Anatomical correlates of articulatory ranges of motion: An EMA study

Daniel Friedrichs[1,2]*, Axel G. Ekström*[3]*, Sam Kirkham*[4]*, Steven Moran*[2,5]
[1] *Department of Computational Linguistics, University of Zürich, Switzerland*
[2] *Institute of Biology, University of Neuchâtel, Switzerland*
[3] *Speech, Music & Hearing, KTH Royal Institute of Technology, Sweden*
[4] *Department of Linguistics & English Language, Lancaster University, UK*
[5] *Department of Anthropology, University of Miami, US*
daniel.friedrichs@ucl.ac.uk, axeleks@kth.se, s.kirkham@lancaster.ac.uk,
steven.moran@unine.ch

Abstract

This study investigates the relationship between anatomical features and articulatory ranges of motion in speech production using electromagnetic articulography (EMA) data. Utilizing canonical correlation analysis (CCA), we identified significant associations between anatomical dimensions – such as vocal tract length and mandible length – and the movements of the tongue, jaw, and lips. The results indicate that longer vocal tracts and mandibles are linked to greater vertical tongue movements but smaller vertical mandibular movements. Additionally, short mandible lengths are associated with extended horizontal lower lip movements, suggesting a form of biomechanical adaptation. Furthermore, the analysis revealed an association between the tongue's swallowing range and various articulatory variables. These findings highlight the role of anatomical structures in shaping articulatory patterns, offering insights into biomechanical constraints and adaptations in speech production.

Introduction

The study of how vocal tract anatomy influences speech production has garnered significant attention in the field of phonetics. Research suggests that individual anatomical variations, such as vocal tract length, palate shape, and mandibular dimensions, are critical in shaping articulatory patterns (e.g., Johnson, 2023; Lammert, Proctor, & Narayanan, 2013; Simpson, 2001).

Variations in anatomy have been shown to constrain strategies for vowel and consonant production and could potentially even bias the realization of suprasegmental features. As early as the 1970s, Lindblom and Sundberg (1971) presented an articulatory model demonstrating the acoustical consequences of variations in lip, tongue, jaw, and larynx movements, underscoring the synergistic role of jaw position in vowel production. More recently, Brunner, Fuchs, and Perrier (2009) demonstrated that palate shape significantly influences articulatory behavior, leading to variations in tongue placement and movement. Specifically, they found that speakers with flat palates reduce their articulatory variability to maintain consistent acoustic output. Fuchs, Winkler, and Perrier (2008) explored how vocal tract geometry impacts vowel production strategies, finding that speakers with a longer pharynx produce larger displacements between low back and high front vowels. Honda, Maeda, Hashi, Dembowski, and Westbury (1996) and Winkler, Fuchs, Perrier, and Tiede (2011) examined the consequences of anatomical constraints on articulatory variability, showing that anatomical limitations can lead to distinct speech patterns among individuals. Perkell, Matthies, Svirsky, and Jordan (1995) emphasized how goal-based speech motor control varies with anatomical differences, suggesting that speakers with different anatomical features adopt unique strategies to achieve phonetic targets. Weirich and Fuchs (2013) showed that palate morphology could affect the realization of phonemic contrasts, influencing how specific sounds are articulated and perceived. Most recently, Johnson (2023) found that speakers with larger vocal tracts exhibit greater vertical tongue motion during vowel production, whereas those with smaller vocal tracts show increased vertical jaw movement. These findings suggest a complex interaction between anatomical features and speech production mechanisms.

Our study utilizes electromagnetic articulography (EMA) to investigate how various anatomical dimensions affect articulatory dynamics. By correlating the vertical and horizontal ranges of tongue, lip, and jaw movements with anatomical features, we aim to elucidate the role of anatomical variation in shaping phonetic diversity and provide new insights into the biomechanical basis of speech production.

Methods

Participants

Fifteen native German speakers (7f, 8m) participated in this study. Participants were aged between 20 and 30 years (mean age: 23.57 years, SD = 2.98) and had no history of speech, hearing, or neurological disorders. All participants provided informed consent prior to their participation in the study.

Data Collection and Processing

Articulatory data were recorded at a sampling rate of 1250 Hz using a Carstens AG-501 EMA system. Sensors were attached midsagitally to various points, including the vermilion borders of the upper and lower lips, upper and lower incisors, tongue dorsum, tongue body, and tongue tip. To determine the precise contact points for tongue sensor placement, we applied a line of ink midsagittally to the palate and asked participants to produce [t] and [k]. This resulted in contact points on the tongue tip for [t] and the tongue dorsum for [k]. The tongue body sensor was positioned midway between the [t] and [k] contact points. The sensors were then glued using Epiglu and Ketac cement. Reference sensors were placed on the bite plane for calibration and on the mastoid processes and nasion to correct for head movement. The raw EMA data were processed using Carstens' built-in Calcpos and Normpos applications and custom MATLAB scripts.

