European Archives of Oto-Rhino-Laryngology
https://doi.org/10.1007/s00405-026-10395-6

OTOLOGY

Dual-task speech-in-noise performance in children with unilateral cleft lip and palate and minimal conductive hearing loss

Yuşa Başoğlu[1] · Reyhan Hürriyetoğlu[2,3] · İlayda Başoğlu[4]

Received: 12 May 2026 / Accepted: 7 June 2026
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2026

Abstract

Purpose Children with cleft lip and palate (CLP) frequently experience middle-ear dysfunction and conductive hearing loss, which may affect functional listening in noise. This study examined speech-in-noise recognition and visual-task latency cost during concurrent auditory–visual task performance in children with unilateral CLP and minimal conductive hearing loss.

Methods Twenty-five children with unilateral CLP and minimal conductive hearing loss and 28 age-matched typically developing controls, aged 7–15 years, completed audiological assessment, a single-task speech-in-noise condition (SIN), a concurrent auditory–visual condition combining speech recognition in noise with a randomized visual saccade task (SIN-WS), and a saccade-only condition. Word recognition accuracy and mean saccadic latency were recorded. Articulation was assessed in the CLP group using a standardized cleft speech assessment.

Results Children with CLP showed poorer word recognition accuracy than controls in both SIN and SIN-WS conditions. Model-based analysis showed significant main effects of group and condition, whereas the Group × Condition interaction for word recognition was not significant. Mean saccadic latency increased during the concurrent-task condition in both groups, with a significantly greater increase in the CLP group. Within the CLP group, higher articulation error counts were associated with poorer word recognition accuracy and greater saccadic-latency cost.

Conclusion Children with unilateral CLP and minimal conductive hearing loss showed reduced speech-in-noise recognition and greater visual-task latency cost during concurrent auditory–visual task performance. These findings support the clinical value of assessing functional listening beyond isolated auditory thresholds in children with CLP.

Keywords Cleft lip and palate · Conductive hearing loss · Speech-in-noise recognition · Saccadic latency · Auditory–visual task performance

Introduction

Cleft lip and palate (CLP) is a common congenital craniofacial condition with important otologic, audiological, and communicative consequences. In addition to structural and speech-related challenges, children with CLP are at high risk for Eustachian tube dysfunction, otitis media with effusion, and conductive hearing loss because of altered palatal and middle-ear mechanics [1–4]. Although hearing loss in this population is often described as mild or minimal, even small reductions in auditory access may be clinically meaningful during childhood, particularly in listening environments that require accurate perception of speech in background noise [5, 6]. For children with CLP, such listening difficulties may occur alongside cleft-related speech production differences, making functional communication

✉ Yuşa Başoğlu
ybasoglu@medipol.edu.tr

1 Department of Audiology, Faculty of Health Sciences, İstanbul Medipol University, İstanbul, Türkiye

2 Department of Audiology, Institute of Health Sciences, Ankara Medipol University, Altındağ, Ankara, Türkiye

3 Department of Audiology, Institute of Health Sciences, Hacettepe University, Altındağ, Ankara, Türkiye

4 Department of Speech and Language Therapy, Faculty of Health Sciences, İstanbul Medipol University, İstanbul, Türkiye

Published online: 05 July 2026

outcomes especially important to evaluate beyond routine pure-tone thresholds.

Speech recognition in noise is a clinically relevant measure of functional listening because it places greater demands on acoustic-phonetic processing than speech recognition in quiet. This is particularly relevant for children with fluctuating or minimal conductive hearing loss, in whom audibility may appear relatively preserved in standard audiological testing but listening performance may deteriorate under adverse acoustic conditions [5, 6]. In everyday environments such as classrooms, children must often understand speech in the presence of competing talkers while simultaneously monitoring visual information, shifting gaze, and responding to ongoing task demands. Therefore, assessment conditions that combine speech-in-noise recognition with an additional nonverbal visual task may provide information about functional listening that is not captured by isolated auditory measures.

The present study was guided by a resource-allocation view of listening. According to this framework, degraded or less reliable auditory input may increase the effort required to maintain speech recognition, leaving fewer resources available for concurrent task performance [7, 8]. Dual-task paradigms have been used to examine how listening performance changes when attention must be shared across simultaneous activities, and secondary-task outcomes may provide an index of the cost associated with maintaining performance under increased listening demand [7]. In children with CLP and minimal conductive hearing loss, this framework is clinically relevant because auditory input may be affected by middle-ear dysfunction, while spoken responses may also be influenced by cleft-related articulation and resonance characteristics. Thus, performance in a concurrent auditory–visual paradigm should not be interpreted as a process-pure measure of auditory perception alone, but rather as a functional measure reflecting the combined influence of auditory access, visual-attentional demand, and speech-production factors.

Speech production characteristics are also important when interpreting speech-recognition tasks in children with CLP. Because many speech-in-noise tests require overt verbal repetition, performance may reflect not only recognition of the target word but also the ability to produce the response clearly. Cleft-related articulation errors, including obligatory distortions and compensatory productions, may therefore influence repetition-based speech-recognition scores [9–11]. For this reason, articulation error burden was included in the present study as an exploratory clinical correlate of speech-in-noise recognition and visual-task cost within the CLP group.

Although the otologic and speech-related characteristics of children with CLP have been widely described, less is known about how children with unilateral CLP and minimal conductive hearing loss perform when speech recognition in noise is combined with a concurrent visual-attentional task. Clarifying this issue may help determine whether children with CLP show functional performance costs under combined auditory and visual task demands, even when hearing loss is minimal. The aim of this study was therefore to examine speech-in-noise recognition and visual-task latency cost in children with unilateral CLP and minimal conductive hearing loss compared with age-matched typically developing peers. We expected children with CLP to show poorer overall speech-in-noise recognition and greater visual-task latency cost during concurrent auditory–visual task performance. We further explored whether greater articulation error burden would be associated with poorer speech recognition and greater visual-task cost within the CLP group.

