The articulatory dynamics of pre-velar and pre-nasal /æ/-raising in English: An ultrasound study[a)]

Jeff Mielke,[b)] Christopher Carignan,[c)] and Erik R. Thomas
Department of English, North Carolina State University, Campus Box 8105, Raleigh, North Carolina 27695-8105, USA

Most dialects of North American English exhibit /æ/-raising in some phonological contexts. Both the conditioning environments and the temporal dynamics of the raising vary from region to region. To explore the articulatory basis of /æ/-raising across North American English dialects, acoustic and articulatory data were collected from a regionally diverse group of 24 English speakers from the United States, Canada, and the United Kingdom. A method for examining the temporal dynamics of speech directly from ultrasound video using EigenTongues decomposition [Hueber, Aversano, Chollet, Denby, Dreyfus, Oussar, Roussel, and Stone (**2007**). in *IEEE International Conference on Acoustics, Speech and Signal Processing* (Cascadilla, Honolulu, HI)] was applied to extract principal components of filtered images and linear regression to relate articulatory variation to its acoustic consequences. This technique was used to investigate the tongue movements involved in /æ/ production, in order to compare the tongue gestures involved in the various /æ/-raising patterns, and to relate them to their apparent phonetic motivations (nasalization, voicing, and tongue position). © *2017 Acoustical Society of America*. [http://dx.doi.org/10.1121/1.4991348]

[CGC]   Pages: 332–349

I. INTRODUCTION

Most dialects of North American English exhibit /æ/-raising in some phonological contexts. This includes raising with a falling trajectory before nasals (e.g., [beən] *ban*) over much of North America, and a less widespread raising pattern with a rising trajectory before /g/ (e.g., [bejg] *bag*). Both the conditioning environments and the temporal dynamics of the raising vary from region to region. The geographical distribution of various acoustic patterns is well documented, but comparatively less is known about why particular raising patterns have arisen in particular geographic locations. The articulatory basis of raising is also unclear. While it is reasonable to expect a higher vowel to be articulated with a higher tongue body, it has also been argued that pre-nasal raising could be due to acoustic consequences of nasalization, at least in some speakers (De Decker and Nycz, 2012; Baker *et al.*, 2008). One of the mysteries of /æ/-raising involves the development of raising before /g/ in the the Northern United States and in Canada but not elsewhere. While some Northern U.S. speakers produce /g/ with a more anterior constriction than /k/ (Purnell, 2008), it is not known whether this difference occurs in other /æg/-raising regions and if it is absent from non-/æg/-raising regions.

To explore the articulatory basis of /æ/-raising across North American English dialects, we collected acoustic and articulatory data from a regionally diverse group of 24 English speakers from the United States, Canada, and the United Kingdom. Our main focus is on tongue movements visualized using ultrasound imaging. Ultrasound is an increasingly common part of the toolkit for studies of dialect variation (e.g., Lawson *et al.*, 2011; De Decker and Nycz, 2012; Mielke, 2015), because data collection is convenient and affordable, compared to other lingual imaging techniques. While it is often possible to extract sufficient articulatory information from a single ultrasound frame from each consonant or vowel token under investigation, the nature of /æ/-raising is well suited to analysis methods that capture the dynamic nature of speech. In this study we apply a method for examining the temporal dynamics of speech directly from ultrasound video using EigenTongues decomposition (Hueber *et al.*, 2007) to extract principal components of filtered images and linear regression to relate articulatory variation to its acoustic consequences.

Section I A reviews some of the important facts about North American English /æ/-raising. Subsequently we will investigate the tongue movements involved in /æ/ production, in order to compare the tongue gestures involved in the various /æ/-raising patterns. A central goal of this project is to search for regional differences in the phonetic motivations for raising, which might help account for the observed dialect variation. The first step is to describe the various raising patterns and examine the relationships between the observed /æ/ raising and the apparent phonetic motivations, which involve nasalization, voicing, and tongue position.

A. /æ/-raising in North American English

/æ/ exhibits some of the most complex dialectal patterning of any vowel in North American English, as noted in

---

various studies (see especially Labov, 1994, pp. 503–526, and Labov et al., 2006, pp. 173–184). /æ/ shows breaking (i.e., rising and falling in the vowel space) across most of North America before the anterior nasals /n/ and /m/, as in *pan* and *ham*, but not in other contexts. On the American side of the Great Lakes, /æ/ shows raising in all contexts, though often to a greater degree in some phonetic environments, particularly before nasals. In the mid-Atlantic area stretching from New York City to Baltimore, including Philadelphia, various complicated configurations prevail in which raising occurs before anterior voiceless fricatives (mostly in monosyllabic morphemes, as in *half*, *pass*, and *bath*) and, depending on the community, it may occur before voiced fricatives and voiced stops, sometimes in lexically specific patterns; however, the raising is consistently absent in this region before voiceless stops, in function words such as *as* and auxiliary *can*, and in irregular verb forms such as *ran*. Yet another pattern is /æ/-raising before /g/ and /ŋ/, as in *bag* and *hang*, which is known from certain parts of the Upper Midwest, Canada, and the Pacific Northwest. In addition to raising, there is currently a lowering and/or retraction process underway for /æ/ in contexts not affected by raising. It should be noted that, in all of these cases, the "raising" and "lowering" appellations are based on acoustic data, not on articulatory data.

The presence of /æ/-raising before /g/ in only some North American English dialects is puzzling because these dialects do not present an obvious phonetic motivation for raising in this context (compared to dialects where /æ/ raising is greater before /d/ than before /g/). Zeller (1997) reported that younger speakers from the Milwaukee, Wisconsin area merged /æg/ with /ejg/ (e.g., *hag = Haig*). Labov et al. (2006) (p. 181) reported the same merger for some speakers in Wisconsin, Minnesota, and central Canada; they also noted that /æ/ tended to be higher before /g/ than before /d/ over a somewhat wider area. Pre-velar raising involves a rising vowel trajectory (Bauer and Parker, 2008; Benson et al., 2011, p. 286), while other observed raising patterns involve a falling trajectory (Labov et al., 2006, pp. 177–178; Jacewicz et al., 2011). Bauer and Parker (2008), using acoustic trajectories, durations, and ultrasound measures, found that speakers from Eau Claire, Wisconsin, raised /æg/ but did not merge it with /ejg/ or /ɛg/. Bauer and Parker's ultrasound data show that the tongue body is raised in /æg/, although /æ/ before /g/ still remains distinct from other front vowels. Wassink (2015) concludes that /æg/ and /ɛg/ are raised in Seattle, but not necessarily merged with /ejg/ or with each other. Rosen and Skriver (2015) do not address potential mergers, but they show that southern Albertans have much higher /æ/ realizations before /g/ than before other obstruents and that Mormons show a lesser degree of the raising than non-Mormons.

In a further complication, Purnell (2008), using x-ray microbeam data, found that Wisconsin subjects articulated /g/ and /k/ differently following /æ/. /g/ was characterized by lip protrusion that /k/ lacked, by greater upward jaw movement than for /k/, and, most importantly, by a more anterior tongue constriction than /k/ exhibited. Purnell also obtained his articulatory data at multiple time points and showed that the dynamic articulatory data corresponded closely with acoustic patterns.

There are a number of potential phonetic motivations for pre-velar raising. Palatal-induced upgliding has occurred at other times in the history of English, mostly before voiced stops and fricatives (and mostly not before voiceless stops). Palatal [ç] conditioned upgliding in Middle English (e.g., Old English *eahta* [æɒxtɑ] > *[æçtə] > Middle English *eight* [aiçt]). Similarly, /g/ = [ɟ], /ŋ/ = [ɲ], /ʃ/, and /ʒ/, as in *bag*, *hang*, *cash*, and *azure*, respectively, condition upglides in various American dialects (see, e.g., Kurath and McDavid, 1961, pp. 103, 104; Hartman, 1969; and Thomas, 2001, pp. 22, 23). Analogous conditioning of /æg/ vs /æk/ is hard to find in other languages because low front vowels are somewhat uncommon, many languages lack a contrast between voiced and voiceless dorsal stops, and voiced palatal stops are prone to devoicing or weakening (to, e.g., [dʒ], [ʝ], or [j]).

Hyperarticulation before voiceless obstruents is another potential factor affecting the realization of /æk/ vs /æg/. There is some evidence that vowels can show more extreme articulations before voiceless obstruents than elsewhere (e.g., Wolf, 1978; Summers, 1987; Moreton, 2008). For low vowels, this means that F1 values are higher before voiceless obstruents than before voiced obstruents (so that the vowel reaches a lower position before voiceless obstruents). Moreover, lower F1 values are also generally associated with voiced obstruents (e.g., Lisker, 1986). Expansion of the pharynx has been directly observed in voiced stops (Kent and Moll, 1969), and serves the function of maintaining the difference between subglottal and intraoral pressure that is necessary for voicing. Pharynx expansion by advancement of the tongue root should lower F1 because F1 is a Helmholz resonance when a tongue constriction is present. Ahn (2015) observes this relationship in ultrasound data for English onsets.

English /æ/-raising occurs in pre-nasal contexts as well, and pre-nasal raising is in fact widespread in North American English (Labov et al., 2006, pp. 174, 175). An apparent phonetic motivation is that nasalization has a strong effect on F1-lowering in low vowels, altering their perceived height (and may also raise F2 for low vowels—see Krakow et al., 1988). The articulatory basis of F1-raising in pre-nasal /æ/ was explored by De Decker and Nycz (2012), who studied /æ/-raising in four speakers from New Jersey. Two of the speakers appeared to show tongue raising in the raised vowel in *pan*, while two appeared to show only acoustic raising. De Decker and Nycz (2012) interpreted the acoustic raising for the latter two speakers to be a consequence of nasalization rather than tongue position. However, the acoustic and articulatory measurements in their study were obtained at different time points in the vowel interval: acoustic measurements were made at the vowel midpoint, while tongue traces were obtained from the most retracted position of the tongue (i.e., an articulation which does not necessarily occur at the vowel midpoint).

In their acoustic/airflow study of variation in /æ/-raising, Baker et al. (2008) offer a second reason for incongruity between F1 realization and tongue position in /æŋ/. While

J. Acoust. Soc. Am. **142** (1), July 2017

Mielke *et al.* 333

the coupling of the nasal to the oral tract (i.e., velo-pharyngeal coupling) results in independent acoustic consequences associated with the nasal cavity, it also alters the shape of the oral cavity. Baker *et al.* argued that F1-lowering in /æŋ/ could be due to velum lowering increasing the effective tongue height by lowering the top of the oral cavity. They noted that /æg/-raising is apparently never observed in the absence of /æŋ/-raising, which would be consistent with the idea that raising in /æg/, which does not require velum lowering, has a proper subset of the phonetic motivations for raising in /æŋ/. In order to separate the relative contribution of tongue position to the acoustic realization of /æ/-raising from the contribution of other sources (such as nasalization), lingual configuration must be monitored in a way that can be related to the acoustics, e.g., with ultrasound imaging.

We have described three different conditions for contextually raised /æ/: (1) raising before anterior nasals /m n/, which is widespread in North America; (2) raising before anterior voiceless fricatives and certain other consonants, which is observed in the Mid-Atlantic region; and (3) raising before voiced velars /g ŋ/, which is associated with the Midwest, Northwest, and parts of Canada. We will address two related questions about these raised /æ/s. First, we are concerned with how the different types of raised /æ/s are produced, and second, we are concerned with how the various /æ/ raising patterns might have arisen in different parts of North America. We will present data from one Philadelphia speaker with raising before anterior voiceless fricatives and some /d/s, but we will focus primarily on pre-nasal and pre-velar raising.

All these types of raising are conditioned by the following context, but only pre-velar raising involves a rising trajectory in the vowel. We expect pre-velar raising to involve different tongue movements than raising before /m n/, which involves a rising-falling trajectory. It is also reasonable to expect that nasalization will contribute to the acoustic raising effect before nasals, and that the tongue will be lower than would otherwise be expected for such an acoustically high vowel, as argued by De Decker and Nycz (2012) and Baker *et al.* (2008). The most obvious explanation for the development of /æ/ raising before nasals is that the acoustic effects of nasalization were transphonologized to tongue position. Accordingly, we will look for evidence of this effect in similar non-raised vowels (e.g., /æ/s produced by speakers without pre-nasal raising, and other vowels produced by speakers with raising).