Speech Material

Participants produced a range of speech utterances, including trisyllabic nonsense words with balanced articulatory trajectories, e.g., [piːtaːkuː], [kuːtaːpiː], [taːkuːpiː], at comfortable (*N*=180) and maximally fast (*N*=120) production rates; minimal pairs containing all long German vowels (*N*=16), words with a schwa vowel in the final syllable (*N*=6); all long German vowels in isolation (*N*=16); and recordings of a read passage ("Nordwind und die Sonne") (*N*=2). Analyses were conducted on entire speech sequences rather than on annotated and labeled data, incorporating 500ms before and after the acoustic onsets and offsets to ensure the inclusion of full articulatory movements.

Anatomical Measurements

Table 1 shows the anatomical measurements analyzed for each participant. A digital caliper was used to measure distances between cephalometric landmarks on the jaw and mouth. Vocal tract length (VTL) was estimated using different acoustic-to-anatomical conversion techniques based on formant frequencies (Lammert & Narayanan, 2015; Flego, 2018). A mean VTL from these estimates was calculated for each talker. Talkers were also recorded during tongue protrusion and swallowing trials. The maximum tongue protrusion and swallowing were quantified by calculating the Euclidean distance between the furthest points of movement for the tongue dorsum, body, and tip. The mean of these distances was then calculated to represent the overall measure for each talker.

Table 1: Anatomical Measures

Jaw	***Mandible Length:*** Distance between condylion (Co) and gnathion (Gn). ***Mandible Perimeter:*** Overall side length of the triangle with corners at Co, gonion (Go), and Gn. ***Maximum Jaw Opening:*** Vertical distance between upper and lower incisors at maximum jaw opening.
Mouth	***Maximum Vertical Mouth Opening:*** Distance between mid-points of the upper and lower lips at maximum jaw opening. ***Maximum Horizontal Mouth Opening:*** Distance between the lip corners at maximum jaw opening.
Tongue	***Maximum Tongue Protrusion*** ***Swallowing***
VT	***Vocal Tract Length:*** Estimated based on formant frequencies.

Articulatory Measurements

In this study, we focused on a range of motion analysis. Therefore, vertical and horizontal ranges of motion were calculated for all sensors attached to the tongue, lips, and jaw (see Table 2), considering only data between the 10th and 90th percentiles to exclude outliers.

Table 2: Articulatory Parameters*

Jaw	*Mandibular motion* Sensor position: lower incisor (LI)
Lips	*Upper lip (UL) motion* *Lower lip (LL) motion*
Tongue	*Tongue tip (TT) motion* *Tongue body (TB) motion* *Tongue dorsum (TD) motion*

*Vertical (v) and horizontal (h) motions were analyzed for each sensor.

Data Analysis

Canonical Correlation Analysis (CCA) with L2 regularization and cross-validation was performed to investigate the relationships between anatomical and articulatory properties. The method follows the approach described by Johnson (2023). In this analysis, positive loadings indicate that as the anatomical measure increases, the related articulatory canonical variable also increases. Conversely, negative loadings indicate that as the anatomical measure increases, the related articulatory canonical variable decreases. The loadings also highlight which individual variables contribute the most to the canonical variables, with higher absolute values indicating stronger contributions.

Results

The analysis revealed two significant canonical correlations: the first canonical correlation (CC1) was 0.8972, and the second canonical correlation (CC2) was 0.7854, indicating a strong relationship between anatomical and articulatory properties. These values indicate a strong relationship between the sets of anatomical and articulatory properties. Wilks' lambda was used to test the significance of the canonical correlations. For CC1, the test yielded a Wilks' lambda of 0.0148, with a chi-square statistic of 1150.7 (df = 96, $p < 0.001$). For CC2, Wilks' lambda was 0.0702, with a chi-square statistic of 730.5 (df = 77, $p < 0.001$). These results confirm a statistically significant relationship between the anatomical and articulatory properties. Canonical loadings were examined to understand the contributions of the original variables to the canonical variables. For the anatomical properties, the highest loadings for CC1 were observed for