Methods

Participants

The study was conducted between January and December 2022 in [blinded for review]. Children with cleft lip and palate were recruited from the [blinded for review] database, a regional referral center for individuals with CLP. The participant recruitment, allocation, and analysis process is illustrated in Fig. 1. To reduce anatomical heterogeneity, inclusion in the CLP group was restricted to children with complete unilateral cleft lip and palate involving both the primary and secondary palate, based on the Kernahan classification [12]. Recruitment was based on anatomical diagnosis and audiological eligibility rather than speech status or treatment history; therefore, the CLP sample represented a clinical range of speech outcomes.

A total of 25 children with unilateral CLP and minimal conductive hearing loss were included in the CLP group. This group consisted of 13 boys and 12 girls aged 7–15 years, with a mean age of 8.72 ± 2.92 years. The control group consisted of 28 age-matched typically developing children, including 17 boys and 11 girls, aged 7–15 years, with a mean age of 9.25 ± 2.38 years.

Inclusion criteria for the control group were as follows: (1) age between 7 and 15 years, (2) normal hearing, defined as a four-frequency pure-tone average (PTA4) of ≤ 15 dB HL at 0.5, 1, 2, and 4 kHz, (3) no known cognitive or psychiatric disorder, confirmed by the Turkish version of the Modified Paediatric Mini Mental Scale together with parent report and clinical observation, and (4) ability to complete the test procedures [13].

Inclusion criteria for the CLP group were as follows: (1) age between 7 and 15 years, (2) diagnosis of complete

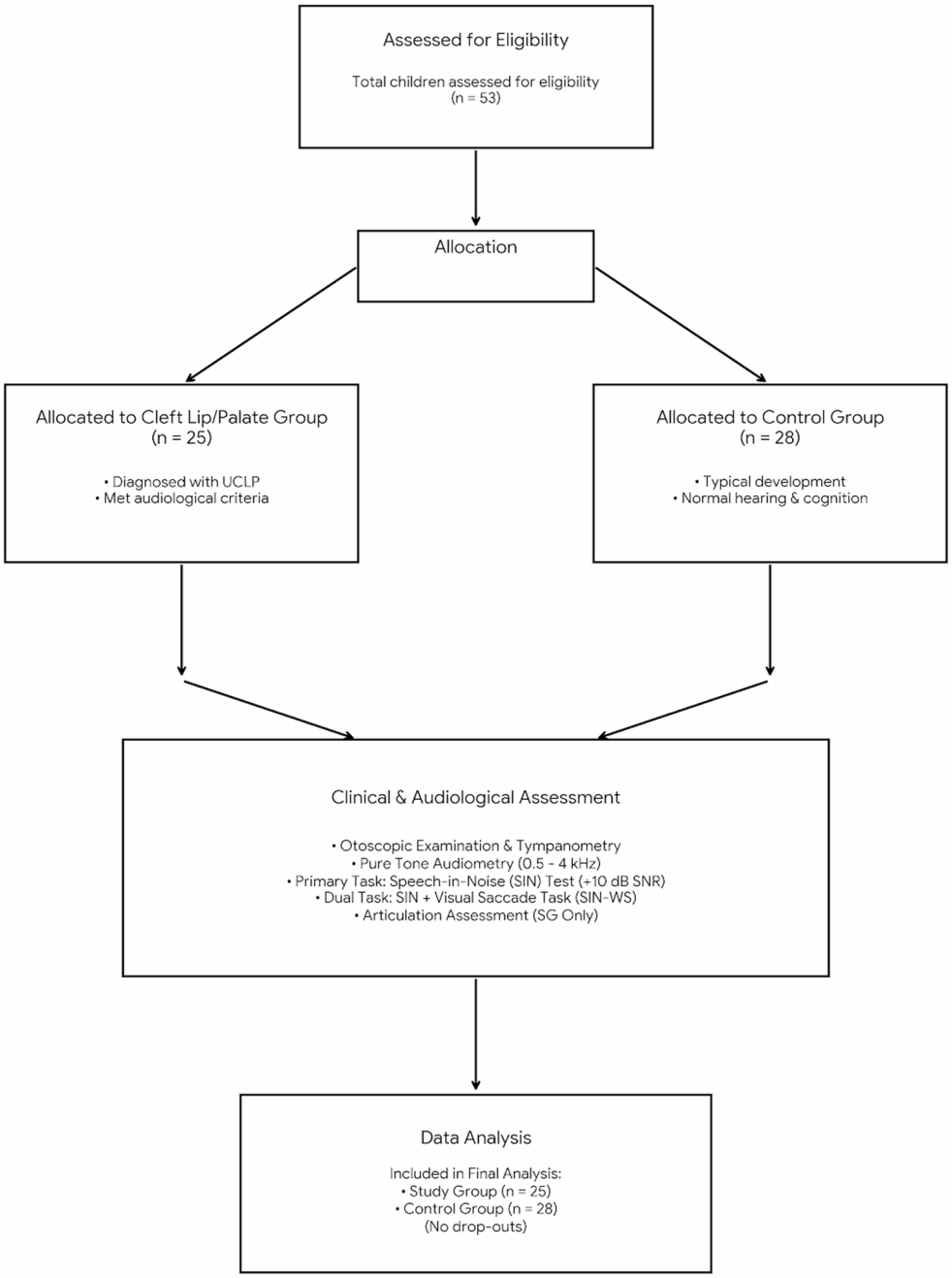


Fig. 1 STROBE flow diagram of the study illustrating the recruitment, allocation, assessment, and analysis process

unilateral cleft lip and palate, (3) minimal conductive hearing loss, defined as an air-conduction PTA4 between 16 and 25 dB HL at 0.5, 1, 2, and 4 kHz, bone-conduction thresholds ≤ 15 dB HL, and an air-bone gap of ≥ 10 dB at two or more test frequencies [14, 15], (4) no known cognitive or psychiatric disorder, confirmed by the Turkish version of the Modified Paediatric Mini Mental Scale together with parent report and clinical observation, and (5) ability to complete the test procedures.

