



# Mastery of Echoics in Chinese Establishes Bidirectional Naming in Chinese for Preschoolers with Naming in English

Yu Cao<sup>1,2</sup>  · R. Douglas Greer<sup>3</sup>

Published online: 8 January 2019

© Association for Behavior Analysis International 2019

## Abstract

The onset of the verbal behavior developmental cusp of bidirectional naming (BiN) in a second language makes it possible for monolingual English-speaking children to learn names of things in a second language incidentally. We conducted 2 experiments to identify why monolingual English-speaking children cannot demonstrate BiN in another language when they demonstrated BiN in their native language. In Experiment I, using a group design ( $n = 32$  preschoolers), we identified Chinese speech sounds that monolingual English-speaking children with BiN in English for familiar stimuli could not echo. In Experiment II, using a multiple-probe design, we investigated if mastery of echoics with the speech sounds identified in Experiment I would result in BiN in Chinese with 6 participants from Experiment I. The dependent variable was untaught responses to the probe stimuli presented following the naming experience based on the echoic stimuli from Experiment I. The results showed that echoic training was functionally related to the establishment of BiN in the second language. It appeared that the emission of accurate echoics might be the key to second-language BiN and that emergent correspondence between producing and hearing that occurs with the mastery of the echoic responding may be the source of reinforcement.

**Keywords** Naming · Bilingual naming · Verbal behavior developmental cusp · Fast mapping · Echoic training for naming · Familiar and nonfamiliar stimuli and naming

How young children acquire language, especially words, has attracted researchers' interests in both cognitive psychology (Carey & Bartlett, 1978) and verbal behavior

---

✉ Yu Cao  
yc2687@tc.columbia.edu

<sup>1</sup> Department of Health and Behavior Studies, Teachers College, Columbia University, 525 W. 120th St., New York, NY 10027, USA

<sup>2</sup> Present address: Gotham Children, New York, NY, USA

<sup>3</sup> Department of Health and Behavior Studies, Teachers College, Columbia University, 529 A Thorndike, 525 W. 120th St., New York, NY 10027, USA

analysis (Greer & Longano, 2010; Horne & Lowe, 1996; Miguel, 2016; Skinner, 1957). Although the majority of studies from both perspectives were conducted to understand the acquisition of one's first language, or English, some researchers attempted to apply the two theories to the acquisition of the second language or with bilingual populations (Kan & Kohnert, 2008; Kan, Sadagopan, Janich, & Andrade, 2014; Mosca, 2015).

Skinner's (1957) analysis of verbal behavior established the foundation for a full behavioral analysis of communicative behavior, including the lexicon and the nonlexicon effects of a speaker on the behavior of a listener and vice versa. One of the most important aspects of his theory was the contention that the listener and speaker are initially developed independently, and the listener and speaker are joined as a function of experiences. The point at which one demonstrates the speaker as one's own listener signifies the establishment of naming (Horne & Lowe, 1996). Miguel (2016) proposed the term *bidirectional naming* (BiN) to describe the bidirectional relation between speaker and listener behaviors involved in naming. The establishment of BiN allows the words learned as a listener (unidirectional naming or UiN) under certain naming experiences to be available as a speaker without instruction, or the learning of the listener and the speaker simultaneously as a result of the naming experiences (Greer & Du, 2015a; Greer & Longano, 2010; Greer & Ross, 2008; Greer & Speckman, 2009; Miguel, 2016).

The term *naming experience* refers to certain conditions where a child is exposed to word-object or word-objects-actions as an assessment of whether the child can learn the relations without direct instruction. Different naming experiences have been used to assess the presence or the absence of BiN. Many of the studies with preschoolers have used a naming experience that includes saying the name of a novel stimulus as the child is required to match the stimulus with the correct stimulus and is reinforced for matching (e.g., Fiorile & Greer, 2007; Gilic & Greer, 2011; Greer, Stolfi, Chavez-Brown, & Rivera-Valdes, 2005). This is done to ensure that the young children are indeed observing the stimuli named. Observing and listening to others tacting visual stimuli may function as naming experiences for individuals with a well-established naming repertoire (e.g., Carnerero & Pérez-González, 2014; Lo, 2016; Longano & Greer, 2014; Woolslayer, 2013). In a recent unpublished dissertation, Orlans (2017) found no difference in the two procedures using an experimental- and control-group design.

The importance of BiN has been studied extensively (see Miguel, 2016); however, considerable evidence suggests that BiN does not emerge without design for all children (Fiorile & Greer, 2007; Greer et al., 2005; Greer & O'Sullivan, 2007). Researchers have identified environmental contingencies that lead to the establishment of BiN. Multiple-exemplar instruction (MEI) was reported to be an effective procedure to induce naming when it was absent (Fiorile & Greer, 2007; Gilic & Greer, 2011; Greer et al., 2005; Greer, Stolfi, & Pistoljevic, 2007; Helou-Care, 2008; Woolslayer, 2013). MEI across listener and speaker responses involves the rapid rotation of listener and speaker responses, as well as multiple sets of stimuli. The procedure facilitates the joining of speaker and listener responses within the same skin. It is thought that providing concentrated instruction with alternating presentations of stimuli across training sets of speaker and listener responses results in the transformation of stimulus function across speaker and

listener responses for novel stimuli (Fiorile & Greer, 2007; Gilic & Greer, 2011; Greer et al., 2005, 2007).

Other interventions have also been reported to establish BiN, such as the intensive tact-teaching protocol (Greer & Du, 2010; Pistoljevic & Greer, 2006). The onset of BiN has also been correlated with training general auditory identity matching (Choi, Greer, & Keohane, 2015). These findings suggest a more fundamental process may be the source of BiN, leading to a series of experiments by Longano and Greer (2014). They reported that MEI did not result in the acquisition of BiN for four participants in one experiment. In the next study, they found the echoic intervention was successful for three out of four participants. The participant that did not demonstrate BiN received an intervention in which visual stimuli and speech stimuli were conditioned as reinforcers for the separate observing responses, looking and listening, as determined by preference measures. After the intervention, the participant demonstrated BiN. In the final experiment, four participants who did not demonstrate BiN received the conditioning intervention alone and subsequently demonstrated BiN. It appears that increases in reinforcement strength in the intervention sessions correlated with the emission of voluntary echoics; that is, incremental improvement in transformation of stimulus control from listener to speaker was closely correlated with increases in voluntary echoics during training sessions to establish BiN. This suggested that the auditory product of one's echoic behavior, an auditory and listener stimulus, may reinforce the bidirectional relation between speaker and listener responses, as Horne and Lowe (1996) theorized. However, the Longano and Greer (2014) data suggested that the product of the echoic (i.e., the speech sound) appeared to be a conditioned reinforcer. Their participants demonstrated BiN as a function of conditioning the echoic product as a reinforcer and conditioning the visual stimuli as reinforcers for looking. In both responses, looking and listening, the reinforcement stimulus control may have become embedded in the auditory stimuli as the visual stimuli also reinforced the looking response. At present, it is possible that the fundamental source of BiN might be the conditioned reinforcement for observing responses for visual and speech stimuli (the spoken word) simultaneously. This simultaneous reinforcement then leads to the joining of the listener, speaker, and looking. The automatic reinforcement that resulted from echoics (speech sound), in which there is correspondence between the auditory product of the echoic and the auditory stimuli heard in the naming experience, might be the reinforcement for the echoic that then joins the reinforcement for looking at the visual stimulus.

