Applied Psycholinguistics (2026), **47**, e9, 1–18
doi:10.1017/S0142716426100472

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

Listening to different talkers and its effect on word identification for children using hearing devices

Nan Xu Rattanasone[1], Katie Neal[2], Greg Leigh[3,4], Mridula Sharma[5] and Katherine Demuth[6]

[1]Department of Linguistics, Macquarie University, Australia, [2]The Shepherd Centre, Australia, [3]NextSense Institute, Australia, [4]School of Education, Macquarie University, Australia, [5]Macquarie University Department of Linguistics, Australia and [6]Linguistics, Macquarie University, Australia
Corresponding author: Nan Xu Rattanasone; Email: nan.xu@mq.edu.au

(Received 9 September 2024; revised 21 December 2025; accepted 14 January 2026)

Abstract

Whether listening to different talkers improves or impedes word identification has important implications for theory and practice. Yet, past research on children with hearing devices shows discrepant findings. This study tested 22 children with typical hearing (mean 5;0) and 20 with hearing devices (mean 4;11) on a remote, online 4-alternative forced-choice task (with a 4-picture display) delivered on iPads, with blocks containing 1 vs. 6 different talkers. All words were familiar to young children and were minimal pairs contrasting in voicing and place of articulation in the word-initial and word-final positions. Word identification was worse for place contrasts occurring word finally when listening to different talkers, but no effect was found for voicing contrasts. A consistent position effect was also found, where word identification was poorer across all word-final contrasts. However, no group differences were detected. These results suggest that even when listening to familiar words in good listening environments, the word-final position remains vulnerable to word misidentification, which can be further impeded by listening to different talkers. These effects impact children with and without hearing devices to a similar degree.

Keywords: hearing loss; preschoolers; lexical identification; multi-talker; minimal pairs

Young children using hearing devices, such as hearing aids (HA) and cochlear implants (CIs), receive an altered acoustic signal that is either amplified or reduced in acoustic information. Hearing devices cannot restore typical hearing, and many children will require early intervention to support speech and language development. These children can also often have limited experience listening to different talkers during early development. For example, intervention is often delivered in clinics by a single clinician and reinforced in the home by the primary caregiver

(often the mother). This contrasts with the listening environment in preschools and schools, where there are multiple talkers and sources of language input (e.g., teachers and peers). Compared to listening to a single talker, listening to multiple talkers introduces acoustic variations that are not relevant to understanding word meaning, such as information about the speakers' sex, age, and dialect. Given the hearing device limitations and experiences of these young children (based on our discussions with clinicians), receiving intervention through a single clinician supported by a single carer at home, the acoustic variation inherent in listening to many different and unfamiliar speakers may pose challenges to word identification and comprehension for young children with aided hearing. This study addresses this issue by investigating the effect, if any, of listening to different talkers on word identification for children with aided hearing.

Development of phonological categories

Variation in production is ubiquitous, both within and across different talkers, because speech is easily influenced by anatomical differences, speaking rate, discourse intent, and social context. Therefore, tuning in to phonological constancy despite acoustic-phonetic variation is an important part of acquiring language (Mulak & Best, 2013). This process begins at the very onset of language development, with infants tuning into their native phonological categories before their first year of life. For example, whereas 6–8-month-olds exposed to English can discriminate non-native Hindi dental vs. retroflex stops, whereas older 11–13-month-olds cannot (Werker & Lalonde, 1988). This shows that infants begin to ignore acoustic variation that is not phonemically contrastive in their native language. Across different languages, language-specific categories are established early during infancy, i.e., for lexical tones (e.g., Cantonese and Mandarin) at around 4 months, vowels at 6 months, and consonants at 8 months (Polka & Werker, 1994; Werker & Lalonde, 1988; Yeung et al., 2013). By the middle of the second year of life, infants can associate a change in native phonological category with a change in word meaning (Werker et al., 1998). However, research evidence on language development is primarily based on input from the main carer, predominantly the mother (Dailey & Bergelson, 2022).

As children develop, they must fine-tune their perception of phonological categories to accommodate acoustic variations that can differ considerably across talkers depending on the physiology of the talker's articulators, e.g., vocal tract size, sex, age, the nature of the target audience, e.g., child vs. adults, the linguistic background of the speaker, e.g., regional dialects and cross linguistic accents, and speaker discourse intent, e.g., focus. Different talkers can implement acoustic cues with substantial variation for both within and across phonological category contrasts, leading to the well-established problem of many-to-many mappings between the acoustic implementation of phonological categories and their perception (Magnuson et al., 2021). These talker differences do not contrast word meaning, and developing this understanding is essential for word identification.

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

Acoustic variation and lexical representations

One way this ability may be established is by learning which acoustic variations are not important to word meaning, as demonstrated in research on word learning during infancy and in second-language (L2) learners. Infants trained on novel syllables paired with novel objects demonstrated better learning (and, in some instances, could only learn novel pairings) when listening to many different talkers (Höhle et al., 2020; Rost & McMurray, 2009, 2010). Adult L2 learners who listened to many different talkers during training also performed better in perceiving L2 categories compared to single talkers, and more importantly, only the different talkers condition generalized to new and untrained items (Lively et al., 1993; Sadakata & McQueen, 2013; Zhang et al., 2021). These findings suggest that listening to many different talkers exposes the listener to acoustic variation, which helps learners discover relevant acoustic cues for distinguishing phonemic categories to establish phonological constancy (Apfelbaum & McMurray, 2011; Rost & McMurray, 2010). However, despite the advantages of word learning through exposure to different talkers, this is often not possible during early development, where young children typically receive input from a very limited range of carers. And for children with aided hearing who are receiving intervention, e.g., on vocabulary development, this is always delivered by a single clinician. This raises questions about whether and how listening to different talkers might affect their lexical representations and word identification, especially as they move into preschools and schools where they will encounter many different talkers.

