Running head: *Say /hVC/ again*

Is /h/ phonetically neutral?

Michael P. Robb
Department of Communication Disorders
University of Canterbury
Christchurch, New Zealand

Yang Chen
Department of Speech-Language Pathology
Duquesne University
Pittsburgh, PA

Abstract

Use of /h/ in the phrase, "*Say /hVC/ again*" has been tacitly assumed to provide a neutral phonetic context in which to study the articulatory characteristics of speech either preceding or following /h/ articulation. Yet, assessment of the stability or neutrality of /h/ has gone untested. The current study sought to determine whether articulation of /h/ differs according to sex and language accent, as well as to examine its influence on subsequent vowel articulation. Selected acoustic features of */hVC/* were measured in 40 speakers of American English (AE) and 40 speakers of Mandarin accented English (MAE). Results of an analysis of /h/ duration revealed no sex differences within each language group, however considerable variation was found according to accented versus unaccented English. Clear sex differences were found for the production of /ɦ/, occurring more often among male speakers regardless of language variety. Considerable variation in production of /ɦ/ was found between language groups. Analysis of vowel formant frequencies immediately following /h/ articulation indicated minimal coarticulatory effects for both AE and MAE speakers. The present results appear to support the suggestion that /h/ is not exclusively sex-linked and may indeed vary according to non-biological factors. In spite of these variations, /h/ articulation appears to have a negligible influence on neighboring vowel articulation.

Key words: acoustics analysis, speech production measurement, voice, fricative, sex differences, English, Mandarin

Introduction

The consonant /h/ is a fricative with the point of constriction in the vocal tract being the glottis. Although /h/ is typically classified as a voiceless consonant (Shriberg & Kent, 2003), voicing of /h/ (i.e., /ɦ/) is occasionally observed in connected speech since this sound is often surrounded by voiced sounds enabling vocal fold vibration to continue without interruption. Speakers do not need to maintain the same amount of articulatory control for production of /h/ compared to other consonants. The only required movement is the approximation of the vocal folds, controlled by the laryngeal abductors and adductors to achieve a breathy voice quality (Borden, Harris, & Raphael, 2003).

A common framework for acoustic evaluation of vowels is use of /h/ in a context surrounding a vowel (V) or consonant (C), such as a /hV/, /VhV/, or /hVC/ syllable. Indeed, use of the carrier phrase "*Say /hVC/ again*" is pervasive in speech and voice research (e.g., Bohn, 2004; Cervera, Miralles, & Gonzalez-Alvarez, 2001; Cox, 2006; Flege, 1992; Hillenbrand et al., 1995; Katz & Assmann, 2001; Kurowski, Blumstein, & Alexander, 1996; Lane et al., 2005; McCaffrey & Sussman, 1994; Nearey, 1997; Peterson & Barney, 1952; Peterson & Lehiste, 1960; Pols, van der Kamp, & Plomp, 1969; Schiavetti, et al., 2004; Steinlen, 2005; Svirsky & Tobey, 1991; van Wieringen & Wouters, 1999; Xue & Hao, 2003). The */hVC/* context has also been used to evaluate speech production in non-English languages such as Danish and German (cf., Steinlen, 2005), Spanish (Cervera et al., 2001); Japanese (Nishi & Rogers, 2002), and Swedish (Reuter, 1971). Yet, in all of these studies, the /h/ portion of the acoustic waveform is discarded with analysis typically restricted to the vowel following /h/. Although a rationale for constructing syllables containing /h/ is seldom provided, the assumption is that /h/ is produced with an open vocal tract posture involving minimal supralaryngeal involvement, thereby allowing for clear inspection of neighboring vowel production. In essence, articulation of /h/ supposedly provides a neutral context in which to evaluate articulatory behavior in a sample of connected speech.