We have described a variety of possible causes of raising before voiced velars (palatal-induced upgliding, hyperarticulation before voiceless obstruents, and lower F1 next to voiced obstruents). We may expect the populations of speakers with pre-velar raising to produce /g/ with a more palatal place of articulation, which would be expected to favor raising more than a more posterior place of articulation. The other two motivations involve voicing directly. We may also expect pre-velar raising to be associated with voicing-related differences in tongue root position, e.g., speakers exhibiting more tongue root advancement in /g/ are likely to have lower F1 frequency. This is also related to the /g/ fronting factor, as speakers with less tongue root advancement might be

more likely to facilitate voicing in /g/ by moving the dorsal constriction location forward. Another possible scenario would be that advancing the tongue root is most easily accomplished by moving the whole tongue forward, which would move the constriction forward as well.

B. Deriving articulatory trajectories from ultrasound images

Most of the questions just described require dynamic information about tongue movement during the production of vowel-consonant sequences. Dynamic properties are crucial for characterizing vowels, in particular, not only because of consonantal transitions but also because of vowel-inherent spectral changes (e.g., Nearey and Assmann, 1986; Fox and Jacewicz, 2009; Nearey, 2013). The dataset reported here includes 68 277 ultrasound frames just of /æ/ vowels (from 57 /æ/ words repeated 3 times by 24 speakers), so an automated way to analyze these frames is an important part of an analysis of tongue movement throughout /æ/ vowel intervals.

Contour-based ultrasound image analysis often involves the selection of single representative image from a target segment, followed by tongue surface contour tracing. Examples of articulatory imaging methodologies that preserve temporal information include pellet trajectories from x-ray microbeam data (e.g., Westbury, 1994; Purnell, 2008) and coil sensor trajectories from electromagnetic articulography (e.g., Hoole and Zierdt, 2010). Sequences of traced ultrasound tongue contours can also form part of dynamic analysis (e.g., Proctor, 2011; Mielke, 2012; Zharkova *et al.*, 2014). Articulatory signals can be obtained manually from ultrasound images by deriving time-series data from measured tongue contour tracings (Gick *et al.*, 2006; Falahati, 2013, Chap. 5). M-mode ultrasound imaging also lends itself well to studying dynamic information for slices of the tongue image (Campbell *et al.*, 2010).

Various techniques have been described for extracting dynamic information from whole ultrasound images. One is optical flow analysis, a technique that Moisik *et al.* (2014) have applied to ultrasound images of the larynx. Articulatory signals can be generated by applying a dimensionality reduction or feature extraction method such as principal component analysis (e.g., EigenTongues decomposition; Hueber *et al.*, 2007), Discrete Cosine Transforms (Cai *et al.*, 2011), or Gabor Jets (Berry, 2012, Chap. 3), and treating the resulting vectors as time series. Techniques for relating the output of these methods to linguistic categories or phonetic dimensions include rotation, linear discriminant analysis of PC scores, support vector machines (Berry, 2012, Chap. 3), neural networks (Berry, 2012, Chap. 3), and hidden Markov models (Hueber *et al.*, 2007). Here we apply a relatively simple dimensionality reduction technique, EigenTongues decomposition, which involves the application of principal component analysis to pixel data from filtered ultrasound images. We then use linear regression to derive acoustically relevant articulatory signals from PC score matrices. The model predictions from linear regressions are arranged

334   J. Acoust. Soc. Am. **142** (1), July 2017

Mielke *et al.*

sequentially to create a time-varying signal that represents lingual articulatory correlates of acoustic raising.

II. METHODS

A. Participants

Ultrasound and acoustic data were collected from 24 speakers (15 male, age range 20–72), of which 22 were from regions of North America known to exhibit distinct regional patterns of /æ/-raising. Nine speakers were from the North and the Northwest of the United States, where /æ/ raising before /g/ has been observed. Four speakers were from Canada, where similar pre-velar raising has been observed. We keep these two regions separate because it is not clearly demonstrable that the pre-velar raising patterns reported for both regions are related. Nine speakers were from the South and Mid-Atlantic regions of the United States. We have grouped these two regions together because they both are described as having pre-nasal but not pre-velar raising. We will comment on other differences on a speaker-by-speaker basis. The other two speakers are intended as controls: a male from Newfoundland[1] and a female from England. Neither of these two speakers is expected to display /æ/-raising. The participants and their demographic information are listed in Table I.[2]

B. Materials

The stimuli consisted of 170 English words and English-like nonwords, each of which was presented three times in the experiment. These included 41 stimuli with /æ/ followed by a range of consonants, and in most cases

TABLE I. Participant information.

Subject	Sex	Year of Birth	City, State/Prov	Region
C1	M	1992	Casselman, ON	Canada
C2	F	1970	Ottawa, ON	
C3	F	1991	Barrie, ON	
C4	M	1987	Woodstock, NB	
N1	M	1976	Vancouver, WA	North/Northwest
N2	M	1982	Olympia, WA	
N3	F	1988	Burnsville, MN	
N4	M	1987	Altoona, WI	
N5	M	1981	Fargo, ND	
N6	F	1990	Farmington Hills, MI	
N7	M	1965	Johnstown, OH	
N8	F	1947	Batavia, NY	
N9	M	1950	Buffalo, NY	
S1	M	1992	Broadway, NC	South/Mid-Atlantic
S2	M	1993	Harrisburg, NC	
S3	M	1986	Wilmington, NC	
S4	F	1990	Hickory, NC	
S5	F	1988	Woodbridge, VA	
S6	M	1992	Arlington, TX	
S7	M	1985	Havertown, PA	
S8	M	1941	Philadelphia, PA	
S9	F	1954	Cobbs Creek, VA	
NL	M	1985	Lewisporte, NL	Other
UK	F	1987	Prees, Shropshire, UK	

TABLE II. Stimuli [variable following consonant, (usually) labial preceding consonant].

Context	/æ/	/ɛ/	/e/	/ɑ/	/ɔ/
/p/	app, bap	pep	ape	bop	
/t/	pat, bat	pet	bait	bot	bought
/k/	pack, back	peck	bake	bock	hawk
/b/	bab, ab	ebb	babe	bob	
/d/	pad, bad, fad, sad	bed	bade	bod	pawed
/g/	hag, bag, sag, flag	beg	Hague, vague, plague	bog	hog
/m/	bam, ham, Pam	hem	aim	bomb	
/n/	ban, fan, pan	Ben	bane	bond	pawn
/ŋ/	bang, hang, sang				bong
/f/	half, staff	F			
/θ/	bath, path	Beth		Hoth	
/s/	bass, pass	Bess	pace		
/ʃ/	bash, ash, dash	esh		posh	
/v/	halve, have	Bev	pave		
/z/	as, has	Pez	pays	ahs	
/tʃ/	batch	etch	H	botch	
/dʒ/	badge	hedge	page	hodge	
/l/	pal	bell	pail		ball

preceded by a labial consonant or no consonant (to avoid coarticulatory effects of an additional lingual gesture on the target /æ/). These /æ/ stimuli were matched with 53 stimuli with /ɛ e ɑ ɔ/ in similar contexts. Another 16 stimuli had /æ/ preceded by a variety of consonants, and followed by a labial (in most cases). These were designed to be the mirror images of a subset of the other /æ/ stimuli. A similar set of 21 stimuli had /ɛ e ɑ/ in mirror-image contexts. An additional 39 stimuli were distractors for the purpose of this study, but included items of interest for other research questions. All of the stimuli are listed in Tables II–IV.

C. Procedures

Data collection occurred at two similarly equipped research sites: 20 people participated at the Phonology Laboratory at North Carolina State University in Raleigh, North Carolina, USA, and four people (participants C1, C2, C3, and NL) participated at the Sound Patterns Laboratory at

TABLE III. Stimuli [variable preceding consonant, (usually) labial following consonant].

Context	/æ/	/ɛ/	/e/	/ɑ/
/t/	tab	Tep	tape	top
/k/	cab	Kep	cape	cop
/b/	dab	Depp	dape	dop
/g/	gap, gab, gas	gepp	gape	gopp
/m/	map	mepp	mape	mop
/n/	nab	nepp	nape	nopp
/f/	fab			
/s/	sap	sepp	sape	sop
/ʃ/	shab, shad			
/z/	zap			
/tʃ/	chap			
/dʒ/	jab			
/l/	lap			

J. Acoust. Soc. Am. **142** (1), July 2017

Mielke *et al.* 335

TABLE IV. Stimuli (others).

bin, boat, bold, both, coal, cole, crop, far, foal, free, frog, geese, gold, him, hold, hole, keep, mold, more, mower, oath, only, pa, pea, peel, pill, Poe, pooh, pool, pull, purr, shriek, snarl, south, squirrel, three, throw, trot, whirl

the University of Ottawa in Ottawa, Ontario, Canada. In both labs, data collection occurred inside a sound-attenuated booth, with ultrasound image acquisition occurring on a Terason t3000 ultrasound machine, running Ultraspeech 1.2 (Hueber *et al.*, 2008), recording in direct-to-disk mode, generating 320×240 pixel bitmap images at 60 frames per second. A microconvex array transducer (8MC3 3–8 MHz in Raleigh and 8MC4 4–8 MHz in Ottawa) was used to image with a 90° field of view. Articulate Instruments headsets were used for probe stabilization (Scobbie *et al.*, 2008).

Audio was collected using a head-mounted omnidirectional microphone (an Audio-Technica AT803 lavalier microphone mounted to the headset with an AT8418 instrument mounting clip in Raleigh, and a Shure Beta 53 headset mic in Ottawa), recorded through a SoundDevices USBPre2 preamplifier in Audacity, and synchronized with the ultrasound data afterward.[3] Stimuli were presented one-per-page in a PDF document on a computer screen, advanced by a remote control held by the participant. Each participant saw one of three different randomized orders of stimuli.

Prior to reading the word list, each participant held a mouthful of water in order to generate ultrasound images of the palate (which were not used in the analysis), and held a tongue depressor or plastic utensil between their teeth and pressed their tongue against it in order to generate ultrasound images showing the occlusal plane. Stimuli were presented to participants in three blocks with pauses in between. Each block was begun with an extra word (*dog*, *cat*, *mouse*) in order to avoid having atypical production of a target word.

D. Data analysis

1. Speech segmentation and acoustic analysis

A phone-level segmentation of each audio recording was made using the Penn Phonetics Lab Forced Aligner (P2FA, Yuan and Liberman, 2008). Closure intervals of stops were hand-corrected as necessary, using expected changes in formant structure as an indication of segment boundaries. This correction was most necessary for the boundary between [æ] and [ŋ]. A total of 40 tokens (across 24 speakers) were discarded due to gaps larger than 20 ms in the articulatory data.

The frequencies of the first three formants were measured at 5 ms intervals during all vowel intervals using a PRAAT script that automatically selected the best measurement parameters for each vowel token based on the similarity of the measured formant frequencies and bandwidths to a set of previous measurements. This is based on the procedure described by Evanini (2009, Chap. 4) but it included F3 and considered measurements from two time points, and its models were based on previous measurements of recordings from the Raleigh Corpus of interviews (Dodsworth and

Kohn, 2012). Formant frequencies were normalized using the Lobanov (1971) technique, and where applicable, normalized values were rescaled back into Hertz using parameters measured from the Raleigh Corpus.

2. Image processing and articulatory principal component extraction

The process of deriving principal components from a region of interest within filtered ultrasound images of the tongue (EigenTongues decomposition) is described by Hueber *et al.* (2007). All ultrasound images within one second of segmented speech were included in the analysis. These images were filtered to reduce image noise and to increase contrast between the tongue surface and the rest of the image area, thus improving the ability of the PCA model to explain image variance related to changes in tongue position by reducing spurious image variance (and also variance related to intrinsic lingual muscle tissue). The data were filtered according to the following sequence: anisotropic speckle reduction (edge-sensitive noise reduction; Yu and Acton, 2002; Hueber *et al.*, 2007), median filtering (localized noise reduction), Gaussian filtering (global noise reduction), and Laplacian filtering (edge contrast enhancement). The resolution of the filtered images was then reduced to 30% of the original resolution via bicubic interpolation, to reduce the dimensionality of the input data.

For each speaker's data set, a region of interest was defined in order to reduce the amount of variance not related to tongue surface movement. The region of interest was a polygon surrounding the bounds of the movement of the tongue surface, based on a sample of images spanning the entire length of the recording. The region of interest mask was then applied to each image in the data set, and these masked images were rotated to make the speaker's occlusal plane horizontal. The pixel sites within the region of interest were transposed to a single vector for each image, and a matrix was created from all of the image vectors in each speaker's data set.