While information about different talkers might appear irrelevant for word meaning, there is evidence that talker information is represented at the word level together with phonemic information. This is because acoustic variation arising from different talkers can provide important indexical information about speaker characteristics and is an essential part of communication (Hay & Drager, 2007). The distributed representation account argues that both phonemic and speaker indexical information, along with other grammatical features, are all associated at the word level (Elman, 2004). In fact, eye-tracking shows that phonemic and indexical information are simultaneously activated with similar time courses (during a visual world task) using different talkers (Creel et al., 2008). This account suggests that talker information is not simply ignored because it is not relevant for word meaning, and is also consistent with findings that show performance costs to *word identification* when listening to different talkers, where the coactivation of indexical information must require additional attentional resources.

Indeed, studies on children and adults show slower, less accurate, or both types of performance costs in word identification when listeners must switch between different talkers, compared to listening to a single talker (where presumably indexical information is not coactivated). While adults are slower to respond during word identification (Magnuson & Nusbaum, 2007—using a word-monitoring task), children are less accurate in word identification (Ryalls & Pisoni, 1997—using a 6-alternative forced-choice task with pictures) when listening to different talkers. These performance costs suggest that sensitivity to acoustic/phonetic information associated with different talkers, while useful for

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

learning words, impedes word identification. The effect on children has real-world implications for both learning and social development. Until their language system matures, listening to different talkers will continue to impact children's language performance. However, children with aided hearing have altered and reduced access to acoustic information. Hearing devices might therefore impact perceptual sensitivity to the more subtle and unpredictable acoustic variation inherent in indexical information. If so, they may be less negatively impacted by switching among different talkers.

Impact of talker variation on aided hearing

Research on school-aged children with CIs has shown that they cannot reliably identify unfamiliar talkers (Cleary & Pisoni, 2002). A recent systematic review found that those who were prelingually deaf performed worse than their typically hearing peers at identifying different talkers (Colby & Orena, 2022). These studies suggest that hearing devices reduce talker-indexical information, likely by reducing acoustic variation. This would suggest that children with aided hearing might be less affected by talker variation during word identification. However, research on the effects of listening to different talkers in populations with aided hearing is limited, with conflicting results. A study on a small sample of 6 *post-lingually* deafened adult CI users found that listening to different individual talkers sequentially led to less accurate word identification performance on words with different vowel contrasts (i.e., hVd environments, where V represents 12 different vowels) (Chang & Fu, 2006). Similarly, a larger sample of *prelingually* deafened Mandarin-speaking 10- to 20-year-olds with CIs was also less accurate in word identification when listening to different talkers for words that contrasted in lexical tones and vowels (Chang et al., 2016). Together, these studies suggest that it is not just indexical information that is affected by hearing loss and altered through hearing devices, but also phonemic information.

However, contrary to the above, one study examining prelingually deafened primary school-aged children with CIs found that they were more accurate at word identification when listening to multiple talkers than when listening to a single talker (Kirk et al., 2000). This finding is not only counter to previous research but also to research on typical hearing samples. The results could suggest that reduced acoustic information may lead to reduced sensitivity to talker variation and affect word identification. It is also possible that in younger children with less exposure to different talkers, word identification is not impacted by talker variation. However, this research did not include typically hearing children as a comparison group. Also unclear is whether known hard-to-distinguish acoustic contrasts in children with aided hearing may be more affected by different talkers, as the current Kirk et al. (2000) study was more concerned with controlling for word-frequency effects (neighborhood density). Addressing these two issues will help provide a better understanding of whether and how acoustic variations associated with listening to different talkers might affect word identification for children using hearing devices, support clinicians in more informed counselling discussions with parents and teachers, and inform intervention practice.

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

Voicing and place contrasts

Two pervasive challenges in terms of speech production for children using hearing devices are voicing and place contrasts. Several studies across languages have reported voicing errors by children using hearing devices based on both perceptual ratings and phonetic transcriptions of spontaneous speech (e.g., Baudonck et al., 2011; Elfenbein et al., 1994). However, a recent Australian study reporting on a sample of preschoolers (3–5-year-olds) using hearing devices showed that they can produce acoustically distinct voicing categories across all three places of articulation (bilabial, alveolar, and velar), but these were less systematic in the word-final compared to the word-initial position (Bruggeman et al., 2021). Given these results, we might also expect word identification to be more challenging when contrasts occur word finally, where there is greater phonetic variation. Indeed, this is consistent with the listening hierarchy used in clinical practice (Erber, 1982; Estabrooks et al., 2020), which proposes that voicing and place contrasts in word-final position are among the hardest to master for children using hearing devices.

The present study

In Australia, universal newborn hearing screening is available nationally. Once a child is identified as having possible hearing loss, their parents are offered further clinical support. This includes the option to receive different types of hearing device(s) depending on the type and severity of their child's hearing loss. HA are typically recommended for mild to moderate hearing loss, while a CI is recommended for severe to profound hearing loss. Many families also access early oral language intervention through government-subsidized programs. This project is conducted in collaboration with researchers from two such programs, NextSense and The Shepherd Centre. While our industry collaborators provided theoretical and applied contributions to the research, e.g., identifying a gap in the evidence base and assisting with participant recruitment, they were not involved in data collection, analyses, or interpretation of the results.

This study investigated the effect of listening to different talkers on word identification using minimal pair words by young children using hearing devices. Minimal pair words differ by a single segment and phonological feature, e.g., ***b****in* vs. ***p****in* (voicing contrast) and ***p****ea* vs. ***k****ey* (place contrast). When there is limited semantic and discourse information, the listener must rely on acoustic cues to identify the correct word. Therefore, minimal pair words provide a great method for isolating the effect of increased acoustic variation from different talkers on word identification accuracy. Voicing and place contrasts in both word-initial and word-final positions were examined. Only words familiar to young children were used to better ensure that their performance reflected word identification rather than novelty effects. The task was delivered remotely online in the children's home, which reflects their typical listening environment. To ensure that children were familiar with the task, we used a 4-alternative forced-choice paradigm similar to other vocabulary tests (e.g., the Peabody Picture Vocabulary Test (PPVT)) that are already familiar to children using hearing devices who have received speech and oral language support.

