JASA Express Letters

MAY 01 2026

Auditory acuity to vowels is more refined in adulthood compared to childhood

Katharina M. Polsterer; Nikki Hoekzema; Thomas B. Tienkamp; Teja Rebernik; Hedwig G. Sekeres; Reihaneh Amooie; Raoul S. S. J. Buurke; Wietse de Vries; Frank Tsiwah; Martijn B. Wieling; Defne Abur

JASA Express Lett. 6, 055201 (2026)
https://doi.org/10.1121/10.0043710

JASA EXPRESS LETTERS ARTICLE asa.scitation.org/journal/jel

Auditory acuity to vowels is more refined in adulthood compared to childhood

Katharina M. Polsterer,[1,2,3,a] Nikki Hoekzema,[1,b] Thomas B. Tienkamp,[4,c] Teja Rebernik,[5,6,d] Hedwig G. Sekeres,[1,e] Reihaneh Amooie,[1,f] Raoul S. S. J. Buurke,[1,g] Wietse de Vries,[1,h] Frank Tsiwah,[1,i] Martijn B. Wieling,[1,2,6,j] and Defne Abur[1,2,k]

[1]*Department of Computational Linguistics, Center for Language and Cognition, University of Groningen, Groningen, The Netherlands*
[2]*Research School of Behavioral and Cognitive Neurosciences, University of Groningen, Groningen, The Netherlands*
[3]*Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria*
[4]*Department of Neurology, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany*
[5]*Laboratoire de Phonétique et Phonologie, CNRS & Sorbonne Nouvelle, Paris, France*
[6]*Brussels Centre for Language Studies, Vrije Universiteit Brussel, Brussels, Belgium*

Abstract: Auditory-perceptual sensitivity to acoustic variation, referred to as auditory acuity, plays a central role in children's speech development. In speech perception, auditory acuity determines the acceptable phonetic variation of a given speech sound. In speech production, auditory acuity determines whether a speech sound is produced as intended and guides error-based motor learning. This study investigated how auditory acuity differs between children (6–11 years old) and adults (20–36 years old). Participants discriminated between recordings of their own speech modified in the first vowel formant (F_1). On average, children showed lower F_1 acuity but similar interindividual variability compared to adults.

[Editor: Yolanda F. Holt] https://doi.org/10.1121/10.0043710

Received: 25 September 2025 **Accepted:** 14 April 2026 **Published Online:** 1 May 2026

1. Introduction

Speech perception is foundational to acquiring a spoken language. Yet, speech perception itself undergoes development,[1,2] including auditory-perceptual sensitivity to acoustic structures, referred to as *auditory acuity*. The more sensitive an individual is to fine-grained acoustic variation, the better their auditory acuity. Auditory acuity is relevant for both speech perception and production development. The current study aimed to characterize how children differ from adults in auditory acuity.

Early in development, infants can discriminate a wide range of phonetic contrasts[1] suggesting high auditory acuity. By six to eight months of age, perception begins to attune to language-specific phonemic categories,[3] although infants remain sensitive to within-category differences.[4–6] The developmental reorganization of perception from global acoustic sensitivity to the formation of language-specific categories (called perceptual attunement[4]) has been observed to continue to the age of 10 to 12 months.[4,7] Language-specific auditory-perceptual representations, therefore, appear to begin in infancy.

Neural and behavioural evidence suggests that speech perception continues to develop well into childhood.[2] Neural pathways supporting auditory perception, particularly for spectral and temporal features, change throughout childhood and in adolescence.[2] Behaviourally, children showed reduced acuity to non-speech[8,9] and speech[10–16] stimuli, often improving gradually with age.[8,10–15] This developmental refinement can occur both between phonemic categories,

