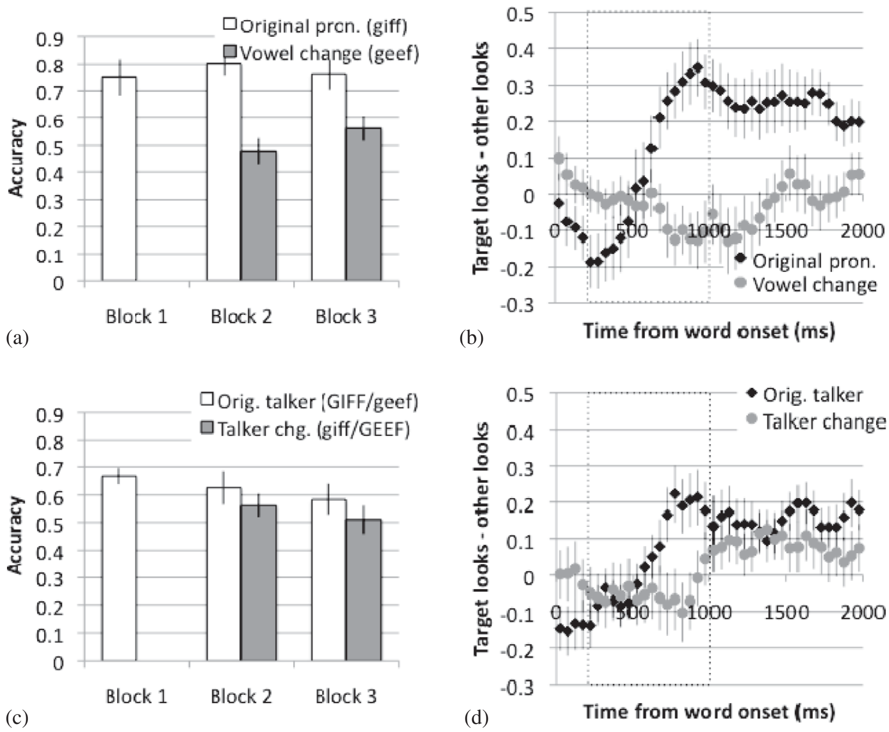


FIGURE 1 Experiment 1, (a) accuracy and (b) target looks minus other looks, Blocks 2–3; Experiment 2, (a) accuracy and (b) target looks minus other looks, Blocks 2–3. All error bars are standard errors. Dashed line indicates tested time window.

Eye tracking

Gaze fixation data (Figure 1b) were pooled into 50-millisecond bins. Looks in Blocks 2 and 3 were analyzed, and were collapsed across Block to boost power. Before analysis, individual trials where a child was not visually fixating either of the two objects at least 50% of the time (14.3% of trials) were dropped. Except where noted, patterns of significance were identical between analyses, with all data and analyses using this trial exclusion criterion. To assess increases in looks to the target from the beginning of the trial, a linear regression analysis was conducted on empirical-logit-transformed target looks minus empirical-logit-transformed other-picture looks, following Barr (2008), to correct for nonnormal distribution of looking-proportion data. This transformed "target advantage" score should be zero when both pictures are fixated equally and positive when the target is fixated more. Note that the empirical logit (e-logit) transformation is, except at extreme values, nearly linear—thus, Figure 1b would look qualitatively similar if plotted in e-logit space. The analysis window started with the bin at 200–250 milliseconds (ms) after target word onset, which is the first point in time where one would expect to see looks driven by the acoustic signal (Hallett, 1986). The window ended with the 950–1000 ms time bin, since

TABLE 3
Summary of All Experiments

	Words learned	*Additional words tested (Tests 2–3 only)*		*Accuracy (SD)*			*Slope (SE)*		
Exp.	*Novel*	*Novel*	*Familiar*	*OP1*	*OP2–3*	*AP*	*OP1*	*OP2–3*	*AP*
1	giff, keeb	Vowels changed	*(none)*	0.750 (0.332)	0.766 (0.242)	0.521 (0.183)	.0024 (.0008)	.0036 (.0006)	−.0008 (.0007)
2	$\text{giff}_F = \text{geef}_M$, $\text{kib}_F = \text{keeb}_M$	Talkers changed accents	*(none)*	0.667 (0.143)	0.604 (0.232)	0.536 (0.15)	−.0005 (.0005)	.0021 (.0006)	.0001 (.0006)
3	giff, kib	Vowels changed	fish, feesh	0.724 (0.27)	0.703 (0.214)	0.615 (0.147)	.0022 (.0008)	.0026 (.0006)	.0014 (.0007)
4	giff, kib	Vowels changed	pizza, pig	0.813 (0.217)	0.724 (0.188)	0.589 (0.182)	.0046 (.0007)	.0018 (.0007)	−.0004 (.0009)
5	$\text{giff}_F = \text{geef}_M$, $\text{kib}_F = \text{keeb}_M$	(No changed pronun- ciations)	pizza, pig_F, peeg_M	0.729 (0.229)	0.643 (0.163)	*(none)*	−.0006 (.0008)	.0009 (.0006)	*(none)*

at this point target looks across experiments appeared to reach asymptote, and because pointing responses began to obscure children's eyes, leading to a drop in data acquisition. This window extended well past the end of the word ($M = 555$ ms, $SD = 118$ ms), even allowing an additional 200–300 ms for eye movement planning. This time window can thus be expected to include looking responses to the entirety of the word. Note that this window starts and ends earlier and is shorter than that often used with infants or very young children (e.g., 367–2000 ms in Swingley & Aslin, 2000, 2002). For uniformity, the same time window was used across all experiments. Nonetheless, the reader should be aware that, in some cases, longer windows of analysis may have yielded different patterns of significance (i.e., slow recognition effects which do not emerge until after 1000 ms).

Factors in the model of target advantage were Time Bin, Pronunciation (AP, OP; both within-participant), and Learned Vowels (between-participants). All factors were mean-centered. Time Bin was significant ($t = 2.92$, $p = .004$), indicating that target advantage increased as the word unfolded. Pronunciation did not reach significance ($t = 1.61$, $p = .11$), indicating no difference in overall looks between OPs and APs.[1] However, Time Bin and Pronunciation interacted ($t = 5.68$, $p < .0001$), because the effect of Time Bin was stronger on OP trials: the slope (increase in target looks) was .0036 ($SE = .0006$: $t = 5.81$, $p < .0001$), greater than the negative, nonsignificant slope on AP trials (slope = -.0008 ± .0007; $t = 1.2$, $p = .23$). There was also a significant interaction of Time Bin and Learned Vowels ($t = 2.40$, $p = .02$), suggesting that the overall slope of increases in looks was greater for children who learned two words with two different vowels than those who learned two words which had the same vowel. This may have resulted from a lower visual-fixation measure baseline in the two-vowel group, from lower altered-pronunciation novelty when those

[1] In analyses without excluding trials, an effect of Pronunciation approached significance ($t = 1.95$, $p = .051$), indicating higher overall target advantage on OP trials.

children had already heard both vowels during learning, or from the greater distance between the two learned words in the two-vowel condition than in the one-vowel condition. However, this did not interact with the Pronunciation effect. No other effects were significant.

