forum acusticum 2026
8-12 September · Graz, Austria

SPECTRAL, TEMPORAL, AND INTENSITY PATTERNS IN SIBILANTS ASSOCIATED WITH MALOCCLUSION

Daniel Aalto[1], Torrey M Loucks[2], Shawna DeGrange[3], Hunter Davis[3], Caroline Garcia[3], Naomi Gurevich[4], Işıl Aras[5], Romão Tormena Jr.[3], Sawsan Tabbaa[3]

[1] *Department of Communication Sciences and Disorders, University of Alberta, Canada*
[2] *Department of Communication Sciences and Disorders, Jacksonville University, USA*
[3] *School of Orthodontics, Jacksonville University, USA*
[4] *Department of Communication Sciences and Disorders, Purdue University Fort Wayne, USA*
[5] *Department of Orthodontics, University of Florida, USA*

This paper is a revised version of the authors' submitted manuscript that was peer-reviewed by at least two anonymous reviewers.

Abstract

Malocclusion impacts articulation of speech sounds, in particular fricatives. The present study compares spectral and prosodic patterns between adults with and without jaw disharmony during production of the English sibilants /s/ and /ʃ/. 36 adult participants were recruited and categorized using the Angle Classification into control (Class I) and experimental (Class II and Class III) groups. Speech recordings were obtained using an iPhone in a clinical setting. The sibilants were analyzed for first and second spectral moment (i.e., center of gravity (COG), and spectral dispersion), duration, and intensity using a linear mixed effects model. Significant differences in each acoustic variable were found between the malocclusion classes including prominent interactions with sex. Class III malocclusion exhibited more prominent acoustic spectral distortions followed by Class II. Class III females had the highest COG for /s/ and /ʃ/ along with longer duration. Class III males showed reduced acoustic distinction between the two sounds along with intensity changes. Compensatory speech strategies involving duration and intensity changes might be learned to mitigate the impact of malocclusion on speech sound distinctions. These findings underscore intricate relationships between jaw alignment and speech articulation.

Keywords: *malocclusion, speech production, sibilants, spectral moments*

Corresponding author: *aalto@ualberta.ca.*

1. Introduction

Anatomical variability of the upper airway shapes speech production and contributes to the uniqueness of a speaker's voice [1], [2], [3]. The sibilants /s/ and /ʃ/ are prominent English sounds formed by a narrow constriction in the oral tract that results in an acoustic jet with a characteristic noise spectrum that is higher for /s/ than /ʃ/. Both sounds show considerable variation in articulatory and acoustic realization due to individual anatomy, communication intent, phonetic context, and sociolinguistic factors like gender [4], [5], [6], [7]. The interaction of these factors with dentofacial disharmony can lead to acoustic distortions of speech sounds [8], [9] that is proportional to the skeletal severity of the malocclusion [10]. The long-term presence of dentofacial disharmony may further induce compensatory articulatory behaviors intended to preserve acoustic distinctions between sounds similar to compensation observed in response to sensory and mechanical perturbations of speech or changes in anatomy [11], [12], [13], [14], [15], [16], [17].

The recent large-scale study by Lathrop-Marshall and colleagues reported Class III patients had significantly higher acoustic center of gravity (COG - first spectral moment) of English consonants /t, ʧ, k, s/ than Class I controls [10]. The Class III patients also had significantly greater spectral dispersion of acoustic energy (second spectral moment) than Class I controls for /t, ʧ, k, s, ʃ/. In other words, there was significant distortion of the center and distribution of acoustic energy in the most common English stop and sibilant consonants with Class III malocclusion. Jhingree et al. reported that the sibilant contrast in COG (/s/-/ʃ/) in 11 Class III patients was perceived as significantly more distorted than in the 8 Class I control speakers [18]. Patients with Class III malocclusion were perceived as showing both auditory and visual distortion

of syllables particularly the sibilant /s/ [10].
The current study extends this line of research by predicting that acoustic spectral moments and prosodic aspects (duration and intensity) of two English sibilant sounds will be altered in adults with Class II and Class III malocclusion compared to Class I malocclusion. The spectral changes are more directly related to jaw misalignment while prosodic differences point towards learned articulatory compensation.