The routine use of /h/ as a neutral phonetic context in acoustic research warrants a more thorough understanding of its possible variable nature. Few quantitative studies exist examining the articulatory characteristics of /h/ (Koenig, 2000; Koenig et al., 2005; Manuel & Stevens, 1989; Yoshioka, 1981). The most comprehensive study to date was performed by Koenig (2000) who examined various acoustic and aerodynamic features of voicing control in men, women and children during production of prevocalic /b, d, p, t, h/. Prevocalic /h/ was selected because it was assumed to provide an open (neutral) vocal tract posture preceding vowel production, compared to the supraglottal constriction associated with stop consonant production. No sex differences were found in regard to the average duration of /h/ or for the stops /b, d, p, t/. However, aerodynamic measures for /h/ articulation were revealing of significant age and sex differences; while no such affect was found for stop consonant production. Specifically, peak airflow rates for /h/ were highest among men and lowest among children, and men were most likely to produce /ɦ/ compared to women. The higher peak airflow and greater occurrence of /ɦ/ among men was attributed to biological differences in glottal width. Johnson (1997) defines glottal width as the area of the glottis combined with the tension in the muscles that close the vocal folds. The fricative /h/ is naturally articulated with the vocal folds abducted; however, as glottal width decreases, the vocal folds begin to vibrate resulting in breathy voicing, i.e., /ɦ/. Male vocal folds tend to be less stiff and more bulky than female vocal folds contributing to a predisposition for less glottal width. In addition, females tend to show an increased glottal width compared to males due to a posterior glottal opening (i.e., "chink") that occurs during vocal fold closing (Bless & Abbs, 1983; Titze, 1989). Languages such as Japanese and Swedish have also shown a greater occurrence of /ɦ/ among male speakers (Lindqvist, 1972; Sawashima, 1968). The more frequent occurrence of /ɦ/ in males compared to females would indicate that anatomical and physiological differences between sexes have a direct impact on articulation of this consonant.

The age and sex differences in /h/ articulation identified by Koenig (2000) suggest that caution is required when using the "*Say /hVC/ again*" context to examine speech and voice behavior. That is, articulation of /h/ may have a varying effect on the articulation of consonants and vowels preceding and following its production. Both anticipatory and perseveratory coarticulation are natural features of speech production, so it seems unlikely that /h/ would be exempt from exerting an influence on neighboring phonemes. In addition to biologically-based influences on /h/ articulation, there may also be non-biological factors influencing production of this phoneme. For example, stop consonant production has been found to vary according to sociolinguistic factors. Measurement of voice onset time (VOT) is routinely performed to estimate the timing and coordination of the laryngeal and supralaryngeal systems to control voicing onset. Results of past VOT studies indicate that children take longer to complete this timing sequence than do adults and adult females take longer than males (Swartz, 1992; Whiteside & Marshall, 2001). The age and sex differences are typically attributed to anatomical and physiological differences in vocal anatomy (Kent & Read, 2002). However, there are also reports of a sociolinguistic influence whereby differences between adult male and female speakers reflect conscious manipulation of speech patterns to overtly convey gender identity (Byrd, 1994; Cheshire, 2002; Robb, Gilbert & Lerman, 2005; Whiteside, Henry, Dobbin, 2004).

If the, "*Say /hVC/ again*" context continues to be used as a framework for evaluating articulatory behavior surrounding the /h/ phoneme, it is important to fully appreciate the articulatory characteristics of /h/. While recent research points to a sex difference in some features of /h/ articulation (Koenig, 2000), there have been no direct attempts to consider the likelihood of non-biological factors as well. The purpose of this study was to evaluate selected acoustic properties of /h/ articulation in adult male and female speakers using the framework of *accented vs. non-accented* English production. Examination of accented and non-accented English allows for a unique opportunity to simultaneously evaluate both biological and non-biological influences on speech production. We specifically sought to (1) confirm past research identifying a sex difference in /h/ articulation, (2) determine whether /h/ articulation was influenced by speaker differences in linguistic background (i.e., a non-biological influence), and (3) determine whether /h/ production has an influence on neighboring vowel production.