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

The first research question addressed whether young children with hearing devices would benefit from listening to different individual talkers during a word identification task. Based on past research, we predicted that performance for these children might be better in the different-talker than the single-talker condition (Kirk et al., 2000). On the other hand, performance for children with typical hearing should be less accurate in the different-talker condition than in the single-talker condition (Ryall & Pisoni, 1997).

The second research question explored the locus of any word-identification challenges (e.g., voicing vs. place contrasts in word-initial vs. word-final positions). We expected that word identification accuracy would be reduced for both groups when the contrast occurred word-finally, where there is typically greater phonetic variation within and across talkers. However, if children with hearing devices are less affected by talker variation, they might also be less affected by variation across word positions than their typically hearing peers.

Method

Participants and design

Twenty-two children with typical hearing and a mean age 5;0 (range: 3;1–7;1; 10 boys and 12 girls) and 20 children with hearing devices and a mean age of 4;11 (range: 3;2–8;0; 9 boys and 11 girls) participated in the study. See Table 1 for the hearing profile of children with hearing devices. All children reported English as their only home language. The study was approved by [university removed for reviewing] Human Research Ethics Committee approval number 52021521731926. All parents provided written consent to participate in the study and received a $30 gift voucher from the research team.

A mixed design was adopted with Group as the between-subjects factor and all other manipulations as within-subjects factors; i.e., all children received all levels within all conditions, including single vs. multiple talkers, all voicing and place-of-articulation contrasts, in both word-initial and word-final positions.

Stimuli

The target minimal pair words were created to contrast in voicing (voiced/voiceless) or place (bilabial-alveolar, bilabial-velar, and alveolar-velar), with half of the targets in word-initial and half in the word-final position (see Tables 2 and 3). All words used were familiar to preschoolers, based on the CBeebies Subtlex-UK database (van Heuven et al., 2014), and depicted in black-and-white line drawings. All children were shown the pictures by their parents the evening before the experiment, and their parents also read the labels for each picture aloud. This ensured that children were familiar with the pictures, especially the two with Zipf[1] scores of less than 4 (considered low-frequency words). Two counterbalanced versions were created so that, for each minimal pair, half the children in each group heard one as the target and the other as the non-target.

The stimuli were recorded in the carrier sentence "Find the X." The sentences were spoken by six native Australian English-speaking female talkers with no reported hearing loss. Five talkers were between 28 and 39 years old, and one was 62

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

Table 1. Characteristics of children with HL

Child	Age (y;m)	Sex	Device type	Type of HL	4-freq PTA		Age (mths) at initial device fitting	
					(R)	(L)	(R)	(L)
1	3;11	M	CI (Bilateral)	Sensorineural	Moderately severe	Moderately severe	22	1
2	3;11	M	HA (Bilateral)	Sensorineural	Moderately severe	Moderately severe	1	1
3	3;2	F	CI (Bilateral)	Sensorineural	Profound	Profound	7	1
4	3;5	M	HA (L)	Conductive	Mild	Moderately severe	–	8
5	4;0	F	HA (Bilateral)	Mixed	Mild to severe	Mild to severe	23	23
6	4;10	F	Bimodal	Sensorineural		Moderate	43	5
7	4;10	F	CI	Sensorineural	Profound	Profound	9	3
8	4;11	F	HA (R)	Sensorineural	Severe	Slight	30	–
9	5;1	M	HA (Bilateral)	Sensorineural	Slight - Mild	Slight–Mild	38	38
10	5;3	M	Bimodal	Sensorineural	Severe	Severe	6	1
11	6;0	F	CI (Bilateral)	Sensorineural	Severe	Severe	6	6
12	6;0	F	CI (Bilateral)	Sensorineural^	Moderately severe	Moderately severe	27	3
13	6;11	M	Bimodal	Sensorineural	Moderate	Severe	45	1
14	6;11	F	HA (Bilateral)	Sensorineural	Mild	Mild	74	74
15	6;9	F	Bimodal	Sensorineural	Severe	Severe	11	1
16	7;4	F	HA (Bilateral)	Sensorineural	Severe	Mild	2	2

Note: 4 children did not provide permission to share their clinical data. CI = cochlear implant, HA = hearing aid, bimodal = using CI and HA, mixed = multiple aetiologies, and 4-freq PTA = pure tone average over 4 frequencies. Degree of hearing loss following ASHA.

(the grandmother of two preschoolers). All six talkers are experienced in working with young children and were asked to speak as they would to a preschooler. The stimuli were recorded in a sound-attenuated room using a Zoom Q8 Video Recorder, recorded at 720p and 1080p (Full HD) with a RODE NTG-1 Microphone. All sentences were excised using Adobe Premiere Pro with audio extracted from video recordings and saved as uncompressed.wav files. All audio recordings were normalized across different talkers and sentences to 60 dB SPL (conversational level) in Praat (Boersma & Weenink, 2021).