[a]Corresponding author: k.m.polsterer@rug.nl
[b]Email: n.hoekzema@rug.nl
[c]Email: thomas.tienkamp@uk-koeln.de
[d]Email: teja.rebernik@cnrs.fr
[e]Email: h.g.sekeres@rug.nl
[f]Email: r.amooie@rug.nl
[g]Email: raoul.buurke@rug.nl
[h]Email: wietse.de.vries@rug.nl
[i]Email: f.tsiwah@rug.nl
[j]Email: m.b.wieling@rug.nl
[k]Email: d.abur@rug.nl

reflected in sharper category boundaries,[10,11,17] and within categories, reflected in increased sensitivity to fine-grained acoustic differences that do not change phonemic identity.[12,13,18] Perception of both between and within-category differences has been studied via discrimination tasks, which require listeners to distinguish two auditory stimuli based on acoustic features.[19,20] However, the extent to which discrimination tasks are informative about pre-categorical auditory perception has been questioned.[12,13] Identification tasks, on the other hand, require listeners to assign auditory stimuli to a category, and have therefore provided insights into categorical boundaries.[19,20] Although speech perception continues to develop into childhood, the interpretation of research findings hence depends on the specific task used.[11,14,15]

Age-related improvements in auditory acuity appear to depend on the percept being investigated. Across different sound categories, particularly consonant aspiration and vowel formants, identification appeared to develop earlier than discrimination.[14] In a plosive identification task, no difference between younger children (from five to six years), older children (from nine to ten years) and adults (mean age of about 25 years) was found.[15] The same groups of speakers differed significantly in a plosive discrimination task, with older children discriminating plosives better than younger children, and less well than adults.[15] Longitudinally (225 children aged 6 to 12 years over the course of the data collection), plosive discrimination became more gradient with age (i.e., shallower categorization slopes) when using a gradient response system, as acuity to fine-grained differences improved.[12] Identification of the semi-vowels /ɹ/ and /w/ (i.e., width of the perceptual category boundaries) was negatively correlated with age (9–14 years),[17] whereas discrimination of the same phonemes did not correlate with age (9–15 years).[11] When discriminating externally produced vocal pitch, 55% of the tested children (6–11 years old) were classified by the authors as having adult-like acuity [i.e., within two standard deviations (SD) of the adult participants' average acuity; adult age: 18–28 years].[16] Alternatively, children might form a separate distribution with lower mean acuity than adults, with the "adult-like" 55% representing those with relatively higher acuity. Children classified with adult-like acuity to vocal pitch did not significantly differ in age from those with lower acuity.[16] Vowel perception, specifically, primarily relies on time-varying spectral cues,[21,22] most notably the first two formants (F_1 and F_2).[21] Converging evidence suggests that children differ from adults in how they weight and integrate spectral cues when perceiving vowels, indicating continued development throughout childhood.[23–25] Other work suggested adult-like identification of vowels at four years of age and between-category discrimination at six years of age.[14] Within-category vowel discrimination may not yet be adult-like in ten-year-olds (adult age: 20–30 years).[14] Taken together, results are mixed with respect to whether and how auditory acuity improves systematically with age, warranting further investigation of its development in childhood.

The literature reviewed so far focuses on children's perception of speech produced by others, whereas the perception of self-produced speech has not been discussed yet. Perceiving self-produced speech constitutes a critical link between perception and production in theoretical models of speech, where auditory perception contributes to speech motor control and learning.[26,27] Within the Directions Into Velocities of Articulators (DIVA) framework,[27,28] auditory feedback (i.e., hearing ourselves speak) is used to establish time-varying acoustic regions, known as auditory targets, which define acceptable ranges of variation for speech sounds and provide a reference for error-based learning of speech motor representations. Other models differ in their implementation of auditory feedback but converge on a similar functional role. In the State Feedback Control (SFC) model,[29] auditory feedback is integrated with information about the current and predicted state of the speech system (i.e., positions and velocities of the articulators) and contributes to updating internal predictions of the system's state. In task dynamics-based accounts,[30] speech is organized around gestural (i.e., articulatory-motor) tasks, whereas auditory feedback plays a secondary, supportive role in informing learning and helping detect errors. Despite architectural differences, these frameworks share the view that auditory acuity to self-produced speech constrains how precisely speech errors can be detected via auditory feedback and how speech motor representations are learned and refined.