Children were excluded if they had active otorrhea, active ventilation tubes at the time of testing, known syndromic clefting, diagnosed neurological disease, uncorrected visual impairment that would interfere with the saccade task, or inability to complete the behavioral test procedures.

Study design

This was a comparative clinical study designed to examine functional speech-in-noise recognition and visual-task latency cost in children with unilateral CLP and minimal conductive hearing loss. All participants underwent otoscopic examination, tympanometry, and pure-tone audiometry before behavioral testing. Speech recognition in noise was assessed under two conditions: (1) a single-task

speech-in-noise condition (SIN) and (2) a concurrent auditory–visual condition in which the same speech-in-noise task was performed while completing a randomized visual saccade task (SIN-WS). A saccade-only condition was also administered to obtain baseline visual-task performance.

This design was selected to evaluate performance under listening conditions that extend beyond isolated pure-tone thresholds and approximate the combined auditory and visual demands encountered in everyday communication environments. The speech-in-noise task was used as the primary functional listening outcome, whereas change in saccadic latency from the saccade-only condition to the concurrent-task condition was used to estimate visual-task cost. Articulation was assessed only in the CLP group to explore whether cleft-related speech production burden was associated with performance in repetition-based speech-recognition tasks and concurrent visual-task cost.

Tympanometry

Otoscopic examination of the external ear canal and tympanic membrane was performed before acoustic immittance testing. Tympanometry was conducted using an Interacoustics Titan/IMP440 impedance analyzer (Interacoustics A/S, Middelfart, Denmark) with a 226-Hz probe tone presented at 85 dB SPL. Only participants with intact tympanic membranes and no active ventilation tubes at the time of testing were included. Tympanometric patterns were classified according to the Jerger system as Type A, Type B, or Type C for descriptive group comparisons [16].

Pure-tone audiometry

Pure-tone audiometry was performed in a double-walled sound-treated booth using an Interacoustics AC-40 clinical audiometer with TDH-39 supra-aural headphones and a RadioEar B-71 bone vibrator. Air- and bone-conduction thresholds were obtained at 0.5, 1, 2, and 4 kHz, and PTA4 was calculated from air-conduction thresholds. Equipment calibration was performed in accordance with ANSI standards [17].

Speech recognition in noise condition

Speech recognition in noise was assessed by a single licensed audiologist in a sound-attenuated booth. Target words were presented via monitored live voice from a loudspeaker positioned 1 m in front of the participant at 0° azimuth at the child's head level. Monitored live voice was utilized because matched, recorded pediatric Turkish speech-in-noise materials were unavailable. To minimize presentation variability, the same audiologist administered all tests, and levels were verified at the participant position before each session.

A list of 25 Turkish monosyllabic phonetically balanced words from the corpus developed by Mungan was presented alongside a four-talker babble masker, including three female and one male adult voices [18, 19]. Monosyllabic words (e.g., *vur*, *kır*, *kan*, *nar*, *saç*) were intentionally selected to minimize contextual linguistic cues. This masker was selected because competing speech can introduce both energetic and informational masking, making it clinically relevant for evaluating speech recognition under adverse listening conditions [20, 21]. Speech and noise were delivered at fixed levels—75 dB SPL and 65 dB SPL, respectively—establishing a + 10 dB signal-to-noise ratio (SNR) [22]. This SNR was selected to impose an adverse listening demand while avoiding floor effects in children with minimal conductive hearing loss. Output calibration was performed at the participant position using a Class 1 sound level meter.

Participants were instructed to repeat each target word as clearly as possible. Responses were scored at the whole-word level based on lexical identity. A response was scored as correct when the intended target word was clearly identifiable. Cleft-related speech production deviations that did not alter lexical identity (e.g., obligatory distortions or audible nasal emission) were not penalized, whereas substitutions or omissions were scored as incorrect. Speech recognition accuracy was calculated as percent correct.

Saccade-only condition

To obtain baseline visual-task performance, participants completed the randomized saccade task without concurrent speech recognition. Saccadic eye movements were recorded using a videonystagmography (VNG) system (Micromedical Technologies, USA). The task was delivered using discrete light targets presented on a horizontal light bar. Targets appeared at random intervals of 2–3 s and at amplitudes of 15°–20°to the right or left of the central fixation point.

Eye movements were recorded continuously through VNG goggles, and only valid target-directed saccades identified by the software were included in the analysis. For inferential analyses, mean saccadic latency averaged across valid rightward and leftward trials was used as the primary visual-task outcome. This condition provided each participant's baseline visual orienting speed, allowing concurrent-task latency cost to be interpreted relative to individual visual-task performance.

Concurrent auditory–visual condition

In the concurrent auditory–visual condition (referred to as SIN-WS), participants performed the speech-in-noise task while simultaneously completing the randomized visual saccade task. Speech, babble noise, and visual task parameters were identical to those used in the single-task conditions (speech at 75 dB SPL, babble at 65 dB SPL, +10 dB SNR). The visual display was positioned directly below the loudspeaker to avoid obstructing the auditory signal.

Participants were instructed to prioritize speech recognition while following the visual targets as quickly and accurately as possible. Word recognition accuracy was scored using the same whole-word lexical criteria as the SIN condition, and mean saccadic latency quantified visual performance under concurrent load. The difference between mean saccadic latency in the SIN-WS and saccade-only conditions was calculated to estimate the visual-task latency cost. For all participants, the SIN condition was administered before the SIN-WS condition using different phonetically balanced word lists; however, list order was not counterbalanced.

Articulation assessment

Articulation was assessed only in the CLP group using the sentence repetition task from the Turkish CLP Assessment Form [23]. This task utilizes 21 sentences targeting consonants across the Turkish phoneme inventory. Both obligatory errors (related to structural or velopharyngeal constraints) and active compensatory errors (e.g., glottal stops, pharyngeal/posterior nasal fricatives) were classified as incorrect articulations, and a total articulation error count was calculated [10, 24]. All assessments were conducted by the same experienced speech-language pathologist. To evaluate scoring reliability, an independent, blinded speech-language pathologist re-scored 30% of the recordings, yielding a high inter-rater agreement (Cohen's $\kappa = 0.91$). Articulation error count was included as an exploratory clinical correlate rather than a causal predictor, accounting for the fact that the repetition-based speech-in-noise task could be influenced by speech production demands alongside auditory recognition.

