Hammill Institute on Disabilities

Communications Disorders Quarterly
34(1) 38–46

DOI: 10.1177/1525740111435114
http://cdq.sagepub.com
SAGE

Acoustic and Perceptual Description of Vowels in a Speaker With Congenital Aglossia

Betty McMicken, PhD[1], Shelley Von Berg, PhD[2], and Khalil Iskarous, PhD[3,4]

Abstract

The goals of this study were to (a) compare the vowel space produced by a person with congenital aglossia (PWCA) with a typical vowel space; (b) investigate listeners' intelligibility for single vowels produced by the PWCA, with and without visual information; and (c) determine whether there is a correlation between scores of speech intelligibility of PWCA speech and the acoustic properties of those speech samples. The main objective of this study was to determine whether a PWCA was able to compensate for the lack of tongue and whether listeners were able to compensate perceptually for the possible atypical acoustics of the PWCA. Cineradiography for this article was limited to observation of gross function of the tongue base and mylohyoid. An audiovisual recording of the PWCA speaker's output was obtained for a series of isolated vowels, diphthongized vowels, and vowels in monosyllables. Production of vowels was analyzed acoustically and perceptually. Vowels were presented to listeners under two conditions: audiovisual and audio only. Paired differences sample tests revealed no statistical differences in intelligibility for the audio versus audiovisual conditions. Mean intelligibility for vowels was 78.5% overall. Intelligibility was a function of vowel position, with the front vowels revealing the least intelligibility and the back vowels revealing the greatest intelligibility. Quantitative analysis of F1–F2 formant data revealed that the speaker's front vowels showed greater distances from the back vowels when compared with the F1–F2 means of normative data.

Keywords

aglossia, perceptual, acoustic, vowels

Introduction

It is unusual to encounter a person presenting with isolated congenital aglossia (PWCA). Only about 12 cases have been reported internationally since it was first systematically described by De Jussieu in 1718 (Higashi & Edo, 1996; Khalil, Dayal, Gopakumar, & Prashanth, 1995; Kumar & Chaubey, 2007; Salles et al., 2008). From the point of view of speech development, congenital aglossia (referred to as hypoglossia in some reports) represents special cases of tongue absence. In congenital aglossia, speech and language develop in conjunction with tongue absence, and in some cases influence orofacial development. This is quite different from loss of tongue later in life (glossectomy), after speech has been acquired.

Of interest to speech language pathologists, clinical linguists, and speech scientists is the consensus among these investigators that the spoken output of aglossic speech is generally quite intelligible. Such findings are theoretically intriguing, as they would support theories of speech production in which acoustic targets can be achieved with radically different articulatory strategies to achieve intelligibility through articulatory compensation (Guenther, Hampson, & Johnson, 1998; Guenther & Vladusich, in press). Rosenthal (1932) commented that speech, which is very poorly developed in most cases during the first few years in a child with aglossia, may improve considerably when the child learns to use other muscles to substitute for the missing tongue. Indeed, there have been a number of cases in which investigators report that absence of a tongue, or even a rudimentary one, may be compensated for by hypertrophy of the floor of the mouth. Salles et al. (2008) described a case of congenital aglossia whereby a female Brazilian speaker elevated the

[1]California State University, Long Beach, Lake Forest, CA, USA
[2]California State University, Chico, CA, USA
[3]Haskins Laboratory, Hartford, CT, USA
[4]University of Southern California, Los Angelels, CA

Corresponding Author:
Shelley von Berg, California State University, 400 W., First St., Chico, CA 95942, USA
Email: svonberg@csuchico.edu

posterior portion of the floor of the mouth to contact the palate, thus allowing her to develop speech and swallowing functions. They suspected the mylohyoid was the primary muscle of movement. The mylohyoid is one of the suprahyoid elevators. It forms the floor of the mouth, and contraction of the mylohyoid tends to elevate the larynx.

Salles et al. (2008), in describing the spoken output of a PWCA, reported a neutral voice quality and moderate impairment of nasal resonance, but did not specify hypo, hyper, or cul-de-sac. They noted marked distortions for phonemes /t/, /d/, /n/, /s/, /z/ and omission of the liquids, all of which are made by tongue contact with the anterosuperior aspect of the oral cavity, notably the alveolar ridge. They did not report the quality of vowel production.

Eskew and Shepard (1949) described the spoken output of a young speaker of Chinese origin who presented with congenital aglossia. While they did not identify the mylohyoid per se, they reported that the smooth floor of the mouth elevated to contact the incisal edges of the maxillary anterior teeth. They suggested that the lingual-alveolar productions of /t/, /d/, and /l/ were made with mylohyoid contact with the alveolar ridge and that lingua-velar sounds were produced by articulating the buccinator muscles with the lateral molars. These researchers noted that the speaker produced most vowels clearly, with the exception of the /ae/ and /i/, which are both front vowels.