Age effects

Each experiment tested a broad age range of children, raising the question of improvement with age. Accordingly, models of both accuracy and fixations were computed with mean-centered age as a continuous factor, to assess whether learning improved or changed with age. Additionally, data in Experiments 3–5 were compared to children's vocabulary sizes as measured by the Peabody Picture Vocabulary Test – IV (PPVT-IV; Dunn & Dunn, 2007). Effects tended to be positive but were weak, and not significant in any of the studies. (Across all studies combined, there was a trend toward increased accuracy with age: estimate $= .0006$, $t = 1.86$, $p = .06$.) This should not be taken as an implication that there is no improvement in word learning over age. Rather, the lack of clear age effects may mean that, by age 3, age group is a bit less relevant than other factors—that children at this point have diverged sufficiently in their learning abilities that, at least for this task and with these sample sizes, age is not a significant predictor.

Discussion

Preschool children learned words to a reasonable degree of accuracy, and were quick to fixate originally pronounced targets but did not generalize to "accented" forms of the words. Instead, children reacted to mispronounced newly-learned words as though they were not the words learned during training. This contrasts with the results of Schmale et al. (2011), who found that younger children (30 months) readily generalized to an accented word-form. On the face of it, it is puzzling that older children do not show *better* performance. However, Schmale et al.'s task was somewhat easier: children learned only a single word, not two words. Further, the accented variant in Schmale et al. may have been more acoustically similar to the original, in that the words used contained sounds (/i/ and /u/) found in both English and Spanish (the accent used). Thus, children in the current experiment may be doing worse than Schmale et al.'s younger children due to task difficulty and a larger difference between OP and AP word-forms.

In the Introduction, I outlined multiple possibilities for how children would treat accented forms: treating them as the same word and ignoring differences; treating two forms as completely different words; or some intermediate variant, where differences are recognized but are not regarded as important. Experiment 1's results suggest that, at least when children have learned only one variant of a word, they do *not* consider the two forms to be the same due to partial match—nor do they appear to regard the two words as phonological variants of each other. However, this outcome may be due to lack of exposure to the specific variant forms. That is, children have likely heard *familiar* words in variant pronunciations, particularly in a multicultural city, but they have never heard these experimentally learned words in variant pronunciations. This predicts that *training* children with both pronunciations of each word might lead to more flexible recognition. On the other hand, learning two forms for the same word could be a challenging task: if children cannot ignore vowel differences, or cannot grasp the equivalence between two word forms, they will essentially be learning *four* labels (two per picture) instead of two.

EXPERIMENT 2

This experiment asked whether children could learn words in two simultaneous accents. Each talker produced the words with a different vowel than the other talker so that children heard each individual form half as often as children in Experiment 1. If children can merge the two vowels (or word-forms) or simply extract the consistent consonants-to-pictures mapping, then accuracy should be high and looks, rapid. However, if children do not merge, then they will essentially be learning *four* words, not two, decreasing accuracy and speed (looks to target).

Method

Participants

$N = 24$ new monolingual English-speaking preschoolers (mean age = 4.4 years, range: 3.5-6.1; 11 female) took part.

Stimuli

These were the same as in Experiment 1, but were assigned differently within-subjects. Here, each talker had a different "accent"—one talker labeled the objects *giff* and *kib*, while the other talker labeled the objects *geef* and *keeb* (all stimuli in Table 2, task outline in Table 4). Recall that /ɪ/ only occurred in the novel words themselves, so that children never heard the second talker produce the /ɪ/ vowel. However, /i/ occurred in the carrier phrases as well, so that children heard the /i/-/ɪ/ talker use both vowels. For instance, the /i/-/ɪ/ talker might say "Do you see the giff" (containing both /i/ and /ɪ/) while the /i/ talker might say "Do you see the geef?" (containing only /i/). This is consistent with the first talker distinguishing /i/ and /ɪ/, and the second talker merging them.

Procedure

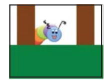
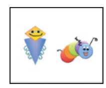
The procedure was much like Experiment 1, except that this time, during training trials, one talker labeled the objects *giff* and *kib* and the other labeled the same objects *geef* and *keeb.* This labeling scheme was maintained in all training trials and in test block 1. Tests 2 and 3 also contained talker-inconsistent trials where each talker used the other's accent, to test whether children were sensitive to the relationship between talker and word form.

Results

Accuracy

Children appeared less accurate in this experiment for heard pronunciations overall (63.5% ± 16.3%; Figure 1c). A logistic regression with Block (second, third) and Pronunciation (OP = each talker's original accent, AP = talker with the "wrong" accent) was conducted. A significant intercept indicated that accuracy exceeded chance overall (57% ± 16.5%; estimate = 0.32;

TABLE 4
Sentence Frames and Trial Sequences in Experiment 2

Phase	_Sample list 1_	_Sample list 2_	_Display_
Train 1a (8 trials)	_Look, a **gif**! That's a **gif**!_ Wow, a **geef**! I see a **geef**! _A **kib**! Do you see the **kib**?_ A **keeb**! Where's the **keeb**?	_Look, a **keeb**! That's a **keeb**!_ Wow, a **kib**! I see a **kib**! _A **geef**! Do you see the **geef**?_ A **gif**! Where's the **gif**?	 
Two distractor trials (no words)			
Train 1b (8 trials)	(Same as train 1a)	(Same as train 1a)	(Same as Train 1a)
Test 1 (8 trials)	_Find the **gif**!_ Show me the **geef**! _Can you find the **kib**?_ What's the **keeb**?	_Find the **keeb**!_ Show me the **kib**! _Can you find the **geef**?_ What's the **gif**?	 _or_ 
Two distractor trials (no words)			
Train 2 (8 trials)	(Same as train 1a)	(Same as train 1a)	(Same as Train 1a)
Test 2 (8 trials)	_Can you find the **gif**?_ _What's the *__geef__?_ Show me the **geef**! Find the ***gif**! _What's the **kib**?_ _Can you find the *__keeb__?_ Find the **keeb**! Show me the ***kib!**	_Show me the **keeb**!_ _Find the *__kib__!_ Can you find the **kib**? What's the ***keeb**? _Find the **geef**!_ _Show me the *__gif__!_ What's the **gif**? Can you find the ***geef**?	(Same as Test 1)
Train 3 (8 trials)	(Same as train 1a)	(Same as train 1a)	(Same as Train 1a)
Test 3 (8 trials)	(Same as Test 2)	(Same as Test 2)	(Same as Test 1)

Note. *indicates changed pronunciations. _Underlined italicized_ text is the male talker, standard font is the female talker. For greater visibility, displays are not to scale.