Procedure

The online experiment was created and hosted by the Gorilla Experiment Builder (www.gorilla.sc; Anwyl-Irvine et al., 2020. All experiments were conducted using an

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

Table 2. Minimal pair words with voicing contrasts in the word-initial (onset) or word-final (coda) position with (Zipf)

	Voiced		Voiceless	
Onset	Bin	(4.51)	Pin	(4.30)
	Bath	(4.94)	Path	(4.47)
	Dough	(4.43)	Toe	(4.58)
	Deer	(4.28)	Tear	(4.61)
Coda	Seed	(4.73)	Seat	(4.65)
	Pod	(5.20)	Pot	(5.07)
	Log	(4.56)	Lock	(4.24)
	Bag	(5.46)	Back	(6.24)

Table 3. Minimal pair words with place contrasts in the word-initial (onset) or word-final (coda) position with (Zipf)

	Bilabial		Alveolar		Velar	
Onset	Pool	(4.79)	Tool	(4.12)		
	Pen	(4.96)	Ten	(5.37)		
	Pearl	(3.69)			Curl	(4.06)
	Pea	(4.45)			Key	(4.59)
			Tape	(4.86)	Cape	(4.27)
			Tea	(5.31)	Key	(4.59)
Coda	Map	(4.84)	Mat	(4.37)		
	Cup	(5.04)	Cut	(5.28)		
	Cape	(4.27)			Cake	(5.46)
	Lip	(3.66)			Lick	(4.16)
			Bite	(4.55)	Bike	(4.88)
			Bat	(4.74)	Back	(6.24)

Apple iPad. This is to ensure some consistency with the display and audio presentation. For participants who did not have an iPad, one was sent to them in the mail with a return-addressed envelope. Parents were asked to conduct the experiment in a quiet room, with no other electronic devices connected to the internet during the task. Before the experiment began, the parents were asked to download two apps onto their iPhones (each iPad package also included an iPod with both apps already installed). The first app (i.e., *ListenApp for Schools*, NSW Department of Education, 2016) was used to ensure that the background noise in the room was comparable to standards for a standard Australian classroom.[2] If the background noise is acceptable, the app shows a green light. Parents then started the Gorilla

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

experiment using a link emailed to them or by clicking an app on the University-provided iPads. The experiment began by playing a continuous pink noise. Parents were instructed to use the second app (i.e., *NIOSH Sound Level Meter*, EB LAB, 2017) to measure the sound output of the iPad and to change the volume accordingly to ensure that the output level for the pink noise was at 70 dB (+/−1 dB), which is just above conversational level.

The experiment then began with 2 practice trials, which were not analyzed but identical in presentation to the test trials. Each trial began with a set of four pictures. After 1 second, the audio stimuli was played. Once the participant selected a picture, the experiment progressed to the next trial. Each trial contained four yoked minimal-pair words: two with initial contrasts (e.g., Bin-Pin) and two with final contrasts (e.g., *Seed-Seat*). The experiment was blocked, and children either received the 1-talker or 6-talker block first. The order was counterbalanced across children. Within each talker version, voicing and place contrasts were randomized. In each version, children were presented with each target six times, by the same talker in the 1-talker block and by six different talkers in the 6-talker block (including the talker from the 1-talker block). All children received both conditions. A further counterbalance was implemented so that, for each contrast, half the children saw the initial contrast in the top two pictures and half in the bottom two pictures, and for half the children, one pair was the target contrast, while the other half received the other pair as the target contrast.

Results

To test the word identification accuracy of the two child groups when listening to single vs. different talkers, two generalized linear models with a logit link were constructed using the Jamovi GAMLj module and R (Gallucci, 2019; R Core Team, 2021; The Jamovi project, 2021). The *first model* compared the performance of children with typical hearing vs. those using hearing devices in identifying words contrasting in Voicing (voiced vs. voiceless) in two different word Positions (initial vs. final) and spoken by 1-talker vs. 6-talkers. Hearing Group, Talker condition, Voicing, and Position were entered as fixed effects. Child and Item were entered as random effects variables with random intercepts, with children allowed to vary by random slopes for Group, Talker, Voicing, and Position. The *second model* compared the group performance on minimal pair words with three Place contrasts (bilabial-alveolar, bilabial-velar, and alveolar-velar entered as a repeated comparison) in two Positions (word initial vs. final) and spoken by 1-talker vs. 6-talkers. Child and Item were entered with random intercepts. In *both models*, Age and Word frequency (Zipf) were entered as covariates, and all interaction terms (up to four way) were included. More maximal models did not converge. Alpha was set at 0.05 with Bonferroni corrections to *p*-value for multiple comparisons. No overdispersion was reported for either model (close to 1).

The results of the first model (see Table 4 and supplementary material 1) on Voicing contrasts detected a significant main effect of Position only, with no other significant effects. This suggests that all children's word identification was significantly worse when voicing contrasts occurred word-finally (M = 87%, SD = 34%)

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

Table 4. Fixed effects parameter estimates for normal hearing (NH) and child with hearing loss (HL) in single vs. different talker conditions with voiced vs. voiceless contrasts in the word-initial vs. word-final positions. Significant results are in bold

				95% Exp(B) CI				
Names	**Estimate**	**SE**	**exp(B)**	**Lower**	**Upper**	***z***	***p***	
(Intercept)	3.222	0.297	25.079	14.013	44.883	10.850	<0.001	
Age (years)	**0.998**	**0.148**	**2.712**	**2.030**	**3.624**	**6.747**	**<0.001**	*******
Word freq	0.183	0.355	1.200	0.599	2.407	0.514	0.607	
Group	−0.370	0.503	0.690	0.258	1.850	−0.737	0.461	
Talker	0.067	0.248	1.069	0.657	1.740	0.270	0.787	
Voicing	0.313	0.343	1.367	0.698	2.677	0.912	0.362	
Position	**−0.890**	**0.393**	**0.411**	**0.190**	**0.887**	**−2.265**	**0.024**	*
Group × Talker	−0.304	0.496	0.738	0.279	1.951	−0.613	0.540	
Group × Voicing	0.073	0.375	1.075	0.516	2.242	0.194	0.846	
Talker × Voicing	−0.081	0.267	0.922	0.546	1.555	−0.305	0.760	
Group × Position	0.165	0.397	1.179	0.542	2.568	0.416	0.678	
Talker × Position	0.070	0.272	1.072	0.630	1.826	0.257	0.797	
Voicing × Position	0.418	0.626	1.518	0.445	5.180	0.667	0.505	
Group × Talker × Voicing	−0.793	0.504	0.452	0.169	1.214	−1.575	0.115	
Group × Talker × Position	−0.027	0.507	0.974	0.361	2.629	−0.052	0.958	
Group × Voicing × Position	−0.273	0.500	0.761	0.286	2.029	−0.546	0.585	
Talker × Voicing × Position	−0.424	0.500	0.654	0.245	1.745	−0.848	0.397	
Group × Talker × Voicing × Position	−0.323	0.983	0.724	0.106	4.966	−0.329	0.742	