Simpson and Meinhold (2007) described similar phoneme distortions for a female presenting with congenital aglossia. Stimuli were monosyllabic and polysyllabic words representing the German vowel and consonant system. Findings revealed that the speaker employed the labial structures rather than the mylohyoid to generate the majority of productions normally involving the tongue. For example, they noted that the /t/, /d/, and /n/ phonemes were produced by the bottom lip making contact with the alveolar ridge. Posterior phonemes such as the /k/ and /g/ were made by medializing the glossopalatal arches and articulating them with the uvula and the floor of the mouth.

While these findings have provided a much-needed foundational description of the articulatory adjustments made by individuals presenting with congenital aglossia, no studies to date have explored in a systematic fashion spoken output using acoustic and perceptual analysis. It is also not clear from these studies whether high intelligibility is due to the individual's ability to use compensatory gestures or whether higher aspects of linguistic and pragmatic knowledge are used by listeners to compensate for the production of aglossic speech. Hence, the main objective of this study was to determine whether the PWCA was able to compensate for the lack of tongue and whether listeners were able to compensate perceptually for the possible atypical acoustics of PWCA. When the oral portion of the tongue is lacking, it is potentially still possible to compensate in speech production to achieve more typical acoustics by using the mylohyoid, glossopalatal arches, and base of tongue (Salles et al., 2008; Simpson & Meinhold, 2007). However, it is also conceivable, based on the acoustic theory of speech production that control of the volume of the pharyngeal resonator and/or lip spreading could affect F1 and F2 at the output to make them more typical (Stevens, 1998). Lip compensation would be visible and could affect observer judgments.

This quantitative and qualitative paper presents an acoustic and perceptual description of vowels produced by an American-English speaker presenting with congenital aglossia. It is critical that vowels be investigated, because it has been suggested that listeners can identify with considerable accuracy even single vowels produced by any speaker (van Bergem, Pols, & Koopmans-van Beinum, 1988).

This article posed several foundational questions. First, what are the characteristics of the vowel space of a PWCA, as compared with typical speakers? Second, does the vowel intelligibility of a speaker presenting with micrognathia and aglossia and subsequent compensatory strategies alter with and without visual cues? This question was posed because past studies have suggested that compensatory placements for difficult phonemes produced by dysarthric speakers may actually detract from the intelligibility of speech (De Feo & Schaefer, 1983; Von Berg, McColl, & Brancamp, 2007). Third, do acoustic analyses of target vowels agree with listener perceptions of vowels? In other words, is it possible to explain the perceptual results by comparing the acoustic properties of an aglossic speaker with the acoustics from normative data?

Method

Speaker

The speaker for this study was a 16-year-old female who was referred in 1986 to the PI's cleft palate, craniofacial, and hospital-based head and neck center by the speaker's mother for a speech and craniofacial status assessment. The client's mother, in the presence of the PI, signed release of information forms, including acknowledgment that audiovisual samples might be used for future research and education purposes. During the 25-year time lapse between the assessment and this study, those hospital records were purged. An intensive Web search failed to reveal the location of the PWCA. Therefore, no current additional information pertaining to the speaker was available at the time of initiation of the present article. The PWCA and her mother presented to the clinic to receive information on cosmetic options for "her small jaw." They were interested in a mandibular advancement procedure. At the time of assessment, the PWCA was not being followed by a pediatrician. Limited medical history was available from the PWCA and her mother; however, the mother did provide sufficient history to determine that her daughter had required minor adaptations for feeding as an infant, in the form of widening the bottle nipple lumen for feeding. Speech and language milestones were reported as unremarkable. The PWCA said that she was able to eat a regular diet as long as she cut solids into small pieces and had a liquid wash. She also related, and her mother confirmed, that she was able to discriminate taste quite well and enjoyed cooking specialized dishes for her family.

As a part of the 1986 assessment, audio/video recordings and cineradiographic films were collected. These cineradiographic films were reviewed in 2010 by the same radiologist initially involved in the data collection in 1986. It was the radiologist's interpretation that the speaker presented with Klippel–Feil syndrome (KFS), with fusion of C3 to C4, involving the bodies and posterior aspects. On the frontal views, the radiologist suggested that scoliosis was present. The speaker demonstrated full labial closure anteriorly and demonstrated almost complete obliteration of the oral cavity by elevating the mylohyoid. Therefore, it was suspected that the PWCA would have been capable of using her lips and mylohyoid for compensation. However, a slight failure to fully elevate the mylohyoid was apparent as she sipped water and needed to blot her lip. This represented the residual liquid bolus that was not propelled posteriorly. Posterior closure of the oral cavity was achieved using the remnant of posterior third of tongue and significant elevation of the hypopharynx and hyoid. The velum functioned well to achieve posterior closure.