Principal component analysis was then applied to each speaker's ultrasound image matrix, to identify principal components (PCs) that represent independent axes of variation within the whole image set.[4] Thus, the PCA model yields a set of PC scores for each frame of a video that indicate how strongly the frame is correlated with each axis of variation. Tongue movements can then be described as trajectories in the space defined by the principal components. Since the articulatory interpretation of the PCs varies from speaker to speaker, it is necessary to transform them in a way that is comparable across speakers. Therefore, formant measurements and linear regression are used in order to separately transform each speaker's PC score vectors into meaningful articulatory parameters, as explained next.

3. Articulatory signal generation and analysis

Two acoustically informed articulatory signals were used in the analysis of /æ/ raising. One articulatory signal is based on the front diagonal of the acoustic vowel space (normalized F2 − normalized F1, or Z2-Z1, where "Z" refers to

Z-scores), with the idea that much of the acoustic variation between raised and unraised /æ/ is along this axis. See Labov *et al.* (2013) for use of a similar acoustic diagonal measure. The articulatory signal related to acoustic Z2-Z1 will be referred to as "lingual Z2-Z1" because it represents the lingual component of movement along the front diagonal of the vowel space. The other articulatory signal ("lingual F1") is based directly on F1, in order to examine the roles of tongue posture and nasalization in the F1 frequency of nasalized vowels. F1 is used here because hypotheses regarding the effect of nasalization on F1 are much more straightforward than hypotheses about its effect on F2 or a combination of F1 and F2. Each signal was generated by a separate regression model for each speaker.

For each signal, the PC scores (which are at 16.7-ms intervals due to the 60 Hz ultrasound frame rate) were linearly interpolated at the time points of the vowel formant measurements (which are at 5-ms intervals). Measurement points at which the bandwidth of F2 was more than 300 Hz were excluded as likely bad measurements. For each speaker's data set, a linear regression was performed with an acoustic measure (either Z2-Z1 or F1) as the dependent variable and independent variables PCs 1–20, for every time point during a vowel lying on the front diagonal [ɑ æ ɛ ej ɪ i], e.g., Z2-Z1 ∼ PC1 + ⋯ + PC20. This cutoff at 20 principal components was established by considering the amount of variance accounted for by subsets of PCs, and by their ability to predict the relevant acoustic variables. The first 20 PCs explain 66%–80% (mean: 73.95%) of the variance in each speaker's image set.[5]

The model predictions at each of the original image time points were combined into a signal representing the lingual analog of the acoustic dimension used as the dependent variable. The time dimension was not included in the PCAs or the linear regressions, but frames that are close together in time are likely to have similar lingual Z2-Z1 values due to similar tongue configurations.

The lingual Z2-Z1 articulatory signal was based on speech intervals segmented as [ɑ æ ɛ ej ɪ i]. For any given ultrasound frame, a higher lingual Z2-Z1 score indicates that the tongue is configured in a manner that correlates more

strongly with a high (acoustic) Z2-Z1 among these vowels. Among these vowels, [ɑ] typically has the lowest Z2-Z1 and [i] typically has the highest. The lingual F1 articulatory signal was based on speech intervals segmented as [ɑ æ ɛ ej ɪ i], excluding vowels produced before /m n ŋ l ɹ/. Excluding nasal contexts means that the F1 values are predicted for an oral vowel, and therefore lingual F1 and acoustic F1 are expected to diverge for nasalized vowels whose F1 is measurably influenced by nasalization.

The units for Z2-Z1 and lingual Z2-Z1 are standard deviations of the original formant measurements (specifically, the number of standard deviations above the mean F2 plus the number of standard deviations below the mean F1). Normalized F1 was rescaled back into Hz before creating the articulatory signal, so the units for (normalized) F1 and lingual F1 are both effectively Hz. An important consequence of the normalization of the acoustic data based on each speaker's entire vowel space is that the values of the articulatory signals are comparable across speakers.

Figure 1 illustrates the lingual Z2-Z1 signal with a sample token of the word *ban* produced by speaker N1. The vertical lines through the spectrogram show the points at which the four sample ultrasound images were extracted. The contour underneath the spectrogram is the lingual Z2-Z1 signal. As the tongue goes up, the distance between F1 and F2 (directly related to acoustic Z2-Z1) gets larger in the spectrogram, and the lingual Z2-Z1 value increases. The lingual Z2-Z1 signal also increases slightly as the tongue blade raises for /n/, demonstrating that the signal reliably depicts articulatory change even in the absence of acoustic information.

Smoothing-spline analysis of variance (SSANOVA), using R's GSS package, was used to make within-speaker comparisons of lingual Z2-Z1 trajectories for different contexts. To compare lingual Z2-Z1 trajectories for groups of speakers in various segmental contexts, a generalized additive model (GAM) was created using the bam function from the R package MGCV. Time was normalized with the start and end of the vowel interval as (0,1), and times in the interval (–0.1,1.2) were included in the model (i.e., including some of the preceding and following consonant intervals). The GAM included an interaction variable (region × context)
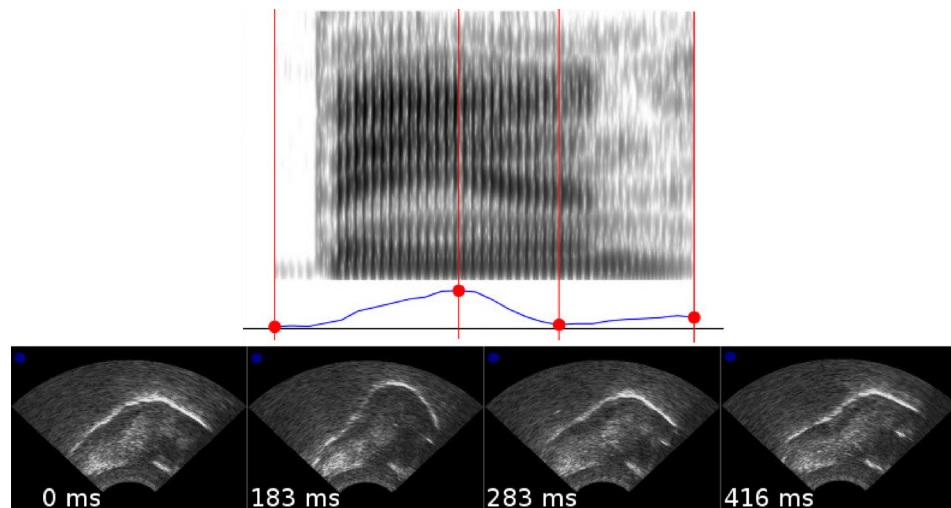

FIG. 1. (Color online) Articulatory signal demo including four ultrasound frames from one token of *ban*. Below the spectrogram is a time-aligned lingual Z2-Z1 signal, with vertical lines indicating the four time points shown in the ultrasound images. The tongue tip points to the right in the ultrasound images.

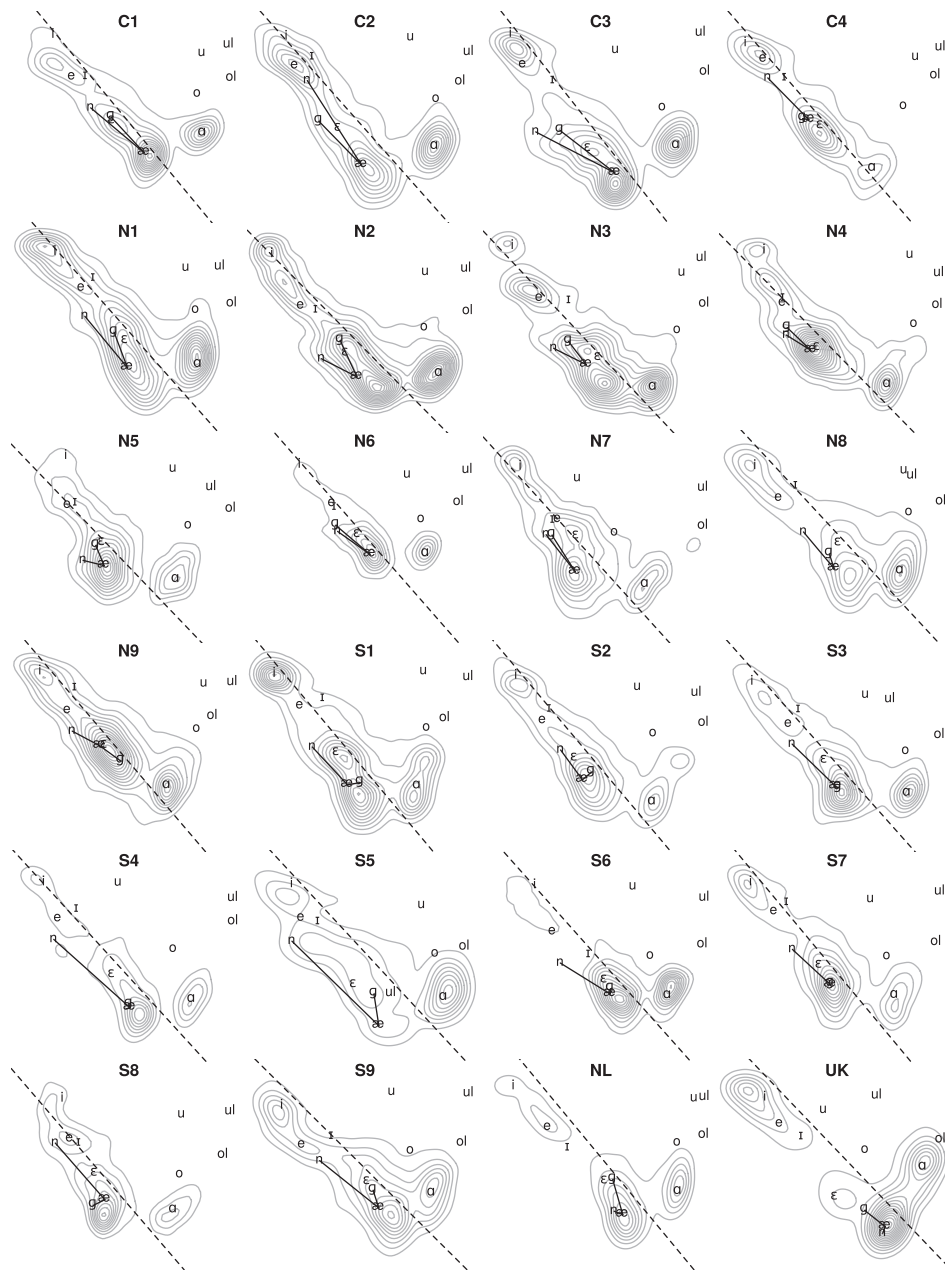0 ms   183 ms   283 ms   416 ms

FIG. 2. Distribution of vowel measurements on the front diagonal. Dashed lines represent slope of acoustic Z2-Z1. Solid lines connect /æ/ in non raising contexts (æ) to its allophones occurring before /g/ (g) and anterior nasals (n).

consisting of all 68 combinations of the four regions (Canada, North/Northwest, South/Mid-Atlantic, Other) and the 17 following contexts /p t k b d ɡ m n ŋ tʃ dʒ f θ s ʃ v z/. The dependent variable lingual Z2-Z1 was modeled with an intercept for region/context, a smooth for normalized time by region/context, and a random smooth for subject. Predicted smooths and confidence intervals were generated for each of the 68 levels of region/context using the get_predictions function from the package ITSADUG. In order to contextualize the /æ/ lingual Z2-Z1 trajectories with trajectories for other front vowels, another GAM or SSANOVA was created with /ɑ ɛ i/ data instead of /æ/ and a smooth for region/vowel instead of region/context.