Our understanding about language development in a first language might shed light on how children learn a second language. Researchers have proposed multiple theories to understand the variables that affect learning a second language. Those who hold the "critical period hypothesis" (Johnson & Newport, 1989; Lenneberg, Chomsky, & Marx, 1967) argue that age is a significant variable that affects the proficiency an individual could achieve in a second language. Others have focused on the role of first language and language-specific experiences (Hansen Edwards & Zampini, 2008; Jeong et al., 2007; Kisilevsky et al., 2003; Krueger & Garvan, 2014; Moon, Lagercrantz, & Kuhl, 2013). The first-language experience affects the speech perception (i.e., listener responding) and speech production (i.e., speaker responding) in a second language around the same time, which occurs around 8 to 12 months (Kuhl, 2004; Werker & Tees, 1984; Zhang & Wang, 2007). The similarities between the second and first

language facilitate the acquisition of a second language, whereas the differences between the two languages make the acquisition of a second language more challenging (Jeong et al., 2007). However, the differences between the two languages related to speech perception and production could be overcome by intensive training and speech practice (i.e., echoic training; Iverson & Evans, 2009; Kan & Kohnert, 2008; Kuhl, Tsao, & Liu, 2003; Nishi & Kewley-Port, 2007).

If BiN is the behavioral developmental cusp and the new learning capability that allows a child to learn his or her native language incidentally, can BiN be extended to the learning of a second language? Mosca (2015) conducted three experiments to answer this question. According to the findings, Swedish and English bilingual children demonstrated balanced responding in both languages. That is, if the bilingual children did or did not demonstrate BiN in one language, they also did or did not demonstrate BiN in the other language. However, for 30 monolingual English-speaking children who demonstrated BiN in English, when tested for BiN in Swedish, they demonstrated UiN only in Swedish. In the third experiment, 30 monolingual English-speaking adults were tested and demonstrated BiN in English as expected but only UiN in the second language (Swedish). What is missing that prevents the monolingual English-speaking children or adults from demonstrating transformation of stimulus function from listener to speaker in the new language?

The current study aimed to test if the mastery of speaking sounds that are idiosyncratic to a new language would result in the transformation of stimulus function from controlling the listener relation alone, UiN, to controlling the speaker as well, or BiN. Was the source of the problem not speaking with point-to-point speaker correspondence with what the children had heard during the naming experience? That is, once the point-to-point echoic is mastered, will emitting this spoken correspondence and simultaneously hearing the match result in BiN in the new language?

## **Experiment I: Identification of Echoic Repertoire and Idiosyncratic Sounds in Chinese**

### **Method**

**Selection of Participants** We recruited 32 monolingual English-speaking preschoolers with and without disabilities as participants. There were 10 females and 22 males, with ages ranging from 2 years to 4 years 10 months at the time of the study. The criterion for selection was the demonstration of an accurate echoic repertoire and BiN in English for familiar visual stimuli. Participants' existing repertoires were previously assessed by the classroom teachers using Greer's (2014) *Comprehensive Application of Behavior Analysis to Schooling International Curriculum and Inventory of Repertoires for Children From Preschool Through Kindergarten* and a list of 100 English words, which included single-syllable to multisyllable words of different consonant-vowel (CV) combinations. All the participants attended a publicly funded and privately run preschool for children with and without disabilities located outside of a metropolitan area. All instruction was based on principles of behavior analysis. None of the participants had been exposed to Chinese prior to the study.

**Materials and Setting** Experimenters used black pens and data sheets designed for this study to record responses using the International Phonetic Alphabet (IPA; International Phonetic Association, 1999) for English and Chinese. A MacBook was used to record all probe sessions.

The study took place in a tutoring room in the participants' preschool. There was a child-sized desk with chairs placed against the wall; the experimenter, who is a Chinese native speaker, and the participant sat side by side at the desk facing the wall. The tutoring room was free of distractions.

**Dependent Variable** The dependent variable was the echoic responses emitted by the participants. A correct echoic response was defined as emitting a point-to-point IPA correspondence with the antecedent provided by the experimenter within 3 s. Correct Chinese intonation was not required, even though the experimenter presented the echoic antecedents using appropriate intonation. An incorrect response was defined as a response that did not demonstrate point-to-point correspondence with the experimenter's spoken antecedent, including approximations (i.e., emitting a partially similar English sound for a Chinese phoneme), and no response.

## Data Collection

**Transcription of Phonemes** A total of 52 Chinese phonemes, excluding 4 back-nasal sounds (i.e., weng, wang, ying, yang), and 41 English phonemes, excluding two *r*-controlled vowels (i.e., /ə-/), were transcribed into IPA format to compare the phoneme pronunciations (International Phonetic Association, 2015). These sounds were excluded because they were present in both languages but were only considered as phonemes in one of the languages. The back-nasal sounds are considered phonemes in Chinese but not in English, even though there are back-nasal phonemic combinations in English. Similarly, the *r*-controlled vowels are considered phonemes in English but not in Chinese. As a result, even though the excluded phonemes sound almost identical in the two languages, they were represented by different IPA symbols. The experimenter divided the Chinese phonemes into two groups: distinctive Chinese sounds and Chinese-English sounds, which included Chinese sounds with approximations in English. Phonemes with the same IPA symbols were categorized as similar sounds, whereas phonemes with different IPA symbols were categorized as distinctive sounds.

**Group Assignments** Participants were randomly assigned to one of the two groups by drawing numbers from Research Randomizer (<https://randomizer.org>). Both groups received the same echoic probes for all the English phonemes. Group I also received probes with Chinese-English phonemes, whereas Group II received echoic probes with distinctive Chinese phonemes. Within each group, the participants were randomly assigned to counterbalance the order of presentation for the two phonemic conditions: echoic probe in Chinese first versus echoic probe in English first.

**Echoic Probe** During the probe session, the experimenter said the phoneme once, and the participant echoed or repeated the sound. A correct response was recorded as a plus (+), and an incorrect response was recorded as a minus (-). The experimenters counted

the numbers of correct responses and then converted the numbers into the percentage of correct responses in Chinese and English. Responses during probe sessions did not receive feedback; however, the participants received social reinforcement for attending to the experimenter and participating in the probe on a variable-ratio schedule of every five responses.

**Interobserver Agreement** An independent second observer who was bilingual in Chinese and English scored the correct and incorrect responses of the participants point by point by independently watching the video with audio recordings conducted during probe sessions. The intraclass correlation coefficient (ICC) was calculated using SPSS (Version 23) for 59% of all probe sessions for consistency. ICC for echoic probes of Chinese phonemes was .90 and ICC for echoic probes of English phonemes was .93.

**Design** Experiment I was an experimental- and control-group design with random assignment to phoneme conditions. The experiment tested the accuracy of echoics when the speech sounds of the stimuli were similar to both languages versus speech sounds distinctive to Chinese. Both groups received echoic probes in English phonemes also; therefore, the design was a 2 (English, Chinese)  $\times$  2 (Chinese-English phonemes, distinctive Chinese phonemes) design.