The development of auditory acuity and its relation to speech motor learning during childhood have been supported by empirical studies. For example, it was found that children aged three to nine years adapt their vowel productions in response to experimentally shifted auditory feedback, to a similar extent as (younger) adults (17–22 years)[31] or to a greater degree than (slightly older) adults (18–29 years).[32] Larger adaptation of vowel productions in response to an auditory feedback shift may reflect an overcompensation for speech errors in children due to less specifically formed auditory targets (i.e., lower auditory acuity) compared to adults.[32] Additionally, the degree of adaptation to F_1 changes in the auditory feedback was found to be enhanced by auditory-perceptual training to improve auditory acuity to F_1 in five to seven-year-olds.[33] Reduced auditory acuity has been linked to delayed speech production development.[34] Other work has found that children (aged 9–14 years) with acoustically more distinct productions of speech sounds also had better auditory acuity compared to those with less distinct productions, suggesting that perception and production are also developing together.[17] Together, these findings indicate that acuity to auditory feedback of self-produced speech supports speech production development during childhood.

Auditory acuity is commonly assessed via just-noticeable differences (JNDs), which estimate the smallest detectable change in a given acoustic dimension.[11,16,35,36] Listeners compare multiple pairs of stimuli differing in a target parameter to determine the threshold at which they can no longer distinguish the stimuli. JNDs have been used to assess auditory acuity to externally produced speech features, including vocal pitch[16] and formants determining consonantal distinctions.[11] Prior research on auditory acuity has focused on externally generated sounds, rather than *self-produced* speech, in childhood. Self-produced speech may be recognized differently in the perceptual system, as it relates to speech motor control and learning, and this therefore warrants further investigation.[26] The current study thus assessed auditory acuity to self-produced vowels (via JNDs to F_1) by (1) comparing children and adults, (2) characterizing the development within childhood, and (3) evaluating variability across both groups. Given prior work that found higher auditory acuity in adults than children for consonants,[11] and continued refinement of vowel perception through ten years of age,[14] we predicted that adults would exhibit higher auditory acuity to vowel F_1 (i.e., smaller JNDs) compared to children. Motivated by findings that sensitivity to fine-grained acoustic variation increases between six and twelve years of age,[12,17] we explored whether JNDs would decrease with age within the child group. We further predicted greater variability in JNDs within the group of children, consistent with previous findings on vocal pitch[16] and evidence that trial-to-trial variability of phoneme categorization decreases across middle childhood.[12]

2. Methods

The Research Ethics Review Committee of the Faculty of Arts of the University of Groningen approved the present study (#82182577). Prior to participation, participants read an information letter and gave written consent. For participants under 13 years of age, parents gave informed consent, while children received a child-friendly information letter. Participants above 16 years of age gave informed consent themselves. The research study complies with the Declaration of Helsinki.

2.1 Participants

We collected data from 49 children (6–11 years of age, mean = 8.5 +/− 1.4 years; 25 girls, 24 boys) and 45 adults (20–36 years of age, mean = 27.1 +/− 3.6 years; 24 women, 20 men, one non-binary). All 94 participants were native speakers of Dutch and passed a hearing screening, indicating that they could hear pure tones of 250 Hz to 4 kHz at 25 dB hearing level. Initially, participants were recruited and tested at two public events in Groningen (Netherlands): Noorderzon Festival of Performing Arts & Society and Zpannend Zernike, a science festival for children. Additional data from adults ($N = 39$) were collected at the University of Groningen to obtain similar group sizes.

2.2 Study setup

All data were collected in our mobile laboratory SPRAAKLAB, which contains a sound booth dampened by 30 dB sound pressure level (SPL).[37] The participants wore Sennheiser HD280 Pro headphones (Sennheiser, Wedemark, Germany) and a Shure MX153 (Shure, Niles, IL) over-the-ear microphone, which was positioned 7 cm away from the mouth at a 45-degree angle.[38] The microphone and headphones were attached to a Focusrite Scarlett Solo soundcard, 2nd generation (Focusrite, Wycombe, UL). Parents stayed outside of the sound booth, where their child could still see them through a window. They accompanied the child in the booth only if the child wished for this.