$z = 2.12$; $p = .03$). However, no other effects approached significance, including the effect of Pronunciation (estimate = 0.34; $z = 1.46$; $p = .15$), indicating no difference between the original accents and the swapped accents. However, taken individually, only the OP trials exceeded chance (60.4% ± 23.2%; estimate = 0.48; $z = 2.26$; $p = .02$); swapped-accent trials did not (53.6% ± 15.0%; estimate = 0.15; $z = 0.14$; $p = .31$).

To compare results to Experiment 1, where children learned consistent OPs, a logistic regression model on accuracy included Experiment, Pronunciation, and Block as factors. In this and all following models that compare experiments, only effects and interactions involving Experiment are reported for brevity of presentation. A marginal effect of Experiment (estimate = 0.48; $z = 1.95$; $p = .05$) was qualified by an Experiment x Pronunciation interaction (estimate = 1.08; $z =$

2.97; p =.003).[2] The interaction resulted from higher accuracy for OPs in Experiment 1 vs. Experiment 2 (estimate = 0.98; z = 2.85; p =.004), but no difference in accuracy of APs (estimate = 0.06; z = 0.31; p =.76). This suggests that children are better at encoding single pronunciations than variable pronunciations.

Eye tracking

Eye tracking data (Figure 1d) were analyzed as in Experiment 1. Due to a computer error one child's eye movement data were lost, so only 23 participants were included. Trials were excluded using the same criterion as in Experiment 1 (20.4% of trials). A model of transformed target advantage with Time Bin and Pronunciation as factors showed no difference in overall looks as a function of Pronunciation (t = 1.22, p = .22), but yielded an effect of Time Bin (t = 2.81, p = .005) and a Time Bin x Pronunciation interaction[3] (t = 2.33, p = .02). The interaction resulted from significant increases in looks for OP trials (slope = .0021, SE = .0006, t = 3.79, p = .0002) but a slope no different from 0 for AP trials (slope = .0001, SE = .0006, t = 0.16, p = .87). Recall that the OPs here are *each talker's* original pronunciations, while APs are each talker using the other talker's pronunciations. This interaction suggests a talker-specific accent effect: children's looks increase more rapidly when each talker uses their original accent.

A combined model with Time Bin, Pronunciation, and Experiment as centered predictors yielded a Time Bin x Pronunciation x Experiment interaction (t = 2.04, p = .04),[4] indicating a larger difference between OP and AP slopes in Experiment 1 than in Experiment 2.

Discussion

Accuracy data suggest weaker learning in the current experiment, where children learned under word-form inconsistency, relative to Experiment 1, where children learned on consistent word forms. This suggests that learning two similar words for one object does not lead children to merge over or ignore variability but may be like learning two separate words, along the lines of Sumner and Samuel's (2009) bidialectal listeners, and Sebastián-Gallés et al.'s (2005, 2009) Catalan speakers. This greater difficulty of learning multiple forms is reminiscent of Floccia et al. (2012): they found that 20-month-olds completely failed to recognize one of two possible accent variants for words containing *r* sounds. This study too suggests that learning multiple similar forms may be quite difficult.

However, the eye tracking data provide some evidence consistent with context tracking: children fixated targets more when talkers used their original accents. This could result from children tracking the context of each accent variant—context in this case being the person speaking. It could also result from talker-specific storage of word forms (see Creel, Aslin, & Tanenhaus, 2008; Creel & Tumlin, 2011; Goldinger, 1996, 1998), without any awareness of the relationship

[2]This interaction was significant when considering just the one-vowel participants or just the two-vowel participants from Experiment 1.

[3]Marginal with all trials included (t = 1.89, p = .058).

[4]This interaction was not significant comparing Experiment 2 just to the one-vowel or two-vowel participants.

between two accent-variants. Nonetheless, this result is consistent with at least some degree of context-dependent representation or recognition of accent variants, potentially forming the basis for context-specific storage and recognition of accents more generally.

Experiment 2 thus suggests that presenting learners with two accent variants during learning does not aid them in recognizing variant forms, and in fact may hurt them. This may explain poor generalization in Experiment 1. That is, children may have shown poor generalization in Experiment 1 simply because they did not recognize the similarity between the differently accented forms at all (they do not think *gif* and *geef* are alike), though this is inconsistent with preschool children's ability to recognize slightly mispronounced *familiar* words (Creel, 2012). However, a different explanation for why children in Experiment 1 generalized poorly is not difficulty perceiving similarity, but a *pragmatic bias* to select novel things. That is, during training children are being asked to map unfamiliar words to unfamiliar visual objects. This experience might bias them to treat the accented forms in the test phase as further novel words and seek out novel referents, even if they have some awareness of the similarity to the original pronunciations. Jarvis, Merriman, Barnett, Hanba, and Van Haitsma (2004) have shown that young children who are initially guided to select similarly-named familiar objects rather than novel objects for novel words—such as being taught that "plasses" are a type of glasses—make fewer novel-object identifications for further wordlike names, such as inferring than "cardon" refers to a car rather than a novel object. This suggests that pragmatic cues may enhance word generalization to a similar but not-previously experienced word form.

The next three experiments addressed these alternative explanations for weak accent learning and generalization. All three experiments attempted to counter children's novelty bias by mixing together novel-word test trials with real-word recognition trials to encourage word-form/picture matching (see Fennell & Waxman, 2010, Experiment 2, for a similar approach). If children in Experiments 1 and 2 in the present study were biased by the task to look for novel words and novel referents, then including real-word trials should decrease this tendency, boosting looks to the object matching the word form. Further, children may specifically need stronger evidence that small word alterations should be interpreted as the word itself. If this is the case, then including *accented* familiar-word test trials (Experiment 3) should increase accent-variant recognition more than unaccented familiar-word trials (Experiment 4). On the other hand, children may fail to pick up on the similarity of the accent-variant to the novel word simply because newly-formed word representations are too weakly activated by a partial match. If so, including familiar-word trials should not increase accent-variant accuracy at all.

EXPERIMENT 3

Experiment 3 taught children the words from Experiment 1 but then presented a talker using variable pronunciations of *familiar* words containing /i/ and /ɪ/ during test trials. Note that previous researchers have asked whether exposure to an accent *prior* to word learning facilitates comprehension of a word in that accent later; Schmale et al. (2012) suggest that this is effective for 24-month-olds. A slightly different question is at hand here: did children fail to generalize to the new accent because the task promoted a novelty response bias? Presenting familiar words during test trials, *after* children had already learned the words, aimed to steer children into a pragmatic set of identifying known words rather than looking for new ones; that is, the manipulation was geared toward higher-level reasoning rather than encoding.