## Results

Descriptive data for all study variables are presented in Table 1. An independent samples *t*-test showed that there were no significant group differences in age or gender. Considering the results of the preliminary analysis, age was therefore excluded as a covariate in subsequent analyses.

An independent samples *t*-test showed that the two groups did not differ significantly in the numbers of correct echoic responses in English ( $M_1 = 39.56$ ,  $M_2 = 39.38$ ),  $t(30) = .24$ ,  $p = .82$ . The mean numbers of correct echoic responses in English are demonstrated in Fig. 1. An independent samples *t*-test was

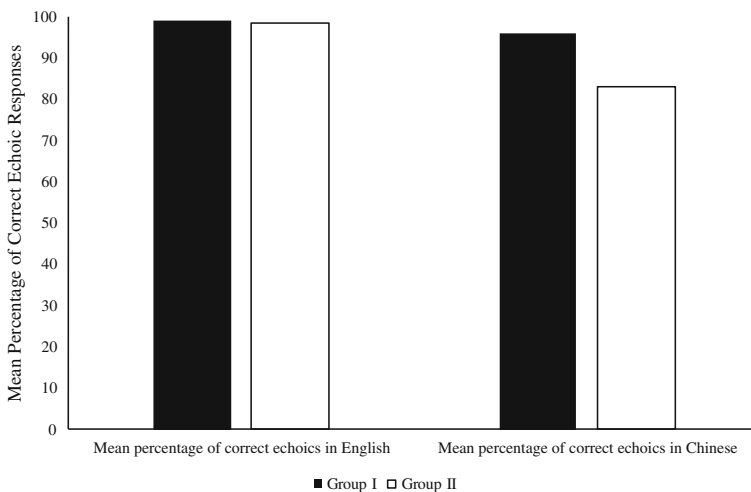
**Table 1** Descriptive data of mean age and mean percentage of correct echoic responses in chinese and english for Experiment I

Variable	Group I ( $n = 16$ )			Group II ( $n = 16$ )			Total sample ( $n = 32$ )		
	Mean	<i>SD</i>	Range	Mean	<i>SD</i>	Range	Mean	<i>SD</i>	Range
Mean percentage correct echoic responses in Chinese	96%	.05	87%–100%	83%	.09	68%–95%	89%	.10	68%–100%
Mean percentage correct echoic responses in English	96%	.03	90%–100%	96%	.07	73%–100%	96%	.05	73%–100%
Age	3.99	.60	2.83–4.83	3.94	.82	2.00–4.83	3.96	.71	2.00–4.83

conducted to determine the group differences of correct echoic responses in Chinese and revealed a significant group difference,  $t(30) = 17.19, p < .01$ . The mean percentage of correct echoic responses in Chinese for Group I ( $M_1 = 96\%$ ) was significantly higher than Group II ( $M_2 = 83\%$ ; Fig. 1). It showed that the participants had more difficulties producing correct echoic responses when the phonemes were unique to Chinese.

The results showed that the participants emitted significantly fewer correct echoic responses to distinctive Chinese phonemes. The results of Experiment I identified the speech sounds in Chinese that were difficult to echo, suggesting that these should be the target echoics to train in order to test if mastery of production of these speech sounds would result in the emission of the speaker component of BiN in Chinese. The results of Experiment I were also used for identification of the sounds that could be used within the naming trials without echoic training.

In Experiment II, we attempted to determine the relation between echoic training and the establishment of BiN in a second language. In the intervention, we taught the participants to produce the distinctive Chinese sounds in various CV combinations using a stringent IPA point-to-point echoic criterion. The targeted sounds were selected from distinctive Chinese sounds tested in Experiment I that had been shown to be difficult to echo by the participants. We also tested if children who demonstrated full BiN in English with familiar visual stimuli also demonstrated naming in Chinese with familiar and unfamiliar visual stimuli. Differences in novel but familiar stimuli and novel but unfamiliar stimuli have been found in several recent studies investigating variables affecting the establishment of BiN (Greer & Du, 2015b). Nedelcu, Fields, and Arntzen (2015) reported differences in effects for familiar and unfamiliar stimuli on the emission of derived relations. Thus, we tested for both types of stimulus control.



**Fig. 1** Mean percentage of correct echoes in English and Chinese for Group I and Group II. Group I received echoic probes in Chinese-English sounds, whereas Group II received echoic probes in distinctive Chinese phonemes

## Experiment II: Effects of Echoic Training on the Acquisition of BiN in a Second Language

### Method

In Experiment II only the setting was the same as Experiment I.

### Participants

Six participants were assigned to three dyads. See Table 2 for a description of the participants' ages, genders, and their diagnostic test scores. These participants were chosen because they demonstrated BiN in English with familiar visual stimuli, which was a prerequisite for Experiment II; however, they did not demonstrate BiN in Chinese with familiar visual stimuli.

### Materials

Twelve sets of two-dimensional stimuli (pictures) were used for naming probes; each set contained five different visual stimuli (see Table 3 for an example). All the stimuli were novel to the participants as determined by the preexperimental screening. There were four visual variations of the exemplars for each target stimulus, which required participants to respond to abstractions for the visual stimuli. The unfamiliar visual stimuli consisted of 20 Chinese characters (see Table 4). The familiar-appearing but novel visual stimuli consisted of pictures of objects that the participants were likely to encounter in their environment. See Table 4 for sample stimuli sets used during naming-probe trials. Phonemes and sounds that are necessary for naming the probe stimuli but not trained during the intervention were the ones that the participants can produce independently.

**Table 2** Description of participants

Participant	Age	Gender	Language at home	Developmental status	PLS-5	IQ
1	3.08	F	English	Typical	–	–
2	3.42	F	English	Typical	–	–
3	3.83	F	English	Preschooler with a disability	Auditory Comprehension 101 Expressive Communication 92	Full Scale 98
4	3.83	F	English	Typical	–	–
6	3.00	M	English	Preschooler with a disability	–	–
7	4.67	M	English	Preschooler with a disability	–	–

PLS-5 Preschool Language Scale–5 (Zimmerman, Steiner, & Pond, 2011)



**Table 3** Sample Chinese familiar stimuli

Visual Stimuli	Text in Pinyin	Text in IPA
Picture of three pomegranates	cui	ts <sup>h</sup> ɔɪ
Picture of the head of a red panda	xiong	ɛɪɔŋ
Picture of the cartoon character Mulan	jiang	tɛɪaŋ
Picture of a toilet bowl brush	qiu	tɛ <sup>h</sup> iŋ
Picture of a staple remover	jia	tɛɪa

### Dependent Variables

The dependent variables were the listener (i.e., the untaught point-to or selection responses) and speaker responses (i.e., the untaught tact responses) during pre- and post-echoic-training probe sessions, as well as the numbers of echoics emitted during probe sessions. A correct listener response was defined as the participant pointing to the correct stimulus when it was presented in an array of three stimuli (i.e., one correct stimulus and two incorrect stimuli) after being given a vocal antecedent (i.e., “Point to [stimulus].”). An incorrect listener response was defined as the participant pointing to the nontarget stimulus or failing to point to any stimulus. A correct tact response was defined as the participant vocally emitting the name of the stimulus when presented with a visual stimulus and the experimenter’s vocal antecedent, “What is this?” An incorrect tact response was defined as the participant emitting an incorrect name to the target stimulus or failing to name the target stimulus. A correct echoic response was defined as the participant emitting a sound with point-to-point correspondence to the antecedent provided by the experimenter. An incorrect echoic response included not emitting any sound or emitting a sound without point-to-point correspondence to the sound provided in the antecedent. Responses to the probe trials did not receive any feedback. The participants received social reinforcement for attending and participating during probe sessions on a variable-ratio schedule of five trials. The criterion for the acquisition of naming was set at 80% accuracy or above (i.e., 16 correct responses out of 20 possible) to the probe trials across both listener and speaker responses.