4. Contour tracing in ultrasound images

In order to validate the articulatory signals analysis and to investigate questions about pre-velar raising in more detail, tongue surface contours were analyzed in a small subset of the ultrasound images (one image each from the middle of the closure interval of the velar consonants following /æ/ in the words *bag*, *hag*, *back*, *pack*, *bang*, and *hang*). These images were selected from the midpoint of the consonant closure (as determined from forced alignment followed by visual inspection of spectrograms) and traced using EdgeTrak (Li *et al.*, 2005). The contours were rotated in order to make the occlusal plane horizontal and downsampled to 25 points before statistical comparisons were performed. Comparisons of tongue contours were made using polar coordinates (Mielke, 2015), so that comparisons are made mostly perpendicular to the tongue surface, instead of mostly perpendicular to the *x* axis. This is especially important for comparisons at the tongue root, where differences are mostly parallel to the *x* axis. The origin for each polar coordinate comparison was located just below the lowest point in any tongue contour, and at an *x* position

(a) S8 /æ/ before /n d g/: trajectories

(b) S8 /æ/ before /n d g/

(c) S8 /æ/ before nasals

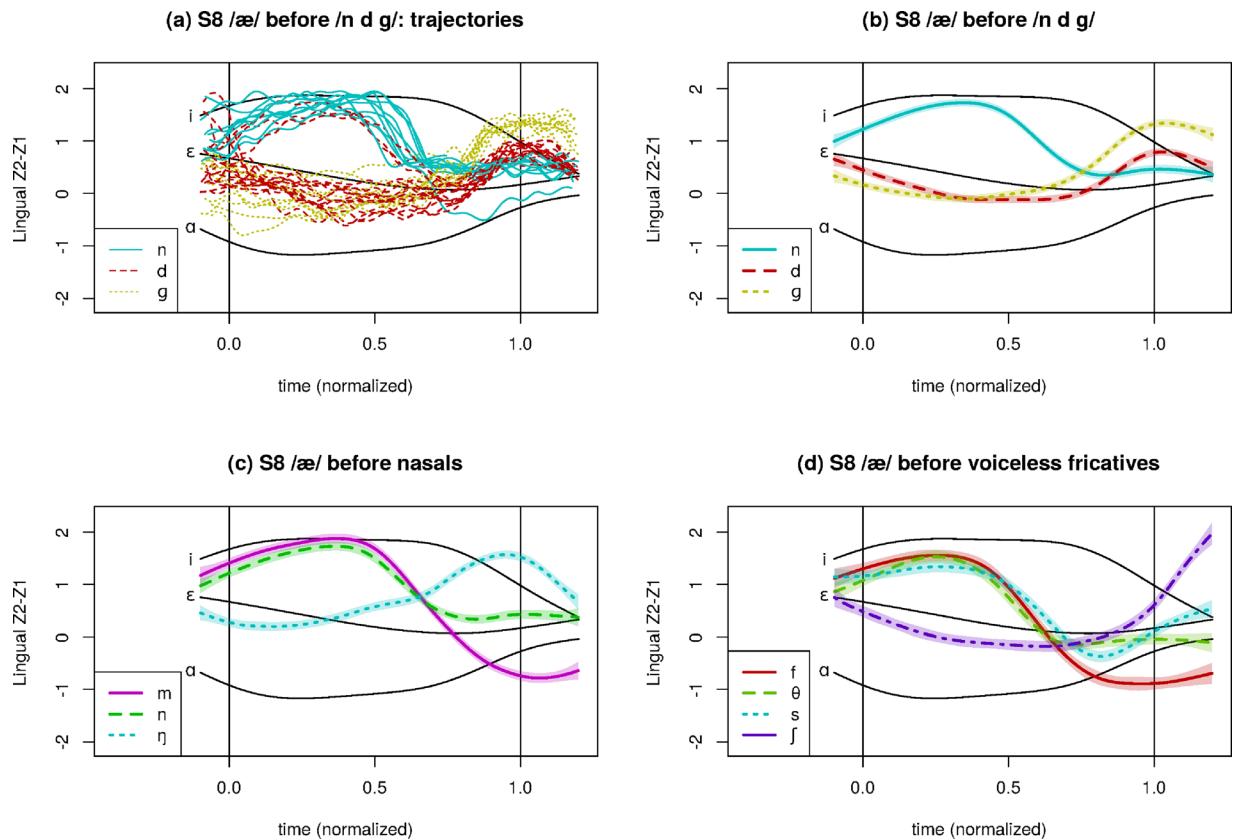**(d) S8 /æ/ before voiceless fricatives**

FIG. 3. (Color online) Comparisons of /æ/ before various consonants for Philadelphia speaker S8.

2/3 of the way from the median of the posterior endpoints of the tongue contours to the median of the anterior endpoints. The origin was placed manually for three speakers whose tongue traces did not extend very far down the tongue root or tongue blade (N4, N8, and N9).

III. RESULTS

A. Overview

The formant measurements reported in this section generally show the expected regional patterns of /æ/ raising. Figure 2 shows the distribution of front vowel measurements for the 24 participants. Measuring vowel formants at 5 ms intervals resulted in an average of 13 591 vowel measurements per speaker (range 8644-21 423). The contour lines in each sub-figure indicate the distribution of these measurements in the F1-F2 plane. In each sub-figure, the dashed line has the slope of acoustic Z2-Z1 and passes through the center of gravity of the (F1, F2) points. This alignment approximates the front diagonal of each speaker's vowel space. The phonetic symbols indicate median (F1, F2) values at 25% of the duration for reference vowels included in the front diagonal (/i ɪ e ɛ ɑ/) as well as back vowels before /l/ and elsewhere (/ol ul o u/). The "æ" symbol indicates the median (F1, F2) values for /æ/ before consonants that are not expected to trigger raising for any of the speakers (/p t k b ʃ tʃ dʒ/). The /n/ indicates /æ/ before anterior nasals /m n/ to indicate pre-nasal raising, and the /g/ indicates /æ/ before /g/, measured at the 75% point in the vowel, to indicate the raising that occurs before /g/ (pre-velar raising).

The three Ontario speakers (C1–3) have considerable raising before both anterior nasals and /g/. The New Brunswick speaker (C4) has a more raised /æ/ in general, but especially before anterior nasals. The first seven North/ Northwest speakers (N1-2 from the Northwest and N3-7 from the Midwest) also have pre-nasal and pre-velar raising, as expected. The other two (N8-9 from upstate New York) show only pre-nasal raising. The South/Mid-Atlantic speakers (S1-9) have pre-nasal raising and in most cases no pre-velar raising. These speakers also have lower and backer /æ/ overall than the North/Northwest speakers. Finally, the speakers from Newfoundland (NL) and England (UK) show no pre-nasal /æ/-raising.

Turning to the articulatory data, we begin with speaker S8. This speaker was selected for this demonstration because he is the only speaker in the sample who exibits the classic Philadelphia /æ/ system described by Ferguson (1975); Payne (1980); Labov (1989); and Labov (1994) (p. 516), *inter alia*. Singling out this speaker gives us an opportunity to examine raising in various Philadelphia contexts alongside the pre-nasal raising that is much more widespread in our sample of speakers. Throughout the results section, examination of the interaction of consonants and vowels in /æ/-raising will proceed by analysis of the lingual Z2-Z1 signal across the duration of the vowel and part of surrounding consonants. Figure 3(a) shows traces for individual tokens of S8's lingual Z2-Z1 trajectories for tokens of /æ/ produced before the consonants /n d g/. The *x* axis shows normalized time, with 0 and 1 indicating the start and end of the vowel. Larger positive values on the *y*-axis indicate tongue postures

J. Acoust. Soc. Am. **142** (1), July 2017
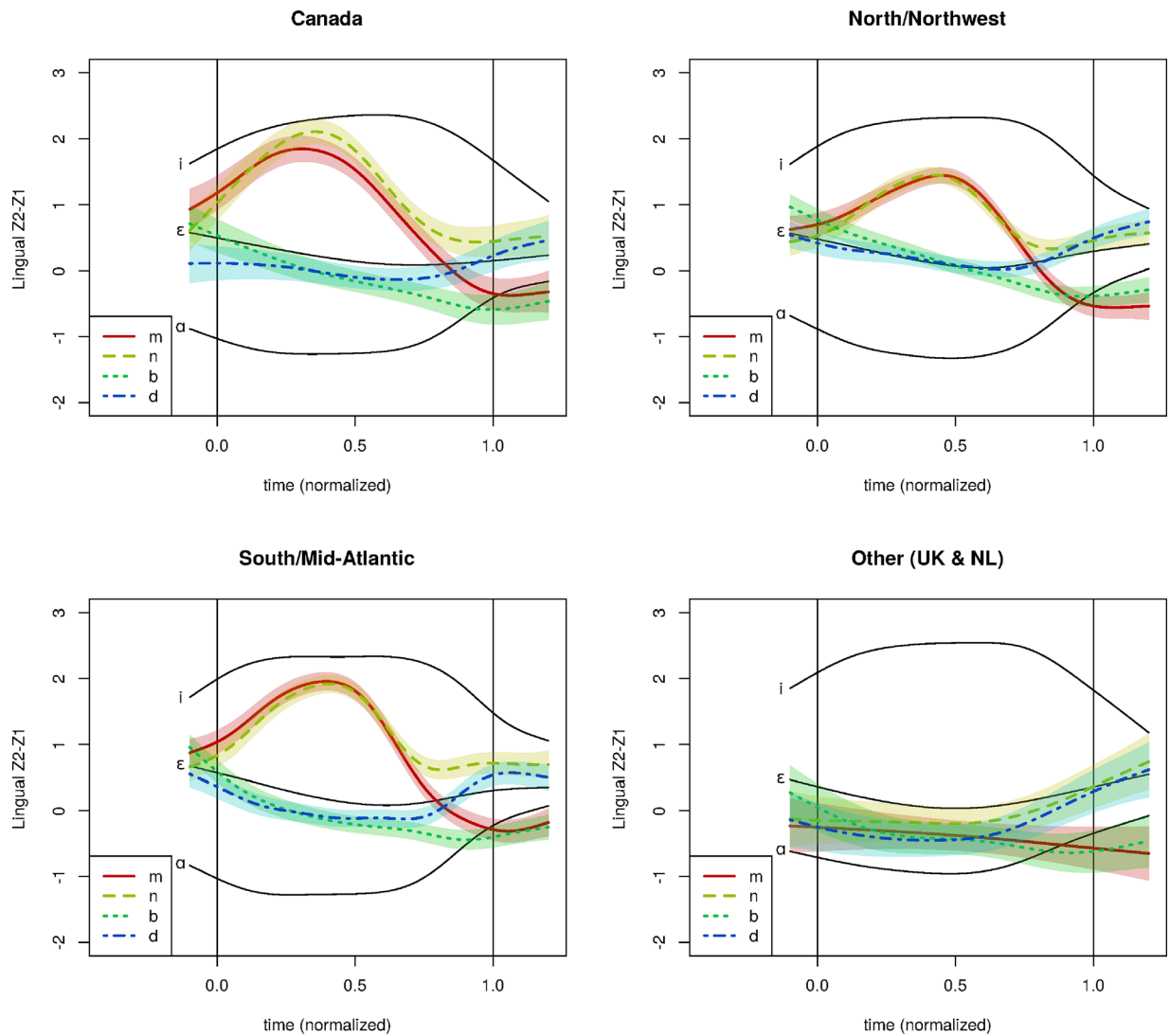
Mielke *et al.*   339

FIG. 4. (Color online) Comparisons of lingual Z2-Z1 for /æ/ before anterior nasals and voiced oral stops by region.

associated with higher and fronter vowel quality. For reference, lingual Z2-Z1 fits for vowels /ɑ ɛ i/ are indicated with labeled contours. Figure 3(b) shows a SSANOVA (Gu, 2002) comparison of these trajectories. Dark contours indicate category fits, and shading indicates 95% confidence intervals around the mean. Two categories are considered to be significantly different at time points for which the confidence intervals do not overlap.

The tongue position for /æ/ before /d g/ is above /ɑ/ and slightly below /ɛ/, as expected, and it is significantly higher in /æ/ before /n/, with a peak approaching the height of /i/ that occurs at about 40% of the vowel's duration. This figure also shows the tongue rising at the end of the pre-/g/ vowel and remaining there during the consonant closure (beyond time 1), while alveolar /n d/ have tongue positions that are more similar to each other. The individual traces in Fig. 3(a) make it clear that there is a lexical distinction within the /æd/ category: *mad*, *bad*, and *glad* are known to have raised /æ/ in Philadelphia, while *sad*, etc., have non-raised /æ/. For most of the vowel interval, three tokens of /æd/ (the word *bad*) are indistinguishable from /æn/, and the others (*pad*, *fad*, *sad*) are indistinguishable from /æg/. The tokens of

bad are excluded from the SSANOVA comparison in Fig. 3(b).