If children in Experiment 1 were hindered in their comprehension of accent variants either by lack of evidence for pronunciation alterations, or by task pragmatics pushing them toward novelty responding, they should now show above-chance accuracy and looks to accented variants. However, if children are simply unable to appreciate the similarity between the learned form and the familiar form, they will still fail to identify or look to accented variants.

Method

Participants

N = 24 new preschool participants (mean age = 4.6 years, range = 3.3–5.3; 5 female) from the same pool as before took part.

Procedure

Training and tests mirrored Experiment 1, with a few changes that attempted to match characteristics of an additional experiment (unreported because learning was poor and results were inconclusive). One change from Experiment 1 was that children always learned two words that had the same vowel (*gif* and *kib*, or *geef* and *keeb*). As analyses of Experiment 1 showed, having words with the same vowel (*gif, kib*) or two different vowels (*gif, keeb*) during training did not affect the overall pattern of accuracy. The key change was that, during Tests 2 and 3, 8 additional real-word trials were added per block (16 total; half OPs, half APs). Real-word trials were included to assess whether children would respond more accurately to unusual pronunciations when they were cued to respond on a familiarity basis. This familiarity basis was evoked by familiar-word test trials, where children had to select from one of two familiar objects. Twelve children each heard *bridge, fish*, **breedge*, and **feesh; or feet, keys,* **fit,* and **kizz*.[5]

Results

Accuracy

Accuracy on familiar words was high for both OPs and APs (OP: 93.8% ± 14.3%; AP: 82.3% ± 19.8%), indicating that, as predicted, children mostly parsed these accented forms as real words. That is, preschool children are quite good at compensating for the novel accent on familiar words (as in Creel, 2012). Did this induce a real-word interpretation of *novel*-word APs, in contrast to Experiment 1? Children were highly accurate on OP words (Figure 2a) in Block 1 (72.4% ± 27.6%; estimate = 1.49; $z = 3.71$; $p = .0002$) and in Blocks 2 and 3 (70.3% ± 21.4%; estimate = 0.98; $z = 4.31$; $p < .0001$). Unlike Experiment 1, AP accuracy also exceeded chance

[5]Due to talker availability, only the female talker produced the real words. This provided an opportunity to assess talker-specific accent processing: if children improve their recognition of accented novel words, will they improve only for the female talker, or would they generalize to the male talker as well? Half the children learned from and were tested on novel words with a single talker (male *or* female), and the other half learned from and were tested on both talkers. Preliminary analyses indicated no effects or interactions of number of training talkers (one vs. two) or talker gender, so these factors were dropped from further analyses.

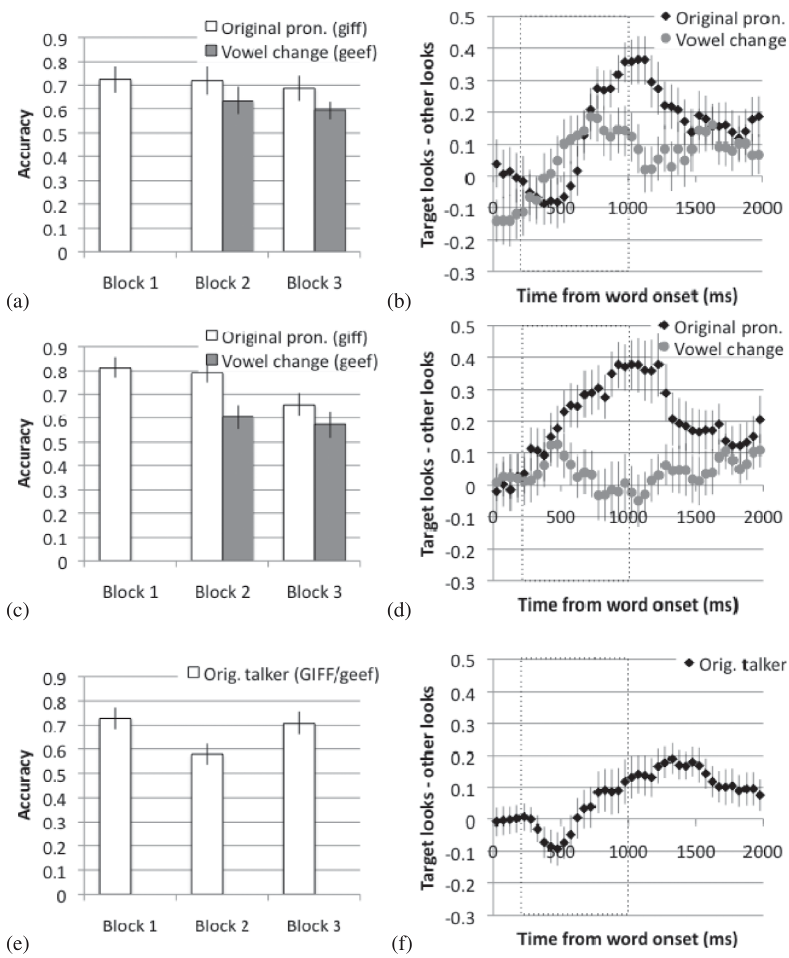


FIGURE 2 Experiment 3, (a) accuracy and (b) target looks minus other looks, Blocks 2–3; Experiment 4, (c) accuracy and (d) target looks minus other looks, Blocks 2–3; Experiment 5, (e) accuracy and (f) target looks minus other looks, Blocks 2–3.

(61.5% ± 14.7%; estimate = 0.47; $z = 3.15$; $p = .002$). A logistic regression model on accuracy with Block and Pronunciation as factors detected a marginal effect of Pronunciation (estimate = 0.53; $z = 1.89$; $p = .06$), indicating slightly greater accuracy on OPs than APs.

Another logistic regression model compared Experiment 3 to Experiment 1, which was nearly identical aside from the familiar-word trials, with Experiment (1, 3), Block, and Pronunciation as factors. There was no effect of Experiment (estimate = .02, $z = .23$, $p = .82$), but there was an interaction of Experiment x Pronunciation (estimate = 0.30; $z = 2.35$; $p = .02$). This resulted from a smaller decrement in accuracy from OPs to APs—that is, better generalization to accented forms—in the current experiment than in Experiment 1.

Eye tracking

Transformed target advantage (Figure 2b) was analyzed as in previous experiments (15.4% of trials excluded). The intercept term was significant ($t = 2.61$, $p = .009$), indicating above-chance looks toward the target overall. There was an effect of Time Bin ($t = 4.90$, $p < .0001$), indicating that target looks increased over time (slope = .0020, $SE = .0004$). However, there was no effect of Pronunciation ($t = 0.11$, $p = .91$) nor an interaction ($t = 1.17$, $p = .24$), suggesting that neither looks nor increases in looks over time differed greatly as a function of pronunciation in the time window analyzed. An additional model with Experiment (1, 3), Pronunciation, and Time Bin as factors yielded an Experiment x Pronunciation x Time Bin interaction ($t = 2.54$; $p = .01$). This indicated a smaller difference between AP and OP looking increases over time in Experiment 3—consistent with a decreased effect of Pronunciation.