Ethical approval

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Non-Invasive Clinical Research Ethics Committee of İstanbul Medipol University (Decision No: 1191/2021; Date: 13/12/2021; Document No: E-10840098-772.02-6298). Written informed consent was obtained from the parents or legal guardians of all participants prior to enrollment.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, version 27.0 (IBM Corp., Armonk, NY, USA). The primary speech outcome was word recognition accuracy in the SIN and SIN-WS conditions, and the primary visual outcome was mean saccadic latency. Visual-task latency cost was calculated as the difference between SIN-WS and saccade-only conditions. Descriptive statistics are presented as mean ± standard deviation or median (interquartile range), as appropriate.

Normality was assessed using the Shapiro–Wilk test and Q–Q plots. Due to ceiling effects and non-normal distributions in speech scores, Mann–Whitney U tests were used for between-group comparisons and Wilcoxon signed-rank tests for within-group comparisons. Categorical variables were analyzed using chi-square tests.

Word recognition accuracy was additionally modeled using generalized estimating equations (GEE) with a binomial distribution and logit link. Scores were converted to counts (correct = percent × 25/100). The model included group, condition, their interaction, and mean-centered age and PTA_4 as covariates. Results are reported as odds ratios (ORs) with 95% confidence intervals (CIs).

Saccadic latency was analyzed using linear mixed-effects models with participant-level random intercepts, including group, condition, their interaction, and age as fixed effects. Skewed distributions were log-transformed when necessary. Associations between articulation error count and behavioral outcomes were examined using Spearman's correlation (ρ). Statistical significance was set at $\alpha = 0.05$ (two-tailed).

Results

Demographic and audiological characteristics are presented in Table 1. The CLP group and control group did not differ significantly in age ($p = .471$). Air-conduction PTA4 values were higher in the CLP group than in the control group for both ears (both $p = .001$), whereas bone-conduction PTA4 values did not differ significantly between groups. Tympanometric findings are shown in Table 2. All ears in the control group showed Type A tympanograms, whereas the CLP group showed Type A, B, and C patterns. Tympanogram distributions differed significantly between groups for both the left ear, $\chi^2(2) = 28.03$, $p < .001$, and the right ear, $\chi^2(2) = 30.53$, $p < .001$.

Word recognition outcomes in the single-task speech-in-noise condition (SIN) and the concurrent-task condition (SIN-WS) are presented in Table 3. The CLP group performed more poorly than the control group in both conditions (both $p < .001$). Mann-Whitney U tests confirmed significant

Table 1 Demographic characteristics and pure-tone averages of the CLP and control groups

Variable	CLP group mean ± SD	CLP group median (IQR)	Control group mean ± SD	Control group median (IQR)	*p* value
Sex (F: M)	12:13	—	11:17	—	0.523 (χ^2 test)
Age (years)	8.72 ± 2.92	7.00 (7.00–10.00)	9.25 ± 2.38	9.00 (7.00–11.00)	0.471
Air-conduction PTA4, right ear (dB HL)	19.28 ± 4.87	20.00 (17.00–21.00)	10.82 ± 4.61	12.00 (9.00–13.00)	0.001
Air-conduction PTA4, left ear (dB HL)	17.44 ± 6.72	18.00 (14.00–20.00)	9.92 ± 4.58	10.00 (5.00–13.50)	0.001
Bone-conduction PTA4, right ear (dB HL)	4.08 ± 2.19	5.00 (3.00–6.00)	2.67 ± 2.89	3.00 (0.00–5.00)	0.780
Bone-conduction PTA4, left ear (dB HL)	3.68 ± 2.57	4.00 (2.00–5.00)	2.28 ± 2.66	2.50 (0.00–5.00)	0.640

Values are reported as mean ± SD and median (IQR) for continuous variables and as counts for sex. CLP = cleft lip and palate; F = female; M = male; PTA4 = four-frequency pure-tone average; dB HL = decibels hearing level; SD = standard deviation; IQR = interquartile range. Mann–Whitney U tests were used for continuous variables, and a χ^2 test was used for sex. Two-tailed $\alpha = 0.05$

Table 2 Tympanogram distributions by group (left and right ears)

Group	Type A *n* (%)	Type B *n* (%)	Type C *n* (%)
Left ear			
Control Group ($n = 28$)	28 (100)	0 (0)	0 (0)
CLP group ($n = 25$)	8 (32)	11 (44)	6 (24)
Group comparison	$\chi^2(2) = 28.03$		$p < .001$
Right ear			
Control Group ($n = 28$)	28 (100)	0 (0)	0 (0)
CLP group ($n = 25$)	7 (28)	9 (36)	9 (36)
Group comparison	$\chi^2(2) = 30.53$		$p < .001$

CLP = cleft lip and palate. Values are presented as n (%). χ^2 = chi-square test

between-group differences for both SIN ($U = 71.5$, $p < .001$) and SIN-WS ($U = 36.0$, $p < .001$). Within-group comparisons showed lower performance in the concurrent-task condition than in the single-task condition for both groups (control group: $W = 0.0$, $p < .001$; CLP group: $W = 0.0$, $p < .001$).

Model-based analysis of word recognition accuracy is shown in Table 3, Panel B. In the GEE model, there were significant main effects of Group and Condition. Relative to the control group, the CLP group had lower odds of correct word recognition overall (OR = 0.34, 95% CI [0.21, 0.53], $p < .001$). Across groups, performance was lower in the concurrent-task condition than in the single-task condition (OR = 0.55, 95% CI [0.46, 0.65], $p < .001$). The Group × Condition interaction was not significant (OR = 1.04, 95% CI [0.83, 1.31], $p = .724$), indicating that the reduction in word recognition from SIN to SIN-WS did not differ significantly between groups.