2.3 Experiment

Auditory acuity to F_1 was assessed via a JND as done in previous research.[35] The experiment consisted of two parts: (1) a speaking task, and (2) a listening task. Participants produced consonant-vowel-consonant (CVC) words and then listened to their own productions in a JND task. All participants underwent a practice round for both parts to make sure the instructions were clear before completing the full experiment.

In the first part, the participants produced three words, three times each: *duur* /dyɑr/ (expensive in Dutch); *deur* /dùɑr/ (door in Dutch); *daar* /daɑr/ (there in Dutch). This resulted in nine productions per participant. These stimuli differ mostly in F_1, with *duur* being produced with the lowest and *daar* being produced with the highest F_1.[39] *Duur* and *daar* were distractors, while *deur* was the target word. The participants were instructed to make the vowel longer than they usually would. All participants saw the stimuli written on a screen in front of them. The children were additionally shown drawings of each stimulus.

For the second part, a perceptual discrimination paradigm was applied to assess JNDs. To this end, F_1 was tracked using linear predictive coding (LPC). To minimize potential effects of differences in formant tracking quality,[40] LPC order was set to 17 for all male participants, and to 15 for all female and child participants. Additionally, prior values of F_1 and F_2, which are used by the formant tracking algorithm, were customized by the same groups. The individual participant's production of *deur* with the median value of F_1 was chosen. This recording was increased in F_1 and played back to the participant in pairs of two. For both tokens, F_1 was increased to exclude the effects of comparing an original production to a manipulated production. One of these two tokens served as a reference and was increased in F_1 by 1%. The second token was increased to a varying degree as follows. The manipulations were applied in an adaptive one-up, two-down staircase procedure. That is, after each pair, participants needed to indicate whether the vowels in the two auditory stimuli were the same or different by pointing their thumb up (same) or down (different). The task was started with 50% increase in F_1, meaning participants heard the reference recording followed by a shifted version with a 50% F_1 increase. The F_1 increase in the shifted token was reduced by three percent following two consecutive correct responses.

The F_1 increase was raised by three percent following an error. The task stopped after 60 trials or ten reversals. For 20% of the trials, the original vowel recording was presented twice to ensure attention to the task ("catch trials").

The JND, reflecting auditory acuity to F_1, was calculated for each participant as the percentage of difference in F_1 at which participants were still able to differentiate two vowels. Specifically, this was done by averaging over the last six reversals of the adaptive staircase paradigm described previously.[35]

2.4 Statistical analysis

To compare JND values of adults to children, a Mann-Whitney U test was run due to the data not being normally distributed. JND values were used as the dependent variable, while age group (i.e., children of 6 to 11 years of age; adults of 20 to 36 years of age) was used as the independent variable. By means of a Levene's test, a potential between-group difference in the variability of the data was tested. To investigate the development of auditory acuity across childhood, a linear regression model was fit only to the data of the children. This model had JND values as the dependent variable and age as a continuous independent variable. We tested sex as a further independent variable, but this addition did not improve the model, so it was omitted.

3. Results

The average JND was 49% F_1 difference (SD = 26% F_1 difference) in children and 33% F_1 difference (SD = 21% F_1 difference) in adults. The Mann-Whitney U test revealed significantly lower JND values in adults than children ($W = 1525$, $p = 0.001$), with a medium effect size ($r^2 = 0.11$). The Levene's test indicated that the group of children and the group of adults did not significantly differ in variability of JND values [$F(1, 92) = 1.72$, $p = 0.19$].[41] Figure 1 presents the results for both groups.

Additionally, we tested for a possible effect of age on JND for vowel F_1 within the group of children only. A linear regression model did not reveal a statistically significant relation between age and JND in the latter group ($\beta = -0.028$, $p = 0.3$) as visualized in Fig. 2.