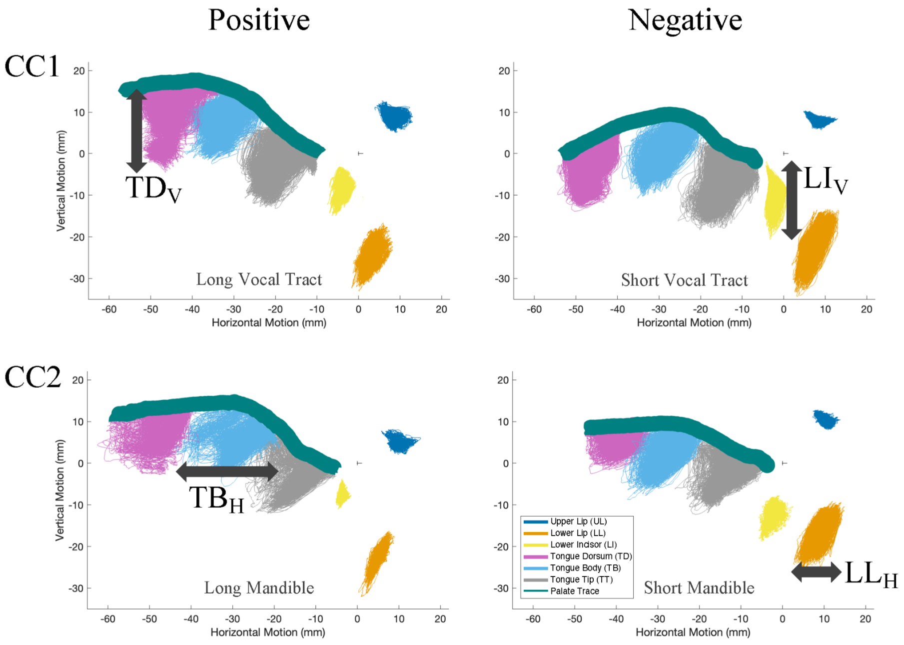

Figure 1. EMA Trajectories of Four Talkers Demonstrating Variations in Vocal Tract Length, Mandible Length, and Canonical Correlation Scores. The top row shows articulatory patterns for the first canonical correlation (CC1). The left panel illustrates the pattern for a talker with a long vocal tract (17.05 cm) and a positive CC1 loading, while the right panel shows the pattern for a talker with a short vocal tract (13.26 cm) and a negative CC1 loading. The bottom row displays articulatory patterns for the second canonical correlation (CC2). The left panel illustrates the pattern for a talker with a long mandible (13.85 cm) and a positive CC2 loading, and the right panel shows the pattern for a talker with a short mandible (10.11 cm) and a negative CC2 loading. Arrows indicate the extended ranges of motion of the respective articulators. The positional data were aligned such that the sensor attached to the upper incisor is positioned at the (0,0) coordinate.

vocal tract length, mandible length, and maximum vertical jaw opening. For the CC2, the highest loadings were observed for vocal tract length, maximum vertical jaw opening, and mandible length. These results suggest that these anatomical features play a significant role in shaping the articulatory movements. For the articulatory properties, CC1 had high positive loadings for tongue back vertical range and tongue back horizontal range, while negative loadings were found for mandible vertical range and mandible horizontal range. This indicates that the tongue's movements, both vertically and horizontally, are strongly related to the anatomical features highlighted, while mandible movements are inversely related. For CC2, high positive loadings were found for tongue body vertical range and tongue body horizontal range, whereas negative loadings were observed for lower lip vertical range and lower lip horizontal range. This suggests that the vertical and horizontal movements of the tongue blade are positively associated with the anatomical variables, whereas the lower lip movements show a negative association.

Additionally, our analysis showed that the tongue’s swallowing motion range had noteworthy contributions to both CC1 and CC2, indicating its involvement in the broader relationship between anatomical and articulatory properties.

Fig. 1 shows EMA trajectories illustrating the articulatory kinematics for four speakers with varying vocal tract and mandible length, illustrating the main findings of CC1 and CC2.

Discussion

The findings of this study provide evidence that anatomical features such as vocal tract length and mandible size significantly influence articulatory patterns in speech production. The strong canonical correlations observed in our results underscore the crucial role these anatomical characteristics play in shaping tongue, lip, and jaw movements during speech. Consistent with Johnson (2023) and Johnson, Ladefoged, and Lindau (1993), we found that larger vocal tracts exhibit greater vertical tongue movement, while smaller vocal tracts show increased vertical jaw movement. This supports the idea that anatomical constraints not only shape individual articulatory patterns but also influence the overall biomechanical strategies employed during speech. However, it must be noted that VTL was estimated from formant frequencies, which inherently links VTL and tongue movement. This methodological aspect suggests a degree of interdependence between VTL and the articulatory movements. Therefore, incorporating direct anatomical measurements of VTL, such as imaging techniques or pharyngometry, in future research could further validate (or refute) these findings and thus provide additional insights. The significant associations between jaw length and tongue movement reinforce the hypothesis that larger vocal tracts and greater tongue mobility might contribute to more expansive vowel spaces, potentially leading to richer vowel inventories. This aligns with the findings of Fuchs, Winkler, and Perrier (2008), who demonstrate that speakers with longer pharynxes produce larger articulatory displacements between certain vowel pairs, suggesting that a larger vocal tract allows for greater articulatory flexibility and thus more expansive vowel spaces. Conversely, smaller vocal tracts with larger jaw movements could potentially affect syllable structure complexity due to the relationship between jaw opening and speech amplitude envelope patterns. MacNeilage (1998) suggested that jaw movements are crucial to the temporal organization of speech, which varies with vocal tract size and dynamics. In smaller vocal tracts, larger jaw movements may require adjustments in speech timing and sequencing, leading to more complex syllable structures. Speakers might use additional articulatory features, such as increased lip use, to maintain clarity. Limited space in smaller vocal tracts can impact consonant production, necessitating exaggerated or precise articulatory gestures to ensure clear differentiation and achieve the desired acoustic outcomes.