Visual-task performance is summarized in Tables 4 and 5. Mean saccadic latency increased from the saccade-only condition to the concurrent-task condition (SIN-WS), and this increase was descriptively greater in the CLP group than in the control group (Table 4). In the covariate-adjusted mixed-effects model, there was a significant Condition effect in the control group, indicating longer saccadic latencies during SIN-WS than during the saccade-only condition ($\beta = 9.75$ ms, $p < .001$). The Group × Condition interaction was also significant ($\beta = 11.11$ ms, $p < .001$), indicating that the increase in saccadic latency from single-task to concurrent-task performance was greater in the CLP group than in the control group. Age was a significant covariate ($\beta = -3.66$ ms/year, $p < .001$), with shorter latencies at older ages.

The distribution of speech production characteristics in the CLP group is presented in Table 6. Exploratory correlational analyses are shown in Table 7. Within the CLP group, higher total articulation error counts were associated with poorer word recognition in both SIN ($\rho = -0.443$, $p = .027$) and SIN-WS ($\rho = -0.514$, $p = .009$). Higher articulation error counts were also associated with greater mean saccadic latency cost between the saccade-only and concurrent-task conditions ($\rho = 0.420$, $p = .030$).

Discussion

The present study examined speech-in-noise recognition and visual-task latency cost during concurrent auditory–visual task performance in children with unilateral CLP and minimal conductive hearing loss. Three main findings emerged. First, children with CLP showed poorer word recognition accuracy than typically developing peers in both the single-task speech-in-noise condition and the concurrent auditory–visual condition. Second, word recognition accuracy decreased under concurrent-task conditions in both groups, but the Group × Condition interaction was not significant, indicating that children with CLP did not show a disproportionate decline in word recognition accuracy when the visual task was added. Third, saccadic latency increased during concurrent auditory–visual task performance in both groups, and this increase was significantly greater in the

Table 3 Word recognition accuracy (%) in speech-in-noise under single-task (SIN) and concurrent-task (SIN-WS) conditions **Panel B. Model-based analysis of word recognition accuracy (binomial logit model; GEE)**

Outcome	CLP group Mean ± SD	CLP group Median (IQR)	CLP group Min-max	Control group Mean ± SD	Control group Median (IQR)	Control group Min-max	*p* value (Mann-Whitney U)	Effect size: Hedges g (CLP - Control), 95% CI
Panel A. Descriptive statistics and nonparametric comparisons								
SIN	82.08 ± 8.41	80.00 (76.00–88.00)	64–100	94.57 ± 3.95	94.00 (92.00–97.00)	88–100	< 0.001	−1.91 [−2.56, −1.25]
SIN-WS	72.48 ± 8.89	72.00 (68.00–76.00)	56–92	90.28 ± 4.38	92.00 (88.00–92.00)	80–100	< 0.001	−2.55 [−3.28, −1.82]
Panel B. Model-based analysis of word recognition accuracy (binomial logit model; GEE)								
Fixed effect	β (log-odds)	Robust SE	Wald z	OR	95% CI (OR)	*p*		
Intercept (CG, SIN; mean age and mean PTA4)	2.748	0.144	19.10	15.61	11.77–20.70	< 0.001		
Group (CLP vs. CG)	−1.092	0.235	−4.65	0.34	0.21–0.53	< 0.001		
Condition (SIN-WS vs. SIN)	−0.596	0.088	−6.77	0.55	0.46–0.65	< 0.001		
Group × Condition	0.041	0.115	0.35	1.04	0.83–1.31	0.724		
Age (centered; per 1-year increase)	0.001	0.029	0.03	1.00	0.95–1.06	0.973		
Bilateral mean air-conduction PTA4 (centered; per 1-dB HL increase)	−0.030	0.022	−1.40	0.97	0.93–1.01	0.162		

Panel A presents descriptive statistics and between-group comparisons using Mann–Whitney U tests. Hedges g is provided as a standardized mean difference for between-group comparisons. Panel B presents a generalized estimating equations analysis with a binomial distribution and logit link, modeling word recognition as the number correct out of 25 trials. Percent-correct scores were converted to counts by dividing by 4 because each list included 25 words. Reference categories were the control group and the SIN condition. Age and bilateral mean air-conduction PTA4 were mean-centered. CLP = cleft lip and palate; CG = control group; SIN = speech-in-noise; SIN-WS = speech-in-noise with concurrent visual saccade task; GEE = generalized estimating equations; OR = odds ratio; CI = confidence interval; SD = standard deviation; IQR = interquartile range; PTA4 = pure-tone average; dB HL = decibels hearing level

CLP group. Within the CLP group, greater articulation error burden was associated with poorer speech-in-noise recognition and greater visual-task latency cost. Taken together, these findings suggest that children with unilateral CLP and minimal conductive hearing loss may show reduced functional listening performance in noise together with increased visual-task cost when auditory and visual task demands are combined.

The poorer speech-in-noise recognition observed in the CLP group is consistent with the known otologic vulnerability of children with CLP. Eustachian tube dysfunction, otitis media with effusion, recurrent acute otitis media, and conductive hearing loss are common in children with cleft palate and CLP and may reduce the quality and consistency of auditory input during communication development [3, 4, 6, 25]. Although the hearing loss in the present CLP group was minimal by pure-tone criteria, the results indicate that even mild reductions in air-conduction thresholds may be associated with measurable functional listening difficulties when speech is presented in background noise. This finding is clinically important because pure-tone thresholds alone may underestimate everyday listening difficulties in children with fluctuating or minimal conductive hearing loss, particularly in environments with competing speech [5].

The tympanometric profile of the CLP group further supports the otologic interpretation of the speech-in-noise findings. Whereas the control group showed normal tympanometric patterns, the CLP group included a substantial proportion of Type B and Type C tympanograms. This pattern is consistent with middle-ear involvement and suggests that reduced middle-ear efficiency may have contributed to poorer access to acoustic-phonetic information in noise. From an otorhinolaryngological perspective, these findings reinforce the importance of evaluating functional listening outcomes in children with CLP rather than relying only on standard audiometric thresholds. Recent evidence from pediatric otology also supports the relevance of otitis media history for later speech-in-noise performance, further emphasizing the value of functional listening measures in children with middle-ear vulnerability [26].