R-Code: Response ~ Word_freq + Age_yrs + Group × Talker × Voicing × Position + (1 + Group + Talker + Voicing + Position | Child) + (1 | Item).
* $p < 0.05$, *** $p < 0.001$.

than when they occurred word-initially (M = 92%, SD = 28%). The model also detected a significant age effect, indicating that children's performance improved with age in both groups. Importantly, no Group differences or effects of Talker number were detected. Therefore, we did not detect any detrimental effects of listening to different talkers on children's word identification accuracy, nor did we detect differences in children with hearing devices compared to their typical-hearing peers.

The second model (results see Table 5) on Place contrasts detected three significant main effects of Talker, word Position, and Place of articulation, and a significant Talker by Position interaction. The model also detected a significant age effect, indicating that children's performance improved with age in both groups. However, no significant Group or other higher-level interactions were detected (see Table 5).

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

Table 5. Fixed effects parameter estimates for normal hearing (NH) and children with hearing loss (HL) in single vs. different talker conditions with place contrasts in the word-initial vs. word-final positions

				95% Exp(B) CI				
Names	**Estimate**	**SE**	**exp(B)**	**Lower**	**Upper**	**z**	**p**	
(Intercept)	2.933	0.271	18.789	11.039	31.980	10.810	<0.001	
Age (years)	**0.572**	**0.175**	**1.771**	**1.256**	**2.497**	**3.261**	**0.001**	***
Word freq	0.434	0.236	1.543	0.972	2.452	1.838	0.066	
Group	0.425	0.486	1.530	0.590	3.966	0.875	0.381	
Talkers	**−0.218**	**0.099**	**0.804**	**0.663**	**0.975**	**−2.216**	**0.027**	**
Position	**0.445**	**0.147**	**1.561**	**1.169**	**2.083**	**3.019**	**0.003**	**
Place1 (Bilabial-Alveolar)	0.171	0.302	1.187	0.657	2.143	0.568	0.570	
Place2 (Bilabial-Velar)	**0.683**	**0.244**	**1.980**	**1.227**	**3.197**	**2.796**	**0.005**	**
Group × Talkers	0.244	0.194	1.277	0.873	1.866	1.260	0.208	
Group × Position	−0.212	0.136	0.809	0.619	1.056	−1.559	0.119	
Talkers × Position	**0.321**	**0.139**	**1.379**	**1.049**	**1.812**	**2.306**	**0.021**	*
Group × Place1	0.260	0.226	1.296	0.833	2.017	1.150	0.250	
Group × Place2	0.078	0.229	1.081	0.690	1.694	0.342	0.733	
Talkers × Place1	−0.071	0.230	0.931	0.594	1.461	−0.309	0.757	
Talkers × Place2	0.343	0.236	1.410	0.888	2.238	1.455	0.146	
Position × Place1	−0.607	0.395	0.545	0.251	1.181	−1.539	0.124	
Position × Place2	0.033	0.424	1.033	0.450	2.372	0.078	0.938	
Group × Talkers × Position	0.064	0.274	1.066	0.623	1.825	0.235	0.814	
Group × Talkers × Place1	0.393	0.453	1.481	0.609	3.601	0.867	0.386	
Group × Talkers × Place2	−0.289	0.462	0.749	0.303	1.854	−0.625	0.532	
Group × Position × Place1	0.149	0.319	1.161	0.622	2.167	0.468	0.640	
Group × Position × Place2	0.065	0.323	1.067	0.567	2.008	0.201	0.841	
Talkers × Position × Place1	−0.241	0.326	0.786	0.415	1.490	−0.737	0.461	
Talkers × Position × Place2	−0.365	0.330	0.695	0.364	1.327	−1.103	0.270	
Group × Talkers × Position × Place1	−0.382	0.644	0.682	0.193	2.409	−0.594	0.552	
Group × Talkers × Position × Place2	−0.412	0.650	0.662	0.185	2.365	−0.635	0.526	

R-code: Response ~ WordFeq + Age_yrs + Group × Speakers × Position × Place + (1 | ID) + (1 | Item).
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Post hoc analyses with Bonferroni corrections to *p*-values conducted for the Talker by word Position interaction (see Table 6) showed that children's word identification performance was likely to be worst when listening to 6-talkers where place contrasts occurred word finally ($M = 84\%$, SD = 36%) as compared to the

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

Table 6. Post hoc analysis with Bonferroni adjustment to p-values for talker (1 vs. 6) by position in word (initial vs. final), with significant results in bold

Speakers	Position		Speakers	Position	exp(B)	SE	z	$p_{bonferroni}$	
1	Final	–	1	Initial	0.669	0.156	−1.722	0.510	
1	**Final**	**–**	**6**	**Final**	**1.561**	**0.195**	**3.567**	**0.002**	**
1	Final	–	6	Initial	0.663	0.153	−1.780	0.450	
1	**Initial**	**–**	**6**	**Final**	**2.334**	**0.537**	**3.684**	**0.001**	***
1	Initial	–	6	Initial	0.991	0.151	−0.059	1.000	
6	**Final**	**–**	**6**	**Initial**	**0.425**	**0.097**	**−3.764**	**0.001**	***