One interesting finding is the relatively strong positive association between the swallowing range of the tongue and various articulatory variables. This suggests that the dynamics of swallowing might play a role in speech articulation, indicating a biomechanical linkage between these two processes.

Conclusions

This study demonstrates that vocal tract anatomy, particularly vocal tract length and mandible size, significantly influences articulatory patterns in speech production. Future research should also consider cross-linguistic comparisons to determine if the observed relationships between anatomical features and articulatory patterns hold across different languages. This would provide valuable insights into the universality versus language-specificity of these biomechanical constraints and adaptations and their implications for phonetic and phonological theory. Overall, this study underscores the importance of considering anatomical factors in speech production research and highlights the complex interplay between anatomy and articulation in shaping human speech.

Acknowledgments

This study was supported by the NCCR Evolving Language, funded through the Swiss National Science Foundation (SNSF) Agreement #51NF40_18088. DF and SM received further funding through the SNSF (Grant No. PCEFP1_186841). We also thank Tiena Danner for useful feedback. The results of this work and the tools used will be made more widely accessible through the national infrastructure Språkbanken Tal under funding from the Swedish Research Council (2017–00626).

References

Brunner, J., Fuchs, S., & Perrier, P. (2009). On the relationship between palate shape and articulatory behavior. *Journal of the Acoustical Society of America,* 125(6), 3936-3949. doi: 10.1121/1.3125313

Flego, S. (2018). Estimating vocal tract length by minimizing non-uniformity of cross-sectional area. *Proceedings of Meetings on Acoustics,* 35(1), 060003. doi: 10.1121/2.0001000

Fuchs, S., Winkler, R., & Perrier, P. (2008). Do speakers' vocal tract geometries shape their articulatory vowel space? In *8th International Seminar on Speech Production* (pp. 333-336). Strasbourg, France.

Honda, K., Maeda, S., Hashi, M., Dembowski, J., & Westbury, J. (1996). Human palate and related structures: Their articulatory consequences. In *Proceedings of the International Conference on Spoken Language Processing (ICSLP),* Vol. 2, pp. 784-787. doi: 10.21437/ICSLP.1996

Johnson, K. (2023). Individual differences in speech production: What is "phonetic substance"? In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the International Congress of Phonetic Sciences 2023* (pp. 1102–1106). Guarant International.

Johnson, K., Ladefoged, P., & Lindau, M. (1993). Individual differences in vowel production. *Journal of the Acoustical Society of America,* 94, 701–714. doi: 10.1121/1.406887

Lammert, A., & Narayanan, S. (2015). On short-time estimation of vocal tract length from formant frequencies. *PLoS ONE,* 10.

Lammert, A., Proctor, M., & Narayanan, S. (2013). Morphological variation in the adult hard palate and posterior pharyngeal wall. *Journal of Speech, Language, and Hearing Research,* 56(2), 521-530. doi: 10.1371/journal.pone.0132193

Lindblom, B. E. F., & Sundberg, J. E. F. (1971). Acoustical consequences of lip, tongue, jaw, and larynx movement. *Journal of the Acoustical Society of America,* 50(4), 1167-1179. doi: 10.1121/1.1912750

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences,* 21(4), 499-546. doi: 10.1017/s0140525x98001265

Perkell, J. S., Matthies, M. L., Svirsky, M. A., & Jordan, M. I. (1995). Goal-based speech motor control: A theoretical framework and some preliminary data. *Journal of Phonetics,* 23(1-2), 23-35. doi: 10.1016/S0095-4470(95)80030-1

Simpson, A. P. (2001). Dynamic consequences of differences in male and female vocal tract dimensions. *Journal of the Acoustical Society of America,* 109(5), 2153-2164. doi: 10.1121/1.1356020

Weirich, M., & Fuchs, S. (2013). Palate morphology can influence speaker-specific realizations of phonemic contrasts. *Journal of Speech, Language, and Hearing Research,* 56(3), 894-908. doi: 10.1044/1092-4388(2013/12-0217)

Winkler, R., Fuchs, S., Perrier, P., & Tiede, M. (2011). Biomechanical tongue models: An approach to studying inter-speaker variability. In *Proceedings of Interspeech 2011* (pp. 273-276). Florence, Italy. doi: 10.21437/Interspeech.2011-108