A central point in interpreting the present findings is that the reduction in word recognition accuracy from the single-task condition to the concurrent auditory–visual condition was observed in both groups. The non-significant Group × Condition interaction indicates that the CLP group did not show a selectively greater decrement in speech recognition accuracy when the visual task was added. Therefore, the speech-recognition results should be interpreted as evidence

Table 4 Saccade latency in single-task and concurrent-task conditions

Outcome (ms)	Control group ($n=28$) Mean ± SD	Control group Median (IQR)	CLP group ($n=25$) Mean ± SD	CLP group Median (IQR)
Single-task saccadic latency, rightward	275.3 ± 16.8	277.0 (264.5–289.0)	265.6 ± 24.3	265.0 (244.0–282.0)
Single-task saccadic latency, leftward	267.2 ± 17.4	269.5 (254.0–282.0)	258.5 ± 23.4	255.0 (241.0–275.0)
Concurrent-task saccadic latency (SIN-WS), rightward	284.5 ± 17.8	287.5 (271.8–299.2)	286.6 ± 24.6	288.0 (264.0–305.0)
Concurrent-task saccadic latency (SIN-WS), leftward	276.5 ± 18.2	278.0 (262.8–292.2)	279.3 ± 23.4	276.0 (262.0–298.0)
Saccadic latency cost, rightward	9.2 ± 3.4	10.0 (9.0–11.0)	21.0 ± 2.2	21.0 (20.0–23.0)
Saccadic latency cost, leftward	9.3 ± 3.4	10.0 (9.0–11.0)	20.8 ± 1.7	21.0 (19.0–22.0)
Mean saccadic latency cost	9.2 ± 3.3	10.0 (9.5–10.5)	20.9 ± 1.6	21.0 (19.5–22.0)

Saccadic latency cost was calculated as latency in the concurrent-task condition (SIN-WS) minus latency in the single-task condition. CLP = cleft lip and palate; SIN-WS = speech-in-noise with concurrent visual saccade task; SD = standard deviation; IQR = interquartile range

Table 6 Distribution of speech production characteristics in the CLP group

Articulation error type	*n* (CLP group)	% of CLP group ($n=25$)	Description of clinical observations
Passive (obligatory) errors	16	64%	Sibilant distortions (e.g.,/s/,/z/) attributable to dental/anatomical constraints.
Audible nasal emission	9	36%	Audible nasal air escape during production of high-pressure consonants (e.g.,/p/,/t/,/k/).
Active (compensatory) errors	4	16%	Compensatory productions such as glottal stops and pharyngeal fricatives.

Percentages are calculated based on the number of participants in the CLP group ($n=25$). Participants may present more than one articulation error type; therefore, percentages may sum to > 100%. "Passive (obligatory) errors" correspond to anatomically/dentally constrained errors

of poorer overall speech-in-noise performance in the CLP group, not as evidence of disproportionate dual-task interference in the speech-recognition measure itself.

In contrast, the visual-task results showed a different pattern. Saccadic latency increased during the concurrent

Table 5 Covariate-adjusted mixed-effects model of saccadic latency

Fixed effect	β (ms)	SE	95% CI	*p* value
Intercept (Control, single-task; mean age)	272.15	3.44	265.41 to 278.88	< 0.001
Group (CLP vs. Control) in the single-task condition	−11.47	5.02	−21.31 to −1.64	0.022
Condition (concurrent-task vs. single-task) in the Control group	9.75	0.47	8.83 to 10.67	< 0.001
Group × Condition (additional concurrent-task effect in the CLP group)	11.11	0.69	9.77 to 12.45	< 0.001
Age (centered; per 1-year increase)	−3.66	0.96	−5.53 to −1.78	< 0.001

Dependent variable. Mean saccadic latency (ms) = (leftward + rightward)/2

Data structure. Two observations per participant (single-task and concurrent-task conditions)

Fixed effects. Group (CLP vs. control), Condition (concurrent-task vs. single-task), Group × Condition, and Age (centered)

Random effect. Participant-level random intercept

β = unstandardized regression coefficient; SE = standard error; CI = confidence interval; CLP = cleft lip and palate; ms = milliseconds. Participant was included as a random intercept, and age was entered as a centered covariate. A positive Group × Condition coefficient indicates a greater concurrent-task-related increase in saccadic latency in the CLP group than in the control group

Table 7 Associations between total articulation error count and task outcomes within the CLP group

Variable	Mean ± SD	Median (IQR)	Min–max
Panel A. Descriptive statistics (CLP group; $n=25$)			
Total articulation errors (count)	12.44 ± 4.39	12 (9–15)	5–19
Panel B. Correlations (CLP group; Spearman's ρ)			
Association	Spearman's ρ	95% CI	*p* value
Total articulation error count and SIN word recognition accuracy (%)	−0.443	[−0.713, −0.058]	0.027*
Total articulation error count and SIN-WS word recognition accuracy (%)	−0.514	[−0.756, −0.149]	0.009**
Total articulation error count and mean saccadic latency cost (ms)	0.420	[0.050, 0.680]	0.030*

ρ = Spearman's rank correlation coefficient. SIN and SIN-WS represent word recognition performance in the single-task and concurrent-task conditions, respectively. Mean saccadic latency cost was calculated as mean latency in the concurrent-task condition (SIN-WS) minus mean latency in the single-task condition. CLP = cleft lip and palate; CI = confidence interval; SD = standard deviation; IQR = interquartile range. *p < .05, **p < .01

auditory–visual condition in both groups, but the increase was significantly greater in the CLP group. This suggests that, although children with CLP did not show a greater speech-recognition decrement under concurrent-task conditions, they showed a greater cost in the secondary visual task. One possible interpretation is that children with CLP may have allocated more resources to maintaining speech recognition in noise, resulting in slower visual orienting during the concurrent task. This interpretation is consistent with a resource-allocation framework, in which degraded or less reliable auditory input increases the demand required to maintain listening performance, leaving fewer resources available for concurrent task performance [7, 8]. Recent pediatric evidence also indicates that background noise can increase listening effort, although the magnitude and detectability of this effect depend on the task and measurement method used [27].