Multiple anomalies are associated with KFS. The most common are a short and webbed neck, a low hairline, and occipital cervical anomalies. The radiologist surmised that the insult occurred about the 4th or 5th week of gestation. He speculated that the complete failure of the development of the anterior 2/3 of the tongue and the other anomalies all occurred about this time. The radiologist observed that the speaker compensated very well for swallow by elevation of the floor of mouth combined with the closure of teeth, lips, and gums for complete anterior closure. Cervical spine stenosis and an anomalous omovertebral bone are also reported with KFS, but were not seen in this case. Hearing anomalies and loss, and cleft palate are reported associations; however, these were not observed in the speaker (See figures 1, 2 and 3).

The speaker presented with micrognathia and microsomia. All upper and lower teeth were present, but had collapsed medially due to the absence of the tongue. She presented with a Class II malocclusion. She was receiving orthodontic intervention by way of braces. Intraoral inspection revealed a tongue rudiment in the region of the tongue root. The absence of the tongue was compensated for by the fact that the floor of the mouth was smooth and its posterior portion could be elevated to contact the palate. Similar to a case study by Salles et al. (2008), this muscle-mass to palate contact allowed the speaker to develop speech and swallowing functions.

At the time of initial assessment, speech was deemed intelligible when the subject matter was highly predictable; however, speech was characterized by marked distortions of vowels and phoneme combinations, notably for single- and two-word utterances. It was suspected that the reduced mylohyoid (pseudotongue) range of motion to the anterior aspect of the oral cavity was associated with these distortions.

Listeners

Prior to the recruitment of listeners, the California State University, Chico, and Long Beach Institutional Review Boards approved the study, and informed consent was obtained from the listeners. Listeners were 16 females and 4 males between the ages of 19 years and 5 months and 62 years and 5 months ($M = 32.54$, $SD = 12.56$). All listeners were native English speakers and were undergraduate students, or acquaintances thereof, at the California State University, Chico, Communication Sciences and Disorders program. Vision was determined adequate for the task through an informed consent task. Listeners were instructed to read aloud the letter of informed consent. In addition, listeners were instructed to report any difficulty viewing the speaker on the computer monitor. None of the listeners reported difficulty seeing the speaker. Hearing acuity was assessed under headphones according to the American National Standards Institute Specifications (2007). Acuity for all listeners was within normal limits at 25 dB HL for the frequencies 500 Hz, 1,000 Hz, 2,000 Hz, and 4,000 Hz.

Stimuli

A DVD of the speaker's spoken output was obtained by implementing an imitative task for a series of isolated vowels, monosyllabic and polysyllabic words, and phrases containing all of the phonemes of the English language (see appendix). Only the isolated vowels and vowels in monosyllables were analyzed in this study. The stimuli were collected in 1986 at the time of the initial assessment. The original recordings were collected using a Sony BetaMax 900 recorder, Sony professional camera, and unidirectional microphone. The VHS was converted to DVD employing the audio Cube AC5. The software was effective in reducing extraneous ambient noise. The audio from the original VHS was digitized at 24 bits and 48 kHz sample by a professional audio studio; the subsequent acoustic data were used for perceptual and spectrographic analysis. A master data spreadsheet was generated for all of the stimuli. This data sheet was used to quickly and accurately record listener responses for the perceptual phase of the study.

Tasks

Perceptual. An experienced researcher conducted the perceptual data collection. Procedures for presentation of stimuli to observers and recording of listener responses were adapted from previous perceptual speech research models (Ross, Saint-Amour, Leavitt, Javitt, & Foxe, 2007; Vatakis & Spence, 2006). Isolated vowels, words, and phrases were presented to listeners individually, randomly assigned, under two conditions: (a) audiovisual and (b) audio only. For the first listening session, 10 listeners were presented with the audiovisual condition and 10 listeners were presented with the audio condition. Listeners were able to view the entire face of the PWCA during the stimuli presentation. Stimuli were presented to listeners in a quiet office on a nonclinic day to control for ambient noise.