The other two subfigures show raising before two sets of anterior consonants. Figure 3(c) compares /æ/ before the three nasals, and /æ/ shows almost identical tongue raising before the two anterior nasals /m n/. However, /æŋ/ is realized with a completely different articulatory pattern: a rising trajectory that is dynamically similar to what was seen in /æg/, but which starts earlier in the vowel and involves a higher tongue position throughout the entire vowel interval. Figure 3(d) compares /æ/ before voiceless fricatives, and it is clear that the tongue is raised before anterior /f θ s/ but not before /ʃ/, consistent with the described acoustic pattern. Moreover, the trajectory observed before anterior voiceless fricatives reveals a gesture similar to /æ/ before /n/ and in *bad*, and the trajectory of /æʃ/ is similar to the non-raised contexts in Figs. 3(a) and 3(b). They diverge primarily in the transitions to the coda consonants.

In summary, there is a clear articulatory distinction between raised and non-raised /æ/ for this speaker. /æŋ/, while traditionally categorized as non-raised, does show a rising trajectory and a raised tongue body position. Of the
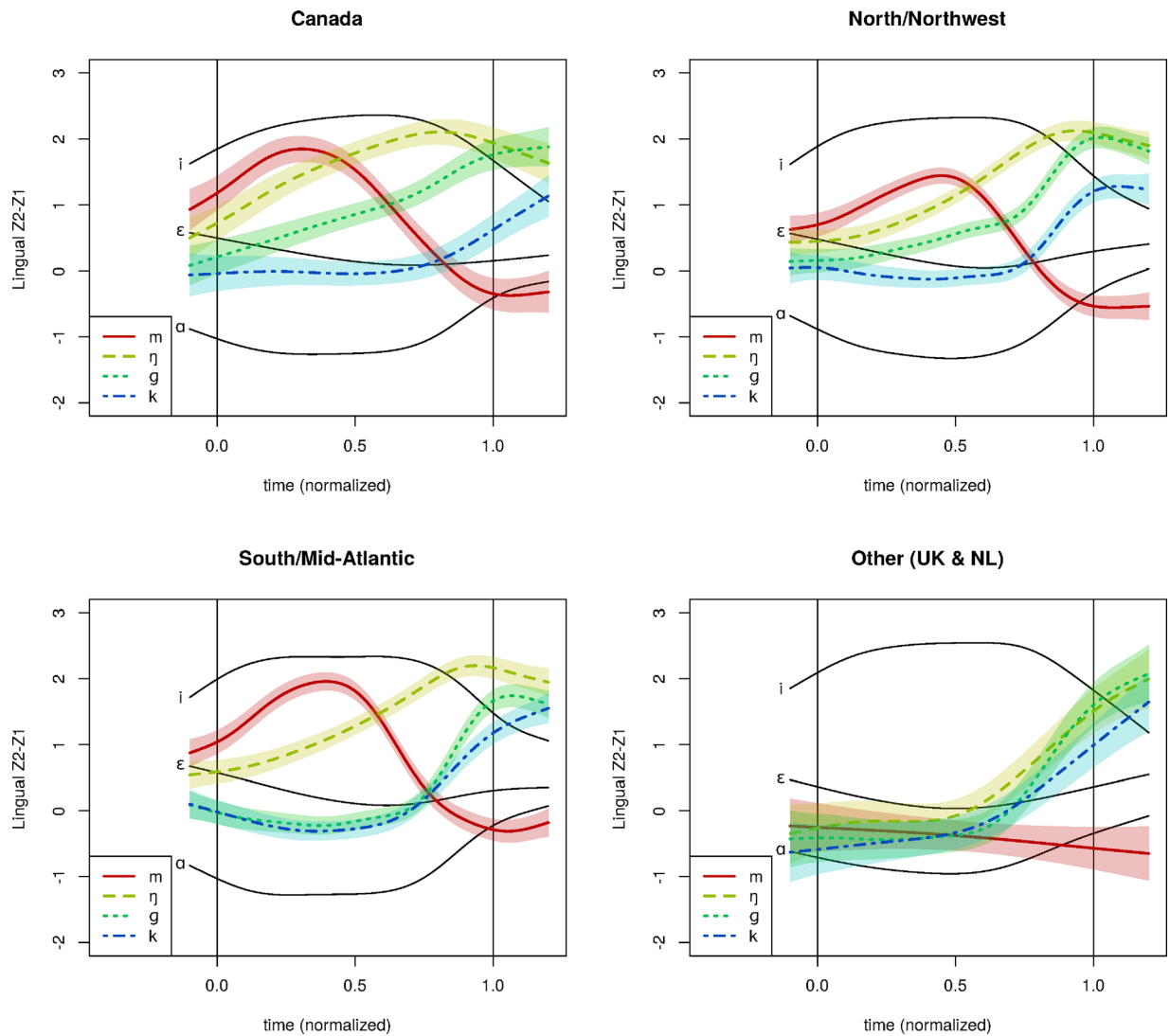
FIG. 5. (Color online) Comparisons of lingual Z2-Z1 for /æ/ before /m/ and velars /k ɡ ŋ/.

contexts in which raised /æ/ is observed for this Philadelphia speaker, the following /m n/ is the only context in which the rising-falling pattern is widespread in North America or in our sample. Figure 4 shows lingual F1 trajectories (GAM predictions) for /æ/ before /m n b d/. The Canada, North/Northwest, and South/Mid-Atlantic groups all show prominent lingual Z2-Z1 peaks just before the midpoint of /æ/ before both of the nasals. There are no peaks in /æ/ before the corresponding voiced oral consonants /b d/. The other group (speakers from the UK and Newfoundland who are not expected to have pre-nasal raising) show no peaks at all, and no significant differences between /m b/ or /n d/, meaning that their /æ/s are articulated with the same lingual gestures preceding both oral and nasal codas. Not surprisingly, all four groups of speakers show significant differences between the alveolar consonants /n d/ and the bilabial consonants /m b/ beginning slightly before the start of the consonant interval. The elevated lingual Z2-Z1 during the alveolar consonant intervals is due to the tongue blade raising involved in producing an alveolar consonant. It is clear that three of these groups of speakers have a lingual raising gesture that is conditioned by the following consonant, and this

gesture is independent of the oral articulation of the following consonant. The height of the pre-nasal peak is noticeably lower for the North/Northwest group than for the Canada and South/Mid-Atlantic groups.

It was apparent above in Fig. 3(d) that S8's /æŋ/ exhibits a rising lingual trajectory, and this pattern is explored in Fig. 5, where /æŋ/ is compared to /æm/, which has the pre-nasal raising pattern just described, and the other two pre-velar contexts. It is useful to consider the three velar consonant contexts together, because all of them are produced with tongue raising that reaches the velum at the end of the vowel, but they exhibit different trajectories during the vowel. The vowels are compared in normalized time, and English vowel duration varies systematically by following consonant, so it is important to consider how much of the differences between trajectories are due to time normalization. The grand means of the vowel duration for /æ/ before velar consonants are as follows: 255 ms before /k/, 307 ms before /ɡ/, and 284 ms before /ŋ/. These are all produced in utterance-final syllables, so the differences in duration are small, and they clearly do not account for the differences between the trajectories in Fig. 5.
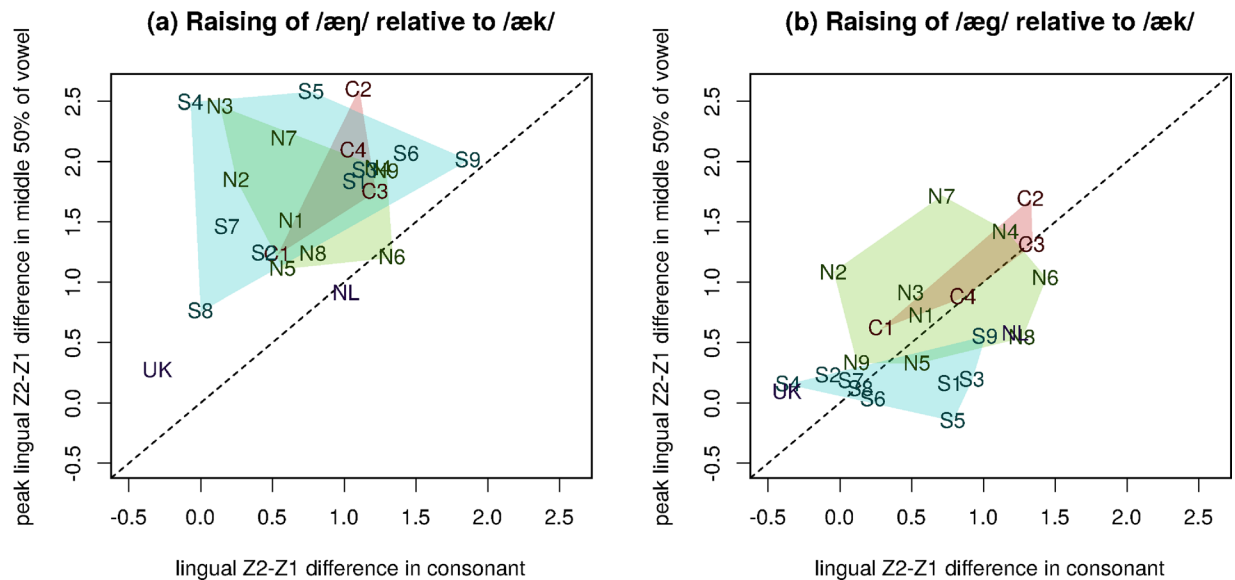
J. Acoust. Soc. Am. **142** (1), July 2017

Mielke *et al.* 341

FIG. 6. (Color online) F1 vs lingual F1 in /ɑ æ ɛ i/ before /m/ and /b/ and /æ/ before /ŋ ɡ k/. Vowel categories are indicated by symbols and redundantly by color. Error bars indicate one standard error in each direction. Lines connect the same vowels across different contexts.

For the three groups with pre-nasal raising, /æŋ/ involves a tongue raising gesture that is similar in magnitude to raising in other pre-nasal contexts, but distinct in timing. The /æm/ and /æŋ/ trajectories cross just after the peak of the /æm/. The lingual Z2-Z1 values during the coda consonant intervals are similarly high for all three velar consonants, but the signal reaches this height much earlier before /ŋ/. The Other group shows no difference between the pre-velar /æ/s, and the only difference is between these three and /æm/, simply due to the difference in consonant place of articulation.

High lingual Z2-Z1 at the end of the vowel before /k ɡ ŋ/ is unsurprising, since the velar consonants are produced with a dorsal constriction. What happens earlier in the vowel interval varies widely between the four groups. For the Canada, North/Northwest, and South/Mid-Atlantic groups, /æŋ/ involves higher lingual Z2-Z1 than /æk/,

throughout the vowel's duration. For the Canada and North/Northwest speakers, /æɡ/ is produced with lingual Z2-Z1 that is significantly higher than /æk/ but lower than /æŋ/, for most of the vowel's duration. These large differences throughout most of the vowel interval can be attributed to the phonological pattern of pre-velar raising.

For the South/Mid-Atlantic group, /æɡ/ is identical to /æk/ except for a small difference near the start of the consonant interval. For the Other group, there is a similar but non-significant difference between the /æɡ/ and /æŋ/ fits and /æk/. This is consistent with a general phonetic effect that is present even in groups without pre-velar raising, but only significant for the South/Mid-Atlantic group, which has more speakers (9) than the Other group (2). Assuming that the tongue achieves a dorsal stop closure for these stop consonants, the higher lingual Z2-Z1 during the /ŋ/ and /ɡ/ closures suggests they are produced with a more anterior tongue

FIG. 7. (Color online) Differences in lingual Z2-Z1 in raising contexts (/æŋ/ and /æg/), relative to /æk/, with consonant difference on the *x* axis and vowel difference on the *y*-axis.

position than /k/ by at least some speakers. This consonant difference could be a consequence of raising during the vowel, or it may indicate that the consonants that condition /æ/ raising are produced with a more anterior place of articulation. This varies within the four groups, and it is explored in detail below. First we turn to the role of nasalization in pre-nasal /æ/ raising.

B. Phonetic motivations for pre-nasal /æ/-raising

It is clear that the Canada, North/Northwest, and South/Mid-Atlantic groups produce /æ/ with tongue raising gestures before all three nasals. To examine the relative importance of tongue raising and nasalization in the acoustic realization of pre-nasal /æ/, we compare F1 frequency with lingual F1 for an assortment of vowels before oral and nasal codas. Most comparisons will focus on vowels followed by /m/ and /b/, to avoid lingual consonant gestures that may interfere with predicted F1.