2. Materials and methods

Thirty-six adults (24 female) were recruited from patients seeking orthodontic evaluation at the Jacksonville University School of Orthodontics (JUSO). This study protocol and methods were approved by the Jacksonville University Institutional Review Board #2023-048. Each participant provided informed and voluntary written consent. The inclusion criteria were: 1) age 18 - 65 years; 2) Class I, Class II, or Class III malocclusion based on Angle Classification; 3) no previous orthodontic treatment; and 4) American English as dominant language. The exclusion criteria were: 1) history of TMJ disorder; 2) any craniofacial deformity; and 3) current pregnancy. The distribution of patients according to malocclusion status, age and sex is shown in 1.

Table 1: Participant characteristics listed by independent variables. The total sample size is 36 participants.

Independent variable	Level	Frequency
Sex	Female	24
	Male	12
Malocclusion Class	I	19
	II	6
	III	11

The orthodontic assessment was conducted by a board-certified dentist in the JUSO graduate program under the supervision of a board-certified orthodontist. The assessment included exposure to a panoramic radiograph and a CBCT 3D scan (to extract a lateral cephalogram) including an intra-oral iTero scan. Malocclusion class was determined from the iTero scan using the Angle Classification, which is based on the first molar occlusion in the anteroposterior direction [19].
Each participant was seated in a quiet room and orally read single words displayed on a computer screen at a rate of approximately one word every 3 seconds (308 words presented in a randomized order). The task took approximately 25 minutes including four brief rest intervals. The participant's speech was recorded using a custom SpeechTracker iPhone app with the microphone consistently placed 55cm from the speaker with a sampling rate of 44.1 kHz [20]. The words are from the FITI word list that comprehensively covers most all phonetic contexts in Standard American English [21]. The recordings were immediately transferred to a secure computer via cable connection and erased from the experiment iPhone, which was never connected to a network or Wi-Fi so patient records remained secure and anonymous. Each recording was imported into Praat 4.6 [22] and analyzed by a speech scientist and research assistant.
The target stimuli were the initial /s/ from the words - 'say', 'searched', 'subject', and 'suburb' - and the initial /ʃ/ from the words - 'shoulder', 'show', and 'shrug' for a total of 4 /s/ tokens and 3 /ʃ/ tokens per participant). The stressed word-initial prevocalic position is optimal because it is particularly prominent and thus important for intelligibility [23]. The words were identified in the acoustic record and displayed with the spectrographic record. The onset and offset of each sibilant were demarcated and annotated by both raters for inter-rater reliability. The onset and offset times from both raters were compared using a 2*2 Anova to test if the onset and offset times from the two raters differed. No differences were found between the raters for onset or offset times so the annotation record from the Speech Scientist was retained for all subsequent analyses. Four dependent variables were used to characterize the sibilants: 1) center of gravity (COG) also known as centroid frequency or first spectral moment; 2) spectral dispersion also known as the second spectral moment; 3) duration, defined as the time interval between the onset marker and the offset marker (ms); 4) acoustic intensity. The spectral moments were extracted from the midpoint of the demarcated sound using a customized Praat script [24].
Four mixed effects models corresponding to the four dependent variables (COG, spectral dispersion, duration, intensity) were fitted separately using lme4 and lmerTest packages in R [25], [26]. Each model had three independent variables with all the interaction terms included: malocclusion class (Class I, Class II, Class III), sex (female, male), and sound (/s/ or /ʃ/) including random intercepts for participants and words [27]. Residuals were visually inspected using scatter plots (heteroscedasticity) and quantile-quantile plots (normality) for violations of model assumptions.

3. Results

3.1. Spectral Variables

The mixed effect results for COG are presented in Table 2 and visualized in Figure 1. The intercept for COG (7293 ± 594 Hz) was significantly altered by malocclusion type (+925 Hz, Class III, $t = 2.9$, $p = .005$), sex (-1257 Hz, male, $t = 4.0$, $p < .001$), and sound (-2470 Hz for /ʃ/, $t = 10.0$, $p < .001$) along with prominent interactions. Class III female speakers had higher overall COG but maintained the acoustic distinction between /s/ and /ʃ/. Male speakers produced a less distinct /s-ʃ/ contrast in COG compared to female speakers (+554 Hz, $t = 2.1$, $p = .03$) with this distinction being particularly reduced in Class III males (+1427 Hz, $t = 3.3$, $p = .0009$). Descriptively, the COG of /ʃ/ for Class II males was reduced com-