4. Discussion

The aim of the current study was to better understand auditory acuity in childhood speech development. Specifically, we assessed how auditory acuity to vowel F_1, assessed via JNDs, (1) differs between children and adults, (2) develops within childhood, and (3) varies across both groups. As predicted, young adults (20–36 years of age) showed significantly lower JND values than children (6–11 years of age), indicating higher auditory acuity. Contrary to our expectations, there was no statistically significant relationship between JNDs and age within the child group. Finally, the two groups did not significantly differ in the variability of JNDs. Thus, the current study provides no evidence for a systematic age-related change in auditory acuity to self-produced vowels between 6 and 11 years, while it does indicate differences between adults and children.

The observed group difference in JNDs between adults and children lends support to our hypothesis that auditory acuity is more refined by adulthood than by childhood. From the perspective of speech perception, this finding is in line with previous work suggesting that childhood speech development involves increasing acuity to within-category acoustic details[12,13] as no category labelling was required. From the perspective of speech production, higher auditory acuity may imply more refined auditory targets,[17] and thus more likely internal errors due to small (noticeable) deviations of F_1.[27,28] Overall, the ability to distinguish between two tokens perceptually, even if they belong to the same auditory target category, seemed to be more refined in adults than children, which may impact cognitive representations

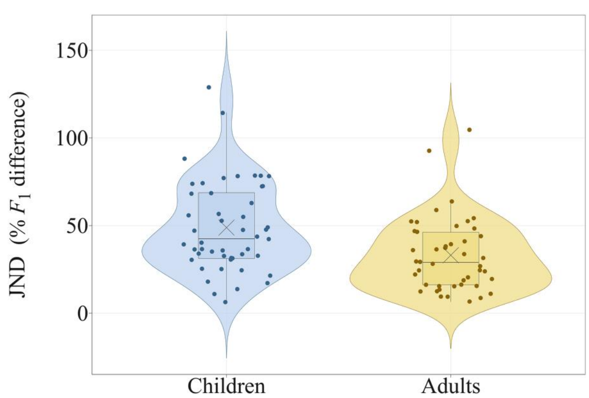

Fig. 1. JNDs for children (left; blue) compared to adults (right; yellow) with dots representing individual data.

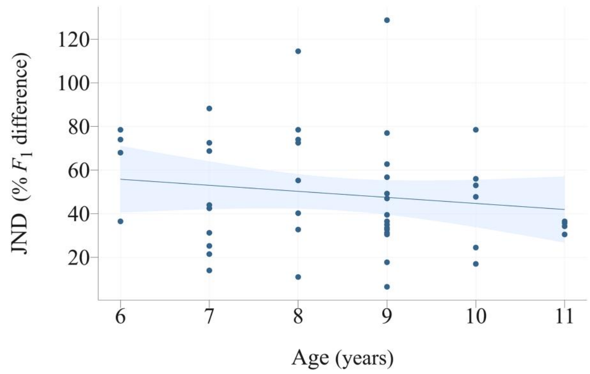

Fig. 2. JND in children across age with dots showing individual results.

(i.e., auditory targets) of speech sounds for production.[27,28] Thus, auditory acuity may continue to refine beyond childhood, likely shaping both speech perception and production.

Although adults showed overall higher acuity to self-produced F_1 than children in the current data, interindividual variability did not significantly differ between the groups. At the individual level, some children showed adult-like acuity (i.e., auditory acuity similar to that of adults), while others were less sensitive. This pattern is similar to prior findings for acuity to externally produced vocal pitch, where children (6–11 years of age) were classified as either adult-like or less sensitive relative to adults, which did not appear to be related to their age.[16] Our results thus align with those previous findings in that auditory acuity is not uniform even after early childhood. The absent age effect in the current study contrasts with other studies that found age-related increases in auditory acuity to externally produced formant-based consonant distinctions (in children aged nine to 15 years)[11] and pure tone frequencies (in children aged three to six years).[9] Similarly, perceptual acuity to externally produced vowels was found to refine during childhood, beyond the age of ten years,[14] and the continued development of vowel production is highly interrelated with auditory acuity to vowels.[33] Given that there was still a group-level difference in auditory acuity between adults and children in the current study, auditory acuity for vowels might continue to develop even later than middle childhood (i.e., between the ages of 11 and 20 years). Based on the present results, it seems possible that the development of auditory acuity to vowel formants does not follow a strictly homogenous or linear trajectory.

