

## How an aglossic speaker produces an alveolar-like percept without a functional tongue tip

Asterios Toutios, Melissa Xu, Dani Byrd, Louis Goldstein, and Shrikanth Narayanan

Citation: *The Journal of the Acoustical Society of America* **147**, EL460 (2020); doi: 10.1121/10.0001329

View online: <https://doi.org/10.1121/10.0001329>

View Table of Contents: <https://asa.scitation.org/toc/jas/147/6>

Published by the *Acoustical Society of America*

---

### ARTICLES YOU MAY BE INTERESTED IN

[Predicting ultrasound tongue image from lip images using sequence to sequence learning](#)

*The Journal of the Acoustical Society of America* **147**, EL441 (2020); <https://doi.org/10.1121/10.0001328>

[A landmark-cue-based approach to analyzing the acoustic realizations of American English intervocalic flaps](#)

*The Journal of the Acoustical Society of America* **147**, EL471 (2020); <https://doi.org/10.1121/10.0001345>

[Passive geoacoustic inversion in the Mid-Atlantic Bight in the presence of strong water column variability](#)

*The Journal of the Acoustical Society of America* **147**, EL453 (2020); <https://doi.org/10.1121/10.0001333>

[Phonetic categorization and discrimination of voice onset time under divided attention](#)

*The Journal of the Acoustical Society of America* **147**, EL484 (2020); <https://doi.org/10.1121/10.0001374>

[No evidence for a link between noise exposure and auditory temporal processing for young adults with normal audiograms](#)

*The Journal of the Acoustical Society of America* **147**, EL465 (2020); <https://doi.org/10.1121/10.0001346>

[Importance of phase guides from beamformed data for processing multi-channel data in highly scattering media](#)

*The Journal of the Acoustical Society of America* **147**, EL447 (2020); <https://doi.org/10.1121/10.0001330>

---



**Advance your science and career  
as a member of the**

**ACOUSTICAL SOCIETY OF AMERICA**

LEARN MORE



# How an aglossic speaker produces an alveolar-like percept without a functional tongue tip

Asterios Toutios,<sup>1,a)</sup> Melissa Xu,<sup>2</sup> Dani Byrd,<sup>3</sup> Louis Goldstein,<sup>3</sup>  
and Shrikanth Narayanan<sup>1</sup>

<sup>1</sup>*Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, California 90089, USA*

<sup>2</sup>*Keck School of Medicine, University of Southern California, Los Angeles, California 90033, USA*

<sup>3</sup>*Department of Linguistics, University of Southern California, Los Angeles, California 90089, USA*  
*toutios@usc.edu, melissax@usc.edu, dbyrd@usc.edu, louisgol@usc.edu, shri@siipi.usc.edu*

**Abstract:** It has been previously observed [McMicken, Salles, Berg, Vento-Wilson, Rogers, Toutios, and Narayanan. (2017). *J. Commun. Disorders, Deaf Stud. Hear. Aids* 5(2), 1–6] using real-time magnetic resonance imaging that a speaker with severe congenital tongue hypoplasia (aglossia) had developed a compensatory articulatory strategy where she, in the absence of a functional tongue tip, produced a plosive consonant perceptually similar to /d/ using a bilabial constriction. The present paper provides an updated account of this strategy. It is suggested that the previously observed compensatory bilabial closing that occurs during this speaker's /d/ production is consistent with vocal tract shaping resulting from hyoid raising created with mylohyoid action, which may also be involved in typical /d/ production. Simulating this strategy in a dynamic articulatory synthesis experiment leads to the generation of /d/-like formant transitions. © 2020 Acoustical Society of America

[Editor: Anders Lofqvist]

Pages: EL460–EL464

Received: 4 March 2020 Accepted: 11 May 2020 Published Online: 2 June 2020

## 1. Introduction

Congenital aglossia is a rare syndrome in which an individual is described as having been born without a tongue. Less than 50 cases of aglossia or hypoglossia have been reported worldwide since De Jussieu in 1718 discussed a case of hypoglossia in a 15-year-old Portuguese girl (Simpson and Meinhold, 2007; Thorp *et al.*, 2003). The subject of the present paper was a 47-year-old (at time of recording) female with isolated congenital aglossia. Studies using electropalatography (McMicken *et al.*, 2014c), cineradiography (McMicken *et al.*, 2014b), and real-time Magnetic Resonance Imaging (MRI; McMicken *et al.*, 2017) have attempted to uncover the compensatory mechanisms that the subject has developed by herself, without external training, in order to produce, overcoming her structural deficiency, speech that has been reported as intelligible, albeit with some vowel and consonant distortions (McMicken *et al.*, 2014a, 2012).