### Independent Variable

The independent variable was the mastery of producing the distinctive Chinese sounds of the training sets; that is, for the independent variable to be in effect, the participants had to demonstrate mastery of the CV speech sounds. This required the participants to accurately produce the speech sounds in Chinese without any echoic models for all the

**Table 4** Chinese unfamiliar visual stimuli

Visual Stimuli	Text in Pinyin	Text in IPA
么	iao	ɪaʊ
西	xi	ɕ
约	iue	öʊ
日	ri	ɹ̥
七	qi	tɕʰ
吁	yu	ö
鸡	ji	tɕ
兹	ci	tʂʰ
哇	wa	ʊa
佣	yong	ɪʊŋ
去	qu	tɕʰu
姐	jie	tɕjɛ
人	ren	ɹ̥en
心	xin	ɕin
华	hua	hua
从	cong	tʂʰʊŋ
蕊	ruì	ɹ̥ui
囧	jiōng	tɕiŋ
小	xiao	ɕiao
巡	xun	ɕön

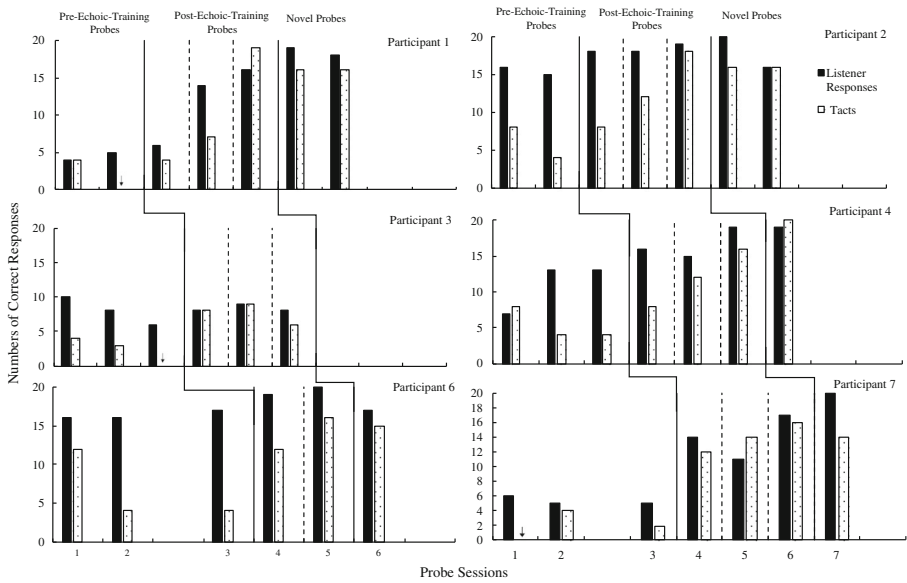
sounds in the training sets in three consecutive sessions. The experimenter trained the participants to echo various CV combinations of Chinese phonemes (see Table 5 for sounds used for echoic training). There were five target CV combinations in each training set, and the targeted distinctive phonemes in each CV combination were identical to those used during naming probes in Chinese; however, the CV combinations were different. For instance, /tsi/ was one of the phoneme combinations used during the naming probe in Chinese phonemes, whereas /tsai/ was one of the phoneme combinations used during echoic training. Both combinations had the common phoneme /ts/, but the CV combinations were different because the vowel was different. The training sets consisted of all the phonemes that were targeted in the probe sets. The intervention consisted of three components: echoic teaching, within-instructional-trials probes, and post-instructional-trials probes. Within-instructional-trials probes and post-instructional-trials probes were used to ensure the acquisition of the independent variable, which was the mastery of producing target Chinese sounds. The intervention is represented by dotted lines in Figs. 2 and 3 after the first post-echoic-training probe. Upon the mastery of each training set, a post-echoic-training probe was conducted.

**Table 5** Echoic training stimuli

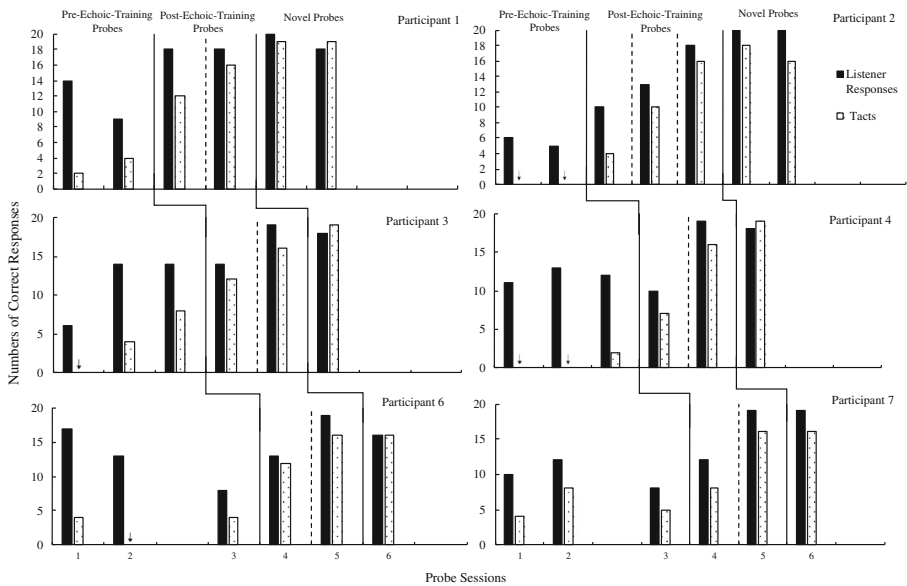
Chinese CV Combinations	Text in IPA
cai	ts <sup>h</sup> aɪ
ju	tɕö
xuan	ɕöɛn
qiong	tɕ <sup>h</sup> joŋ
ru	ɹu
cuo	ts <sup>h</sup> uo
jiao	tɕjao
xu	ɕ <sup>h</sup> u
lüe	löœ
qia	tɕ <sup>h</sup> ja
ca	ts <sup>h</sup> a:
jue	tɕöœ
xia	ɕja
rong	ɹoŋ
qun	tɕ <sup>h</sup> öN

**Echoic Teaching** The echoic training began with the echoic teaching phase, during which the experimenter taught the participant to echo each target CV combination in the training set, one at a time. The experimenter gained the participant's attention, enunciated the sound once, and waited 5 s for the participant's response. Participants' correct echoic responses were reinforced by vocal praise and tokens, whereas their incorrect responses or omission of responding resulted in the correction procedure, during which the experimenter re-presented the sound and again had the participant echo the sound until the participant emitted a correct response. After three attempts in a correction procedure, a new sound was presented.