Future research should expand our understanding of how auditory acuity to self-produced speech develops by including younger children (i.e., less than six years of age) and adolescents. Combining auditory acuity assessments with variability in production could clarify developmental trajectories in speech perception and production. Longitudinal studies are needed to examine individual development. Individual variability could also be explored by assessing lexical knowledge, which may influence phonemic category boundaries,[42,43] or by using a more gradient response system (rather than binary same/different responses) to capture gradual changes to auditory perception, as done in recent prior work.[12] In general, it seems important to consider that measuring children's formants can be more challenging than measuring adults' formants. Children typically have higher fundamental frequencies and larger harmonic spacing than adults, making vowel spectrum peaks likely less defined.[40] While in the current study, formant settings were customized by sex, customizing them per speaker or token might help overcome this challenge.[44] Together, these suggested approaches could further clarify how auditory acuity develops and why individual differences arise.

5. Conclusion

In the current study, auditory acuity to the first vowel formant was directly compared between adults (20–36 years of age) and children (6–11 years of age) in a sample of 94 participants in total. The results showed that, on average, adults had higher auditory acuity than children, suggesting less refined auditory acuity to self-produced vowels in the latter group. Within the group of children, we found variability similar to that of adults, with individual children demonstrating more adult-like F_1 acuity and others being less sensitive to vowel F_1 distinctions. This variability within children was not explained by age. Follow-up research across age, including adolescence, is needed to determine when adult-like acuity is achieved robustly and further clarify the causes of interindividual variability.

Supplementary Material

See supplementary material for a list of the participants' median productions of *deur* in Hz.

Acknowledgments

This work was supported by a University of Groningen Behavioral and Cognitive Neurosciences Research Seed Grant Award (D.A.) and a Ph.D. in the Humanities Grant of the Dutch Research Council (NWO; K.M.P., D.A., M.B.W.). Data collection was enabled by the Noorderzon Festival of Performing Arts & Society and the science festival Zpannend Zernike held in Groningen, the Netherlands. The authors would like to especially thank all the participants who contributed to this research.

Author Declarations

Conflict of Interest

The authors state that they have no conflicts of interest to disclose.

Ethics Approval

For the current research, ethical approval was obtained at the Research Ethics Review Committee of the Faculty of Arts of the University of Groningen with the ID No. 82182577. From all participants or their caregiver in the case of minors, informed consent was obtained prior to participation in this study.

Data Availability

The data of participants who gave consent to their pseudoanonymized data being shared are openly available in the Open Science Framework: https://osf.io/z8yj5/.