In particular, McMicken *et al.* (2017) used real-time MRI data collected by the group of the present paper's authors (Narayanan *et al.*, 2004; Toutios *et al.*, 2019) to examine the articulatory configurations used by this subject that allowed for intelligible perception of different consonants. It was found that, in the absence of a functional tongue tip, the subject formed bilabial closures to produce intelligible alveolar stops, such as /t/ and /d/, which would be produced by forming apicoalveolar closures in typical speech production (McMicken *et al.*, 2017). These bilabial closures created for canonical alveolars displayed an increased anteroposterior extent compared to the ones employed for the productions of canonical /p/ and /b/. McMicken *et al.* (2017) hypothesized (without experimental verification) that the reason for the alveolar-like percept was that the particular shaping of the elongated constriction created a cavity that filtered transient and friction sources during release similarly with typical alveolar plosives.

The present paper offers a re-evaluation of the same data in McMicken *et al.* (2017) to further describe the articulatory strategy. The real-time MRI videos were re-examined to encompass the shaping of the entire vocal tract and not just the shaping of the constriction, which was the focus of McMicken *et al.* (2017). An articulatory synthesis experiment verified the criticality of a widening of the pharynx and concomitant narrowing of the tract at the palatal region for the production of alveolar-like formant transitions. Additionally, novel insights on the anatomical description of the subject are included based on previously unpublished T2-weighted MRI data (Töger *et al.*, 2017) that were recorded in the same session as the real-time data.

<sup>a)</sup> Author to whom correspondence should be addressed. ORCID: 0000-0003-3193-2241.

## 2. Anatomical description

The subject has been previously described as being born without a tongue, having instead a wart-like tongue rudiment in the region of the tongue root, and having a hypertrophied mylohyoid and base of tongue (McMicken *et al.*, 2017). The aglossic speaker also presents with micrognathism and a hypertrophied lower lip.

Figure 1 shows midsagittal slices of T2-weighted MRI data from the aglossic speaker, and a non-aglossic 29-year-old female speaker for comparison, which enable some additional anatomical observations. First, the aglossic speaker presents with a severely hypoplastic anterior part of the genioglossus. Second, the superior longitudinal muscle, which would normally connect directly the tongue tip to the hyoid bone, appears to be missing, which suggests that the speaker cannot have functional control of tongue tip raising. Third, the mylohyoid/geniohyoid complex is proportionally more voluminous for the aglossic speaker compared to the non-aglossic speaker (given their respective overall vocal tract dimensions), which may suggest hypertrophy.

## 3. Articulatory strategy

A typical production of English /d/ involves a full closure of the vocal tract between the tongue tip and the alveolar ridge (Fig. 2; see also multimedia file Mm. 1). Lacking a tongue tip, the aglossic speaker has developed a compensatory strategy to produce a sound that has been reported to be perceived as /d/ (Mm. 2) by using a bilabial closure as described in McMicken *et al.* (2017): the bilabial closure that the aglossic speaker uses for the production of /d/ is qualitatively different from the one she uses for her production of a canonical /b/ (Mm. 3; see also Mm. 4 for a typical non-aglossic /b/). In particular, it has a longer anteroposterior extent, so that the posterior end of the constriction is close to the front edge of the palate, thus roughly approximating the constriction location of /d/ for a non-aglossic speaker.

In addition, the aglossic speaker's /d/ is distinguished from /b/ by a widening of the pharynx with a concomitant narrowing at the palatal region. This effect is also seen in the normal speaker's articulation and is presumably a consequence (in typical speakers) of the fronting of the entire tongue body (including the root) in a back vowel context in order to produce the alveolar constriction. Stevens (2000) and Iskarous *et al.* (2010) have shown, both theoretically and by statistical analysis of point-tracking data, that such fronting in back vowel contexts is responsible for their characteristic  $F_2$  transitions that contribute to their perceptual identification.

Mm. 1. Real-time MRI video of a production of /ada/ by the non-aglossic speaker. This is a file of type "mp4" (543 KB).