**$p < 0.01$, ***$p < 0.001$.

Table 7. Post hoc analysis with Bonferroni adjustment to p-values for place (bilabial, alveolar, and velar), with significant results in bold

Place			Place		exp(B)	SE	*z*	$p_{bonferroni}$	
1	Bilabial	–	2	Alveolar	0.843	0.254	−0.568	1.000	
1	**Bilabial**	**–**	**3**	**Velar**	**0.505**	**0.123**	**−2.795**	**0.016**	**
2	Velar	–	3	Alveolar	0.599	0.171	−1.797	0.217	

*$p < 0.05$.

other three conditions: 6-talkers with place contrasts occurring word initially ($M = 91\%$, SD = 29%) and 1-talker with place contrasts occurring word initially ($M = 91\%$, SD = 28%) or finally ($M = 89\%$, SD = 31%), which were not significantly different from each other.

Post hoc analysis on place of articulation contrasts (see Table 7) revealed that Children's word identification was worse in words starting with Bilabial ($M = 86\%$, SD = 35%) than Velar targets ($M = 92\%$, SD = 27%), but neither was significantly different to alveolar targets ($M = 89\%$, SD = 31%).

Discussion

We examined whether the accuracy of word identification by young children using hearing devices could be improved by listening to different talkers for voicing and place contrasts across word positions. First, our results show that all children performed well in word identification using familiar words (90% for children with typical hearing and 88% for children using hearing devices). Across all analyses, we did not detect any group differences, suggesting that despite the distortion and poverty of acoustic information delivered through hearing devices, children with aided hearing are sensitive to acoustic cues associated with different talkers, and for those who have received early intervention, they are not more affected than their TH peers.

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

In terms of the two types of contrasts that were tested, the results on *voicing contrasts* showed no evidence that listening to many different talkers affected word identification. However, we found a word-position effect: children's word identification was worse for all targets in the word-final position than in the word-initial position. Children's performance was lowest for *place contrasts* when listening to different talkers making contrasts in the word-final position than in the word-initial position. Children were more accurate when listening to a single talker, regardless of whether place contrasts occurred in the word-initial or word-final positions. This suggests that the word-final position is very vulnerable to misunderstandings for young children, more so than phonetic variation from different talkers. The challenge with the word-final position could reflect earlier acquisition of onsets than codas (Bruggeman et al., 2021). This has practical and clinical implications for children using hearing devices—a point we come back to below.

We also found that children's word identification accuracy varied across the place of articulation targets, with the best performance observed for velar targets and the worst for bilabial targets. This variation can be explained by voice onset time (VOT) differences, with velar stops having the longest VOT, followed by alveolar and bilabial stops (Millasseau et al., 2021), suggesting that longer VOTs may provide more robust and reliable place-of-articulation cues across talkers. Bilabial stops also tend to have less acoustically salient burst release (lower amplitude) compared to the other two stops (Repp, 1984). Therefore, visual information may be more useful, especially for bilabials, and when phonetic information is less reliable. This has implications for clinical intervention, and future investigations should examine whether adding visual information reduces ambiguity and improves word-identification performance.

These results suggest that while listening to different talkers can be challenging, the word-final position is consistently challenging and most vulnerable to misunderstandings. In English, fricatives and stops (most affected by hearing loss) are often produced with greater variation in word-final position, leading to greater uncertainty in speech perception for young children still learning the phonetic variations of English. These findings add to our understanding of the challenges children with hearing devices face in speech perception. Our results also provide empirical evidence for the listening hierarchy used in clinical practice (Erber, 1982; Estabrooks et al., 2020), which proposes that place contrasts in word-final positions are the hardest contrast to master for young children with hearing loss. This finding has important implications for supporting current clinical practice and those working/caring for children using hearing devices, such as teachers and parents. Perceiving the ends of words and listening to different talkers could be targeted as key areas for early intervention and communication counselling. Our results also suggest that the reduction in performance for familiar words is small, indicating the benefits of a larger vocabulary (more familiar words) and that early intervention targeting this may provide a protective effect for children using hearing devices.

Our findings support and differ from past research in several ways. First, we found an effect of different talkers on place-of-articulation contrasts only, not on voicing contrasts, and only in the word-final position. Second, we used words known to preschoolers in a 4-alternative-choice task similar to the PPVT. Although Ryall and Pisoni (1997) used a similar task, they presented a more complex 6-

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

picture display with high-frequency words to adults but not to children. This more challenging task may have boosted the effect of different talkers. However, both our findings and those of Ryall and Pisoni (1997) contrast with Kirk et al. (2000), which showed facilitation in word identification when listening to different talkers. Key differences between our study and Kirk et al. (2000) were that we had a larger sample of children (20 vs. 14), our children were younger (4;11 vs. 8), had earlier implantation/fitting (most under 2 years vs. 5 years), and we had a control group of children with typical hearing. It is unclear how typically hearing children would have performed on the Kirk et al. (2000) task. In addition, the linguistic manipulations also differ. Our study examined minimal-pair words familiar to young children, whereas Kirk et al. (2000) examined the effects of word frequency and neighborhood density. Word frequency (familiarity) and acoustic similarity may be two dimensions of word representation that are impacted differently by talker variability. And in terms of acoustic similarity, children may also struggle to perceptually normalize two sources of large acoustic variation simultaneously: talker (indexical) and phonemic (linguistic) information. For example, coda stops are highly variable in production, which, in our study, led to generally poorer identification accuracy for word-final minimal pairs. But some contrasts are more vulnerable to talker variation, and we observe that place contrasts when spoken by different talkers are more prone to identification inaccuracies when occurring in the word-final position. But this needs to be addressed by future studies comparing different contrasts across word positions, with single and multiple talkers, in familiar and unfamiliar words.