The present researchers had access to a database of speakers of Mandarin accented English (MAE) in which to examine sex differences and non-physiological influences on /h/ articulation. The original database was established to evaluate features of vowel articulation and sentence stress production in MAE (Chen, Robb, Gilbert & Lerman, 2001a, 2001b). As part of the original data collection protocols, participants were sampled using the "*Say /hVC/ again*" context. Examination of MAE provides a unique context to evaluate non-biological influences upon /h/ articulation. The Mandarin and American English (AE) languages differ along several linguistic domains; including prosody, syllable structure, and phonetic inventory (see Chen et al., 2001a for further details). Interestingly, /h/ is common to the consonant inventories of both Mandarin and English (Sun, 2006). Accordingly, MAE speaker's production of /h/ in English should not differ from their production of /h/ in Mandarin. This is particularly the case among individuals who are not *simultaneous* bilinguals.[1] Individuals who do not acquire two languages at the exact time are often found to show a pattern whereby phonetic (acoustic) features of the first language are prevalent in the second language, such as voice onset time (Flege, 1987; Sancier & Fowler, 1997). Therefore, if articulation of /h/ is directly attributable to physical factors, such as laryngeal structure and/or aerodynamic quantities (Koenig, 2000), it is predicted that male and female MAE speakers would articulate /h/ in a fashion similar to male and female AE speakers. Alternatively, if the MAE and AE groups differed in their articulation of /h/, support would be provided for a non-physiological influence on articulation of this particular phoneme. Finally, assuming /h/ reflects a phonetically neutral form of articulation, it would be predicted there would be no influence on the articulatory characteristics of adjacent vowel production, regardless of sex and language group. To evaluate the influence of sex and linguistic background on the production of /h/ the following hypotheses were proposed:

1. The duration of /h/ will not differ significantly between males and females speakers.
2. The duration of /h/ will not differ significantly between AE and MAE speakers.
3. Males will produce significantly more /ɦ/ articulations compared to females.
4. The overall occurrence of /ɦ/ articulations will not differ significantly between AE and MAE speakers.

5. The articulation of /h/ will have no significant effect on the subsequent articulatory characteristics (i.e., formant frequencies) of vowel production.

Method

Participants

Two groups of participants were used for the study. The first group included 40 adults (20 males, 20 females) who spoke MAE. The average age of the Mandarin male speakers was 33 years (SD = 5;5 years, range = 30-46 years). The average age of the Mandarin female speakers was 28 years (SD = 4;9 years, range = 21-42 years). Criteria for selection of inclusion in the Mandarin group consisted of: (1) a college education, (2) formal instruction in English, (3) the ability to speak standard Chinese (Beijing) Mandarin as judged by the second author who is a native speaker of Mandarin, (4) the ability to orally read English fluently, and (5) residing in the US for a minimum of 2 years and speaking English a minimum of 30% of their daily conversation determined by self-report. On average, the MAE speakers began learning English as a second language by 12-years of age and had received a minimum of eight years of formal English education throughout their high school and college years in China. The MAE speakers had resided in the US an average of 4.2 years (SD = 7;5 years, range = 2-17 years). The average percentage of daily English usage was 52% (range = 30-90%). All of the MAE participants were considered to be consecutive bilinguals, who had acquired English as a second language. The second group consisted of 40 adults (20 males, 20 females) who spoke AE. The average ages of the American males and females were 33 years (range = 22-46 years) and 27 years (SD = 5;3 years, range = 23-41 years), respectively. All MAE and AE participants were perceptually judged by a speech-language pathologist (M.R.) to have no speech or language disorder. None of the participants reported a hearing disorder.

Data collection

Eleven AE vowels (/i, e, u, o, ɑ, ɪ, ɛ, æ, ʌ, ʊ, ɔ/) were placed in a /hVd/ context. Each /hVd/ syllable was embedded in the carrier phrase: "*Say /hVd/ again*". Each /hVd/ phrase was produced three times in a randomized order for a total of 33 phrases (or /h/ tokens) per speaker. No attempt was made to control for speech tempo. Each participant was asked to speak at a natural, habitual rate, using a comfortable loudness level. All recordings took place in a sound-attenuated booth using high quality audio tapes (Maxell, C60) and a cassette recorder (Marantz, PMD-360) in conjunction with a unidirectional dynamic microphone (Shure, 515SD). A mouth-to-microphone distance of 20 cm was maintained.