The values of each speaker's normalized acoustic F1 and lingual F1 were averaged (separately) for the middle 50% of each vowel token, and then averaged by region. Figure 6 shows the lingual and acoustic F1 for the vowels /ɑ æ ɛ i/ before /m/, where they are nasalized, and before /b/, where they are not nasalized. It also shows some pre-velar vowels, which will be addressed below. In general, vowels with low F1 that is achieved by tongue raising (such as [ejb ejm]) appear in the upper right corner, and vowels with high F1 that is achieved with a low tongue body (such as [æb]) appear in the lower left corner. If lingual F1 perfectly predicted F1 frequency, then all vowels would be arranged along the dotted line (*x* = *y*). Factors that affect F1 frequency but do not contribute to lingual F1 should cause vowels to appear off the diagonal. For example, lip rounding would cause vowels to appear higher in the figure than a comparable unrounded vowel. Since pre-nasal vowels were not included in the data used to generate lingual F1, we expect pre-oral vowels to appear on the diagonal and pre-nasal vowels to appear off it (if nasalization affects their F1 in any

way). More specifically, if nasalization plays a substantial role in lowering F1 in /æ/, then /æm/ should appear above the diagonal, because some of its F1 lowness would not be accounted for by tongue position.

The large distance between /æb/ and /æm/ in both the F1 and lingual F1 dimensions for the Canada, North/Northwest, and South/Mid-Atlantic groups indicates that the large difference in F1 is related to a large difference in tongue posture. /æm/ is below the dotted line for all three groups, meaning that its F1 is higher than what is predicted on the basis of tongue posture, i.e., that non-lingual factors actually make /æ/ before /m/ sound like a lower vowel. The Other group lacks the pre-nasal raising pattern in both articulation and acoustics. For the three groups with raising, tongue position accounts for the observed F1 lowering, with no indication that nasalization itself contributes to the observed acoustic difference. The fact that F1 in /æm/ is higher than expected on the basis of tongue position may be because pre-nasal /æ/ is such a high vowel (due to tongue body position) that its F1 is raised by nasalization, as in high vowels.

All four groups show signs of pre-nasal /ej/ having higher F1 than is expected on the basis of lingual F1 (i.e., they are below the diagonal). Such raising of F1 is predicted for nasalized high vowels (Fujimura and Lindqvist, 1971; Feng and Castelli, 1996). The other two vowels (/ɑ ɛ/) show little difference between the pre-oral and pre-nasal contexts. The lack of difference between [ɑm] and [ɑb] is potentially surprising, given the F1 lowering that is predicted for a nasalized low vowel.

Earlier we described another possible acoustic consequences of velum lowering in /æ/ before /ŋ/, namely, F1 raising due to the change in oral cavity shape (Baker *et al.*, 2008). To investigate this, we compared lingual F1 and F1 frequency for an interval late in the vowel (the interval (0.45, 0.95) of the vowel's duration), for /æ/ before /ŋ/ and /g/. These are also shown in Fig. 6, and /æk/ is also included for
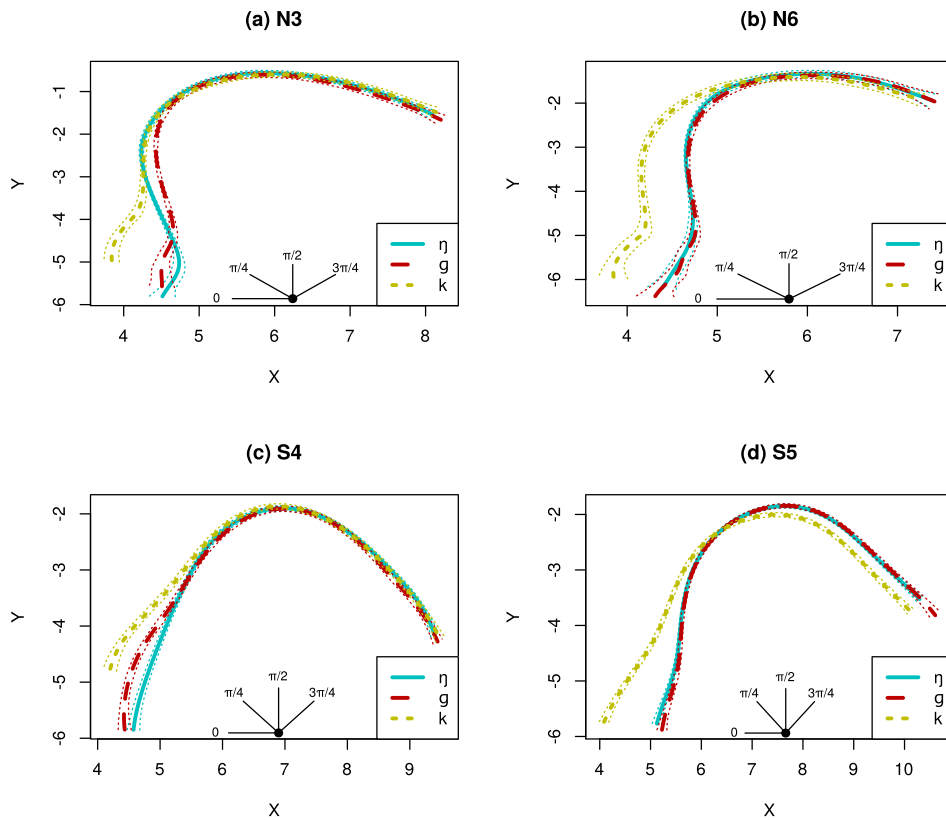
J. Acoust. Soc. Am. **142** (1), July 2017

Mielke *et al.* 343

FIG. 8. (Color online) SSANOVA comparisons of tongue postures during /k g ŋ/ after /æ/. The tongue tip points to the right.

reference. It is clear that /æŋ/ involves more tongue raising and lower F1 (relative to /æg/ and /æk/) for the Canada, North/Northwest, and South/Mid-Atlantic groups of speakers, and once again the nasalized vowel is below the diagonal, indicating that tongue raising again accounts for the observed F1 lowering (because lingual F1 is even lower than actual F1), and there is no evidence that the change in oral cavity shape caused by velum lowering reduces F1 beyond what is accounted for by tongue position.

/æk/ and /æg/ are both included as reference points for /æŋ/ because /æg/ is involved in raising in the Canada and North/Northwest groups. One additional interesting fact is that /æk/ consistently appears below the diagonal. This could be due to differences in tongue root position that affect F1 but are not fully captured by ultrasound imaging (and therefore not by lingual F1 either). The difference is in the direction that is expected if the voiced obstruents are produced with tongue root advancement that lowers F1 of the preceding vowels, but /k/ is produced with a more retracted tongue root position, causing F1 raising that is not captured by lingual F1. Tongue root differences will be explored in more detail in Sec. III C.
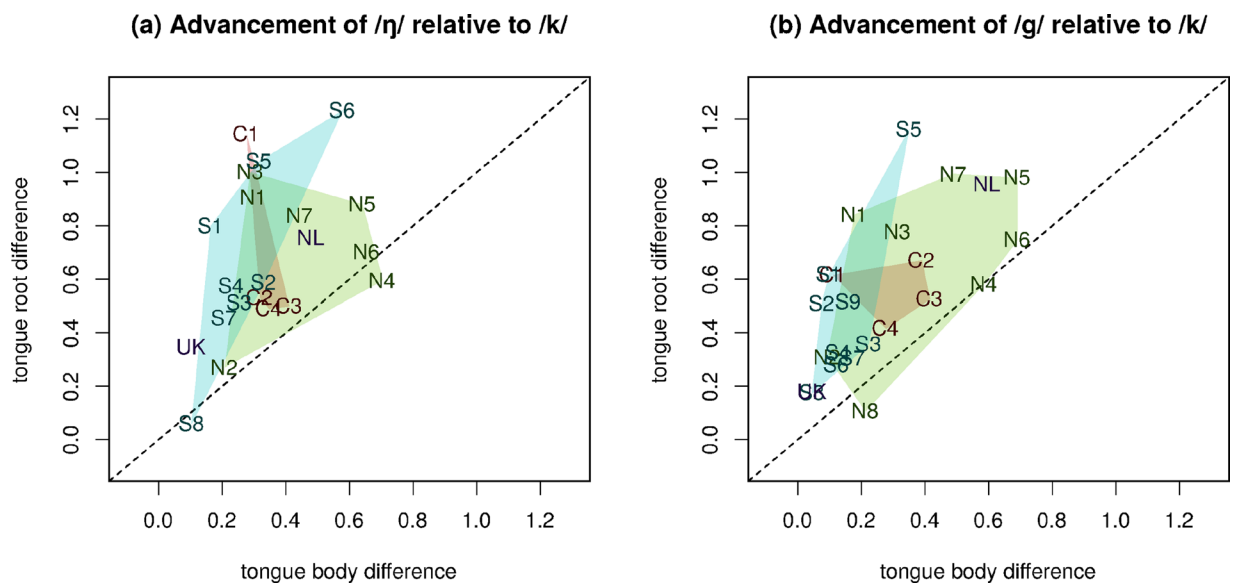
FIG. 9. (Color online) Differences in SSANOVA comparisons between velar consonants produced after /æ/. Tongue root advancement difference (between each voiced velar and /k/) is shown on the *y*-axis, and a measure of tongue body advancement is shown on the *x* axis.

The purpose of this comparison between F1 and lingual F1 was to look for signs that tongue position does not fully account for the F1 lowering observed in pre-nasal /æ/. Instead, we have found that the lowering of the velum into the oral cavity does not appear to lower F1 in /æŋ/, and that velo-pharyngeal coupling may actually raise F1 in /æ/ before nasals. This is consistent with raised /æ/ being articulated with a high enough tongue position, and having such a low F1, that the effect of nasalization on F1 is reversed.

C. Phonetic motivations for pre-velar /æ/-raising

We have already seen that there is /æ/-raising, with a rising lingual trajectory, before /ŋ/ and /g/ in the Canada and North/Northwest regions, and just before /æŋ/ in the South/Mid-Atlantic region. Additionally, we have seen that all groups of speakers show at least some signs that the voiced velars /g ŋ/ are produced with a more anterior tongue position than /k/. Figure 7 illustrates raising in /æŋ/ and /æg/ for each speaker, comparing the magnitude of raising during the vowel and consonant intervals. These comparisons treat /æk/ as a baseline. The values plotted here are based on the differences between the SSANOVA fits for subject-by-subject comparisons of /æŋ æg æk/, i.e., the individual version of the pre-velar comparisons in Fig. 5. In Fig. 7(a), the y-axis shows the maximum difference between the lingual Z2-Z1 fits for the middle 50% of the vowel intervals in /æŋ/ and /æk/. The x axis shows the difference during the consonant interval (measured during an interval between the 1.1 and 1.2 time points in Fig. 5, i.e., shortly after the start of the consonant). Speakers are represented by their codes (as shown in Table I, with C = Canada, N = North/Northwest, and S = South/Mid-Atlantic), and the figure regions enclosed by convex hulls for each dialect region (excluding "Other") are shaded. If there were no lingual difference between /æŋ/ and /æk/, all speakers would be at (0,0). All but the UK speaker show significant raising during the vowel interval, and all have y values well above 0 (which corresponds to no raising in /æŋ/ relative to /æk/). The smallest amount of raising is shown by the Philadelphia speaker (S8), whose /æŋ/ trajectory is shown above in Fig. 3. For nearly all of the speakers, there is a greater difference in the vowel interval than in the following consonant interval (visible in the fact that nearly all speakers are above the diagonal line $x = y$), indicating that raising in /æŋ/ is a property of the vowel, a phonological pattern that does not appear to vary greatly across the three North American regions.

Figure 7(b) shows the same information for /æg/ vs /æk/. The regional difference in /æg/ raising is apparent in the fact that all of the North/Northwest and Canada speakers are higher along the y-axis than all but one of the South/Mid-Atlantic speakers. This means that the North/Northwest and Canada speakers have considerable raising before /g/, and the South/Mid-Atlantic speakers do not. The most extreme vowel differences are comparable with the raising in /æŋ/, and no speaker shows more vowel raising in /æg/ than in /æŋ/. In contrast to /æŋ/, many of the speakers' /æg/ appears below the diagonal, indicating a greater difference in the consonant than in the vowel. If all North/Northwest speakers exhibited the pattern

shown by N6 and N8 (some vowel difference but larger consonant difference), that would suggest that raising in /æg/ is a coarticulatory effect driven by a consonant difference. However, the group of North/Northwest /æg/ raisers includes N2 and N7, who show a large vowel difference and smaller (or no) consonant difference. The other North/Northwest and Canada speakers fall closer to the diagonal, between these two extremes.