Our study also detected a significant age effect, suggesting that, as they aged, all children performed better at word identification, regardless of whether they listened to a single or multiple talkers. What we cannot tease apart in this study are developmental versus vocabulary effects, as they are likely very highly correlated. While we used words that are familiar to young children, familiarity, i.e., input and use, is likely to increase with age. Our study further suggests that, for both children with and without aided hearing, phonemic and indexical information are represented lexically and that listening to different talkers has similar effects on word identification in both groups. Together, these findings further suggest that engaging with a range of people should not hinder vocabulary development in children with aided hearing, provided they have a sufficient vocabulary. While it is not possible in clinical settings to engage multiple clinicians during intervention sessions, our results suggest that supporting children with aided hearing to develop a larger vocabulary (a broader range of familiar words) might improve word comprehension when listening to different speakers. Clinicians should also encourage children with aided hearing to interact more with diverse family, extended family, and friends to increase exposure to different talkers before they enter school. However, given our sample, these findings may be most relevant for children who have received hearing devices and early oral intervention.

Limitations of the study

First, while our sample reflects the heterogeneous populations of children with hearing devices, we could not evaluate the different effects of device type and degree

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

of hearing loss on performance. Poorer spectral information delivered through HA (compared to CIs) may modify the degree and type of impact on performance when listening to different talkers. This should be addressed in future studies.

Second, past studies with different talkers have all used a balanced set of male and female talkers, whereas our study used only female talkers with experience working with children, so the stimuli were spoken in a familiar register (but at a higher pitch) to young children. It is unclear whether there are different effects for male and female talkers in Ryall and Pisoni (1997) and Kirk et al. (2000). This should be investigated in future studies, because while adults are not affected by the sex of the talkers (Magnuson & Nusbaum, 2007), young children might be. Our study examined only female talkers, as young children with aided hearing are often exposed to female voices during early education and clinical intervention. This is very different to later academic environments, and as they mature, they will increasingly encounter more male speakers. A better understanding of talker effects on word identification could help inform education and clinical practice.

Third, we could not monitor children's reaction times. Given that adults show slower reaction times while listening to different talkers (Magnuson & Nusbaum, 2007), it is reasonable to expect the same to be true for children. This is especially important for children using hearing devices because the accumulation of slower processing time over longer stretches of discourse can lead to fatigue, miscommunication, and poorer learning outcomes. Also, listening to recorded speech might pose greater challenges for children with hearing devices than listening to live speech, in both accuracy and processing time. Both should be considered in future studies.

Finally, to increase the generalizability of the findings to more varied learning and social situations, the effect of word identification with low-frequency words or newly learned/taught words across different developmental stages should be investigated.

Conclusion

In partial support of past research on typically hearing children, word identification was found to be worse when listening to different talkers, though only for place contrasts occurring word finally. More consistent for both voicing and place contrasts is the finding that word identification was worse for all contrasts occurring word finally (compared to those occurring word initially). Together, these findings suggest that the word-final position is more vulnerable to misunderstandings. No overall group differences in word identification accuracy were found between children with typical hearing and those using hearing devices. Both groups showed good performance, with different talkers and contrasts in the word-final position contributing to small reductions in performance. This suggests that despite having very different linguistic experiences and later access to spoken language, children using hearing devices can perform like their typical hearing peers when dealing with greater acoustic variability for *familiar* words in *good* listening environments. However, when listening to newly learned or low-frequency words, word position and the number of talkers are likely to pose comprehension challenges.

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0142716426100472.

Data availability statement. All deidentified data, R code, and experimental materials are available below for blind review. (https://osf.io/e235z/?view_only = 4bf858ba9df9492598c4ef312873cce6). Below is a description of where these will permanently sit post-review, in accordance with our ethics protocol. While the deidentified datasets, R code, and study materials are available from the first author's institutional repository, the hearing profiles of children using hearing devices are not publicly available due to small populations with specific combinations of clinical characteristics that could make some children easily identifiable. However, these data can be made available by the corresponding author on reasonable request after consultation with the second and third authors and their centers.

Acknowledgements. This study was funded under the ARC Linkage Project LP180100534 awarded to the last two authors, with industry partners represented by the second and third authors. We thank Inge Kaltenbrunn for her contributions throughout this project, including recruitment, clinical data support, and input on the stimuli selection. We thank our many speakers: Elise Tobin, Elsa Whelan, Dr Isabel O'Keeffe, Linda Buckley, Dr Louise Ratko, and Dr Rebecca Holt, as well as our many student interns: Olivia Cicco, Rachel Jackson, and Neve Ward for stimuli development; Freja Warner van Dijk and Andrew Plant for experiment setup in Gorilla and piloting.

Notes

1 Zipf is a standardized measure allowing for comparisons of word frequencies reported across different corpora. The measure is a straightforward unit that separates low-frequency from high-frequency words (Van Heuven et al., 2014).

2 The suggested values for good rating are unoccupied ANL < 35 dBA, occupied BNL < 50 dBA, SNR > +15 dB, RT < 0.4s (unoccupied), and STI > 0.75.

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: an online behavioral experiment builder. *Behavior Research Methods*, **52**(1), 388–407.

Apfelbaum, K. S., & McMurray, B. (2011). Using variability to guide dimensional weighting: associative mechanisms in early word learning. *Cognitive Science*, **35**(6), 1105–1138.

Baudonck, N., Lierde, K. V., D'haeseleer, E., & Dhooge, I. (2011). A comparison of the perceptual evaluation of speech production between bilaterally implanted children, unilaterally implanted children, children using hearing aids, and normal-hearing children. *International Journal of Audiology*, **50**(12), 912–919.