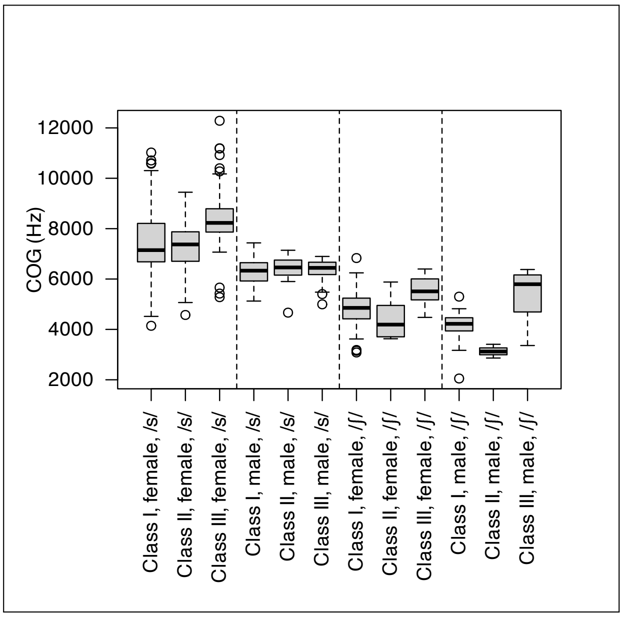

Figure 1: CoG distributions across speakers are illustrated with boxplots grouped by malocclusion type, speaker sex, and phoneme.

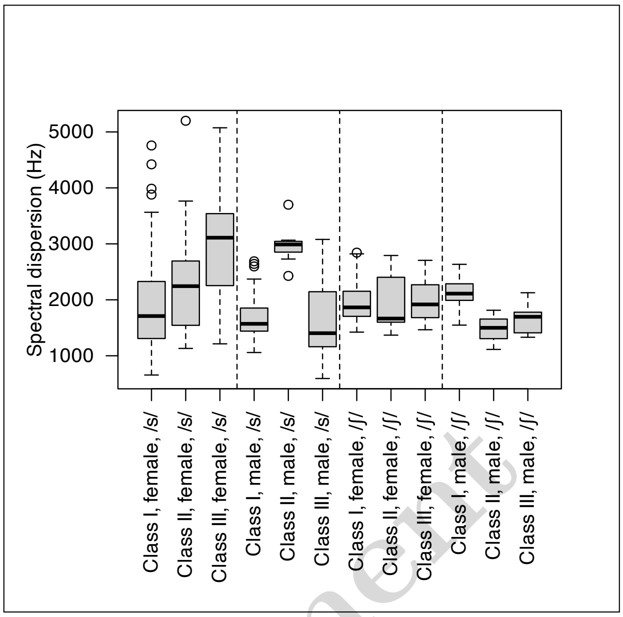

Figure 2: Spectral distortion distributions across speakers are illustrated with boxplots grouped by malocclusion type, speaker sex, and phoneme.

pared to other groups but the three-way interaction was not significant (-905 Hz, $t = 1.3$, $p = 0.16$).

The spectral dispersion data in Figure 3 suggests spread of acoustic energy was greater for /s/ than /ʃ/ as confirmed by Levene's test ($F = 5.7$, $p = 0.017$). As shown in Table 2, the baseline spectral dispersion (intercept: 2096 ± 468 Hz) was significantly affected by malocclusion along with several significant interactions. Class III female speakers had higher spectral dispersion for /s/ but not for /ʃ/ (+1008 Hz, $t = 4.3$, $p < .001$). In contrast, Class II males had a reduced dispersion for /ʃ/ (-1538 Hz, $t = 4.0$, $p < .001$) compared to /s/.