**Within-Instructional-Trials Probe** When the participant had echoed each target CV combination once with point-to-point correspondence, the within-instructional-trials probe began. The experimenter asked the participant to say the five CV combinations independently with no model (i.e., "Now tell me what sounds you remember."). During within-instructional-trials probes, the participants were required to emit sounds that were taught during the echoic teaching. Their correct responses were reinforced with vocal praise and tokens, whereas their incorrect responses did not result in any



**Fig. 2** Numbers of correct listener and tact responses during naming probes in Chinese with unfamiliar visual stimuli for Participants 1, 2, 3, 4, 6, and 7. Dashed lines represent training occurred, whereas solid lines separate pre-echoic-training probes, post-echoic-training probes, and post-echoic-training probes with novel stimuli sets. Arrows represent 0



**Fig. 3** Numbers of correct listener and tact responses during naming probes in Chinese with familiar stimuli for Participants 1, 2, 3, 4, 6, and 7. A dashed line represents that more training occurred, whereas solid lines separate pre-echoic-training probes, post-echoic-training probes, and post-echoic-training probes with novel stimuli sets. Arrows represent 0

correction. The criterion was set as 100% accuracy across three consecutive sessions to receive the post-instructional-trials delayed echoic probes. The experimenter repeated the echoic teaching and within-instructional-trials probes until the participant met the predetermined criterion. Each participant received four to eight sessions per day, and each session lasted approximately 5 min. Sessions consisted of blocks of 20 instructional trials.

**Post-Instructional-Trials Delayed Echoic Probe** Once the participants met criterion for within-instructional-trials probes, the delayed-probe phase began after at least a 2-h time lapse. During the delayed-probe phase, the experimenter did not provide any echoic models but simply told the participant to recall as many of the sounds as possible (i.e., “Tell me the sounds you remember.”). The participants were required to emit the accurate speech sounds for the five CV combinations in the training set with 100% accuracy across three consecutive sessions to master the set. The post-instructional-trials delayed echoic probes always followed the mastery of the within-instructional-trials probes. If the participant failed to meet the criterion for the delayed echoic probe, the experimenter repeated the entire procedure beginning with echoic teaching and within-instructional-trials probes until the participant met criterion on the delayed echoic probe.

## Data Collection

Data were collected on the numbers of correct responses on all dependent measures during pre- and post-echoic-training naming probes. Data were also collected on the numbers of correct echoic responses during echoic teaching sessions, as well as within-instructional-trials and post-instructional-trials probes. The experimenter used data sheets designed for the study and a black pen to collect data. A plus (+) was recorded if the participant emitted a correct response, and a minus (–) was recorded if the participant emitted an incorrect response.

## Interobserver Agreement

Interobserver agreement (IOA) was collected for both the naming experience (described in the [Procedure](#) section) and the probes and calculated trial by trial by two independent observers who were bilingual in English and Chinese. The second observer was naive to the purpose of the experiment and collected data via viewing the recorded video with audio recording of the probe sessions for BiN. IOA was calculated by counting the numbers of point-by-point agreements and disagreements and dividing the number of agreements by the total trials and then multiplying by 100%. IOA was collected for 48% of the sessions of the naming probes in Chinese phonemes with unfamiliar visual stimuli with 97% mean agreement, ranging from 87% to 100%. IOA was obtained for 55% of the sessions for naming probes in Chinese with familiar and novel visual stimuli with a mean agreement of 97%, ranging from 87% to 100%. IOA was obtained for 34% of the echoic training sessions with a mean agreement of 98%, ranging from 91% to 100%.

## Design

A multiple-probe design across three groups of two participants was used to determine the effects of the echoic training procedure on the acquisition of naming in Chinese with familiar and nonfamiliar visual stimuli. That is, the six participants were divided into three groups of two, and multiple-probe designs were implemented across the three dyads.

## Procedure

### Pre-Echoic-Training Naming Probe

There were two phases of the pre-echoic-training naming probe: naming experience and naming probe.

**Naming Experience** The participants received a total of 40 match-to-sample (MTS) trials, with four opportunities for each novel stimulus. During the MTS instruction, the experimenter presented a field of three stimuli, including one target stimulus and two nonexemplars. The experimenter then provided a vocal antecedent, “This is [stimulus]. Match [stimulus] with [stimulus].” The experimenter waited 3 s for the participant to emit a response, which included the participant putting the stimulus on top of the correct stimulus in an array of three stimuli. A correct matching response was reinforced by verbal praise, whereas an incorrect response or no response resulted in a correction procedure. The correction procedure consisted of the experimenter representing the antecedent, “This is [stimulus]. Match [stimulus] with [stimulus],” and pointing to the correct stimulus; then the experimenter repeated the antecedent and gave the participant an opportunity to emit the correct response independently. Verbal praise was not provided for the correction. In all of the BiN probes before the echoic intervention, the naming experience included reinforced MTS as participants heard the word.

The criterion for completing the naming experience was 90% or better correct responses (i.e., 18 correct responses out of 20 possible) across two consecutive sessions. All participants were able to complete the MTS instruction within two sessions of 20-trial blocks. During the MTS instruction, the echoic responses the participants emitted voluntarily were also recorded, but no reinforcement or correction was delivered for the echoic response, although an accurate MTS response was reinforced to ensure the participants were attending to the stimuli for the spoken sounds. The participants were not instructed to echo the stimuli presented by the experimenter.

**Naming Probe** Tests for the presence of BiN or UiN in English and Chinese with nonfamiliar and familiar visual stimuli were conducted 2 h after the naming experience. Each probe session consisted of 40 massed trials, including 20 consecutive listener responses (i.e., “Point to the [stimulus].”), followed by 20 consecutive tact responses (i.e., “What is this?”). Each target stimulus was presented four times under each of the speaker and listener conditions.

## Echoic Training

After having two to three pre-echoic-training probes (Figs. 2 and 3), the participants began with echoic training, which continued until mastery of Chinese speech sounds was demonstrated. See the section on the independent variable for a detailed description of echoic teaching, within-instructional-trials probes, and post-instructional-trials delayed echoic probes.

## Post-Echoic-Training Probes

Post-echoic-training probes were conducted upon the participant's mastery of each training set. They were conducted in the same manner as the pre-echoic-training probes. However, an MTS naming experience was only provided when a novel set was introduced. The participants were required to demonstrate BiN with the probe set used during the second pre-echoic-training probe to move on to the post-echoic-training probe with the novel set.

## Results

The data showed that the echoic training was effective in establishing BiN in Chinese with familiar visual stimuli for six out of six participants and BiN in Chinese with unfamiliar visual stimuli for five out of six participants. Mastery of the echoic responses required extensive training.

**BiN in Chinese with Unfamiliar Visual Stimuli** As shown in Fig. 2, none of the participants demonstrated full BiN in Chinese with unfamiliar visual stimuli during the pre-echoic-training probes. Participants 2 and 6 demonstrated UiN during pre-echoic-training probes, and the remaining participants did not.

Five out of six demonstrated BiN in Chinese with unfamiliar visual stimuli according to our criterion for the establishment of BiN. All Participants showed an ascending trend for correct responses in both listener and speaker responses upon the mastery of each training set. All but Participant 6 and Participant 3 demonstrated BiN in Chinese with unfamiliar visual stimuli after mastering three training sets. Participant 6 acquired BiN in Chinese with unfamiliar visual stimuli after completing two training sets. Participant 3 showed a slight increase in her correct listener and speaker responses after three sets of trainings, but her overall correct responding remained under 50%.