References

[1] D. M. Houston, "Speech perception in infants," in *The Handbook of Speech Perception*, edited by D. B. Pisoni and R. E. Remez (John Wiley & Sons, New York, 2005), Chap. 17.
[2] J. K. Moore and F. H. Linthicum, "The human auditory system: A timeline of development," Int. J. Audiol. **46**, 460–478 (2007).
[3] P. K. Kuhl, K. A. Williams, F. Lacerda, K. N. Stevens, and B. Lindblom, "Linguistic experience alters phonetic perception in infants by 6 months of age," Science **255**, 606–608 (1992).
[4] J. F. Werker, "Phonetic perceptual reorganization across the first year of life: Looking back," Infant Behav. Dev. **75**, 101935 (2024).
[5] A. Götz, A. Krasotkina, G. Schwarzer, and B. Höhle, "Asymmetries in infants' vowel perception: Changes in vowel discrimination in German learning 6- and 9-month-old infants," Lang. Speech **67**, 1135–1149 (2024).
[6] B. McMurray and R. N. Aslin, "Infants are sensitive to within-category variation in speech perception," Cognition **95**, B15–26 (2005).
[7] J. F. Werker and R. C. Tees, "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life," Infant Behav. Dev. **7**, 49–63 (1984).
[8] D. R. Moore, M. A. Ferguson, L. F. Halliday, and A. Riley, "Frequency discrimination in children: Perception, learning and attention," Hear. Res. **238**, 147–154 (2008).
[9] J. Rose, M. Flaherty, J. Browning, L. J. Leibold, and E. Buss, "Pure-tone frequency discrimination in preschoolers, young school-age children, and adults," J. Speech. Lang. Hear. Res. **61**, 2440–2445 (2018).
[10] V. Hazan and S. Barrett, "The development of phonemic categorization in children aged 6–12," J. Phonetics **28**, 377–396 (2000).
[11] S. A. Ayala, A. Eads, H. Kabakoff, M. T. Swartz, D. M. Shiller, J. Hill, E. R. Hitchcock, J. L. Preston, and T. McAllister, "Auditory and somatosensory development for speech in later childhood," J. Speech. Lang. Hear. Res. **66**, 1252–1273 (2023).
[12] E. Kutlu, H. Kim, and B. McMurray, "Longitudinal changes in the structure of speech categorization across school age years: Children become more gradient and more consistent," Dev. Sci. **29**, e70085 (2026).
[13] B. McMurray, "The acquisition of speech categories: Beyond perceptual narrowing, beyond unsupervised learning and beyond infancy," Lang. Cogn. Neurosci. **38**, 419–445 (2023).
[14] Y. Feng and G. Peng, "Development of categorical speech perception in Mandarin-speaking children and adolescents," Child. Dev. **94**, 28–43 (2023).
[15] J. E. Sussman and A. E. Carney, "Effects of transition length on the perception of stop consonants by children and adults," J. Speech. Lang. Hear. Res. **32**, 151–160 (1989).
[16] E. S. Heller Murray and C. E. Stepp, "Relationships between vocal pitch perception and production: A developmental perspective," Sci. Rep. **10**, 3912 (2020).
[17] T. McAllister Byun and M. Tiede, "Perception-production relations in later development of American English rhotics," PLoS One **12**, e0172022 (2017).
[18] B. McMurray, M. K. Tanenhaus, and R. N. Aslin, "Gradient effects of within-category phonetic variation on lexical access," Cognition **86**, B33–B42 (2002).
[19] D. B. Fry, A. S. Abramson, P. D. Eimas, and A. M. Liberman, "The identification and discrimination on synthetic vowels," Lang. Speech **5**, 171–189 (1962).
[20] D. B. Pisoni and J. H. Lazarus, "Categorical and noncategorical modes of speech perception along the voicing continuum," J. Acoust. Soc. Am. **55**, 328–333 (1974).
[21] Y. Yan, F. Chen, and J. Li, "An overview of the impacts of vowels and consonants in speech understanding and their applications," npj Acoust. **1**, 28 (2025).
[22] T. Nazzi and A. Cutler, "How consonants and vowels shape spoken-language recognition," Ann. Rev. Linguist. **5**, 25–47 (2019).
[23] J. E. Sussman, "Vowel perception by adults and children with normal language and specific language impairment: Based on steady states or transitions?," J. Acoust. Soc. Am. **109**, 1173–1180 (2001).
[24] S. Nittrouer and J. H. Lowenstein, "Dynamic spectral structure specifies vowels for adults and children," Lang. Speech **57**, 487–512 (2014).
[25] S. Nittrouer, "Dynamic spectral structure specifies vowels for children and adults," J. Acoust. Soc. Am. **122**, 2328–2339 (2007).
[26] B. Parrell and J. Houde, "Modeling the role of sensory feedback in speech motor control and learning," J. Speech. Lang. Hear. Res. **62**, 2963–2985 (2019).

[27]F. H. Guenther, *Neural Control of Speech* (MIT Press, Cambridge, MA, 2016).

[28]F. H. Guenther, "A neural network model of speech acquisition and motor equivalent speech production," Biol. Cybern. **72**, 43–53 (1994).

[29]J. F. Houde and S. S. Nagarajan, "Speech production as state feedback control," Front. Hum. Neurosci. **5**, 82 (2011).

