The ability to use contextual cues to achieve phonological constancy emerges by 14 months

Ye Feng[1,2,4], René Kager[3], Regine Lai[1], Patrick C. M. Wong[1,2]

1. Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Hong Kong SAR, China.
2. Brain and Mind Institute, The Chinese University of Hong Kong, Hong Kong SAR, China.
3. Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, Netherlands.
4. Department of Linguistics, Beijing Language and Culture University, Beijing, China.

Introduction

- The ability to map similar sounding words to different meanings alone is far from enough for successful speech processing. To overcome variability in the speech signal, young learners must also adapt to the surface variability and achieve phonological constancy.
- Previous studies have shown that infants at 14 months are able to utilize variations in stimulus-internal cues to form phonological categories and to learn words (Apfelbaum and McMurray, 2011; Höhle et al., 2020; Rost and McMurray, 2009, 2010).
- The present study takes into consideration the fact that talker variability can easily lead to acoustic overlap between categories, in which case reliance on stimulus-external or contextual cues is obligatory for successful talker adaptation.

Experiment 1

Method

- Stimuli
  - Cantonese Tone 1 (high level tone) vs Tone 3 (mid level tone) carried by /pi/
  - Non-words in Cantonese
  - 3 tokens x 2 tones x 6 speakers (female native speaker of Cantonese) = 36 tokens

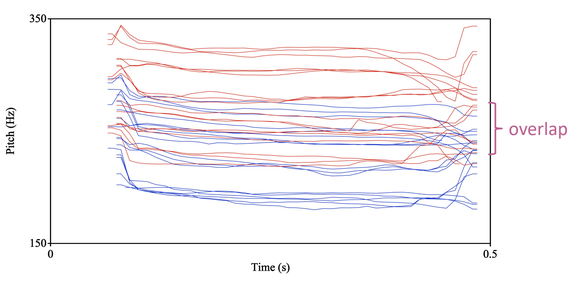


Figure 1. Pitch contours of the 36 tokens in Exp1-3, Tone1 in red and Tone 3 in blue.

Tone	F0 (Hz)	F1 (Hz)	F2 (Hz)
T1	273	342	2767
	33	*51*	*226*
T3	229	356	2797
	29	*47*	*85*

Table 1. Average acoustic measurements and standard deviations (*in italics*) of the vowels of the 36 tokens in Exp1-3.

- Procedures
  - Habituation-based visual-fixation procedure (VFP)
  - Two word-object pairings, the Switch Task

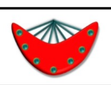
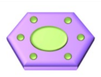
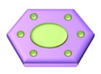
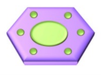
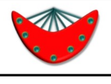
Phases		Visual Stimuli	Auditory Stimuli: Exp1 & 3a	Auditory Stimuli: Exp2 & 3b
Pre-test (1 trial)			/nu2/, /nu2/, […], /nu2/	/nu2/, /nu2/, […], /nu2/
Habituation (6-24 trials)			/pi1/, /pi1/, […], /pi1/	睇呢度。/pi1/。睇唔睇到？/pi1/。你睇下。 /pi1/。 好靚啊！/pi1/。好得意啊！/pi1/。咩嚟架？/pi1/。
			/pi3/, /pi3/, […], /pi3/	睇呢度。/pi3/。睇唔睇到？/pi3/。你睇下。/pi3/。 好靚啊！/pi3/。好得意啊！/pi3/。咩嚟架？/pi3/。
Test	**same** (1 trial)		/pi1/, /pi1/, […], /pi1/	睇呢度。/pi1/。睇唔睇到？/pi1/。你睇下。 /pi1/。 好靚啊！/pi1/。好得意啊！/pi1/。咩嚟架？/pi1/。
	switch (1 trial)		/pi3/, /pi3/, […], /pi3/	睇呢度。/pi3/。睇唔睇到？/pi3/。你睇下。/pi3/。 好靚啊！/pi3/。好得意啊！/pi3/。咩嚟架？/pi3/。
Post-test (1 trial)			/nu2/, /nu2/, […], /nu2/	/nu2/, /nu2/, […], /nu2/

Table 2. A demonstration of visual and auditory stimuli used throughout the procedures.
Note: The order of the two test trials and the switched tone were counterbalanced across subjects. Experiment 1 and 3a, as well as Experiment 2 and 3b, differed in the number of speakers during the habituation and test phases. Gloss on the Cantonese carrier phrases can be found in Figure 3.