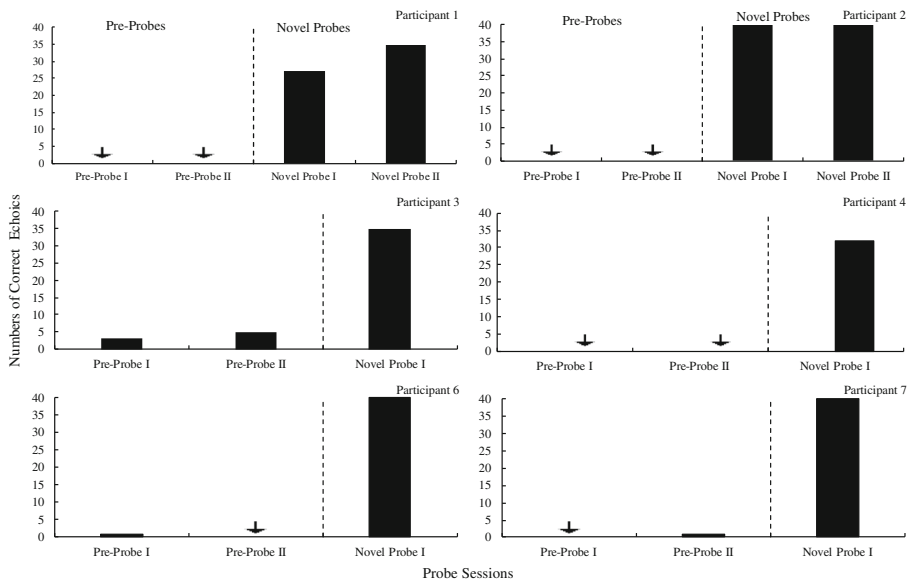
**BiN in Chinese with Familiar-Appearing Novel Stimuli** During pre-echoic-training probes, none of the participants demonstrated BiN in Chinese with familiar-appearing visual stimuli (Fig. 3). All participants who received echoic training acquired BiN in Chinese with familiar visual stimuli as a function of the echoic training. Their correct responding increased after completing each training set. All except Participant 2 acquired BiN in Chinese with familiar visual stimuli after mastering two training sets. BiN emerged for Participant 2 after three training sets. All participants who received echoic training consistently demonstrated BiN in Chinese with familiar visual stimuli during novel probes.

**Echoics Emitted during Naming Experiences** Each participant received three to four naming experiences during naming probes in Chinese with nonfamiliar and familiar-appearing visual stimuli. The experimenter compared the numbers of echoics emitted during these sessions prior to and following the echoic training for the participants. As shown in Figs. 4 and 5, the numbers of echoics emitted during the naming experience in Chinese increased for all participants.

**Echoic Training** All but Participant 6 required three training sets, whereas Participant 6 only needed two training sets for BiN to emerge. Each participant required a different number of sessions to master the training set, ranging from 13 to 44 sessions. See Table 6 for descriptive data of each participant's responding during echoic training sessions.

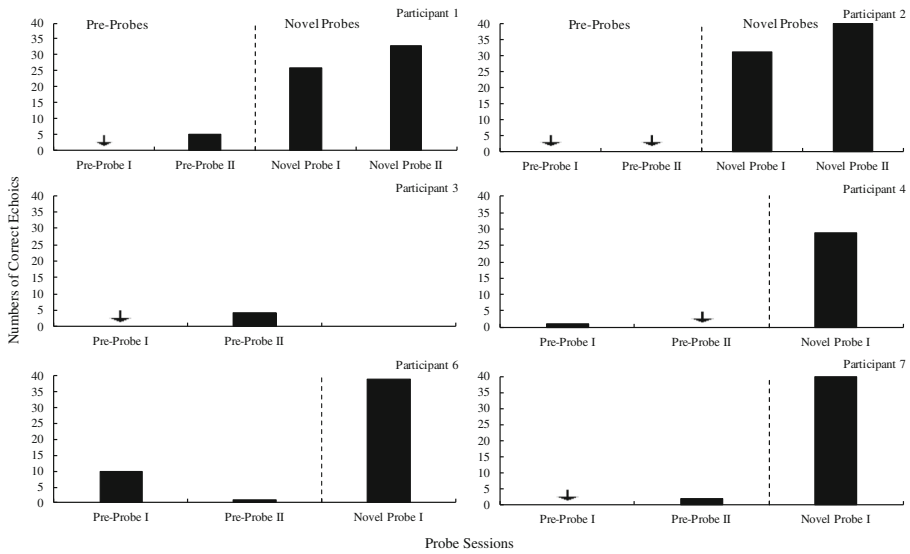
## General Discussion

The results from Experiment I suggested that monolingual English-speaking children emitted more accurate echoics to Chinese-English phonemes compared to echoics to distinctive Chinese phonemes. This finding supports the argument on the role of the first language during second-language acquisition. The similarities in speech sounds and the production of those speech sounds particularly between the two languages often facilitate the acquisition of the second language, whereas the differences between the two languages are more likely to hinder the learning of the second language (Birdsong & Molis, 2001; Hansen Edwards & Zampini, 2008). However, the results from Experiment II demonstrated that with intensive echoic training, monolingual English-



**Fig. 4** Numbers of echoics emitted during probe sessions in Chinese with familiar stimuli for Participants 1, 2, 3, 4, 6, and 7. The dashed lines separate pre-echoic-training probes and post-echoic-training probes with novel stimuli sets. Arrows represent 0





**Fig. 5** Numbers of echoes emitted during the probe sessions in Chinese with unfamiliar visual stimuli for Participants 1, 2, 3, 4, 6, and 7. The dashed lines separate pre-echoic-training probes and post-echoic-training probes with novel stimuli sets. Arrows represent 0

speaking children mastered producing distinctive Chinese phonemes independently, which was functionally related to the establishment of BiN in Chinese. The results reported herein support Kan et al.’s (2014) findings that speech practice, or echoic training, facilitated the initial forming of word-object relations. Moreover, they suggest that the accurate production of the small units of speech sounds is functionally related to the emission of the speaker component of BiN. It is possible that prior to the intervention, the participants were simply not able to produce the Chinese sounds, thus preventing them from emitting correct speaker responses during the preintervention naming probes. The echoic training may have resulted in the mastery of producing distinctive Chinese sounds that then allowed the participants to emit correct speaker responses in Chinese during the postintervention probes.

Horne and Lowe (1996) explained the echoic responses as follows:

The sounds and words uttered by parents may function as potent classically conditioned stimuli that have strong emotional effects on the child so that when she hears her replication of these vocal patterns she generates stimuli that have similar strong reinforcing consequences. (p. 198)

**Table 6** Descriptive data of the numbers of sessions each participant needed to master a training set

	Participant 1	Participant 2	Participant 3	Participant 4	Participant 6	Participant 7
Set I	35	23	26	35	21	12
Set II	29	23	35	23	36	22
Set III	17	37	42	9	–	19
Total	81	83	103	67	57	53

In other words, when emitting echoic responses, the children may have been reinforced by the point-to-point correspondence of hearing and saying. The echoic training procedure could have resulted in the children emitting point-to-point echoics in the new language, and hearing that correspondence functioned to reinforce the speaker behavior as it appears to do in primary language BiN (Lo, 2016; Longano & Greer, 2014; Lowenkron, 2006). The conditioned reinforcement of Chinese auditory stimuli could, in turn, then be paired with unfamiliar visual stimuli during naming probes and thus lead to the establishment of conditioned reinforcement for observing unfamiliar visual stimuli, as reported by Du, Broto, and Greer (2015) and in Greer and Han (2015).