Stimuli were presented on a Dell PC with a sound bar attachment (Dell Sound bar AS 501). Frequency response of the audio speakers was 95 Hz to 20 kHz, which encompasses the human speech hearing range. Listeners were seated 24 inches from the monitor (13″ × 22″). The volume was preset across each condition at 75 dB as measured by a sound level meter (EMCO SLM-120). Listeners were instructed to verbally repeat what they thought they had heard. Responses were recorded by the researcher in writing on a master data sheet. Responses that accurately matched the stimulus items produced by the PWCA were noted with a "+." Responses that did not match were transcribed orthographically. Listener responses were simultaneously recorded onto a Marantz portable audio recorder for subsequent intrarater reliability.

Analysis

Perceptual data analysis. The analysis in this article is restricted to the vowels, produced in isolation and monosyllabic words. This was done to gain a basic understanding of the vocalic capability and perceivability of the speaker. The majority of vowels in word samples, 19 of each, were of the vowels /iy/, /ah/, and /uw/. Only a limited and uneven number of other vowels in words were collected. Therefore, the analysis of vowels in words was limited to these samples.

Listener responses were scored for number of vowels identified correctly, overall error rates, and specific vowel errors. Results were analyzed with SPSS, version 17, using crosstabs and paired samples tests. The dependent variable was listener intelligibility, and the independent variables were the two conditions: audio only and audiovisual. Data combined were 67 stimuli × 20 listeners × 2 listening/viewing sessions, totaling 2,680 data points. Interrater reliability was measured by using intraclass correlations. Intrarater reliability was conducted by retranscribing the responses from 2 of the 20 listeners for 10% of the sample.

Acoustic data analysis. For the acoustic analyses, Matlab scripts were used. The vowel formants were extracted from the speech by sampling each vowel at its acoustic midpoint. A window of 40 ms, centered at the midpoint, was then extracted. Linear Predictive Coding Analysis was performed after a Hamming window and preemphasis was applied. A total of 22 poles were used, and the energy maxima of F1 and F2 were extracted by automatic peak-picking. Means across trials of each vowel were calculated and compared by simple linear distance with the means for the normative vowels in Hillenbrand, Getty, Clark, and Wheeler (1995; H95). A bark scaling (psychoacoustic scaling of the critical bands of hearing) of the formants was attempted, but it did not affect the results and is therefore not included.

Table 1. Percent Correct of Isolated Vowel (IV) and Vowel in Word (VIW), Percent of Distortions (DIS) and Substitutions (SUB), and Acoustic Distance (AD) in Hertz

Isolated vowel	Correct (%)	SUB (%)	DIS (%)	AD (HZ)	Vowel in word	Correct (%)	SUB (%)	DIS (%)	AD (HZ)
(IV)	(IV)	(IV)	(IV)	(IV)	(VIW)	(VIW)	(VIW)	(VIW)	(VIW)
iy	20	0	80	853	Iy	48	42	10	694
ih	60	15	25	501					
ei	95	0	5	425					
eh	60	20	20	298					
ae	95	0	5	474					
uh	70	0	30	111					
ah	95	0	5	344	Ah	92	6	2	311
oa	100	0	0	155					
oo	95	5	0	195					
uw	95	5	0	94	uw	93	6	1	496

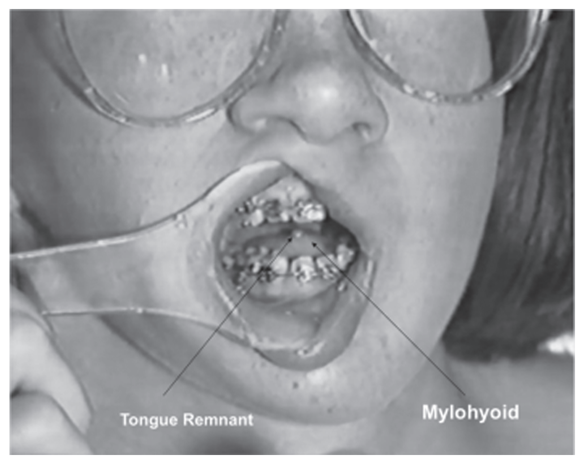

Figure 1.

