Object labeling in”uences infant phonetic learning and generalization

H. Henny Yeung , Thierry Nazzi

CNRS (Laboratoire Psychologie de la Perception, UMR 8242) & Université Paris Descartes, Paris Sorbonne Cité, France

a r t i c l e i n f o

Article history:
Received 6 August 2013
Revised 31 March 2014
Accepted 1 April 2014

Keywords:
Infant
Speech perception
Word learning
Prosody
Lexical stress

a b s t r a c t

Different kinds of speech sounds are used to signify possible word forms in every language. For example, lexical stress is used in Spanish (/•be.be/, • he/she drinks• versus /be.•be/, •baby•), but not in French (/•be.be/ and /be.•be/ both mean • baby•). Infants learn many such native language phonetic contrasts in their “rst year of life, likely using a number of cues from parental speech input. One such cue could be parents• object labeling, which can explicitly highlight relevant contrasts. Here we ask whether phonetic learning from object labeling is abstract„that is, if learning can generalize to new phonetic contexts. We investigate this issue in the prosodic domain, as the abstraction of prosodic cues (like lexical stress) has been shown to be particularly dif“cult. One group of 10-month-old French-learners was given consistent word labels that contrasted on lexical stress (e.g., Object A was labeled /•ma.bu/, and Object B was labeled /ma.•bu/). Another group of 10-month-olds was given inconsistent word labels (i.e., mixed pairings), and stress discrimination in both groups was measured in a test phase with words made up of new syllables. Infants trained with consistently contrastive labels showed an earlier effect of discrimination compared to infants trained with inconsistent labels. Results indicate that phonetic learning from object labeling can indeed generalize, and suggest one way infants may learn the sound properties of their native language(s).

1. Introduction

Experience with one•s native language(s) alters speech perception from early in infancy (Jusczyk, 2000). This developmental phenomenon is well exempli“ed by phonetic attunement, which describes a decline in infants• perception of certain non-native phonetic contrasts and improvement in the perception of many other phonetic contrasts„particularly native ones„a process that begins already in the “rst year of life (e.g., Best, McRoberts, & Goodell, 2001; Eilers, Wilson, & Moore, 1977; Narayan, Werker, & Beddor, 2010; Rivera-Gaxiola, Klarman, Garcia-Sierra, & Kuhl, 2005; Werker & Tees, 1984; Yeung, Chen, & Werker, 2013). Many of the phonetic patterns to which infants attune are determined by the lexical structure of the native language, but it is important to note that attunement begins before infants have acquired lexicons of any substantial size. Recent work has thus focused on learning mechanisms that could drive phonetic attune-of phonetic tokens in acoustic space (Salminen, Tiitinen, & May, 2009; Toscano & Mcmurray, 2010; Vallabha, McClelland, Pons, Werker, & Amano, 2007; Werker et al., 2007). A number of studies have shown that infants• perceptual sensitivities in the speech domain are indeed affected by the distributional characteristics of speech input (Cristia, 2011; Cristia, McGuire, Seidl, & Francis, 2011; Maye, Weiss, & Aslin, 2008; Maye, Werker, & Gerken, 2002; Yoshida, Pons, Maye, & Werker, 2010). Nevertheless, recent work has challenged the notion that distributional learning alone can explain phonetic attunement (see Werker, Yeung, & Yoshida, 2012 for review). For example, it is unclear whether distributional information can support the learning of acoustically dif“cult phonetic contrasts (Cristia et al., 2011; Narayan et al., 2010; Sato, Kato, & Mazuka, 2012), or explain learning when phonetic distributions show substantial overlap (Feldman, Myers, White, Grif“ths, & Morgan, 2013; Swingley, 2009). Indeed, other computational models have suggested that additional learning mechanisms beyond simple token-counting„such as an unsupervised competitive learning process that selects among possible statistical distributions„are needed to successfully learn some phonetic distinctions (McMurray, Aslin, & Toscano, 2009; Toscano & Mcmurray, 2010).

lexical stress when tested on tasks using single words, just like Spanish- or English-learning infants of the same age. However, these French-learning infants (unlike their Spanish- or English-learning peers at this age) fail in tasks that require the abstraction of stress patterns over multiple words (Bijeljac-Babic et al., 2012; Skoruppa, Cristia, Peperkamp, & Seidl, 2011; Skoruppa et al., 2009, 2013). Together, these results show that infants• ability to generalize lexical stress at 9…10 months of age is related to language-speci“c prosodic learning about the use (or lack thereof) of stress to distinguish words in one•s native language.