3.2. Prosodic Variables

Malocclusion impacted sound duration with the patterns modulated by sex and sound (Figure 4; Table 3). The intercept ($\pm$ standard deviation) for duration 149 ± 14 ms was modulated by malocclusion (+47 ms for Class III; $t = 5.4$, $p < .001$) and sex (+19 ms for male; $t = 2.2$, $p = .03$). The significant Class by Sex interaction (-63 ms Class III x female, $t = -4.4$, $p < .001$; -23 ms malocclusion III x sound /ʃ/, t=-2.1, p=.04) reveals that Class III females had significantly longer duration for both sounds, but especially for /s/. The acoustic intensity was greater for /ʃ/ than for /s/ (+2.5 dB, $t = 2.4$, $p = 0.04$) but other main effects were not significant (Figure 5). Interaction terms showed /ʃ/ was modulated by malocclusion class and sex (Table 1). Class III speakers had higher intensity /ʃ/ (+3.5 dB, $t = 4.1$, $p < .001$) while male speakers had lower intensity /ʃ/ compared to females (-2.3 dB, $t = 2.7$, $p = 0.008$).

4. Discussion

We compared spectral and prosodic parameters of English sibilants between three malocclusion classes. As predicted, Class II and Class III adults showed significantly greater spectral distortion of sibilants than Class I adults. In Class III patients, acoustic distortion was observed in both COG and spectral dispersion, whereas Class II showed deviations in spectral dispersion compared with Class I speakers. COG provides the primary distinction between the two sibilants, as /s/ has a characteristically higher COG than /ʃ/. The higher COG of /s/ can be related to a more anterior tongue position along the alveolar ridge, resulting in a small cavity in front of the noise source (i.e., the contact between tongue tip and palate) and hence a higher resonant frequency [28]. Spectral dispersion provides additional information of articulatory behavior that indicates the regularity of the signal-to-noise ratio in acoustic phonetics. The distortion observed in both sibilants, therefore, indicates jaw misalignment impacted anterior and midpalatal articulatory postures and changes the noise profile of these basic sounds. This distortion pattern conforms to recent reports [10], [29], [30] that jaw disharmony compromises a speaker's ability to form the articulatory occlusions that underlie the intelligibility of basic speech sounds. Our spectral results particularly align with the recent large-scale report by Lathrop-Marshall et al. whose patients with Class III malocclusion had significantly higher COG for /s, k, t, ʧ/ than control speakers.

The current results extend these findings by showing prominent interactions between Malocclusion Class and sex. Class III females had significantly higher COG for both sibilants, which was most distinctive

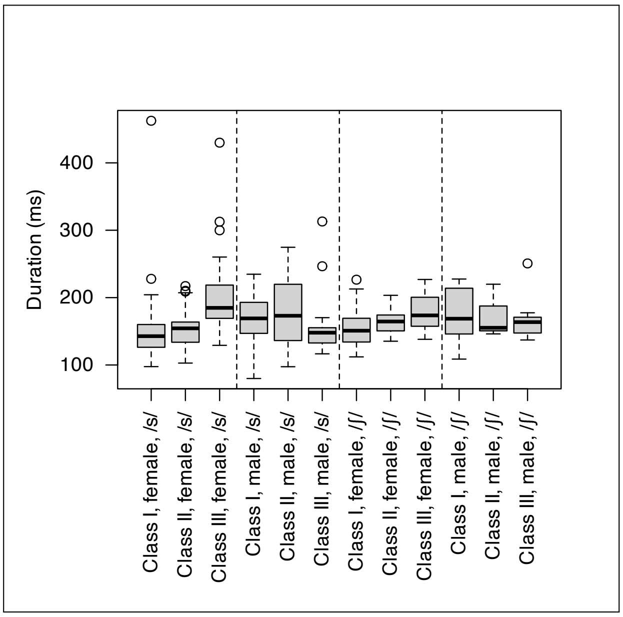

Figure 3: Spectral distortion distributions across speakers are illustrated with boxplots grouped by malocclusion type, speaker sex, and phoneme.