Results

Table 1 presents the percent-correct intelligibility for all isolated vowels and vowels in monosyllables (/iy/, /ah/, / uw/). It also presents listeners error types, identified as substitutions or distortions. The table also shows the acous-

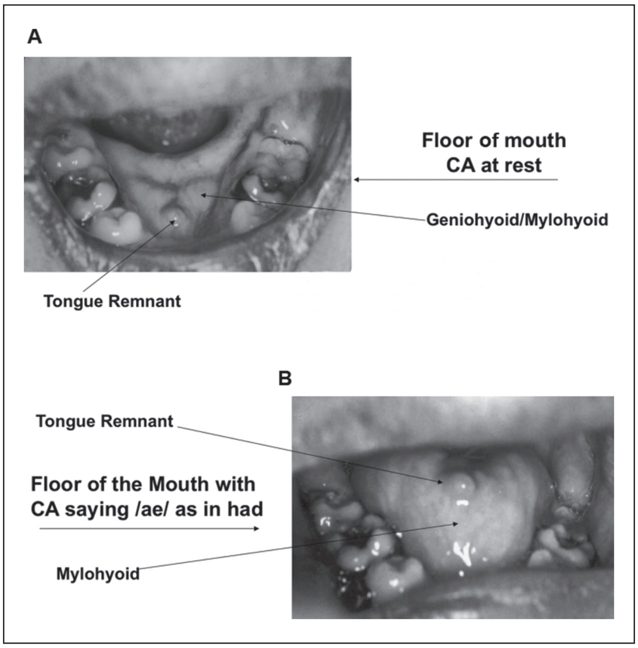

Figure 2.

tic distance between the PWCA vowels and the normative vowels from Hillenbrand et al. (1995).

Overall mean intelligibility for vowels was 78.5%. Paired differences samples tests revealed insignificant levels of difference between the audio and audiovisual conditions ($p > .05$). In addition, there was a moderately strong paired samples correlation (.524, $p < .05$) between the audio and audiovisual conditions. Interrater reliability was measured by using intraclass correlations; both individual (.677, $p < .05$) and group (.977 $p < .05$) ratings were moderate to strong. Intrarater agreement was tested by rescoring the responses from Listeners 15 and 17, for a 97% result in rating agreement.

As can be seen, listener intelligibility for vowels generally increased as a function of vowel location on the vowel quadrilateral, with the front vowels revealing the least intelligibility and the back vowels revealing the greatest intelligibility. Percentage of error types differed across vowel presentation; that is, substitution and distortion errors differed as a function of vowels presented in isolation or in words.

Figure 4 shows the F1 and F2 locations for vowels produced by normal speakers (Hillenbrand et al., 1995; H95) and the F1 and F2 locations for the vowels produced by the PWCA. The vowels in monosyllables (right panel) are presented in a square frame to distinguish them graphically from the isolated vowels, which are not framed (left panel). The centers of the H95 ellipses and the PWCA ellipses are joined by a line. Generally, the PWCA back vowels are relatively close to the H95 ellipses, but the front vowels are far from the H95 ellipses, especially in the F2 dimension. Indeed the PWCA vowel quadrilateral is less wide in the F2 dimension than the H95 vowels, but the height of the PWCA vowel quadrilateral seems to be largely the same as the H95 quadrilateral. The reason for this will be addressed in the Discussion section. There is a significant strong negative relationship between the two variables with Pearson correlation coefficient $r = -.68$, whereby greater perception scores correlate with smaller differences between H95 and PWCA means. It can be seen that the back vowels reveal a high perception score and have relatively low distances to the H95 vowels, whereas the front vowels have smaller perception scores and tend to have a higher H95–PWCA distance magnitude.

Discussion

An audiovisual recording of a PWCA was collected for a series of isolated vowels and monosyllabic words. Vowels were presented to listeners under two conditions: audiovisual and audio only. Data analysis revealed no difference between the audiovisual and audio conditions for listener intelligibility, suggesting that the cues from lip activity did not significantly increase intelligibility.

The mean vowel intelligibility rate of 78.5% is quite high, but further analysis shows that the perception scores for front vowels are low, especially for the frontmost vowel /iy/, suggesting that no extra compensation in production took place. Distortion was the greatest error for the /iy/ in isolation, and substitution was the greatest error for /iy/ in monosyllables. It is suspected that the error types are associated with the erroneous coarticulatory cues listeners gleaned from the /iy/ embedded in monosyllables. This phenomenon will be explored in future papers; however, these findings are strikingly consistent with Salles et al. (2008) and Eskew and Shepard (1949), who found in their experiences with PWCA that vowels requiring a pseudolingual apex and advancement in the oral cavity were less intelligible than those phonemes requiring a pseudolingual dorsum. Eskew and Shepard in particular noted that their speaker produced most vowels clearly, with the exception of the /ae/ and /iy/, which are both front vowels. Moreover, the strong negative correlation between intelligibility and acoustic distance between PWCA vowels and normative vowels suggests that poor intelligibility is predicted by the atypical acoustics. If the listeners used information other than the first two formants to perceptually compensate for the PWCA atypical acoustics, then the correlation would have been low. The strong correlation, therefore, suggests that there was no extra perceptual compensation.