The use of saccadic latency as a secondary-task outcome is relevant in this context because it provides a nonverbal, time-sensitive measure of visual-task performance during concurrent listening. This finding may be particularly relevant for children with CLP because daily communication often occurs in settings where children must listen in noise while simultaneously monitoring visual information, shifting gaze, and responding to environmental demands. The present results suggest that such combined task demands may reveal functional performance costs that are not captured by isolated auditory testing.

The association between articulation error burden and task performance provides an additional clinically relevant finding. Within the CLP group, higher articulation error counts were associated with poorer word recognition accuracy in both speech-in-noise conditions and with greater saccadic-latency cost. Because the speech-recognition task required overt verbal repetition, performance likely reflected not only auditory recognition of the target word but also the child's ability to produce the response accurately and efficiently. This is particularly important in children with CLP, for whom obligatory and compensatory articulation errors are central components of the clinical speech profile [9–11]. Recent work has also shown that conductive hearing loss and cleft speech characteristics may be related in children born with cleft palate, supporting the need to interpret auditory and speech-production factors together in this population [28].

These articulation findings also support a broader interpretation of functional communication in children with CLP. In this population, listening, speech production, and middle-ear status are closely related clinical domains. A child with minimal conductive hearing loss may have reduced access to acoustic detail, while cleft-related articulation differences may influence the accuracy or efficiency of spoken responses. In a repetition-based speech-in-noise task, these factors may interact at the level of functional performance. Accordingly, poorer word recognition scores in the CLP group should not be interpreted as a pure measure of auditory perception alone. Rather, they likely reflect the combined influence of conductive hearing status, speech-production profile, and task demands.

From a clinical standpoint, the present findings suggest that children with unilateral CLP and minimal conductive hearing loss may benefit from functional listening assessment in addition to conventional audiometry and tympanometry. Standard otologic and audiological evaluation remains essential, but pure-tone thresholds may not fully capture the listening challenges encountered in noisy or visually demanding environments. Speech-in-noise testing may provide additional information about functional auditory performance, whereas concurrent-task outcomes may help identify performance costs that become apparent when listening is combined with visual-attentional demands. Such information may be useful for counseling families, planning classroom accommodations, and identifying children who may require closer audiological and speech-language follow-up.

The findings also have implications for multidisciplinary CLP care. Children with CLP are commonly followed by otolaryngologists, audiologists, and speech-language professionals, but functional communication outcomes may fall between these domains. The present results support an integrated approach in which middle-ear status, minimal conductive hearing loss, speech-in-noise performance, and articulation profile are considered together. In particular, the observation that articulation error burden was associated with both speech-recognition performance and visual-task cost suggests that speech production characteristics may be relevant when interpreting functional listening tests that require spoken repetition.

Limitations

Several limitations should be considered. First, the sample size was modest and included a relatively broad pediatric age range; thus, these findings should be confirmed in larger, age-stratified cohorts. Second, although a history of recurrent otitis media with effusion (OME) was documented in all participants within the CLP group, detailed longitudinal data regarding the exact frequency and duration of individual OME episodes were not systematically recorded at baseline clinical assessment, preventing further stratification based on otologic burden. Third, because the study was designed to evaluate the real-world clinical profile of this population, the independent contributions of cleft status, minimal conductive hearing loss, and articulation

characteristics cannot be completely isolated. Fourth, the speech-recognition task utilized monitored live voice and a fixed single-loudspeaker setup, which may constrain direct ecological generalization to complex classroom environments. Additionally, task order was not counterbalanced across conditions. Despite these constraints, this study provides valuable clinical evidence regarding altered functional listening and visual-task cost under combined demands, supporting the need for future replication using standardized recorded speech materials and randomized task orders.

Conclusion

In conclusion, children with unilateral CLP and minimal conductive hearing loss demonstrated poorer speech-in-noise recognition than age-matched typically developing peers and greater visual-task latency cost during concurrent auditory–visual task performance. The absence of a significant Group × Condition interaction for word recognition indicates that the CLP group did not show disproportionate speech-recognition decline under concurrent-task conditions; rather, they showed poorer overall speech-in-noise performance. The stronger concurrent-task effect on saccadic latency suggests that visual-task cost may be a sensitive indicator of altered performance under combined auditory–visual demands in this population. These findings highlight the clinical value of assessing functional listening and task performance beyond isolated auditory thresholds in children with CLP.

Acknowledgements The authors thank Dr. Özlem Ülkü Logacev for providing the Turkish CLP Assessment Form used in this study. The authors also thank the participating children and their families.

Author contributions Yuşa Başoğlu: Conceptualization, Methodology, Writing – original draft, Supervision.

Funding No funding was received for conducting this study.

Data availability The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethical approval This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Non-Invasive Clinical Research Ethics Committee of İstanbul Medipol University (Decision No: 1191/2021; Date: 13/12/2021; Document No: E-10840098-772.02-6298).

Consent to participate Written informed consent was obtained from the parents or legal guardians of all participants.

Consent to publish Not applicable.

Prior presentation This research was presented as an oral presentation at the 3rd Auto-Audiology Congress organized by İstanbul Medipol University on June 4, 2022, and was awarded second prize.

Competing interests The authors have no relevant financial or non-financial interests to disclose.