Participant 3 failed to demonstrate BiN in Chinese with nonfamiliar visual stimuli upon the completion of echoic training. It is possible that she did not acquire conditioned reinforcement for observing nonfamiliar visual stimuli. Because echoic responding only joins listener and speaker responses when visual and auditory stimuli are jointly conditioned, the lack of conditioned reinforcement for observing nonfamiliar visual stimuli may have prevented her from acquiring BiN in Chinese with nonfamiliar visual stimuli.

Findings reported herein added to our understanding of the emergence of naming in a second language. Consistent with Mosca's (2015) findings, results from Experiment II confirmed that naming is language specific for monolingual English-speaking children. These data suggest that the source of the difficulty in the second language is the emission of the correspondence in speech sounds. That is, when participants were provided with intensive language experience and exposures to the visual and auditory stimuli, monolingual English-speaking children produced the speech sounds that did not exist in their first language and, as a result, demonstrated naming in a second language. Thus, it appears that the speech sound differences contribute to the effect of the language differences. It is important to note that the listener component or UiN preceded BiN and is a necessary prerequisite to the emission of the speaker component.

There are several limitations in this study. First, the incorrect echoic responses to distinctive Chinese phonemes were not further analyzed. Second, Participants 1 and 2 did not have a third pre-echoic-training probe with the probe set used during the second pre-echoic-training probes. To remedy this limitation, the two participants received post-echoic-training probes with two novel sets. Third, there was a lack of procedural fidelity data to be reported.

The naming experience is a test of the presence or absence of the necessary stimulus control that constitutes BiN. Reinforcement for the visual MTS would be a part of the independent variable if the participants do demonstrate BiN in the probes that are done as part of the test for BiN in the naming experience. In Experiment 2, during the naming probes, the participants increased their voluntary emission of echoics. The increase corresponded with increased MTS responses. Because we reinforced the participants for MTS responses with social reinforcers to ensure the children were looking at the visual stimuli, it is possible that the same reinforcer may have inadvertently functioned to reinforce the emission of echoics. As a result, the reinforcement of the MTS response during the naming experience may have contributed to the effects. If so, it was an unwitting but crucial ingredient. However, Longano and Greer (2014) did not find any difference in the emergence of BiN under naming experiences with and without reinforcement for visual MTS responses. Orlans (2017), in a group comparison

study, also did not find any differences in the two procedures. Future research should address whether or not this was an active ingredient. The effects of the procedure reported herein should also be tested with monolingual adults or with a different language. Future research may also include a detailed analysis of the error patterns of the echoic responses in a foreign language and investigate the effects of the echoic errors on the acquisition of BiN in the novel language.

Despite the aforementioned limitations, the results of the two experiments reported herein suggest that the reliable demonstration of BiN in a second language required the mastery of accurate echoic responding to the speech sounds that are idiosyncratic to the second language. The accurate production of the sounds may have resulted in reinforcement for correspondence between the children's own speech sounds and what is produced by the children when their speech is accurate (Greer & Du, 2015b; Greer & Longano, 2010). The repeated experiences with the unfamiliar stimuli allowed the existing reinforcers for observing responses to be paired with the unfamiliar visual stimuli, resulting in transfer of reinforcement for observing to the unfamiliar stimuli, as found in prior studies (Du et al., 2015; Greer & Han, 2015).

These data contribute to a growing number of studies on how conditioning histories make it possible for children to contact stimuli they did not contact before those conditioning histories. This study, and others on the development of new stimulus control that makes new cusps possible, points to the importance of stimulus control for conditioned reinforcement for the embedded stimuli that select out observing responses.

## Compliance with Ethical Standards

**Conflict of Interest** The authors declare that there is no conflict of interest.

**Ethical Approval** All procedures performed in this study were reviewed by the Institutional Review Board of Columbia University Teachers College and the Fred S. Keller School and were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Informed Consent** Informed consent was obtained from all individual participants included in this study.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, 44(2), 235–249. <https://doi.org/10.1006/jmla.2000.2750>
- Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, 15, 17–29.
- Carnerero, J. J., & Pérez-González, L. A. (2014). Induction of naming after observing visual stimuli and their names in children with autism. *Research in Developmental Disabilities*, 35(10), 2514–2526. <https://doi.org/10.1016/j.ridd.2014.06.004>
- Choi, J., Greer, R. D., & Keohane, D. (2015). Effects of auditory matching on the intercept of speaker and listener repertoires. *Behavioral Development Bulletin*, 20(2), 186–206. <https://doi.org/10.1037/h010131>