Interestingly, French-learning 10-month-olds• problems with stress generalization are also accompanied by comparatively more effortful processing, even in basic stress discrimination tasks (i.e., with single words). Bijeljac-Babic et al. (2012) tested French-learning infants at this age: Infants were “rst familiarized with (one stress pattern for either 1 (either /•ga.ba/ or /ga.•ba/) or 2 min, and then tested with either the same or a novel stress pattern. Infants discriminated this stress contrast after 2 min of familiarization (looking longer at the novel stress pattern at test), but failed after only 1 min of familiarization. Notably, a group of bilingual infants„who heard at least 50% of another language that had lexical stress„ succeeded at discriminating, and looked longer at the novel stress pattern with just 1 min of familiarization. Moreover, the amount of exposure to the non-French language correlated positively with discrimination. Together, these results show that exposure to French does not just affect generalization abilities. French-learning infants„although still capable of discriminating lexical stress„show more effortful processing of stress information, even in single-word discrimination tasks.

1.3. Overview of the current experiment

Section 1.2 suggests that French-learning 10-month-olds can indeed process lexical stress information, but just do so less robustly than German-, Spanish-, or English-learning infants in failing to generalize stress patterns. We taught a lexical stress contrast to French-learners at 10 months of age, and then studied their ability to generalize this information.

Infants were “rst presented with a training phase, where two CVCV words minimally distinguished by stress (e.g., trochaic /•li.fo/ versus iambic /li.•fo/) were presented as object labels. Two groups of infants were each given a different kind of training: The “rst group of infants received consistent training, where trochaic and iambic word forms statistical learning did not generalize, we would have expected null results after the inconsistent training, and possibly some discrimination in the consistent training from the additional object labeling cues. A more likely possibility is that we would “nd evidence of lexical stress discrimination in both conditions, as infants have previously been shown to generalize based on prior statistical learning (Maye et al., 2008). Under this scenario, we still predicted enhanced discrimination after consistent (statistical and object labeling cues present) versus inconsistent training (only statistical cues present).

2. Material and methods

2.1. Participants

Forty-eight healthy, full-term infants hearing French at least 90% of the time in their environment were tested. Infants were around 10 months of age (24 infants in each training condition; 18 boys, $M = 296$ days, Range= 273 - days…322 days). Thirty additional infants were tested, but not included in the analysis for the following reasons: external noise from building construction ($n_{consistent} = 8$; $n_{inconsistent} = 6$); fussiness ($n_{consistent} = 5$; $n_{inconsistent} = 4$); parental interference ($n_{consistent} = 2$; $n_{inconsistent} = 2$); experimenter/equipment error ($n_{inconsistent} = 2$); prolonged medical intervention at birth ($n_{consistent} = 1$).

2.2. Auditory stimuli

To ensure that any effects were not idiosyncratic to a particular phonetic sequence, each infant heard one of the eight pairs of CVCV disyllables, created with two consonants from the set of /b, f, g, m, l/ and two vowels from the set of /a, i, o, u/. All disyllables were phonotactically legal non-words in French, and pairs were chosen such that the non-word heard during training shared no common vowels (which contain the most important acoustic cues to lexical stress) with the non-word heard during test (Table 1). In addition, three familiar words (i.e., ••clefs,•• keys; ••main,••hand; and ••chien,••dog) were selected for the pre-training, because they are among the earliest acquired monosyllabic words by French-learning 10-month-olds (Kern, 2007).