References

1. Panamonta V, Pradubwong S, Panamonta M, Chowchuen B (2015) Global birth prevalence of orofacial clefts: A systematic review. J Med Assoc Thai 98(Suppl 7):S11–S21
2. Salari N, Darvishi N, Heydari M, Bokaee S, Darvishi F, Mohammadi M (2022) Global prevalence of cleft palate, cleft lip and cleft palate and lip: a comprehensive systematic review and meta-analysis. J Stomatol Oral Maxillofac Surg 123:110–120. https://doi.org/10.1016/j.jormas.2021.05.008
3. Azman A, Manuel AM (2020) Otological outcome in cleft lip and palate children with middle ear effusion. Int J Pediatr Otorhinolaryngol 138:110274
4. Flynn T, Möller C, Jönsson R, Lohmander A (2009) The high prevalence of otitis media with effusion in children with cleft lip and palate as compared to children without clefts. Int J Pediatr Otorhinolaryngol 73:1441–1446. https://doi.org/10.1016/j.ijporl.2009.07.015
5. Keogh T, Kei J, Driscoll C, Khan A (2010) Children with minimal conductive hearing impairment: speech comprehension in noise. Audiol Neurootol 15:27–35. https://doi.org/10.1159/000218360
6. Kuo CL, Lien CF, Chu CH, Shiao AS (2013) Otitis media with effusion in children with cleft lip and palate: A narrative review. Int J Pediatr Otorhinolaryngol 77:1403–1409. https://doi.org/10.1016/j.ijporl.2013.07.015
7. Gagné JP, Besser J, Lemke U (2017) Behavioral assessment of listening effort using a dual-task paradigm: A review. Trends Hear 21:2331216516687287. https://doi.org/10.1177/2331216516687287
8. McFadden B, Pittman A (2008) Effect of minimal hearing loss on children's ability to multitask in quiet and in noise. Lang Speech Hear Serv Sch 39:342–351. https://doi.org/10.1044/0161-1461(2008/032)
9. Hall A, Wills AK, Mahmoud O et al (2017) Centre-level variation in outcomes and treatment for otitis media with effusion and hearing loss and the association of hearing loss with developmental outcomes at ages 5 and 7 years in children with non-syndromic unilateral cleft lip and palate: The Cleft Care UK study. Part 2. Orthod Craniofac Res 20(Suppl 2):8–18. https://doi.org/10.1111/ocr.12184
10. Lohmander A, Persson C (2008) A longitudinal study of speech production in Swedish children with unilateral cleft lip and palate and two-stage palatal repair. Cleft Palate Craniofac J 45:32–41. https://doi.org/10.1597/06-123.1
11. Sell D, Southby L, Wren Y et al (2017) Centre-level variation in speech outcome and interventions, and factors associated with poor speech outcomes in 5-year-old children with non-syndromic unilateral cleft lip and palate: The Cleft Care UK study. Part 4. Orthod Craniofac Res 20(Suppl 2):27–39. https://doi.org/10.1111/ocr.12186
12. Kernahan DA (1971) The striped Y—a symbolic classification for cleft lip and palate. Plast Reconstr Surg 47:469–470. https://doi.org/10.1097/00006534-197105000-00010
13. Kurt M, Savaş D, Tarsuslu Şimşek T, Yiş U (2023) The psychometric properties of the Turkish version of the Modified Paediatric Mini Mental Scale. Child Care Health Dev 49:572–578
14. Clark JG (1981) Uses and abuses of hearing loss classification. ASHA 23:493–500

15. Northern JL, Downs MP (2014) Hearing in children, 6th edn. Plural Publishing, San Diego, pp 136–148
16. Jerger J (1970) Clinical experience with impedance audiometry. Arch Otolaryngol 92:311–324. https://doi.org/10.1001/archotol.1970.04310040005002
17. American National Standards Institute (2010) ANSI S3.6-2010: Specification for audiometers. American National Standards Institute, New York
18. Mungan S (2010) Development of a Turkish monosyllabic word recognition test for adults. Master's thesis, Dokuz Eylül University
19. Auditec (2015) Four talker noise. Audio recording. Auditec, St. Louis
20. Brungart DS (2001) Informational and energetic masking effects in the perception of two simultaneous talkers. J Acoust Soc Am 109:1101–1109. https://doi.org/10.1121/1.1345696
21. Mattys SL, Davis MH, Bradlow AR, Scott SK (2012) Speech recognition in adverse conditions: A review. Lang Cogn Process 27:953–978. https://doi.org/10.1080/01690965.2012.705006
22. Beattie RC, Barr T, Roup C (1997) Normal and hearing-impaired word recognition scores for monosyllabic words in quiet and noise. Br J Audiol 31:153–164. https://doi.org/10.3109/03005364000000018
23. Ünal-Logacev Ö, Kazanoğlu D, Balo E, Nemutlu A (2018) Cleft lip and palate assessment form: Medical history, oral-peripheral characteristics, speech problems. Turk J Plast Surg 26:156–168
24. Bauman-Waengler JA (2016) Articulatory and phonological impairments: A clinical focus, 5th edn. Pearson, Boston, pp 88–102
25. Felisati G, Di Berardino F, Maccari A, Sambataro G, Gaffuri M, Marchisio P, Meazzini MC (2022) Effects of rapid maxillary expansion on hearing loss and otitis media in cleft palate children. Eur Arch Otorhinolaryngol 279:1703–1710. https://doi.org/10.1007/s00405-021-07187-5
26. Reijers SNH, Vroegop JL, Kremer B, van der Schroeff MP (2025) The long-term effect of childhood otitis media on speech-in-noise testing at ages 9 and 13. Laryngoscope Investig Otolaryngol 10:e70077. https://doi.org/10.1002/lio2.70077
27. Spoorthi GN, Uppunda AK, Kalaiah MK, Shastri U (2025) The effect of noise on listening effort in children as measured using different methods: a systematic review and meta-analyses. Eur Arch Otorhinolaryngol 282:2855–2886. https://doi.org/10.1007/s00405-025-09232-z
28. Baker S, Wren Y, Zhao F, Cooper F (2021) Exploring the relationship between conductive hearing loss and cleft speech characteristics in children born with cleft palate. Int J Pediatr Otorhinolaryngol 148:110820. https://doi.org/10.1016/j.ijporl.2021.110820

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.