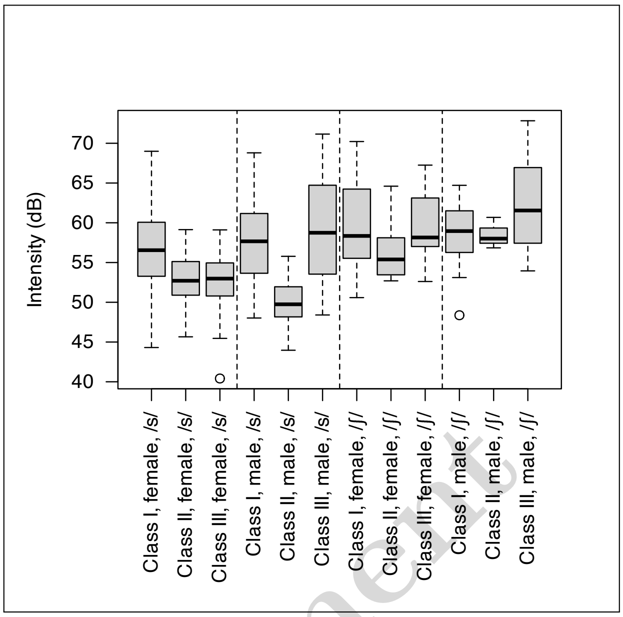

Figure 4: Spectral distortion distributions across speakers are illustrated with boxplots grouped by malocclusion type, speaker sex, and phoneme.

for /s/. A wider spectral dispersion of acoustic energy was also present for /s/ production by Class III females. A considerable body of literature confirms that female adult speakers have higher COG for sibilants than males due to social and anatomical factors [7]. Considering that Class III females show an even more exaggerated COG could indicate their underbite results in an atypically small anterior cavity for the /s/. Alternatively, it could be a compensatory strategy to enhance the acoustic difference between /s/ and /ʃ/. While the relationship between the spectral dispersion and articulation is not well characterized, systematic deviations in these values reflect the combined impact of articulation and anatomy.

The male patients also showed acoustic distortions that contributed to the significant interactions. The COG of /s/ in Class III males did not increase, rather their COG for /ʃ/ increased leading to a reduction in the /s-ʃ/ acoustic distance. The /s-ʃ/ distinction is considered a primary contributor to speech discrimination with adults showing a difference in COG greater than 1000 Hz [31]. Recently, Jhingree and colleagues reported that reduced /s-ʃ/ distinctions in malocclusion were associated with higher listener perceptions of acoustic distortion [18]. We suggest that underbite in the Class III males contributed to a decreased acoustic /s-ʃ/ distinction of $<$ 1000 Hz that could interfere with consistent sound identification. Their underbite could limit lingual range of motion and shift /s/ and /ʃ/ occlusions into a shared articulatory space. Class II males maintained the /s-ʃ/ distinction, but there was a significant decrease in their COG of /ʃ/. These speakers potentially shift their tongue body posteriorly to produce the distinction. The other classes maintained a clear /s-ʃ/ distinction in spectral acoustic terms.

Our other critical findings related to sex differences were identified in the prosodic variables. The Class III females had the longest acoustic duration compared to the Class I and Class II patients, but particularly for /s/. Increased duration could be a compensatory strategy used by the female speakers to maintain the distinctiveness of these sounds but would also extra effort. The lower COG of the class II males may also not be fully determined by overbite but also by a functional adjustment of exaggerated tongue body retraction for /ʃ/ to promote the /s-ʃ/ distinction. Preliminary evidence from acoustic intensity indicated that Class III speakers almost uniformly increased the intensity of the sibilants /ʃ/ relative to the other groups. Anatomical differences due to malocclusion contribute to the acoustic distortion, but there also appear to be compensatory adjustments patients use to distinguish the sibilants reflected in the duration patterns. Other compensatory strategies could involve articulatory postures, but this needs to be investigated with paradigms that manipulate place of articulation and therapeutically induced articulatory compensation.

This research initiative is continuing as we build a larger database of orthodontic patients with a focus on how sex interacts with dental disharmony. Immediate attention is focused on prosodic changes (duration and intensity) as possible functional compensations but also on how speech kinematic patterns change with malocclusion. The results regarding the six Class II participants should be interpreted cautiously and confirmed with a larger cohort. We used a relatively

Table 2: Mixed effects linear regression results for centre of gravity (CoG) and spectral dispersion (SD).

	CoG (Hz)	SD (Hz)
Intercept	7293	2096
Class II	-164	+398
Class III	+925	+1008
Male	-1257	-209
/ʃ/	-2470	-120
II:Male	+236	+889
III:Male	-861	-1053
II:/ʃ/	-155	-408
III:/ʃ/	-227	-1000
Male: /ʃ/	+554	+369
II: Male : /ʃ/	-905	-1538
III: Male : /ʃ/	+1427	+557
Participants (σ)	594	468
Words (σ)	245	217
Residuals (σ)	804	475

Table 3: Mixed effects linear regression results for duration and intensity.