These results allow us to link the anatomical abilities of this PWCA to the acoustic properties of her vowel space. In the case of this PWCA, it appears that the greater typicality of

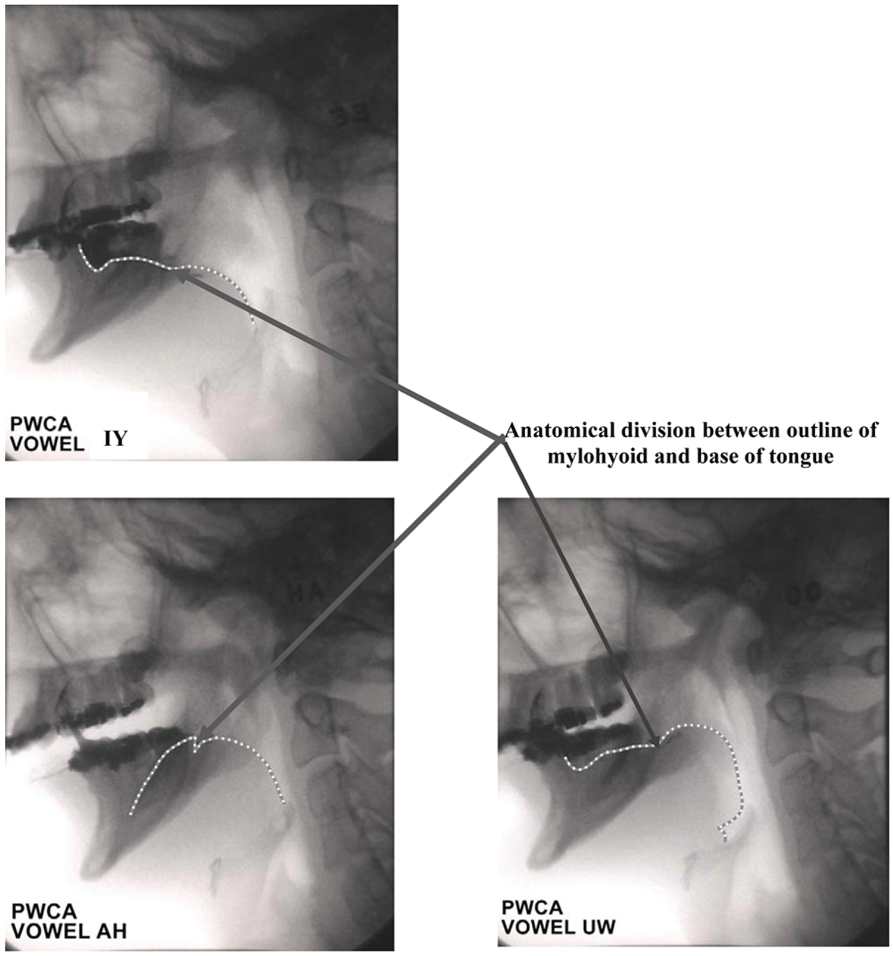

Figure 3.

her back vowels, as opposed to her front vowels, is due to several anatomical factors. Her tongue base has vertical and horizontal movement, which allows the base to be positioned appropriately for the back vowels, resulting in relatively typical acoustics. The less typical acoustic patterns of her front vowels could be due to the following limitations: (a) use of the mylohyoid, which allows for partial, but insufficient, constriction in the front region; (b) backing of the mandible to assist the mylohyoid in placing a front constriction; and (c) the micrognathia limits the size of the oral space, limiting range of motion, particularly in the anterior aspect. These three limitations are suspected to have led to the expected change in formant definition: The reduced size of the front constriction causes F2 to be lower, as seen in Figure 4.

Vocal tract compensation for this PWCA is greater in the back than the front of the vocal tract. Two compensatory maneuvers that the PWCA could have used to engage in relatively radical compensatory ability to achieve equivalent acoustics between PWCA and H95 were lip spreading and pharyngeal expansion. Both of these would increase F2. However, these compensations were not used; there is no acoustic or visual evidence for these compensatory maneuvers. The F2 is a great deal lower for the PWCA than for the typical average, and the lack of significance of the visual factor in perception shows that the speaker did not use an extreme visual cue.