The speakers in the non-raising (South/Mid-Atlantic and Other) groups exhibit two patterns: the several who are clustered around the origin (no raising apparent in the vowel or the consonant), and five (S1, S3, S5, S9, and NL) who show a considerable difference in the consonant but not the vowel. These differences within the consonant closure interval, which are found only among some speakers, suggest a latent articulatory motivation for /æg/ raising. It is necessary to examine the articulatory differences between /g/ and /k/ more closely, by looking at the traced tongue contours from the middle of the consonant closure intervals.

Figure 8 shows SSANOVA comparisons of tongue contours (traced as described in Sec. II D 4) for pairs of North/Northwest speakers and South/Mid-Atlantic speakers who showed large or small differences in lingual Z2-Z1 for /g/. All four show a significant difference in tongue root position, with voiceless /k/ more retracted than voiced /g ŋ/. Speakers N3 and N6 both showed raising in the vowel, but N6 showed considerably more difference in the consonant interval. Here we see that N3 shows a more advanced tongue root in /g/ and /ŋ/, and a comparatively smaller difference in the back of the tongue dorsum for /g/ (relative to both /k/ and /ŋ/). N6 shows much more advanced tongue root and tongue body position in /g/ and /ŋ/ than in /k/. Speakers S4 and S5 both lacked raising in the vowel interval, but S5 showed a lingual Z2-Z1 difference in the consonant interval. Here we see that both have more tongue root advancement in /g/ and /ŋ/, but S5 has a more anterior tongue body position for /g/ and /ŋ/, and a larger difference in tongue root advancement. Each sub-figure shows the polar origin as a black dot.[6] Line segments indicate four angles (shown in radians) that will be involved in between-speaker comparisons.

Figure 9 summarizes the tongue shape differences between the SSANOVA fits for the velar consonants. This is similar to Fig. 7 except that the two axes show two different types of differences in the consonant intervals. In each sub-figure, the y-axis shows the difference in tongue root position between /ŋ/ or /g/ and /k/, measured as the maximum difference in the interval $(0, \pi/4)$, representing the tongue root. The x axis represents the anteriority of the dorsal constriction by subtracting the maximum difference in the interval $(\pi/4, \pi/2)$ from the maximum difference $(\pi/2, \pi)$ and dividing by two. In other words, the y-axis shows how advanced the tongue root is in /ŋ/ or /g/ relative to its position in /k/, and the x axis shows how much farther forward the dorsal constriction is. Tongue root advancement and tongue body fronting are both associated with positive values.

Nearly all of the speakers show greater tongue root advancement relative to /k/ in both of the voiced velars. For the South/Mid-Atlantic speakers, there is a roughly linear

relationship between the amount of tongue root advancement and the amount of tongue body fronting. Of the four South/Mid-Atlantic speakers previously identified as having higher lingual Z2-Z1 in /g/, three (S1, S5, and S9) have more tongue root advancement in /g/ than the other speakers in their region. S2 has a similar amount of tongue root advancement, which was missed by the lingual Z2-Z1 measure, a difference that could be due to the lingual Z2-Z1 comparisons being made earlier in the vowel than the tongue contour comparisons. S3 and S5 have the most tongue body fronting in /g/.

The Canada and North/Northwest speakers have tongue root advancement in the same range as the South/Mid-Atlantic speakers. As a group, the Canada speakers show slightly more tongue body fronting, relative to South/Mid-Atlantic speakers with a similar amount of tongue root advancement. The North/Northwest speakers show a wide range of tongue body differences. The speakers with the largest tongue body differences are N4, N5, N6, and N7 (all from the Midwest), along with NL. These speakers with the largest tongue body differences were also identified as having large lingual Z2-Z1 differences during the consonant closure. Interestingly, the same tongue body differences are observed in /ŋ/, which shows much less variation in the preceding vowel quality. In other words, the tongue body differences observed in the Midwestern speakers and the Newfoundland speaker do not appear to be due to the preceding raised vowel, because similar tongue body differences are not observed in /ŋ/ for the non-Midwestern speakers who also raise before /ŋ/. Three speakers with considerable raising before /g/ (C1, N1, N2) show consonant articulation that is consistent with the South/Mid-Atlantic speakers who do not raise in this context, suggesting that their raising is not a direct result of their velar consonant production.

IV. DISCUSSION

We have seen that all of the Canada, North/Northwest, and South/Mid-Atlantic speakers exhibit a tongue raising gesture in /æ/ before all three nasals. The main difference between the anterior nasals (/m n/) and /ŋ/ is that /æm/ and /æn/ are produced with a rising-falling tongue gesture peaking around the vowel midpoint, while /æŋ/ is produced with a mostly rising trajectory that peaks late in the vowel. These two types of pre-nasal patterns are both observed in all of these speakers, and the magnitude of the lingual peak is similar across the nasal contexts. The magnitude of the raising gesture in /æ/ before /m n/ is somewhat smaller for the North/Northwest speakers than it is for the South/Mid-Atlantic and Canada speakers, but raising before /ŋ/ has about the same magnitude for all three groups. A significant lingual Z2-Z1 difference was observed for /æŋ/ vs /æk/ for all of the North American speakers (including the speaker from Newfoundland) but not the UK speaker.

For the Philadelphia speaker S8, the lingual trajectory for /æ/ before anterior voiceless fricatives /f θ s/ appears identical to the lingual trajectory before anterior nasals. The regionally limited /æ/ raising pattern before /g/ appears very similar to /æ/ raising before /ŋ/, with the two main differences being that it is observed only in the Canada and North/Northwest groups, and that the raising gesture is consistently smaller in magnitude and later in the vowel than raising before /ŋ/.

While it is reasonable to think that the acoustic consequences of nasalization are the original phonetic basis of pre-nasal /æ/ raising, we did not find evidence of any such effects operating in present-day speakers. Tongue raising alone accounts for F1-lowering in pre-nasal /æ/. This finding contradicts what De Decker and Nycz (2012) reported about two of their New Jersey speakers, but it seems likely that this apparent difference is methodological. DeDecker and Nycz measured formants at the vowel midpoint but traced the ultrasound frame which displayed the most posterior tongue position in the vowel interval, with the idea that the constriction location for /æ/ is in the pharynx. As shown here, however, the raised /æ/ is quite high and front, and the most retracted tongue position in the trajectory does not correspond to the peak of the raising gesture, which is near the vowel midpoint.

It is somewhat surprising that no evidence materialized for the predicted acoustic effect of nasalization in pre-nasal /æ/, i.e., an independent F1-lowering due to velo-pharyngeal coupling. One explanation is that the tongue postures for the raised and non-raised /æ/ differ greatly. In fact, the most extreme cases of pre-nasal /æ/-raising were realized quite high in the vowel space, e.g., often peaking near [ĩ]. Since the effect of nasalization on the F1 dimension is a tendency towards mid realizations (i.e., F1 is lowered for nasalized low vowels and raised for nasalized high vowels), one would predict the opposite effect for these high variants of /æ/, namely raised F1, which is indeed what occurred for pre-nasal /æ/ (Fig. 6). This would explain why pre-nasal /ej/ is also manifested with a raised F1 that is not accounted for by lingual position. In other words, F1-lowering is predicted for nasalization of the low vowel [æ], but for cases where /æ/ is actually realized as [ẽ͡ã] or [ẽ͡j̃], F1-raising rather than F1-lowering is predicted. We also did not observe signs of an effect of velum lowering on oral cavity shape, suggested by Baker et al. (2008), which predicted that the F1 difference would exceed the difference found in the lingual analog of F1. If raised /æ/ is sufficiently high that the F1-centralizing effect of nasalization raises its F1 instead of lowering it, then it is possible that the influence of velo-pharyngeal coupling obscures the effect of velum lowering on oral cavity shape. This possibility invites further study.

The two apparent natural sources of F1 lowering late in /æ/ before /g/ are tongue root advancement to facilitate voicing and tongue body raising to achieve the velar closure. In /æ/+velar sequences, nearly all of our speakers produce /ŋ/ and /g/ with a significantly more advanced tongue root than /k/. The magnitude of /g/'s tongue root advancement does not differ across regions (even though raising before /g/ is observed in only two of the regions), and it does not differ from tongue root advancement in /ŋ/ (even though the aerodynamic motivation for tongue root advancement is greater for the obstruent /g/. Within each region, some speakers have considerably more tongue root advancement than other speakers in the same region.

All of the speakers obviously have a dorsal constriction for /g/ and /ŋ/, but four Midwestern speakers (N4-7) and the Newfoundland speaker have a more anterior dorsal constriction for these consonants. Within the South/Mid-Atlantic group, there is a roughly linear relationship between the amount of tongue root advancement and the amount of tongue body fronting, which suggests that the differences in the dorsal constriction are a consequence of the differences in tongue root advancement. The tongue body differences observed in the Midwest and Newfoundland speakers do not appear to be due to the raising pattern, since they are not observed in all speakers with raising before /g/, and they are not observed in /ŋ/ of all the non-Midwest speakers who have even greater vowel raising before /ŋ/ (with the exception of S6, whose /ŋ/ is more anterior than her /g/ and /k/.

It is difficult to draw conclusions from the Newfoundland speaker's consonant differences without a larger sample of speakers from Newfoundland. However, the more advanced dorsal constriction in four Midwest speakers (from Wisconsin, North Dakota, Michigan, and Ohio) may be important to the development of pre-velar raising, which was first noticed in Wisconsin. This is consistent with Purnell's (2008) study of Upper Midwest speakers, which found a more anterior velar constriction location in /æg/ than in /æk/, even in speakers who do not appear to have pre-velar raising. A more anterior constriction location contributes to /æ/ raising by moving the tongue body and tongue root forward at the end of the vowel, leading to a fronter and higher offglide. The presence of a more anterior tongue constriction for /g/ could have been an important phonetic motivation for the development of /æg/ raising in the Midwest and not elsewhere. The phonological phenomenon of /æg/ raising has a wider geographical distribution than the more anterior /g/ constriction in our sample. Most other North/Northwest and Canada speakers who exhibit /æg/ raising have a greater lingual Z2-Z1 difference in the vowel than in the consonant. Some, such as N1-2 (from the state of Washington), N3 (from Minnesota), and C1 (from eastern Ontario) have consonant differences that would be typical of South/Mid-Atlantic speakers, despite vowel differences that are similar to the other North/Northwest speakers.

In other words, if the Midwestern articulatory difference between /g/ and /k/ is what led to the development of /æg/ raising, then the phonological pattern has outrun a major part of its phonetic motivation, because /æg/ raising is found in speakers without the /g/-/k/ articulatory difference. Raising before nasals /m n/ appears to have outrun its phonetic motivation in a different way. The original impetus for raising before nasals may have been the acoustic raising effect of nasalization on a low vowel, and we have no reason to think that this effect is larger in the North American region where the pre-nasal raising phonological pattern is now observed. If pre-nasal /æ/ has in fact raised so high that nasalization no longer has an acoustic raising effect on the vowel, because that effect applies only to low vowels, then this is a clear case of a phonological pattern completely outgrowing its phonetic motivation. This contrasts with raising before /g/, which appears to be motivated by a combination of tongue root advancement that occurs during voiced consonant and a

dorsal constriction.[7] Raising /æ/ does not eliminate either of these motivations for raising, but only Midwestern speakers appear to have the most extreme motivation.

V. CONCLUSIONS

/æ/ raising before anterior /m/ and /n/ (and before anterior fricatives in Philadelphia) is produced with a tongue raising gesture that is aligned near the midpoint of the vowel interval and independent of the following consonant gesture. Raising before /ŋ/ is just as great and just as widespread as raising before other nasals, but it involves a tongue gesture with a rising trajectory that peaks near the end of the vowel, where the velar closure starts. While it remains plausible that the acoustic consequences of nasalization constitute the original phonetic motivation for all pre-nasal raising, the raised variants of /æ/ that we observe appear to be too high (due to tongue raising) for nasalization to have the effect of lowering F1 frequency and making the vowels sound higher.

/æ/ raising before /g/, as observed in the North and Northwest of the United States and parts of Canada, involves a tongue gesture similar to raising before /ŋ/, but the magnitude of the tongue gesture involved in /æg/ raising is consistently less than for /æŋ/ raising. Tongue root advancement is observed for both voiced velars for nearly all speakers in the sample, but only a small group of Midwestern speakers show a more anterior constriction location for /g/ and /ŋ/. Anterior /g/ may have led to the development of /æg/ raising in the Midwest, and this phonological raising pattern may then have spread independently to other regions such as the Pacific Northwest and Canada.