Discussion

When children were presented with accented familiar-word trials amongst novel-word recognition trials, they showed significant recognition of accent-variants of *novel* words. Unlike Experiment 1, which suggested that children did not recognize partial match to the learned forms, this experiment suggests that children recognized the partial match of accent-variants to the original words. This implies that children *are* sensitive to the similarity between the two accented forms, which would allow merging of representations or context-specific representations. This experiment is consistent with two accounts: that familiar-word trials alerted children to the possibility of accented variants; or, that familiar-word trials simply put children in a familiar-word pragmatic "mode" (as opposed to a detecting-novelty mode). The next experiment attempted to tease apart these explanations by again including familiar word trials, but *without* accented pronunciations.

EXPERIMENT 4

This experiment replicated Experiment 3 but presented only unaccented familiar words. If children can generalize to a new accented variant of a learned word but simply need to know to select familiar words rather than always looking for novel pictures, then hearing *un*accented familiar words alone should lead children to show above-chance accuracy and looks to novel APs. However, if children need to know specifically that talkers may use unfamiliar pronunciations, then AP accuracy and looks should be at chance.

Method

Participants

$N = 24$ new preschool participants (mean age = 4.6 years, range = 3.3–5.3; 12 female) from the same pool as before took part.

Procedure

Training and Test 1 mirrored Experiment 3, but the design was simplified so that all children heard both talkers equally often during learning. Tests 2 and 3 contained 16 total real-word trials, all unaccented. Each talker spoke the words *pizza, zebra, pig,* and *milk* twice. Words differed relative to Experiment 3 to provide a broader familiar-word set. The lack of difference between Experiments 3 and 4 suggests that this choice of words was not consequential.

Results

Accuracy

Children were highly accurate on familiar-word trials (94.3% ± 9.6%). They were also accurate on novel OPs in Block 1 (Figure 2c; 81.3% ± 22.1%; estimate = 2.00; $z = 5.43$; $p < .0001$) and Blocks 2–3 (72.4% ± 18.8%; estimate = 1.02; $z = 5.24$; $p < .0001$). Further, they again exceeded chance on novel APs (58.9% ± 18.2%; estimate = 0.47; $z = 3.15$; $p = .002$). A logistic regression model with Block and Pronunciation as factors showed an effect of Pronunciation (estimate = 0.73; $z = 3.02$; $p = .002$), indicating greater accuracy for OPs than APs. There was also an effect of Block (estimate = 0.50; $z = 2.06$; $p = .04$), reflecting a decrease in accuracy from Block 2 (79.1% ± 21.7%) to Block 3 (65.5% ± 24.2%). Why this decrease occurred is not clear; it might represent chance fluctuation, boredom, or even increased confusion after introducing novel-word APs without accompanying real-word APs. Though the decrease appeared to be carried by OPs alone, the interaction did not approach significance.

To compare results with the previous experiment, which contained familiar word APs, a new logistic regression model was computed with Experiment (3, 4), Block, and Pronunciation as factors. There was no main effect of Experiment (estimate = 0.01; $z = 0.06$; $p = .95$), nor an Experiment x Pronunciation interaction (estimate = 0.27; $z = 0.77$; $p = .44$), suggesting that the two experiments showed equivalent results: hearing "accented" familiar words (Experiment 3) in addition to unaccented familiar words did not increase accuracy. A comparison to Experiment 1 was more ambiguous: the Experiment x Pronunciation interaction approached significance (estimate = 0.16, $z = 1.77$, $p = .08$). This resulted from a marginally smaller decrement in accuracy from OPs to APs in the current experiment than in Experiment 1.

Eye tracking

Target advantage (Figure 2d) was analyzed as before (16.4% of trials excluded). Time Bin was not significant ($t = 1.14$, $p = .25$), but Pronunciation was ($t = 2.55$, $p = .01$), reflecting higher overall target advantage for OP trials. A Time Bin x Pronunciation interaction ($t = 2.03$, $p = .04$) resulted from a positive slope on OP trials (slope = .0018, $SE = .0007$, $t = 2.43$, $p = .02$) but a nonsignificant slope on AP trials (slope = -.0004, $SE = .0009$, $t = 0.50$, $p = .62$). In a model comparing this experiment to Experiment 1, which did not have real-word test trials and may have thus generated a novelty-selection bias in children, there was a marginal effect of Experiment (t = 1.93; p = .05), indicating greater looks overall in Experiment 4 than in Experiment 1. This is consistent with overall slightly better recognition in Experiment 4, but there were no interactions with Pronunciation,[6] suggesting that there was not a marked difference in looks relative to Experiment 1. Compared with Experiment 3, which contained AP real-word trials, there was a marginal Experiment x Time Bin interaction (t = 1.82, p = .07) due to a faster increase in looks with time in Experiment 3. This indicates a tendency toward *overall* faster looking when real-word AP trials are included, but given its marginal nature, should be interpreted cautiously. Further, there were no interactions with Pronunciation, suggesting that there were not substantial differences in looking patterns in the two experiments. These comparisons suggest that Experiment 4 occupies an intermediate position between Experiments 1 and 3.

Discussion

In summary, accuracy data from Experiments 3 and 4 suggest that children are more likely to interpret accent-like variants of newly learned words as being the words themselves if the pragmatic context reinforces word form-picture matching. Eye tracking results alone hinted at a slight benefit for hearing *accented* real words. However, Experiment 4 performance was only marginally better than Experiment 1, and was not clearly worse than Experiment 3. This makes it somewhat difficult to pin down what task aspects led to the best accent-variant recognition. It is possible that hearing real-word accent variants is more beneficial than hearing only canonically pronounced real-word trials, but it is also possible that there is no benefit.

A final question is whether a pragmatic benefit extends to situations of word-form inconsistency. If pragmatically cued to select familiar things, will children process words learned in *two* accents more effectively?

EXPERIMENT 5

In this final experiment, children again learned in two accents. As in Experiments 3 and 4, accent-consistent trials were interspersed during test. These trials were included with the idea that low accuracy in Experiment 2 may have been driven not by poor learning but rather by children's construal of the task as a selecting-novel-things task. If children are more willing to recognize the two word forms across accents given familiar-word trials, then accuracy and looks should be well above chance. However, if the difficulty in Experiment 2 was not pragmatic but an actual encoding difficulty, then accuracy and looks should be relatively low. This experiment also provided a modified replication of Experiment 2, retesting the hypothesis that if children are simply learning two word-forms when presented with word-form inconsistency, accuracy and looking time will be relatively poor.