Mm. 2. Real-time MRI video of a production of /ada/ by the aglossic speaker. This is a file of type "mp4" (472 KB).

Mm. 3. Real-time MRI video of a production of /aba/ by the aglossic speaker. This is a file of type "mp4" (471 KB).

Mm. 4. Real-time MRI video of a production of /aba/ by the non-aglossic speaker. This is a file of type "mp4" (545 KB).

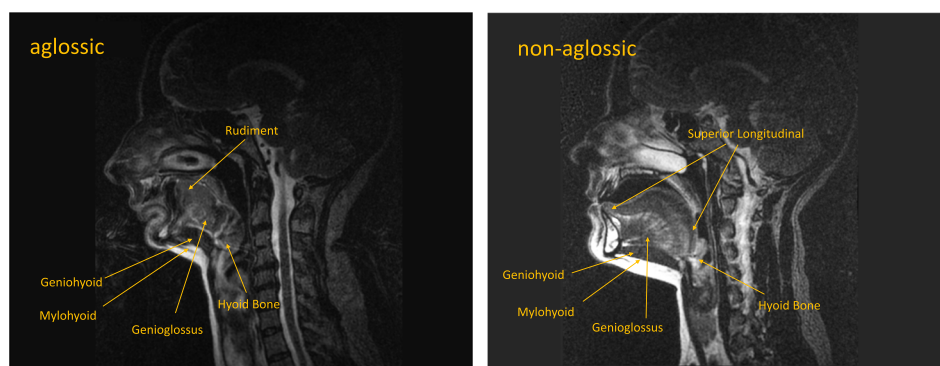


Fig. 1. (Color online) Midsagittal slices from T2-weighted MRI data [aglossic (left) and non-aglossic (right) subjects], which offer significantly more anatomical detail than real-time MR imaging, captured while subjects lay in the scanner at a rest position for several minutes. Structures discussed in the text are annotated.

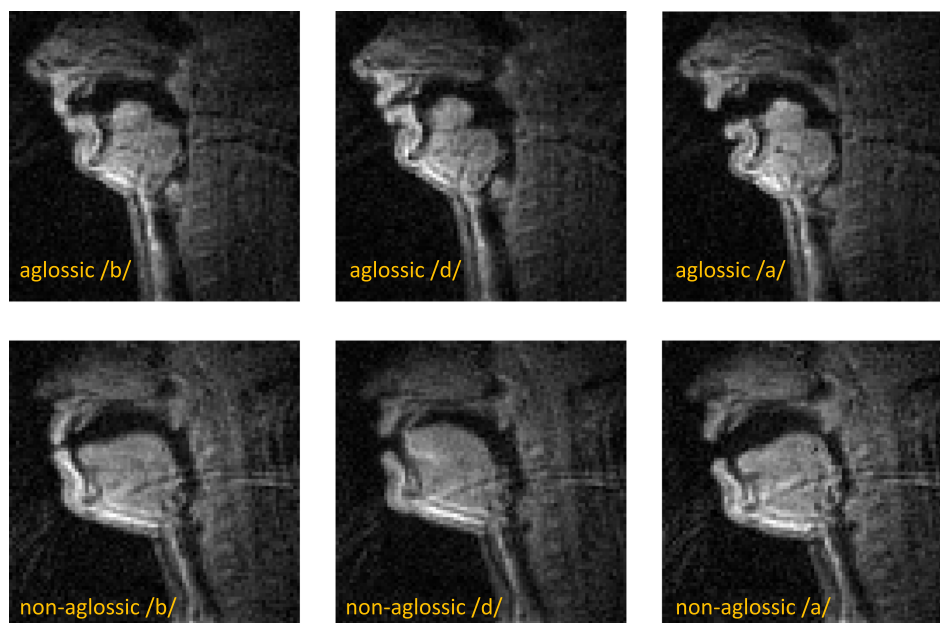


Fig. 2. (Color online) Frames extracted from real-time MRI data (at 83 frames per second) of /ada/ and /aba/ sequences for the aglossic and non-aglossic speakers. The /a/ shown was extracted after release of /b/.

#### 4. Acoustic simulation

We simulated the time-variation of plosive production by testing five dynamically changing tube configurations corresponding to the formation and release of three different oral constrictions (typical bilabial, typical alveolar, long [i.e., spatially-extended] bilabial) with or without concurrent pharyngeal widening and palatal narrowing, all in /aCa/ contexts. These dynamic configurations were used as inputs to Maeda's articulatory synthesizer (Maeda, 1982; Toutios and Maeda, 2012).