An interesting aspect of the results is the difference between the isolated vowels and the vowels in monosyllables. For the vowel /iy/, even though the acoustic distance

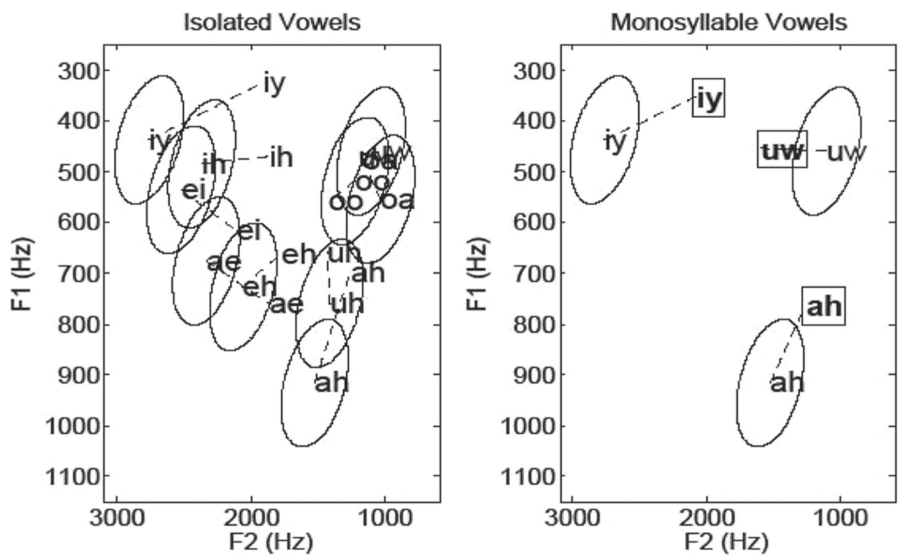

Figure 4. F1 and F2 locations for vowels produced by normal speakers (Hillenbrand et al., 1995; H95) in ellipses and the F1 and F2 locations for the vowels produced by the PWCA.
Note: PWCA = person with congenital aglossia. The center of the H95 ellipses and the PWCA means are joined by a line. In general, front vowels are closer and back vowels are further from the H95 ellipses, notably in the F2 dimension. The vowels in monosyllables are framed and the isolated vowels are not.

between the monosyllable vowel formants to the HB95 normative data is quite high, the intelligibility of the vowels in CVC context is greater than double the intelligibility for the isolated vowel. We believe that the reason for this is that the transition from C to V and V to C in the CVC monosyllables contains an enormous amount of information about vowel identity, as has been previously shown in the literature (Strange, 1989). The listeners seem to be using this transitional information, but that extra information still leaves the intelligibility of /iy/ at about half the rate as that for /ah/ and /uw/. The results for /uw/ may seem quite surprising, as the PWCA average distance from the monosyllabic vowels to the HB95 vowel is quite large as compared with the distance from the PWCA isolated /uw/; however, the intelligibility for the two are about the same. We believe that the reason for the large difference between the PWCA monosyllabic /uw/ and the HB95 data is that the former is averaged across many C_C contexts, whereas the HB95 data is all in h_d context. The formants for /uw/ do seem to vary a great deal according to context; therefore, the averaging across the 19 C_C contexts for /uw/ in the PWCA data makes that average quite different from that for isolated /uw/ by PWCA and the HB95. However, in spite of this large discrepancy, the monosyllabic /uw/ is perceived quite accurately. We believe that this is further evidence for the importance of dynamic transitions for vowel perception, as argued by Strange (1989). The /iy/ and /ah/ center vowel frequencies do not vary as much with the consonantal context, so the same problem does not arise for them.

The main theoretical conclusion to be drawn from these results suggest that even though the PWCA was able to compensate in general for the aglossia, she did *not* seem to use certain available compensatory maneuvers, such as using pharyngeal expansion and/or lip spreading, which would raise F2, to make the front vowels more typical. Therefore, theories that argue that speakers will use *whatever* compensatory maneuvers they have to make their acoustics more typical (Guenther et al., 1998) are not fully supported.