	Dur (ms)	Intensity (dB)
Intercept	149	57.3
Class II	+4	-3.5
Class III	+47	-3.7
Male	+19	1.0
/ʃ/	+6	2.5
II:Male	+7	-4.1
III:Male	-63	5.8
II:/ʃ/	+7	0.2
III:/ʃ/	-23	3.5
Male: /ʃ/	-0.7	-2.3
II: Male : /ʃ/	-18	7.2
III: Male : /ʃ/	-32	-2.0
Participants (σ)	14	4.1
Words (σ)	8	1.2
Residuals (σ)	34	2.7

simple molar classification but other methods, such as Wits appraisal, account for skeletal and dental parameters, specifically the functional occlusal plane. Applying a different classification could result in a different distribution of patients across classes but can also contribute quantitative anatomical predictors of acoustic distortion [10]. Another consideration is the spacing of anterior or posterior teeth, which may compromise the spectral measures, and often accompanies malocclusion. Different approaches to appraisal of spacing are used, but a binary spacing analysis with malocclusion class could help to understand acoustic implications.

5. Conclusion

Class II and III malocclusion are associated with significant distortion of the spectral characteristics of English sibilants in adults. Acoustic distortion differs with malocclusion and sex. Duration and intensity patterns suggest a compensatory strategy. The influence of orthodontic intervention on immediate changes in tongue-jaw coordination and long-term compensation need to be evaluated in adult and pediatric populations.

References

[1] E. B. Holmberg, J. S. Perkell, R. E. Hillman, and C. Gress, "Individual variation in measures of voice," *Phonetica*, vol. 51, no. 1-3, pp. 30–37, 1994.

[2] G. Fant, *Acoustic theory of speech production*. The Hague: Mouton, 1960.

[3] B. Munson, "Variability in /s/ production in children and adults: Evidence from dynamic measures of spectral mean," *J Speech Lang Hear Res*, vol. 47, no. 1, pp. 58–69, 2004.

[4] M. Stone, A. D. Gomez, J. Zhuo, A. L. Tchouaga, and J. L. Prince, "Quantifying tongue tip shape in apical and laminal /s/: Contributions of palate shape," *Journal of Speech, Language, and Hearing Research*, vol. 62, no. 9, pp. 3149–3159, 2019.

[5] B. Lindblom, "Explaining phonetic variation: A sketch of the H&H theory," in *Speech production and speech modelling*, Dordrecht: Springer, 1990, pp. 403–439.

[6] N. Harandi, J. Woo, M. Stone, R. Abugharbieh, and S. Fels, "Variability in muscle activation of simple speech motions: A biomechanical modeling approach," *The Journal of the Acoustical Society of America*, vol. 14, no. 4, pp. 2579–2590, 2017.

[7] M. Sonderegger, J. Stuart-Smith, and J. Mielke, "How variable are english sibilants?" In *Proc. of the 20th International Congress of Phonetic Sciences - ICPhS 2023*, 2023, pp. 3196–3200.

[8] S. Görgülü, D. Sağdıç, E. Akin, S. Karaçay, and N. Bulakbası, "Tongue movements in patients with skeletal class iii malocclusions evaluated with real-time balanced turbo field echo cine magnetic resonance imaging," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 139, no. 5, pp. 405–414, 2011.

[9] L. D. Vallino and B. Tompson, "Perceptual characteristics of consonant errors associated with malocclusion," *Journal of Oral and Maxillofacial Surgery*, vol. 51, no. 8, pp. 850–856, 1993.

[10] H. Lathrop-Marshall et al., "Orthognathic speech pathology: Impacts of class III malocclusion on speech," *European Journal of Orthodontics*, vol. 44, no. 3, pp. 340–351, 2022.

[11] S. L. Hamlet, B. L. Cullison, and M. Stone, "Physiological control of sibilant duration: Insights afforded by speech compensation to dental prostheses," *The Journal of the Acoustical Society of America*, vol. 65, no. 5, pp. 1276–1285, 1979.

[12] D. H. McFarland and S. R. Baum, "Incomplete compensation to articulatory perturbation," *The Journal of the Acoustical Society of America*, vol. 97, no. 3, pp. 1865–1873, 1995.