- Du, L., Broto, J., & Greer, R. D. (2015). The effects of establishment of conditioned reinforcement for observing responses for 3D stimuli on generalized visual match-to-sample in children with autism spectrum disorders. *European Journal of Behavior Analysis*, 16(1), 82–98. <https://doi.org/10.1080/15021149.2015.1065655>
- Fiorile, C. A., & Greer, R. D. (2007). The induction of naming in children with no prior tact responses as a function of multiple exemplar histories of instruction. *The Analysis of Verbal Behavior*, 23(1), 71–87. <https://doi.org/10.1007/BF03393048>
- Gilic, L., & Greer, R. D. (2011). Establishing naming in typically developing two-year-old children as a function of multiple exemplar speaker and listener responses. *The Analysis of Verbal Behavior*, 27(1), 157–177. <https://doi.org/10.1007/BF03393099>
- Greer, R. D. (2014). *CABAS® international curriculum and inventory of repertoires for children from preschool through kindergarten*. Yonkers: Fred. S. Keller School.
- Greer, R. D., & Du, L. (2010). Generic instruction versus intensive tact instruction and the emission of spontaneous speech. *The Journal of Speech and Language Pathology – Applied Behavior Analysis*, 5(1), 1–19. <https://doi.org/10.1037/h0100261>
- Greer, R. D., & Du, L. (2015a). Experience and the onset of the capability to learn names incidentally by exclusion. *The Psychological Record*, 65(2), 355–373. <https://doi.org/10.1007/s40732-014-0111-2>
- Greer, R. D., & Du, L. (2015b). Identification and establishment of reinforcers that make the development of complex social language possible. *International Journal of Behavior Analysis & Autism Spectrum Disorder*, 1(1), 13–34.
- Greer, R. D., & Han, H. S. A. (2015). Establishment of conditioned reinforcement for visual observing and the emergence of generalized visual-identity matching. *Behavioral Development Bulletin*, 20(2), 227–252. <https://doi.org/10.1037/h0101316>
- Greer, R. D., & Longano, J. (2010). A rose by naming: How we may learn how to do it. *The Analysis of Verbal Behavior*, 26(1), 73–106. <https://doi.org/10.1007/BF03393085>
- Greer, R. D., & O’Sullivan, D. (2007). *The incidence of naming for novel two-dimensional stimuli in first graders at the beginning, mid-year, and end of the year*. Paper presented at the annual meeting of the Association for Behavior Analysis International, San Diego.
- Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: Inducing and expanding new verbal capabilities in children with language delays*. Boston: Allyn & Bacon.
- Greer, R. D., & Speckman, J. (2009). The integration of speaker and listener responses: A theory of verbal development. *The Psychological Record*, 59(3), 449–488. <https://doi.org/10.1007/BF03395674>
- Greer, R. D., Stolfi, L., Chavez-Brown, M., & Rivera-Valdes, C. (2005). The emergence of the listener to speaker component of naming in children as a function of multiple exemplar instruction. *The Analysis of Verbal Behavior*, 21(1), 123–134. <https://doi.org/10.1007/BF03393014>
- Greer, R. D., Stolfi, L., & Pistoljevic, N. (2007). Emergence of naming in preschoolers: A comparison of multiple and single exemplar instruction. *European Journal of Behavior Analysis*, 8(2), 109–131. <https://doi.org/10.1080/15021149.2007.11434278>
- Hansen Edwards, J. G., & Zampini, M. L. (2008). *Phonology and second language acquisition*. Amsterdam, Netherlands: John Benjamins.
- Helou-Care, Y. J. (2008). *The effects of the acquisition of naming on reading comprehension with academically delayed middle school students diagnosed with behavior disorders* (Doctoral dissertation). Available from UMI ProQuest Digital Dissertations database. (Publication No. 3317559).
- Home, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, 65(1), 185–241. <https://doi.org/10.1901/jeab.1996.65-185>
- International Phonetic Association. (1999). *Handbook of the international phonetic association: A guide to the use of the international phonetic alphabet*. New York: Cambridge University Press.
- International Phonetic Association. (2015). *International Phonetic Alphabet chart*. Retrieved from [https://www.internationalphoneticassociation.org/sites/default/files/IPA\\_Kiel\\_2015.pdf](https://www.internationalphoneticassociation.org/sites/default/files/IPA_Kiel_2015.pdf)
- Iverson, P., & Evans, B. G. (2009). Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *The Journal of the Acoustical Society of America*, 126(2), 866–877. <https://doi.org/10.1121/1.3148196>
- Jeong, H., Sugiugra, M., Sassa, Y., Yokoyama, S., Horie, K., Sato, S., et al. (2007). Cross-linguistic influence on brain activation during second language processing: An fMRI study. *Bilingualism: Language and Cognition*, 10(2), 175–187. <https://doi.org/10.1017/S1366728907002921>
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60–99. [https://doi.org/10.1016/0010-0285\(89\)90003-0](https://doi.org/10.1016/0010-0285(89)90003-0)

- Kan, P. F., & Kohnert, K. (2008). Fast mapping by bilingual preschool children. *Journal of Child Language*, 35(3), 495–514. <https://doi.org/10.1017/S0305000907008604>
- Kan, P. F., Sadagopan, N., Janich, L., & Andrade, M. (2014). Effects of speech practice on fast mapping in monolingual and bilingual speakers. *Journal of Speech, Language and Hearing Research*, 57(3), 929–941. [https://doi.org/10.1044/2013\\_JSLHR-L-13-0045](https://doi.org/10.1044/2013_JSLHR-L-13-0045)
- Kisilevsky, B. S., Hains, S. M., Lee, K., Xie, X., Huang, H., Ye, H., ... Wang, Z. (2003). Effects of experience on fetal voice recognition. *Psychological Science*, 14(3), 220–224. <https://doi.org/10.1111/1467-9280.02435>
- Krueger, C., & Garvan, C. (2014). Emergence and retention of learning in early fetal development. *Infant Behavior & Development*, 37(2), 162–173. <https://doi.org/10.1016/j.infbeh.2013.12.007>
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5(11), 831–843. <https://doi.org/10.1038/nrn1533>
- Kuhl, P. K., Tsao, F., & Liu, H. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, 100(15), 9096–9101. <https://doi.org/10.1073/pnas.1532872100>
- Lenneberg, E. H., Chomsky, N., & Marx, O. (1967). The Biological Foundations of Language. In *Biological foundations of language*. New York: Wiley.
- Lo, C. (2016). *How the presence of the listener half of naming leads to multiple stimulus control* (Doctoral dissertation). Available from UMI ProQuest Digital Dissertations database. (Publication No. 10100022).
- Longano, J. M., & Greer, R. D. (2014). Is the source of reinforcement for naming multiple conditioned reinforcers for observing responses? *The Analysis of Verbal Behavior*, 31(1), 96–117. <https://doi.org/10.1007/s40616-014-0022-y>
- Lowenkron, B. (2006). Joint control and the selection of stimuli from their description. *The Analysis of Verbal Behavior*, 22(1), 129–151. <https://doi.org/10.1007/BF03393035>
- Miguel, C. (2016). Common and intraverbal bidirectional naming. *The Analysis of Verbal Behavior*, 32(2), 125–138. <https://doi.org/10.1007/s40616-016-0066-2>
- Moon, C., Lagercrantz, H., & Kuhl, P. K. (2013). Language experienced in utero affects vowel perception after birth: A two-country study. *Acta Paediatrica*, 102(2), 156–160. <https://doi.org/10.1111/apa.12098>
- Mosca, K. F. (2015). *Incidental learning of two languages by bilingual Swedish- and English-speaking children, monolingual English-speaking children and monolingual English-speaking adults* (unpublished doctoral dissertation). Department of Health and Behavior Studies, Teachers College, New York: Columbia University.
- Nedelcu, R. I., Fields, L., & Arntzen, E. (2015). Arbitrary conditional discriminative functions of meaningful stimuli and enhanced equivalence class formation. *Journal of the Experimental Analysis of Behavior*, 103(2), 349–360. <https://doi.org/10.1002/jeab.141>
- Nishi, K., & Kewley-Port, D. (2007). Training Japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech, Language, and Hearing Research*, 50(6), 1496–1509. [https://doi.org/10.1044/1092-4388\(2007\)103](https://doi.org/10.1044/1092-4388(2007)103)
- Orlans, S. E. (2017). *The effects of naming experiences and properties of visual stimuli on language acquisition and the relationship between curiosity and naming* (Doctoral dissertation). Available from ProQuest Dissertation Publishing database. (Publication No. 10269736).
- Pistoljevic, N., & Greer, R. D. (2006). The effects of daily intensive tact instruction on preschool students' emission of pure tacts and mands in non-instructional setting. *Journal of Early and Intensive Behavior Intervention*, 3(1), 103–120. <https://doi.org/10.1037/h0100325>
- Skinner, B. F. (1957). *Verbal behavior*. Acton, MA: Copley Publishing Group.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, 7(1), 49–63. [https://doi.org/10.1016/S0163-6383\(84\)80022-3](https://doi.org/10.1016/S0163-6383(84)80022-3)
- Woolslayer, L. (2013). *The functional relation between the onset of naming and the joining of listener to untaught speaker responses* (Doctoral dissertation). Available from ProQuest Dissertation Publishing database. (Publication No. 3560915).
- Zhang, Y., & Wang, Y. (2007). Neural plasticity in speech acquisition and learning. *Bilingualism: Language and Cognition*, 10(2), 147–160. <https://doi.org/10.1017/S1366728907002908>
- Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool language scales* (5th ed.). San Antonio: Pearson.