[1]Newfoundland has a very different settlement history from the rest of Canada, having been primarily settled by people coming directly from southwestern England and southeastern Ireland, whereas the founding group in the rest of English-speaking Canada was loyalists from what became the U.S.; Newfoundland did not join Canada until 1949 and was quite isolated from Canada before that time (Handcock, 1977).

[2]Our sample is a convenience sample of native English speakers available at North Carolina State University and the University of Ottawa. The authors are included as speakers N1, N2, and N7.

[3]Recording audio in Audacity was necessary because Ultraspeech 1.2 could not record sound for the duration of long recordings. Ultraspeech would record a 14-s wav file aligned to the start of the recording (matching the time stamps of the ultrasound images). This short recording was manually aligned with the full audio recording by lining up the same non-speech noise in both recordings, and then the start of the long recording was cropped to match the short one, so that the ultrasound image time stamps would match times in the long audio recording. The current version of Ultraspeech (1.3) can make long audio recordings.

[4]Similar techniques have been described by Story (2007) for point-tracking data, and Carignan et al. (2015) for MRI data.

[5]For reference, the first 10 PCs explain 54%–70% (mean: 62.45%) of the variance, and the first 30 explain 73%–85% (mean: 79.76%). For speaker N1, lingual Z2-Z1 regressions were performed with different subsets of PCs in order to examine the effect of including PCs near the cutoff of 20. With 20 PCs, the adjusted R^2 is 0.8692. With 15 PCs it is 0.8638, with 10 it is 0.8555, with 5 it is 0.7987, with 3 it is 0.7705, with 2 it is 0.5712, and with 1 it is 0.107. We conclude from this that including 20 PCs is more than adequate for our purposes, and that most of these PCs (and the ones over 20 that we did not include) represent image variance that does not contribute substantially to predicting formant frequencies.

[6]The value (0, 0) corresponds to the upper-left corner of the original ultrasound images, and the units on the axes are centimeters.

[7]Note that raising before /ŋ/ has all the motivations for raising before nasals and the motivations for raising before voiced velars. Regardless if which set of motivations were most important, it is still true that only the voiced velar motivation (and not the nasalization motivation) applies to a raised /æŋ/.

Ahn, S. (2015). "Utterance-initial voiced stops in American English: An ultrasound study," J. Acoust. Soc. Am. 138, 1777.

Baker, A., Mielke, J., and Archangeli, D. (2008). "More velar than /g/: Consonant coarticulation as a cause of diphthongization," Cascadilla Proceedings Project, Somerville, MA, http://www.lingref.com/cpp/wccfl/26/paper1656.pdf (Last viewed July 1, 2017).

Bauer, M., and Parker, F. (2008). "/æ/-raising in Wisconsin English," Am. Speech 83, 403–431.

Benson, E. J., Fox, M. J., and Balkman, J. (2011). "The bag that Scott bought: The low vowels in northwest Wisconsin," Am. Speech 86, 271–311.

Berry, J. J. (2012). "Machine learning methods for articulatory data," Ph.D. thesis, University of Arizona, Tucson, AZ, Chap. 3, pp. 55–77.

Cai, J., Denby, B., Roussel-Ragot, P., Dreyfus, G., and Crevier-Buchman, L. (2011). "Recognition and real time performances of a lightweight ultrasound based silent speech interface employing a language model," in Interspeech, pp. 1005–1008.

Campbell, F., Gick, B., Wilson, I., and Vatikiotis-Bateson, E. (2010). "Spatial and temporal properties of gestures in North American English /r/," Lang. Speech 53, 49–69.

Carignan, C., Shosted, R., Fu, M., Liang, Z.-P., and Sutton, B. P. (2015). "A real-time MRI investigation of the role of lingual and pharyngeal articulation in the production of the nasal vowel system of French," J. Phon. 50, 34–51.

De Decker, P. M., and Nycz, J. R. (2012). "Are tense [æ]s really tense? The mapping between articulation and acoustics," Lingua 122, 810–821.

Dodsworth, R., and Kohn, M. (2012). "Urban rejection of the vernacular: The SVS undone," Lang. Var. Change 24, 221–245.

Evanini, K. (2009). "The permeability of dialect boundaries: A case study of the region surrounding Erie, Pennsylvania," Ph.D. thesis, University of Pennsylvania, Philadelphia, PA, Chap. 4, pp. 50–94.

Falahati, R. (2013). "Gradient and categorical consonant cluster simplification in Persian: An ultrasound and acoustic study," Ph.D. thesis, University of Ottawa, Ottawa, Ontario, Canada, Chap. 5, pp. 79–142.

Feng, G., and Castelli, E. (1996). "Some acoustic features of nasal and nasalized vowels: A target for vowel nasalization," J. Acoust. Soc. Am. 99, 3694–3706.

Ferguson, C. A. (1975). "'Short a' in Philadelphia English," in Studies in Linguistics: In Honor of George L. Trager, edited by M. E. Smith (Mouton, the Hague), pp. 259–274.

Fox, R. A., and Jacewicz, E. (2009). "Cross-dialectal variation in formant dynamics of American English vowels," J. Acoust. Soc. Am. 126, 2603–2618.

Fujimura, O., and Lindqvist, J. (1971). "Sweep-tone measurements of vocal-tract characteristics," J. Acoust. Soc. Am. 49, 541–558.

Gick, B., Campbell, F., Oh, S., and Tamburri-Watt, L. (2006). "Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids," J. Phon. 34, 49–72.

Gu, C. (2002). Smoothing Spline ANOVA Models, Springer Series in Statistics (Springer-Verlag, New York).

Handcock, W. G. (1977). "English migration to Newfoundland," in The Peopling of Newfoundland: Essays in Historical Geography, edited by J. J. Mannion, Social and Economic Papers No. 8, Institute of Social and Economic Research, Memorial University of Newfoundland, St. John's, Newfoundland, Canada, pp. 15–48.

Hartman, J. W. (1969). "Some preliminary findings from DARE," Am. Speech 44, 191–199.

Hoole, P., and Zierdt, A. (2010). "Five-dimensional articulography," in Speech Motor Control: New Developments in Basic and Applied Research, edited by B. Maassen and P. van Lieshout (Oxford University Press, Oxford), pp. 331–349.

Hueber, T., Aversano, G., Chollet, G., Denby, B., Dreyfus, G., Oussar, Y., Roussel, P., and Stone, M. (2007). "Eigentongue feature extraction for an ultrasound-based silent speech interface," in IEEE International Conference on Acoustics, Speech and Signal Processing (Cascadilla, Honolulu, HI), pp. 1245–1248.

Hueber, T., Chollet, G., Denby, B., and Stone, M. (2008). "Acquisition of ultrasound, video and acoustic speech data for a silent-speech interface application," in Proceedings of the 8th International Seminar on Speech Production, pp. 365–369.

Jacewicz, E., Fox, R. A., and Salmons, J. (2011). "Vowel change across three age groups of speakers in three regional varieties of American English," J. Phon. 39, 683–693.

Kent, R. D., and Moll, K. L. (1969). "Vocal-tract characteristics of the stop cognates," J. Acoust. Soc. Am. 46, 1549–1555.

Krakow, R., Beddor, P., Goldstein, L., and Fowler, C. (1988). "Coarticulatory influences on the perceived height of nasal vowels," J. Acoust. Soc. Am. 83, 1146–1158.

Kurath, H., and McDavid, R. I., Jr. (1961). The Pronunciation of English in the Atlantic States (University of Michigan Press, Ann Arbor, MI), pp. 103–104.

Labov, W. (1989). "Exact description of the speech community: Short a in Philadelphia," in Language Change and Variation, edited by R. W. Fasold and D. Schiffrin (John Benjamins, Amsterdam), pp. 1–57.

Labov, W. (1994). Principles of Linguistic Change: Internal Factors (Blackwell, Oxford), pp. 503–526.

Labov, W., Ash, S., and Boberg, C. (2006). The Atlas of North American English: Phonetics, Phonology and Sound Change (De Gruyter Mouton, Berlin), pp. 173–184.

Labov, W., Rosenfelder, I., and Fruehwald, J. (2013). "One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis," Language 89, 30–65.

Lawson, E., Scobbie, J. M., and Stuart-Smith, J. (2011). "The social stratification of tongue shape for postvocalic /r/ in Scottish English," J. Socioling. 15, 256–268.

Li, M., Kambhamettu, C., and Stone, M. (2005). "Automatic contour tracking in ultrasound images," Clin. Ling. Phon. 19, 545–554.

Lisker, L. (1986). "'Voicing' in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees," Lang. Speech 29, 3–11.

Lobanov, B. M. (1971). "Classification of Russian vowels spoken by different speakers," J. Acoust. Soc. Am. 49, 606–608.

Mielke, J. (2012). "A phonetically-based metric of sound similarity," Lingua 122, 145–163.

Mielke, J. (2015). "An ultrasound study of Canadian French rhotic vowels with polar smoothing spline comparisons," J. Acoust. Soc. Am. 137, 2858–2869.

Moisik, S. R., Lin, H., and Esling, J. H. (2014). "A study of laryngeal gestures in Mandarin citation tones using simultaneous laryngoscopy and laryngeal ultrasound (SLLUS)," J. Int. Phon. Assoc. 44, 21–58.

Moreton, E. (2008). "Realization of the English postvocalic [voice] contrast in F1 and F2," J. Phon. 32, 1–33.

Nearey, T. M. (2013). "Vowel inherent spectral change in the vowels of North American English," in Vowel Inherent Spectral Change, edited by G. S. Morrison and P. F. Assmann (Springer, New York), pp. 49–85.

Nearey, T. M., and Assmann, P. F. (1986). "Modeling the role of inherent spectral change in vowel identification," J. Acoust. Soc. Am. 80, 1297–1308.

Payne, A. (**1980**). "Factors controlling the acquisition of the Philadelphia dialect by out-of-state children," in *Locating Language in Time and Space*, edited by W. Labov (Academic Press, New York), Vol. 1, pp. 143–178.

Proctor, M. (**2011**). "Towards a gestural characterization of liquids: Evidence from Spanish and Russian," Lab. Phonol. **2**, 451–485.

Purnell, T. C. (**2008**). "Prevelar raising and phonetic conditioning: Role of labial and anterior tongue gestures," Am. Speech **83**, 373–402.

Rosen, N., and Skriver, C. (**2015**). "Vowel patterning of Mormons in southern Alberta, Canada," Lang. Commun. **42**, 104–115.

Scobbie, J. M., Wrench, A. A., and van der Linden, M. (**2008**). "Head-probe stabilisation in ultrasound tongue imaging using a headset to permit natural head movement," in *Proceedings of the 8th International Seminar on Speech Production*, pp. 373–376.

Story, B. H. (**2007**). "Time dependence of vocal tract modes during production of vowels and vowel sequences," J. Acoust. Soc. Am. **121**, 3770–3789.

Summers, W. V. (**1987**). "Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses," J. Acoust. Soc. Am. **82**, 847–863.

Thomas, E. R. (**2001**). *An Acoustic Analysis of Vowel Variation in New World English*, publication of the American Dialect Society (Duke University Press, Durham, NC), Vol. 85, pp. 22, 23.

Wassink, A. B. (**2015**). "Sociolinguistic patterns in Seattle English," Lang. Var. Change **27**, 31–58.

Westbury, J. R. (**1994**). "X-ray microbeam speech production database user's handbook," University of Wisconsin, Madison, WI.

Wolf, C. G. (**1978**). "Voicing cues in English final stops," J. Phon. **6**, 299–309.

Yu, Y., and Acton, S. T. (**2002**). "Speckle reducing anisotropic diffusion," IEEE Trans. Image Process. **11**, 1260–1270.

Yuan, J., and Liberman, M. (**2008**). "Speaker identification on the SCOTUS corpus," in *Proceedings of Acoustics '08*, pp. 5687–5690.

Zeller, C. (**1997**). "The investigation of a sound change in progress: /æ/ to /e/ in Midwestern American English," J. Eng. Ling. **25**, 142–155.

Zharkova, N., Hewlett, N., Hardcastle, W. J., and Lickley, R. J. (**2014**). "Spatial and temporal lingual coarticulation and motor control in pre-adolescents," J. Speech Lang. Hear. Res. **57**, 374–388.

J. Acoust. Soc. Am. **142** (1), July 2017

Mielke *et al.*   349