[6]With all trials included, there was a marginal interaction of Experiment x Pronunciation x Time Bin, t = 1.79, p = .07.

Method

Participants

N = 24 new preschool participants (mean age = 4.6 years, range = 3.4–5.5; 10 female) from the same pool as before took part.

Procedure

Training and Test 1 matched Experiment 2—each talker had a characteristic “accent.” For maximum consistency, each talker continued to use only their own accent throughout Tests 2 and 3 (see Table 2); note that this differs from Experiment 2, which presented talker-inconsistent trials in Tests 2 and 3. During Tests 2 and 3, real-word trials were added (16 total). Each talker’s 8 real-word trials were also accent-consistent: the “accented” talker’s productions contained the /i/-/ɪ/ merger. For instance, the male talker might say novel words *geef* and *keeb,* and real words *pizza, zebra,* **peeg,* and **meelk*—producing both vowels as /i/—while the female talker said *giff, kib, pizza, zebra, pig,* and *milk*, distinguishing the two vowels.

Results

Accuracy

Real-word accuracy was high even for APs (OP: 94.8% ± 13.4%; AP: 92.7% ± 18.8%), suggesting that children interpreted these variants as familiar. Children were above chance (Figure 2e) in recognizing novel words in Block 1 (72.9% ± 19.7%; estimate = 1.08; $z = 5.13$; $p < .0001$) and Blocks 2–3 (64.3% ± 16.3%; estimate = 0.62; $z = 4.21$; $p < .0001$). As this experiment contained no novel APs (deviations from the trained words), the logistic regression model contained only an effect of Block (2, 3). The effect of Block (estimate = 0.64, $z = 2.84$, $p = .004$) indicated greater accuracy in Block 3 (70.8% ± 20.4%) than in Block 2 (57.8% ± 20.0%). It is not clear why accuracy would differ across blocks; it may reflect noise variability, or confusion and subsequent recovery after presentation of several AP real-word trials.

A second model compared accuracy to OP accuracy Experiment 2, which also asked children to learn in two accents simultaneously. The model contained Experiment (2 vs. 5) and Block (2, 3) as factors. An Experiment x Block interaction (estimate = .20, $z = 2.15$, $p = .03$) resulted from the Block effect in Experiment 5 which did not appear in Experiment 2. However, there was no effect of Experiment (estimate = .05, $z = .75$, $p = .45$), indicating that accuracy did not improve overall from Experiment 2 to Experiment 5 as a result of including familiar words. Two additional models compared accuracy in the current experiment to OP accuracy in Experiments 3 and 4, where familiar words were intermingled with novel words but the learned accent was not variable. Compared with Experiment 3, there was a marginal Experiment x Block interaction (estimate = 0.38, $z = 1.95$, $p = .051$), presumably resulting from the significant Block effect in the current experiment. However, overall accuracy (effect of Experiment) did not differ (estimate = 0.14, $z = 1.11$, $p = .27$). Compared with Experiment 4, an Experiment x Block interaction (estimate = 1.32, $z = 3.28$, $p = .001$) occurred due to the previously reported Block effects in opposite directions in the two experiments. The effect of Experiment did not reach significance (estimate $= 0.41$, $z = 1.62$, $p = .11$), despite a numerical tendency toward higher accuracy in Experiment 4.

Eye tracking

Eye movement data were analyzed as before (15.6% of trials excluded). Despite high accuracy on novel-word trials, there was little evidence of looks to the correct picture (Figure 2f) in the observed time window. A regression analysis like those in previous experiments showed a marginal effect of Time Bin[7] (slope $= .0009$, $SE = .0006$; $t = 1.68$, $p = .09$), suggesting that—at least within the observed window—there was not an upswing in target looks. Compared with Experiment 2 (OPs only), no effects related to Experiment approached significance, indicating similar patterns of looking in the two experiments. Compared with Experiment 3 (OPs only), there was an Experiment x Time Bin interaction ($t = 2.08$, $p = .04$), with a faster increase in looks in Experiment 3, with consistent pronunciations. Similarly, compared with Experiment 4 (OPs only), there was an overall effect of Experiment ($t = 3.18$, $p = .001$), due to greater looks in Experiment 4 (again with consistent pronunciations) than in the current experiment.

Discussion

Accuracy in Experiment 5 was not markedly lower than in Experiments 3 and 4, suggesting a possible benefit of pragmatic orientation on recognizing inconsistent word-forms. However, visual fixations suggested that, as in Experiment 2, accent inconsistency affected *speed* of recognition. This replicates the difference observed between Experiments 1 and 2, further suggesting that accent inconsistency slows learning. Note that the current experiment is a particularly conservative test of the difficulty of word-form variability relative to Experiment 2, in that children in the current experiment had two additional pieces of information: accent-consistent familiar-word test trials, and maintenance of talker-consistent accents throughout. This fits with Sebastián-Gallés et al.'s (2005, 2009) account that listeners who distinguish a sound pair must learn two different words for the same referent and is not consistent with merging across inconsistency, which would presumably lead to higher accuracy and faster recognition.

GENERAL DISCUSSION

This study set out to discover how children process accent inconsistency in word learning. First, can children recognize accent-variants (e.g., *gif*) based on partial match to the learned form (*geef*) without ever hearing the accent variant? Overall, the answer seems to be yes: children seem able to recognize an unheard accent variant based on partial match to the learned form. When children learned one pronunciation of each novel word (Experiments 1, 3, and 4), they showed moderate generalization (above-chance accuracy) to accent variants as long as the recognition task included real-word recognition trials (Experiments 3 and 4). While Experiment 1 appeared to suggest that children cannot use partial match for novel words, this appears to be due to children's

[7] Not significant when analyzing all data ($t = 1.58$, $p = .11$).

pragmatic construal of the referent-selection task. When real-word trials were interspersed with newly learned word trials, children showed above-chance recognition of accent variants, without needing explicit training on the other accent. Nonetheless, accuracy for accent variants was lower than for the originally-learned pronunciations, consistent with children's slightly-lower accuracy in recognizing accent variants of familiar words (Creel, 2012).

A second set of questions concerned learning under accent variability (both *gif* and *geef* as labels for the same object). When exposed to variable accents, would children *ignore* inconsistent elements, encode *two forms* for each word (one per accent), or store the *context* in which each word occurs? Experiments where children learned in two accents simultaneously suggested that learning in two accents is more difficult than learning in one accent. Specifically, children were somewhat less accurate in recognizing trained pronunciations (Experiment 2), and were slower to visually fixate targets (Experiments 2 and 5) relative to the single-accent experiments. Assuming that learning more words is harder than learning fewer words, this implies that children may be encoding two forms for each referent. These results suggest that children did not merge across vowel inconsistency, but learned dual word representations from two-accent input. Interestingly, the eye tracking data in Experiment 2 suggested some degree of talker specificity in two-accent learning: looks increased more rapidly when the talkers spoke in their original accents than when the talkers swapped accents. This is consistent with children learning the *contexts* of each accented form. However, it is also consistent with children simply encoding four talker-specific word forms, without necessarily understanding that two accent-variants were related to each other.