A female native bilingual Dutch…French speaker produced several tokens of each non-word and familiar word in adult-directed French, applying lexical stress to either the “rst or second syllable (i.e., a trochaic or iambic pattern, respectively). Adult-directed speech was used, following previous studies of lexical stress perception in infants (e.g., Bijeljac-Babic et al., 2012; Höhle et al., 2009; Skoruppa et al., 2009), and because it is not clear how the prosodic variation inherent in infant-directed speech (e.g., Fernald, 1989; Stern, Spieker, & MacKain, 1982) could affects stress discrimination in infants. Eight tokens of each non-word were selected for use in the experiment, and measurements of duration, peak amplitude, and peak pitch were performed on the vowels of each non-word token (Table 2). These 3 acoustic cues are some of the most prominent for lexical stress (Cutler, 2005 averaged measurements for each non-word in Table 1. Between-items factors included both POSITION [1] (“rst or second vowel) and STRESS (trochee or iamb).

In the test phase, infants were “rst familiarized with repeated trials containing 4 tokens of a new test word with a single stress pattern (e.g., /•ma.bu/) until 1 min of looking time was accumulated (coded online). Four blocks of two trials (8 trials total) were then presented. Each block had one trial containing 4 (new) tokens of the same stress pattern (e.g., /•ma.bu/) and one trial containing 4 tokens of the novel stress pattern (e.g., /ma.•bu/). Longer looking to the for checklist measures, overall scores, $M = 13.08$, $SD = 11.22$, were positively skewed, S.W$(748) = .88$, $p < .001$. The over-

4. Discussion

Infants acquire phonetic patterns de“ned by native lexical contrasts by their “rst year of life. This likely happens as infants pay attention to the distribution of phonetic tokens in acoustic space, but recent work also suggests that additional mechanisms could supplement learning. For example, the lexical structure of the input, and even lexical knowledge itself may contribute to this posit new perceptual constructs (i.e., an association between features, or a novel gestalt grouping). To analogize to the current experiment, consistent training could have either primed attention to the critical acoustic cues during test, or helped infants construct perceptual representations de“ned by a contrastive grouping (i.e., a trochee versus an iamb), allowing test phase stimuli to be perceived through these gestalt groupings. Our data cannot de“nitively distinguish between these two possibilities, but the fact that our effects were somewhat unstable is important. If performance is a true indication of perceptual competence (i.e., if methodological factors were not at play), then results are more indicative of a priming effect that only momentarily increased attention to critical acoustic features of stress in the test phase. Future research must tease apart these two possibilities.

Another outstanding question concerns the precise acoustic cue(s) to which infants were paying attention in the present study. When French-learners hear lexical stress cues as used in a natural language, do they generalize based on all available cues (some combination of duration, intensity, and pitch), or do they preferentially rely on language-salient cues (duration, since syllable-duration is used to indicate the end of a phrase in French, and is the closest native cue to lexical stress)? One possibility is that French-learners would have learned better if stress were implemented using just the single acoustic cue that is already salient in their language (i.e., duration). This issue requires additional study.

Future work will also need to specify whether generalization is constrained, and whether constraints differ according to the learning mechanism in question. For example, the kind of phonetic generalization observed by Maye et al. (2008) violates a pattern seen in the world•s languages. Infants learning certain languages (like Thai) would not bene“t from the kind of feature generalization tested in that paper: Voicing distinctions can exist for some phonetic classes (e.g., labial /b/ and /p/ stop consonants exist in Thai, as do coronal /d/ and /t/ stops), but not for

during the “rst 6 months of life. Developmental Medicine and Child Neurology, 50(9), 678…683. http://dx.doi.org/10.1111/j.1469-8749.2008.03055.x.

Hoff, E. (2006). How social contexts support and shape language development. Developmental Review, 26(1), 55…88. http://dx.doi.org/10.1016/j.dr.2005.11.002 .

Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language speci“c prosodic preferences during the “rst half year of life: Evidence from German and French infants. Infant Behavior and Development, 32(3), 262…274. http://dx.doi.org/10.1016/j.infbeh.2009.03.004.

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-speci“c information in word segmentation by infants. Journal of Experimental Psychology: Human Perception and Performance, 26(5), 1570…1582.

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier & L. P. Lipsitt (Eds.). Advances in infancy research (Vol. 5, pp. 69…95). Westport, CT: Ablex .

Johnson, S. P. (2010). How infants learn about the visual world. Cognitive Science, 34(7), 1158…1184. http://dx.doi.org/10.1111/j.1551-6709.2010.01127.x.