Figure 3 summarizes the dynamic vowel-consonant-vowel (VCV) configurations and the resulting formant trajectories. The two bottom-right plots show the area function for the vowel (a typical /a/) and the dynamics of the critical constriction degree that was input to the synthesizer, both applied to all VCVs synthesized; the latter dynamics can be thought of as governing the transition from vowel to consonant and back to vowel area functions. The other plots show the five different consonant area functions in the simulated VCVs and the synthesized formant trajectories of these VCVs.

C<sub>1</sub> in Fig. 3 (see also multimedia file Mm. 5 for the synthesized VC<sub>1</sub>V audio) corresponds to a typical /b/ (also the aglossic speaker's) and C<sub>2</sub> (Mm. 6) to a typical /d/. C<sub>3</sub> (Mm. 7) corresponds to the aglossic speaker's /d/. C<sub>4</sub> (Mm. 8) and C<sub>5</sub> (Mm. 9) are hybrids that do not correspond to actually observed productions. C<sub>4</sub> shows the effect on the formants of only extending the C<sub>1</sub> lip constriction in a posterior direction. The effect is to make F<sub>3</sub> fall sharply at the release of the stop, while F<sub>2</sub> continues to rise sharply, as it does in C<sub>1</sub>. Neither of these is what is found for alveolar stops in the context of a central low vowel of this kind, either theoretically or empirically [see Stevens (2000)]. The high F<sub>3</sub> during the closure is presumably due to the shortening of the front cavity. When the vocal tract is "straightened" due to anterior movement of the

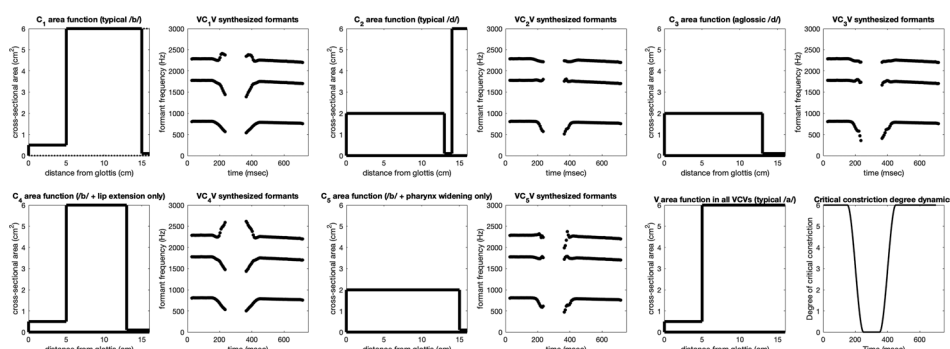


Fig. 3. Dynamic VCV configurations used as inputs to the simulation and resulting acoustics. See text for details and discussion.

tongue (seen both in typical and aglossic /d/ production), as in C<sub>3</sub>, the formant transitions are largely flat, consistent with observations, and modeling of /d/ in the context of a vowel of this kind. A comparison of C<sub>5</sub> and C<sub>3</sub> suggests that the “extension” of the bilabial constriction does not contribute change to the formant transitions substantially. While a detailed analysis of the sensitivity functions of all these shapes could contribute further to our understanding, and to the confidence in the robustness of our results, such an analysis would be of limited utility given the non-ideal acoustic quality of the MRI data currently available.

**Mm. 5.** Synthesized VC<sub>1</sub>V audio. This is a file of type “wav” (31 KB).

**Mm. 6.** Synthesized VC<sub>2</sub>V audio. This is a file of type “wav” (31 KB).

**Mm. 7.** Synthesized VC<sub>3</sub>V audio. This is a file of type “wav” (31 KB).

**Mm. 8.** Synthesized VC<sub>4</sub>V audio. This is a file of type “wav” (31 KB).

**Mm. 9.** Synthesized VC<sub>5</sub>V audio. This is a file of type “wav” (31 KB).

## 5. Conclusion

The aglossic speaker produces differences in the articulation of labial and coronal stops, despite a morphology that does not allow her tongue tip to form a constriction at the alveolar ridge. It is possible that she has learned to engage her mylohyoid muscles in a manner similar to typical speakers to produce a coronal stop in an /a/ context.