[30]E. L. Saltzman and K. G. Munhall, "A dynamical approach to gestural patterning in speech production," Ecol. Psychol. **1**, 333–382 (1989).

[31]E. N. MacDonald, E. K. Johnson, J. Forsythe, P. Plante, and K. G. Munhall, "Children's development of self-regulation in speech production," Curr. Biol. **22**, 113–117 (2012).

[32]F. Van Brenk and H. Terband, "Compensatory and adaptive responses to real-time formant shifts in adults and children," J. Acoust. Soc. Am. **147**, 2261–2270 (2020).

[33]D. M. Shiller and M. L. Rochon, "Auditory-perceptual learning improves speech motor adaptation in children," J. Exp. Psychol. Hum. Percept. Perform. **40**, 1308–1315 (2014).

[34]D. M. Shiller, S. Rvachew, and F. Brosseau-Lapré, "Importance of the auditory perceptual target to the achievement of speech production accuracy," Can. J. Speech-Lang. Pathol. Audiol. **34**, 181–192 (2010).

[35]R. A. Lester-Smith, A. Daliri, N. Enos, D. Abur, A. A. Lupiani, S. Letcher, and C. E. Stepp, "The relation of articulatory and vocal auditory-motor control in typical speakers," J. Speech. Lang. Hear. Res. **63**, 3628–3642 (2020).

[36]D. Abur, A. Subaciute, M. Kapsner-Smith, R. K. Segina, L. F. Tracy, J. P. Noordzij, and C. E. Stepp, "Impaired auditory discrimination and auditory-motor integration in hyperfunctional voice disorders," Sci. Rep. **11**, 13123 (2021).

[37]T. Rebernik, J. Jacobi, R. S. S. J. Buurke, T. B. Tienkamp, D. Abur, and M. B. Wieling, "SPRAAKLAB—Mobile laboratory for speech recorded acoustically and kinematically," Behav. Res. **57**, 205 (2025).

[38]R. R. Patel, S. N. Awan, J. Barkmeier-Kraemer, M. Courey, D. Deliyski, T. Eadie, D. Paul, J. G. Svec, and R. Hillman, "Recommended protocols for instrumental assessment of voice: American peech-Language-Hearing Association expert panel to develop a protocol for instrumental assessment of vocal function," Am. J. Speech. Lang. Pathol. **27**, 887–905 (2018).

[39]P. Adank, R. van Hout, and R. Smits, "An acoustic description of the vowels of Northern and Southern Standard Dutch," J. Acoust. Soc. Am. **116**, 1729–1738 (2004).

[40]D. H. Whalen, W. R. Chen, C. H. Shadle, and S. A. Fulop, "Formants are easy to measure; resonances, not so much: Lessons from Klatt (1986)," J. Acoust. Soc. Am. **152**, 933–941 (2022).

[41]To assess if the variability in JND values was affected by the differing quality of formant tracking (Ref. 40) in children, we conducted a follow-up analysis. For this analysis, a subset of participants was selected (five children with the highest JNDs and five children with the lowest JNDs). For these children, the formants tracked by Audapter were compared to manually tracked formants in Praat. The differences between the Audapter F_1 and manually calculated F_1 were not correlated with JND values (Pearson's correlation coefficient $r = 0.05$, $p = 0.896$). This result, therefore, suggests that the quality of formant tracking did not drive the variability in JNDs within the group of children.

[42]S. Hearnshaw, E. Baker, R. Pomper, K. K. McGregor, J. Edwards, and N. Munro, "The relationship between speech perception, speech production, and vocabulary abilities in children: Insights from by-group and continuous analyses," J. Speech. Lang. Hear. Res. **66**, 1173–1191 (2023).

[43]A. C. Walley and J. E. Flege, "Effect of lexical status on children's and adults' perception of native and non-native vowels," J. Phon. **27**, 307–332 (1999).

[44]S. Barreda, "Fast Track: Fast (nearly) automatic formant tracking using Praat," Linguist. Vanguard **7**, 20200051 (2021).