Johnson, S. P., & Shuwairi, S. M. (2009). Learning and memory facilitate predictive tracking in 4-month-olds. Journal of Experimental Child Psychology, 102(1), 122…130. http://dx.doi.org/10.1016/j.jecp.2008.02.004.

Jusczyk, P. W. (1993). From general to language-speci“c capacities: The WRAPSA model of how speech perception develops. Journal of Phonetics, 21(1…2), 3…28.

Jusczyk, P. W (2000). The discovery of spoken language. MIT Press, p. 314.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants• preference for the predominant stress patterns of English words. Child Development, 64(3), 675…687. http://dx.doi.org/10.1111/j.1467-8624.1993.tb02935.x .

Kern, S. (2007). Lexicon development in French-speaking infants. First Language, 27(3), 227…250. http://dx.doi.org/10.1177/0142723706075789 .

Toscano, J. C., & Mcmurray, B. (2010). Cue integration with categories: Weighting acoustic cues in speech using unsupervised learning and distributional statistics. Cognitive Science, 34(3), 434–464. http://dx.doi.org/10.1111/j.1551-6709.2009.01077.x.Cue.

Vallabha, G. K., McClelland, J. L., Pons, F., Werker, J. F., & Amano, S. (2007). Unsupervised learning of vowel categories from infant-directed speech. Proceedings of the National Academy of Sciences of the United States of America, 104(33), 13273–13278. 10.1073?pnas.0705369104.

van Heugten, M., & Johnson, E. K. (2014). Learning to contend with accents in infancy: Benefits of brief speaker exposure. Journal of Experimental Psychology: General, 143(1), 340–350. http://dx.doi.org/10.1037/a0032192.

Vihman, M. M., Nakai, S., DePaolis, R. A., & Hallé, P. A. (2004). The role of accentual pattern in early lexical representation. Journal of Memory and Language, 50(3), 336–353.

Walley, A. C. (2005). Speech perception in childhood. In D. B. Pisoni & R. E. Remez (Eds.), The Handbook of Speech Perception (pp. 449–468). Boston, MA: Blackwell.

Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. Reading, 16(1), 5–20. http://dx.doi.org/10.1023/A:1021789804977.

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–b68.

Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. Cognition, 103(1), 147–162. http://dx.doi.org/10.1016/j.cognition.2006.03.006.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. Infant Behavior and Development, 7(1), 49–63.

Werker, J. F., Yeung, H. H., & Yoshida, K. A. (2012). How do infants become experts at native-speech perception? Current Directions in Psychological Science, 21(4), 221–226. http://dx.doi.org/10.1177/0963721412449459.

White, K. S., & Aslin, R. N. (2011). Adaptation to novel accents by toddlers. Developmental Science, 14(2), 372–384. http://dx.doi.org/10.1111/j.1467-7687.2010.00986.x.

Wilcox, T., & Chapa, C. (2004). Priming infants to attend to color and pattern information in an individuation task. Cognition, 90(3), 265–302. http://dx.doi.org/10.1016/S0010-0277(03)00147-1.

Wu, R., Gopnik, A., Richardson, D. C., & Kirkham, N. Z. (2011). Infants learn about objects from statistics and people. Developmental Psychology, 47(5), 1220–1229. http://dx.doi.org/10.1037/a0024023.

Yeung, H. H., Chen, K. H., & Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. Journal of Memory and Language, 68(2), 123–139. http://dx.doi.org/10.1016/j.jml.2012.09.004.

Yeung, H. H., Chen, L. M., & Werker, J. F. (2014). Referential labeling can facilitate phonetic learning in infancy. Child Development. doi:10.1111/cdev.12185 (in press).

Yeung, H. H., & Werker, J. F. (2009). Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. Cognition, 113(2), 234–243. http://dx.doi.org/10.1016/j.cognition.2009.08.010.

Yoshida, K. A., Pons, F., Maye, J., & Werker, J. F. (2010). Distributional phonetic learning at 10 months of age. Infancy, 15(4), 420–433. http://dx.doi.org/10.1111/j.1532-7078.2009.00024.x.