[13] M. Niemi, J.-P. Laaksonen, S. Ojala, O. Aaltonen, and R.-P. Happonen, "Effects of transitory lingual nerve impairment on speech: An acoustic study of sibilant sound/s," *International Journal of Oral and Maxillofacial Surgery*, vol. 35, no. 10, pp. 920–923, 2006.

[14] M. Thibeault, L. Ménard, S. R. Baum, G. Richard, and D. H. McFarland, "Articulatory and acoustic adaptation to palatal perturbation," *The Journal of the Acoustical Society of America*, vol. 129, no. 4, pp. 2112–2120, 2011.

[15] G. d. Boer and D. Aalto, "Longitudinal acoustic analysis of/s/in head and neck cancer patients treated with surgery," *Journal of Speech, Language, and Hearing Research*, vol. 69, no. 5, pp. 2007–2019, 2026.

[16] A. S. Lee, T. L. Whitehill, V. Ciocca, and N. Samman, "Acoustic and perceptual analysis of the sibilant sound/s/before and after orthognathic surgery," *Journal of Oral and Maxillofacial surgery*, vol. 60, no. 4, pp. 364–372, 2002.

[17] I. D. Vilanova et al., "Impact of orthognathic surgery on voice and speech: A systematic review and meta-analysis," *European Journal of Orthodontics*, vol. 45, no. 6, pp. 747–763, 2023.

[18] S. Jhingree et al., "Validating spectral moment analysis as a quantitative measure of speech distortions in speakers with class iii malocclusions," *Perspectives of the ASHA Special Interest Groups*, vol. 7, no. 3, pp. 728–740, 2022.

[19] M. I. Katz, "Angle classification revisited 2: A modified angle classification," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 102, no. 3, pp. 277–284, 1992.

[20] K. Armstrong, C. Kincade, M. Osswald, J. Rieger, and D. Aalto, "A prototype smartphone jaw tracking application to quantitatively model tooth contact," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 11, no. 7, pp. 2713–2724, 2024.

[21] N. Gurevich and H. Kim, "Development of novel speech stimuli with phonetic coverage and phonemic balance," *Perspectives of the ASHA Special Interest Groups*, vol. 8, no. 2, pp. 424–437, 2023.

[22] P. Boersma and V. V. Heuven, "Speak and unspeak with praat," *Glot International*, vol. 5, no. 9/10, pp. 341–347, 2001.

[23] N. Gurevich and D. Aalto, "Frequency and context prominence effects on english allophone perception: Identification of/r/and/θ," in *Proc. of Meetings on Acoustics*, Acoustical Society of America, vol. 55, 2024, pp. 1–6.

[24] C. Dicanio. "Spectral moments of fricative spectra script in Praat." cdicanio/scripts.html. Updated 2021. [Online]. Available: `https://www.acsu.buffalo.edu/`.

[25] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[26] A. A. Kuznetsova and P. B. B. R. Christensen, "LmerTest package: Tests in linear mixed effects models," *Journal of Statistical Software*, vol. 82, no. 13, pp. 1–26, 2017.

[27] R. H. Baayen, D. J. Davidson, and D. M. Bates, "Mixed-effects modeling with crossed random effects for subjects and items," *Journal of memory and language*, vol. 59, no. 4, pp. 390–412, 2008.

[28] A. Jongman, R. Wayland, and S. Wong, "Acoustic characteristics of english fricatives," *The Journal of the Acoustical Society of America*, vol. 108, no. 3, pp. 1252–1263, 2000.

[29] M. M. B. Keyser et al., "Impacts of skeletal anterior open bite malocclusion on speech," *FACE*, vol. 3, no. 2, pp. 339–349, 2022.

[30] S. Oliver et al., "Impacts of anterior-posterior jaw disproportions on speech of dentofacial disharmony patients," *European Journal of Orthodontics*, vol. 45, no. 1, pp. 1–10, 2022.

[31] J. S. Perkell et al., "The distinctness of speakers' /s/-/ʃ/ contrast is related to their auditory discrimination and use of an articulatory saturation effect," *Journal of Speech Language and Hearing Research*, vol. 47, no. 6, pp. 1259–1269, 2004.