Specifically it has been speculated (Epstein *et al.*, 2002) that during normal alveolar/dental articulations, the mylohyoid muscle may elevate the hyoid bone (located at the base of the mandible and superior to the larynx), pulling the tongue toward the alveolar ridge, creating more space in the pharynx as a result. If the aglossic speaker is engaging this mechanism, this could result in the characteristics observed in the imaging of her productions of /d/: (i) the closure at the lips (in conjunction with the speaker’s micrognathism and lip hypertrophy, and noting that the lips will follow passively the jaw movement); (ii) the narrowing at the palatal region; and (iii) the pharyngeal widening. It is the synergistic combination of all these three features that produce acoustic patterns similar to those of typical /d/ constrictions in our acoustic simulations, and this provides preliminary support for this analysis.

## Acknowledgments

We thank Professor Betty McMicken, Chapman University, for bringing us in contact with the aglossic speaker and for her foundational work on this case. This study was supported by NIH Grant No. R01DC007124 and NSF Grant Nos. 1514544 and 1908865. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or NSF.

## References and links

- Epstein, M., Hacopian, N., and Ladefoged, P. (2002). “Dissection of the speech production mechanism,” UCLA Working Papers in Phonetics 102.
- Iskarous, K., Fowler, C. A., and Whalen, D. H. (2010). “Locus equations are an acoustic expression of articulator synergy,” *J. Acoust. Soc. Am.* **128**(4), 2021–2032.
- Maeda, S. (1982). “A digital simulation method of the vocal-tract system,” *Speech Commun.* **1**(3–4), 199–229.
- McMicken, B., Berg, S. V., and Iskarous, K. (2012). “Acoustic and perceptual description of vowels in a speaker with congenital aglossia,” *Commun. Disorders Q.* **34**(1), 38–46.
- McMicken, B., Salles, F., Berg, S. V., Vento-Wilson, M., Rogers, K., Toutios, A., and Narayanan, S. S. (2017). “Bilabial substitution patterns during consonant production in a case of congenital aglossia,” *J. Commun. Disorders, Deaf Stud. Hear. Aids* **5**(2), 1–6.
- McMicken, B., Vento-Wilson, M., Berg, S. V., Iskarous, K., Kim, N., Rogers, K., and Young, S. (2014a). “Semantic and phonemic listener confusions in a case of isolated congenital aglossia,” *Commun. Disorders Q.* **35**(2), 74–83.
- McMicken, B., Vento-Wilson, M., Berg, S. V., and Rogers, K. (2014b). “Cineradiographic examination of articulatory movement of pseudo-tongue, hyoid, and mandible in congenital aglossia,” *Commun. Disorders Q.* **36**(1), 3–11.
- McMicken, B. L., Kunihiro, A., Wang, L., Von Berg, S., and Rogers, K. (2014c). “Electropalatography in a case of congenital aglossia,” *J. Commun. Disorders, Deaf Stud. Hear. Aids* **2**, 1–7.
- Narayanan, S. S., Nayak, K. S., Lee, S., Sethy, A., and Byrd, D. (2004). “An approach to real-time magnetic resonance imaging for speech production,” *J. Acoust. Soc. Am.* **115**(4), 1771–1776.
- Simpson, A. P., and Meinhold, G. (2007). “Compensatory articulations in a case of congenital aglossia,” *Clin. Ling. Phonetics* **21**(7), 543–556.
- Stevens, K. N. (2000). *Acoustic Phonetics* (MIT Press, Cambridge, MA).
- Thorp, M., de Waal, P., and Prescott, C. (2003). “Extreme microglossia,” *Int. J. Pediatric Otorhinolaryn.* **67**, 473–477.



- Töger, J., Sorensen, T., Somandepalli, K., Toutios, A., Lingala, S. G., Narayanan, S., and Nayak, K. (2017). "Test-retest repeatability of human speech biomarkers from static and real-time dynamic magnetic resonance imaging," *J. Acoust. Soc. Am.* **141**(5), 3323–3336.
- Toutios, A., Byrd, D., Goldstein, L., and Narayanan, S. (2019). "Advances in vocal tract imaging and analysis," in *The Routledge Handbook of Phonetics*, edited by W. Katz and P. Assmann (Routledge, London and New York), pp. 34–50.
- Toutios, A., and Maeda, S. (2012). "Articulatory VCV synthesis from EMA data," in *Interspeech*, Portland, OR.