Appendix

All Vowels, Words, and Phrases Recorded From PWCA in 1986

VOWEL /oa/	**VOWEL /oo/**	**VOWEL /ei/**
VOWEL /uw/	**VOWEL /ae/**	**VOWEL /eh/**
VOWEL /iy/	**VOWEL /ah/**	**VOWEL /uh/**
VOWEL /ih/	**BAUB**	**BOOB**
BEEB	**CHAUD**	**CHOOD**
CHEED	**DAUD**	**DOOD**
DEED	**DAUM**	**DOOM**
DEEM	**DAUNG**	**DOONG**
DEENG	**DAUN**	**DOON**
DEEN	**FAUD**	**FOOD**
FEED	**GAUG**	**GOOG**
GEEG	**JAUD**	**JOOD**
JEED	**KAUD**	**KOOK**
KEEK	**LAUD**	**LOOD**
LEED	**POP**	**POOP**
PEEP	**RAUD**	**ROOD**
REED	**SAUD**	**SOOD**
SEED	**SHAUD**	**SHOOD**
SHEED	**VAUD**	**VOOD**
VEED	**WAUD**	**WOOWED**
WEED	**YAUD**	**YOOD**
YEED	**ZAUD**	**ZOOD**
ZEED		
	Goo Koo	Goop Koop
Mutt muff	Bean beam	Dime time
I went in a van	Bun bum	Mad mat
Fey, pay my movie fee	Tea time for two	Do you want to eat a banana
Take time to talk	Pig pick	Ida did it
May I have a whip	Gab cab	Tide
Vee fee	Puff of foam	Duet
Debt	Made mate	What time do you have
Hated	Did you have fun when you won	Gam cam
Bud buck	I went in a van	Admit
Dave dove deeply	Fey, pay my movie fee	Goop koop
Tee fee	Vick gave Kim a book about pigs and cows	Dime time
Beet beef	I went in a van	Mad mat
Dated	Fey, pay my movie fee	Dave's bow was too famous
The taffy was tough	An am	Mode moat
Today	Nigh migh	Toe foe
Don't annoy me	Diet	No moe
bated	Too few	I might type it today
View a movie	duty	

Note: PWCA = person with congenital aglossia. The stimuli were recorded in an imitative task presented by a speech pathologist. The responses were presented to listeners in random order. The present study examined vowels presented in isolation and in monosyllables (bold, all caps).

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The third author was supported by NIH-NIDCD Grant DC-02717 to Haskins Laboratories.

References

American National Standards Institute (2007). Speech material used in audiology. Washington, D.C.

De Feo, A. B., & Schaefer, C. M. (1983). *Bilateral facial paralysis in a preschool child*: *Oralfacial and articulatory characteristics*. In W. Berry (Ed.), *Clinical dysarthria* (pp. 165–190). San Diego, CA: College-Hill Press.

De Jussieu, A. (1718). Observations sur la maniere dont une fille sans langue s'acquitte des functions qui dependent de cet organe. [Observation of the manner in which a woman with no tongue accomplishes the functions that depend on that organ.] *Memoires of l'Academie Royale des Sciences de Paris, 1718*, 6–14.

Eskew, H. A., & Shepard, E. E. (1949). Congenital aglossia. *American Journal of Orthodontics, 35*, 116–119.

Guenther, F., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review, 105*, 611–633.

Guenther, F., & Vladusich, T. (in press). A neural theory of speech acquisition and production. *Journal of Neurolinguistics*.

Higashi, K., & Edo, M. (1996). Conductive deafness in aglossia. *Journal of Laryngology and Otology, 110*, 1057–1059.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97*, 3099–3111.

Khalil, K. C., Dayal, P. K., Gopakumar, R., & Prashanth, S. (1995). Aglossia: A case report. *Quintessence International, 26*, 359–360.

Kumar, P., & Chaubey, K. K. (2007). Aglossia: A case report. *Journal of the Indian Society of Pedodontics and Preventive Dentistry, 25*, 46–48.

Rosenthal, R. (1932). Aglossia congenita: A report of a case of the condition combined with other congenital malformations. *American Journal of Diseases in Children, 44*, 383–389.

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., & Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex, 17*, 1147–1153.

Salles, F., Anchieta, M., Costa Bezerra, P., Torres, M. L., Queiroz, E., & Faber, J. (2008). Complete and isolated congenital aglossia: Case report and treatment of sequelae using rapid prototyping models. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontics, 105*, e41–e47.

Simpson, A. P., & Meinhold, G. (2007). Compensatory articulations in a case of congenital aglossia. *Clinical Linguistics & Phonetics, 7*, 543–556.

Stevens, K. (1998). Acoustic Phonetics. MIT press.

Strange, W. (1989). Evolving theories of vowel perception. *Journal of the Acoustical Society of America, 85*, 2081–2087.

van Bergem, D. R., Pols, L. C. W., & Koopmans-van Beinum, F. J. (1988). Perceptual normalization of the vowels of a man and a child in various contexts. *Speech Communication, 7*, 1–20.

Vatakis, A., & Spence, C. (2006). Audiovisual synchrony perception for music, speech and object actions. *Brain Research, 1111*, 134–142.

Von Berg, S., McColl, D., & Brancamp, T. (2007). Moebius syndrome: Measures of listener intelligibility with versus without visual cues in bilateral facial paralysis. *Cleft Palate-Craniofacial Journal, 44*, 518–522.