- Participants
  - Cantonese monolingual families
  - No prior history of perceptual or neurological disorders

Group	N	Range (days)
14 months	24 (12 girls)	392—450
18 months	22 (13 girls)	510—567
24 months	24 (12 girls)	690—747

Results

- When talker variability yielded category over-lap, infants did not benefit from it in word learning, not even 2-year-olds.
- It was also confirmed that infants were not able to resolve the category ambiguity across speakers to reach phonological constancy without extrinsic cues.

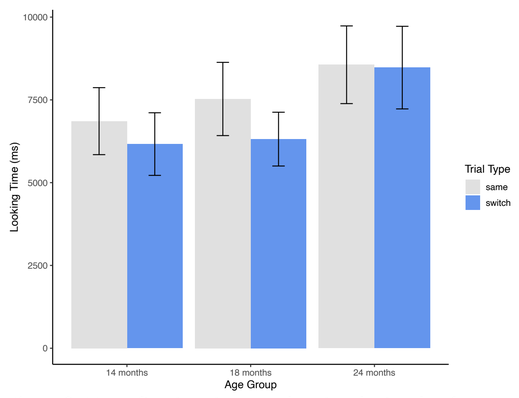
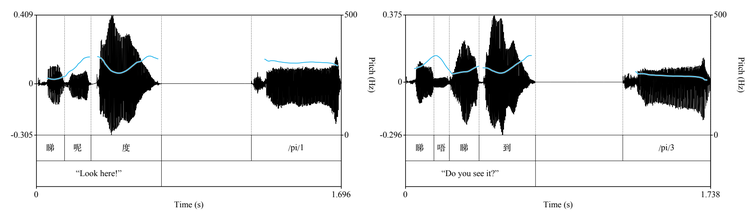


Figure 2. Mean fixation times to the visual stimulus for same and novel trials in the test phase in Experiment 1 divided by age group (error bars: SEM).

Experiment 2

Method

- Stimuli



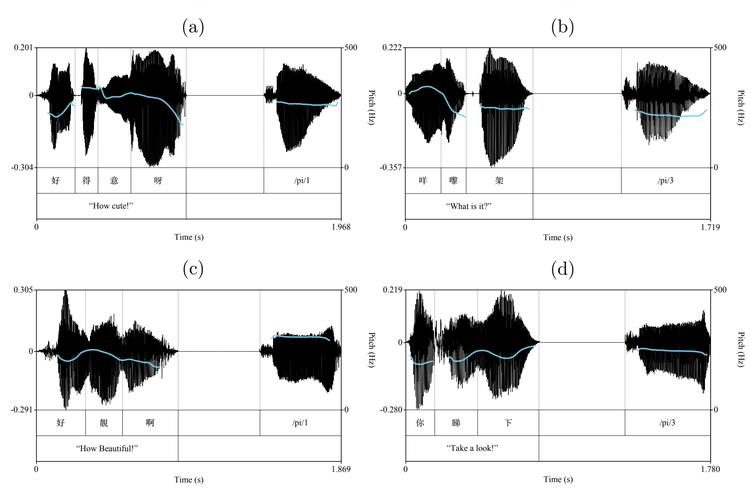


Figure 3. Examples of the final stimuli with sound waves (in black), pitch tracks (in blue), text grids and glosses. (a) and (b) are example stimuli of the speaker-matched condition, where the carrier phrases and the target non-word were from the same speaker (Speaker B in these examples). (c) and (d) are example stimuli of the speaker-mismatched condition, where carrier phrases from Speaker A are mismatched with targets from Speaker D who has a lower mean F0. In this case, the target T1 sounds more like a T3 (3c) and the target T3 sounds lower than normal (3d). (e) and (f) are example stimuli of the speaker-mismatched condition, where carrier phrases from Speaker F are mismatched with targets from Speaker B who has a higher mean F0. Therefore, the target T1 sounds higher than normal (3e) and the target T3 sounds closer to a T1 (3f).

- Target non-words same as those in Exp.1
- Six carrier phrases (adapted from Singh et al., 2016) used as carrier phrases (contextual cues)

- Participants
  - Cantonese monolingual families
  - No prior history of perceptual or neurological disorders

Condition	Group	N	Range (days)
Speaker-matched	14 months	24 (12 girls)	393-447
	18 months	24 (12 girls)	512-569
	24 months	22 (10 girls)	692-745
Speaker-mismatched	14 months	24 (12 girls)	392-449
	18 months	24 (12 girls)	512-569
	24 months	22 (11 girls)	691-746

Results

- When given contextual cues, infants from all age groups (14 to 24 months) showed cognitive capacity for adult-like talker adaptation.
- Infants relied on the phonetic information, rather than referential information, provided in contextual cues to track different speakers' phonetic spaces (tone spaces in this case) and extract the relative pitch height of the target tone produced by each speaker.