Boersma, Paul & Weenink, David (2021). Praat: Doing Phonetics by Computer [Computer program]. Version 6.2.01, retrieved 17 November 2021 from https://www.praat.org/.

Bruggeman, L., Millasseau, J., Yuen, I., & Demuth, K. (2021). The acquisition of acoustic cues to onset and coda voicing contrasts by preschoolers with hearing loss. *Journal of Speech, Language, and Hearing Research*, **64**(12), 4631–4648.

Chang, Y. P., Chang, R. Y., Lin, C. Y., & Luo, X. (2016). Mandarin tone and vowel recognition in cochlear implant users: effects of talker variability and bimodal hearing. *Ear and Hearing*, **37**(3), 271.

Chang, Y. P., & Fu, Q. J. (2006). Effects of talker variability on vowel recognition in cochlear implants. *Journal of Speech, Language & Hearing Research*, **49**(6), 1331–1341.

Cleary, M., & Pisoni, D. B. (2002). Talker discrimination by prelingually deaf children with cochlear implants: preliminary results. *Annals of Otology, Rhinology & Laryngology*, **111**(5_suppl), 113–118.

Colby, S., & Orena, A. J. (2022). Recognising voices through a cochlear implant: a systematic review of voice perception, talker discrimination, and talker identification. *Journal of Speech, Language, and Hearing Research*, **65**(8), 3165–3194.

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**(2), 633–664.

Dailey, S., & Bergelson, E. (2022). Language input to infants of different socioeconomic statuses: a quantitative meta-analysis. *Developmental Science*, **25**(3), e13192.

EA LAB (2017). NIOSH sound level meter [Mobile app]. https://apps.apple.com/us/app/niosh-sound-level-meter/id1096545820.

Elfenbein, J. L., Hardin-Jones, M. A., & Davis, J. M. (1994). Oral communication skills of children who are hard of hearing. *Journal of Speech and Hearing Research*, **37**(1), 216–226.

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences*, **8**(7), 301–306.

Erber, N. P. (1982). *Auditory training/Norman P. Erber*. Alexander Graham Bell Association for the Deaf.

Estabrooks, W., Morrison, H. M., & MacIver-Lux, K. (2020). *Auditory-verbal therapy: science, research, and practice*. Plural Publishing, Inc.

Gallucci, M. (2019). GAMLj: General Analyses for Linear Models. [jamovi module]. Retrieved from https://gamlj.github.io/.

Hay, J., & Drager, K. (2007). Sociophonetics. *Annual Review of Anthropology*, **36**(1), 89–103.

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.

Kirk, K. I., Hay-McCutcheon, M., Sehgal, S. T., & Miyamoto, R. T. (2000). Speech perception in children with cochlear implants: effects of lexical difficulty, talker variability, and word length. *Annals of Otology, Rhinology & Laryngology*, **109**(12_suppl), 79–81.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English/r/and/l/. II: the role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America*, **94**(3), 1242–1255.

Magnuson, J. S., & Nusbaum, H. C. (2007). Acoustic differences, listener expectations, and the perceptual accommodation of talker variability. *Journal of Experimental Psychology: Human Perception and Performance*, **33**(2), 391.

Magnuson, J. S., Nusbaum, H. C., Akahane-Yamada, R., & Saltzman, D. (2021). Talker familiarity and the accommodation of talker variability. *Attention, Perception, & Psychophysics*, **83**(4), 1842–1860.

Millasseau, J., Bruggeman, L., Yuen, I., & Demuth, K. (2021). Temporal cues to onset voicing contrasts in Australian English-speaking children. *The Journal of the Acoustical Society of America*, **149**(1), 348–356.

Mulak, K. E., & Best, C. T. (2013). Development of Word Recognition Across Speakers and Accents. In L. Gogate, & G. Hollich (Eds.), *Theoretical and computational models of word learning: trends in psychology and artificial intelligence* (pp. 242–269). IGI Global.

NSW Department of Education (2016). ListenApp for Schools (v 1.09) [Mobile app]. https://apps.apple.com/au/app/listenapp-for-schools/id981300043.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, **20**(2), 421.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. (Version 4.0) [Computer software]. Retrieved from https://cran.r-project.org. (R packages retrieved from MRAN snapshot 2021-04-01).

Repp, B. H. (1984). Closure duration and release burst amplitude cues to stop consonant manner and place of articulation. *Language and Speech*, **27**(3), 245–254.

Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, **12**(2), 339–349.

Rost, G. C., & McMurray, B. (2010). Finding the signal by adding noise: the role of noncontrastive phonetic variability in early word learning. *Infancy*, **15**(6), 608–635.

Ryalls, B. O., & Pisoni, D. B. (1997). The effect of talker variability on word recognition in preschool children. *Developmental Psychology*, **33**(3), 441.

Sadakata, M., & McQueen, J. M. (2013). High stimulus variability in non-native speech learning supports formation of abstract categories: evidence from Japanese geminates. *The Journal of the Acoustical Society of America*, **134**(2), 1324–1335.

The Jamovi project (2021). jamovi. (Version 1.8) [Computer Software]. Retrieved from https://www.jamovi.org.

Van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: a new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, **67**, 1176–1190.

https://doi.org/10.1017/S0142716426100472 Published online by Cambridge University Press

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental psychology*, **34**(6), 1289.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: initial capabilities and developmental change. *Developmental Psychology*, **24**(5), 672.

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.

Zhang, X., Cheng, B., & Zhang, Y. (2021). The role of talker variability in non-native phonetic learning: a systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research*, **64**(12), 4802–4825.

Cite this article: Xu Rattanasone, N., Neal, K., Leigh, G., Sharma, M., & Demuth, K. (2026). Listening to different talkers and its effect on word identification for children using hearing devices. *Applied Psycholinguistics*. https://doi.org/10.1017/S0142716426100472