The current results suggest that children can recognize accent variants via partial match to a learned form. However, though they can recognize incomplete similarity, children do not seem to merge over or ignore accent variability. Results are more consistent with storing multiple word forms, with some indication of context-specific storage of accent variants. Nonetheless, the results do not clearly distinguish between storage of multiple forms, vs. storing a single form plus knowledge of its realization in different contexts.

Implications for Learning Language with Accent Variability

These results have several implications for accent processing in 3-5-year-olds. First, newly learned words are highly sensitive to accent-like changes: children have difficulty recognizing a word that is changed by a single vowel feature. While this fits accounts of detailed novel-word representations in toddlers (Mani & Plunkett, 2008), it does not lead to strong recognition of newly learned words in a new accent. A pragmatically supportive context (though not necessarily exposure to the accent) aids recognition of accented forms, though accuracy is still lower than for learned pronunciations.

A second implication is that learning from accent-variable input may be more difficult than learning in one accent. It is still possible that, with lengthier exposure, learners might be able to disregard accent variability, or that children might succeed with greater within-talker variability (see Richtsmeier et al., 2011) instead of talkers using internally consistent accents. However, for monolingual English speaking children, brief experimental exposure appears insufficient to generate a merged percept.

An additional possibility raised in the Introduction, and partially supported in Experiment 2, was that learners might ascribe accent variability to the context (say, the talker or talkers with a particular accent), rather than to the word form itself. This would presumably allow quicker word learning because both variants would map to one underlying form. Presumably, extended exposure to two accents would afford development of phonological translation abilities, which would ameliorate effects of word-form inconsistency during learning (see Mattock et al., 2010, for results consistent with this account; though note that Floccia et al., 2012, is inconsistent with this account). A variant of this notion of context encoding, untested here, is that children might be aided in forming context-dependent accent variants by high cross-accent similarity; that is, learning that gif = geef (a phonetically close change) might be easier than gif = keece. Note that children here appeared to have difficulty learning two very similar forms for the same referent, despite high cross-accent similarity. If cross-accent similarity facilitates learning, then children in the current study might have had even greater difficulty in learning multiple variants if the word forms had been further apart. A remaining possibility is that there is a *deleterious* effect of cross-accent similarity, such that it would actually be easier to learn two very different words for the same referent than to learn two very-similar words. This might result if a word's two labels compete with each other for recognition.

A final implication of learning with accent variability concerns novelty interpretation. Results here, like those of Jarvis et al. (2004), suggest that pragmatics influence whether children interpret an accented form as novel or familiar. Perhaps over the longer term, children in multiaccent contexts would become accustomed to familiar-word interpretations of similar variants. This is supported by research on multilingual children, who commonly experience situations where a novel word form is simply a word from another language, not a novel object: Byers-Heinlein and Werker (2009) showed that bilingual and trilingual 18-month-olds appeared not to use mutual exclusivity to constrain the referent of a novel word, while monolingual children did; Houston-Price, Caloghiris, and Raviglione (2010) found similar results in a slightly older sample of 17–22 month-olds. This is consistent with multilingual children *not* learning a bias to assume that novel forms mean novel referents. Children in multiple-*accent* environments might acquire a similar pragmatic bias to interpret partially matching forms as familiar.

Accent Processing Across Age

One goal of the current study was to connect the literature on accent processing in infants and toddlers to that in adults. The current results with preschoolers showed reasonable consistency between implicit (looking) measures, as used with younger children, and in explicit (accuracy) measures, as used with adults. This suggests that these measures may in fact be tapping similar variables. What picture emerges of accent processing across development?

First, generalization of forms to new accents appears to improve with age. Very young infants (9 months) cannot generalize a familiarized word form to a new accent, but they can do so by 12–13 months (Schmale et al., 2010; Schmale & Seidl, 2009). By 14 months, children show partial recognition of mispronounced familiar words in a referential preferential-looking task (Swingley & Aslin, 2002), and by 19 months can acclimate to an accent change (White & Aslin, 2011). Standing in contrast to White and Aslin's study, Floccia et al.'s (2012) study with 20-month-olds suggests that children receiving long-term input in two accents—one rhotic, the other non-rhotic—may actively *tune out* one of the accents. Alternately, Floccia et al.'s participants may be showing some talker specificity in their representations: if only one or two speakers (one or both parents) in the child's environment use a non-rhotic accent, the child may recognize non-rhotic words only for those specific talkers (akin to the talker-specificity seen in Experiment 2 of the current study). At 24 months, children can generalize a single newly-learned word-meaning mapping to an accented form if given preexposure to the accent (Schmale et al., 2012), and by 30 months they can generalize a single new word to an unfamiliar accent *without* exposure (Schmale et al., 2011). The current study shows that 3-5-year-olds can generalize *two* new forms to accented variants provided that they are not guided into a pragmatic set of novelty selection. Finally, adults can recognize accented words via partial match to a known word form (e.g., Bradlow & Bent, 2008). The tasks accomplished at increasing ages increase in difficulty (lower word-form familiarity, more words to learn). Thus, generalization to accented forms appears to increase slowly and steadily with age.

It is less clear how multiple-accent encoding changes with age. Further research is required, though Mattock et al. (2010)'s infant study, Floccia et al.'s work with toddlers, and the current study with preschoolers provide hints that children are at least as affected by word-form inconsistency as adults.

CONCLUSION

Preschoolers learned words in the presence of accent-like variability. When learning in a single accent, they partially transferred word knowledge to a different accent as long as recognition trials included some familiar words. When learning in two accents, performance dropped, implying that word-form inconsistency slowed encoding. This may have been because children were unable to ignore accent variability, so that they were forced to encode additional word forms (one per accent). Together with previous research, this study suggests steady improvement in accent processing over the course of development. Future work should assess effects of different types of longer-term accent exposure on ability to comprehend variant forms.

ACKNOWLEDGMENTS

Portions of this work were presented at the 2012 Boston University Conference on Language Development. Thanks to Heather Pelton and Adrienne Moore for running child participants; to children, parents, and schools for participating; and to Carolyn Quam for comments on an earlier draft. This work was supported by NSF CAREER Award BCS-1057080 to SCC.

REFERENCES

Albareda-Castellot, B., Pons, F., & Sebastián-Gallés, N. (2011). The acquisition of phonetic categories in bilingual infants: New data from an anticipatory eye movement paradigm. *Developmental Science*, *14*(2), 395–401.

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*(4), 457–474.

Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C. A. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science*, *20*(5), 539–542.

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*, 707–729.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Brunellière, A., Dufour, S., Nguyen, N., & Frauenfelder, U. H. (2009). Behavioral and electrophysiological evidence for the impact of regional variation on phoneme perception. *Cognition*, *111*(3), 390–396.

Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: infants' language experience influences the development of a word-learning heuristic. *Developmental Science*, *12*(5), 815–823.

Chambers, K. E., Onishi, K. H., & Fisher, C. (2011). Representations for phonotactic learning in infancy. *Language Learning and Development*, *7*(4), 287–308.

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America*, *116*(6), 3647–3658.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The eyelink toolbox: Eye tracking with MATLAB and the psychophysics toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 613–617.

Creel, S. C. (2012). Phonological similarity and mutual exclusivity: On-line recognition of atypical pronunciations in 3–5-year-olds. *Developmental Science*, *15*(5), 697–713.

Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2008). Heeding the voice of experience: The role of talker variation in lexical access. *Cognition*, *106*, 633–664.

Creel, S. C., & Tumlin, M. A. (2011). On-line acoustic and semantic interpretation of talker information. *Journal of Memory and Language*, *65*(3), 264–285.

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-IV: Peabody picture vocabulary test* (4th ed.). Circle Pines, MN: American Guidance Service.

Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics*, *67*(2), 224–238.

Fallon, M., Trehub, S. E., & Schneider, B. A. (2002). Children's use of semantic cues in degraded listening environments. *Journal of the Acoustical Society of America*, *111*(5), 2242–2249.

Fennell, C. T., & Waxman, S. R. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development*, *81*(5), 1376–1383.

Floccia, C., Delle Luche, C., Durrant, S., Butler, J., & Goslin, J. (2012). Parent or community: where do 20-month-olds exposed to two accents acquire their representation of words? *Cognition*, *124*(1), 95–100.

Fox, R. A., Flege, J. E., & Munro, M. J. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *Journal of the Acoustical Society of America*, *97*(4), 2540–2551.

Gerken, L., Murphy, W. D., & Aslin, R. N. (1995). Three- and four-year-olds' perceptual confusions for spoken words. *Perception & Psychophysics*, *57*(4), 475–486.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(5), 1165–1183.

Goldinger, S. D. (1998). Echoes of echoes?: An episodic theory of lexical access. *Psychological Review*, *105*(2), 251–279.

Hallett, P. E. (1986). Eye movements. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 10.1–10.112). New York, NY: Wiley.

Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy*, *15*(2), 125–150.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*(2), 151–195.

Hudson Kam, C. L., & Newport, E. L. (2009). Getting it right by getting it wrong: when learners change languages. *Cognitive Psychology*, *59*(1), 30–66.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.

Jarvis, L. H., Merriman, W. E., Barnett, M., Hanba, J., & Van Haitsma, K. S. (2004). Input that contradicts young children's strategy for mapping novel words affects their phonological and semantic interpretation of other novel words. *Journal of Speech, Language, and Hearing Research*, *47*, 392–406.

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, *51*(2), 141–178.

Mani, N., & Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Developmental Science*, *11*(1), 53–59.

Mattock, K., Polka, L., Rvachew, S., & Krehm, M. (2010). The first steps in word learning are easier when the shoes fit: Comparing monolingual and bilingual infants. *Developmental Science*, *13*(1), 229–243.

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science*, *32*, 543–562.

Merriman, W. E., & Schuster, J. M. (1991). Young children's disambiguation of object name reference. *Child Development*, *62*(6), 1288.

Nathan, L., Wells, B., & Donlan, C. (1998). Children's comprehension of unfamiliar regional accents: A preliminary investigation. *Journal of Child Language*, *25*(2), 343–365.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127–162.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*(2), 204–238.

Oller, D. K., Cobo-Lewis, A. B., & Eilers, R. E. (1998). Phonological translation in bilingual and monolingual children. *Applied Psycholinguistics*, *19*, 259–278.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*(3), 353–363.

Ramon-Casas, M., & Bosch, L. (2010). Are non-cognate words phonologically better specified than cognates in the early lexicon of bilingual children? In M. Ortega-Llebaria (Ed.), *Selected proceedings of the 4th conference on laboratory approaches to Spanish phonology* (pp. 31–36). Somerville, MA: Cascadilla Press.

Ramon-Casas, M., Swingley, D., Sebastián-Gallés, N., & Bosch, L. (2009). Vowel categorization during word recognition in bilingual toddlers. *Cognitive Psychology*, *59*(1), 96–121.

Richtsmeier, P., Gerken, L., & Ohala, D. (2011). Contributions of phonetic token variability and word-type frequency to phonological representations. *Journal of Child Language*, *38*(5), 951–978.

Schmale, R., Cristia, A., & Seidl, A. (2012). Toddlers recognize words in an unfamiliar accent after brief exposure. *Developmental Science*, *15*(6), 732–738.

Schmale, R., Cristià, A., Seidl, A., & Johnson, E. K. (2010). Developmental changes in infants' ability to cope with dialect variation in word recognition. *Infancy*, *15*(6), 650–662.

Schmale, R., Hollich, G., & Seidl, A. (2011). Contending with foreign accent in early word learning. *Journal of Child Language*, *38*(5), 1096–1108.

Schmale, R., & Seidl, A. (2009). Accommodating variability in voice and foreign accent: Flexibility of early word representations. *Developmental Science*, *12*(4), 583–601.

Sebastián-Gallés, N., Echeverria, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language*, *52*(2), 240–255.

Sebastián-Gallés, N., Vera-Constán, F., Larsson, J. P., Costa, A., & Deco, G. (2009). Lexical plasticity in early bilinguals does not alter phoneme categories: II. Experimental evidence. *Journal of Cognitive Neuroscience*, *21*(12), 2343–2357.

Smith, K., & Wonnacott, E. (2010). Eliminating unpredictable variation through iterated learning. *Cognition*, *116*(3), 444–449.

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*, 381–382.

Storkel, H. L. (2002). Restructuring of similarity neighborhoods in the developing mental lexicon. *Journal of Child Language*, *29*(2), 251–274.

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language*, *60*(4), 487–501.

Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: Bilingual infants keep pace with their monolingual peers. *Cognition*, *108*(1), 232–242.

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147–166.

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*, 480–484.

Waxman, S. R., & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: new evidence from 12-month-old infants. *Cognition*, *95*(3), B59–68.

Weil, S. A. (2001). *Foreign-accented speech: Adaptation and generalization*. (Unpublished doctoral dissertation). Ohio State University, Columbus, OH.

White, K. S., & Aslin, R. N. (2011). Adaptation to novel accents by toddlers. *Developmental Science*, *14*(2), 372–384.

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*(1), 114–132.

Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition*, *85*(3), 223–250.