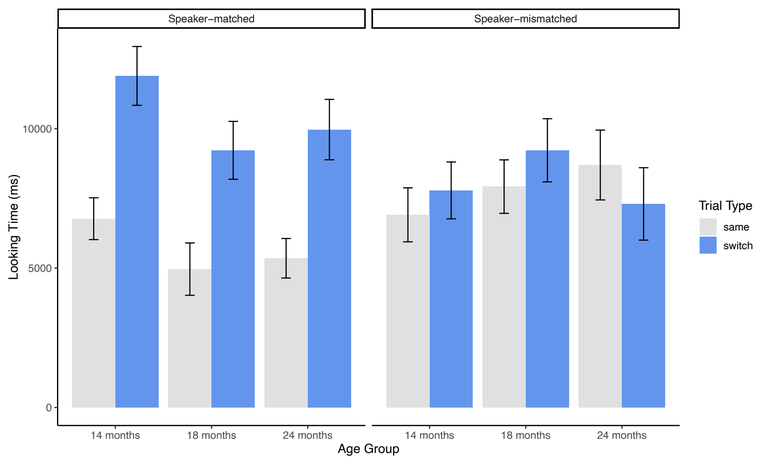


Figure 4. Mean fixation times to the visual stimulus for same and novel trials in the test phase in Experiment 2 divided by age group with speaker-matched condition on the left and speaker-mismatched condition on the right (error bars: SEM).

Experiment 3

Method

- Single-speaker versions of experiments with the same stimuli and procedures as Exp. 1 and 2.
- Tested on the 14 months group

Exp.	Group	N	Range (days)
3a	14 months	24 (12 girls)	391—446
3b	14 months	24 (12 girls)	513—569

Results

- Results of Experiment 3 replicated previous findings on phonological distinctiveness, showing that even when the stimuli were produced by a single speaker, 14-month-old infants cannot reliably integrate Cantonese T1-T3 contrast into different word meanings unless contextual cues were provided, similar to previous results with segmental contrasts.

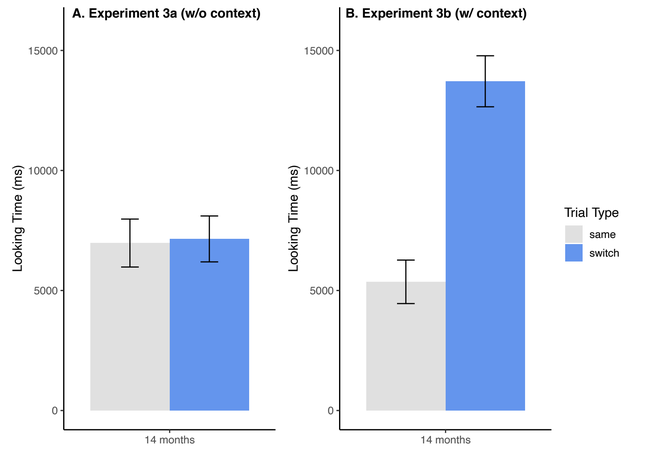


Figure 5. Mean fixation times to the visual stimulus for same and novel trials in the test phase in Experiment 3a (left) and Experiment 3b (right) (error bars: SEM).

Acknowledgements

- This work was supported by the University Grants Committee (HKSAR) (RGC34000118), the Innovation and Technology Fund (HKSAR) (ITS/067/18), Dr. Stanley Ho Medical Development Foundation, and the Global Parent Child Resource Centre Limited.
- This study was completed as part of the first author's doctoral dissertation. We thank all of our infant participants and their caregivers for their invaluable contributions to the study.

References

- Apfelbaum, K. S. and McMurray, B. (2011). Using variability to guide dimensional weighting: Associative mechanisms in early word learning. *Cognitive Science*, 35(6):1105-1138. https://doi.org/10.1111/j.1551-6709.2011.01181.x
- Höhle, B., Fritzsche, T., Meß, K., Philipp, M., & Gafos, A. (2020). Only the right noise? Effects of phonetic and visual input variability on 14-month-olds' minimal pair word learning. *Developmental* Science, 23(5), e12950. https://doi.org/10.1111/desc.12950
- Rost, G. C. and McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, 12(2):339-349. https://doi.org/10.1111/j.1467-7687.2008.00786.x
- Rost, G. C. and McMurray, B. (2010). Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy*, 15(6):608-635. https://doi.org/10.1111/j.1532-7078.2010.00033.x
- Singh, L., Poh, F. L., and Fu, C. S. (2016). Limits on monolingualism? a comparison of monolingual and bilingual infants' abilities to integrate lexical tone in novel word learning. *Frontiers in psychology*, 7:667. https://doi.org/10.3389/fpsyg.2016.00667