Analyses

Duration of /h/. The total number of /h/ tokens available for analysis was 2640 (11 vowels x 3 repetitions x 80 speakers). Each "*Say /hVd/ again*" phrase was digitized at 10,000 Hz using a speech analysis software package (Pratt, ver 4.2). Each phrase was visualized on a computer monitor using a combination of amplitude-by-time waveform and sound spectrographic (narrowband) displays. Vertical cursors were placed at the onset and offset of aperiodic noise for each /h/ token comprising the */hVd/* syllable to determine consonant duration. Duration of /h/ was only measured for tokens containing no visible vocal fold periodicity.

Duration ratio. To compensate for the possibility of speaking rate differences on the overall duration of /h/, a relative measure of /h/ duration was created by also measuring the overall duration of each */hVd/* syllable. Using a combination of amplitude-by-time waveform and narrow band spectrograms, vertical cursors were placed at the onset and offset points of acoustic energy associated with the entire */hVd/* token. The value obtained for word duration was divided by the /h/ duration value to create a duration ratio. The duration ratio was assumed to provide an estimate of the contribution of /h/ to the overall word duration. An example of a typical display for measuring /h/ duration and */hVd/* duration in the present study is provide in Figure 1.

Occurrence of /ɦ/. The occurrence of (voiced) /ɦ/ articulations was tabulated for each participant. In order for a token to be classified as /ɦ/, visible periodicity within the consonant segment was required. This judgment was based on examination of the harmonics as displayed in narrowband spectrograms. On the basis of this analysis, the overall number of /ɦ/ occurrences were noted for each group according to sex. An example of a typical display demonstrating /ɦ/ in the present study is provided in Figure 2.

Vowel formant frequencies. To examine whether /h/ had an influence on subsequent vowel production, the first (F1) and second (F2) formant frequency of each vowel was measured at two distinct locations. The first location was termed, *vowel onset*, and was defined as the first identifiable point of periodicity in the vowel. The second location was termed, *vowel midpoint*, and was defined as the mathematical center of the overall vowel duration. The F1 and F2 frequencies at each location were determined by positioning a 20 msec time window and extracting the formant frequency values using a 12-coefficient linear predictive coding (LPC) analysis. The center frequencies of the first two spectral peaks displayed in the LPC spectra were used to represent F1 and F2, respectively. Based on F1 and F2 frequency values collected at onset and midpoint locations, a difference value was calculated. This value reflected the absolute difference (in Hz) between the onset and midpoint locations. The difference value was assumed to reflect the influence of /h/ at vowel onset compared to the vowel at its approximate steady state (midpoint) location. A low difference score would be indicative of little influence of /h/ (i.e., minimal formant transitions) on subsequent vowel production.

Measurement reliability

Ten percent of the entire */hVd/* data base (264 tokens) was randomly selected across AE and MAE groups (i.e., four AE and four MAE speakers) for assessment of intra- and inter-judge measurement reliability. The second author re-measured the duration of each /h/ token, as well as F1 and F2 frequency for intra-judge reliability. Inter-judge reliability was performed by another individual experienced in acoustic analysis techniques. Average intra-judge error for measurement of /h/ duration was 5.62 ms. The Pearson correlation coefficient for /h/ duration between the first and second measurements was 0.97 (p<0.01).). A t-test was performed to assess whether the intra-judge measurements differed. The test was non-significant ($p = 0.41$). Average inter-judge error for measurement of /h/ duration was 4.93 ms. The corresponding Pearson correlation coefficient between the two judges was 0.96 (p<0.01), and results of a t-test indicated no significant difference between the two measurements ($p = 0.37$). The identification of /ɦ/ occurrences was also determined using the same random sample of 264 tokens. Both intra-judge and inter-judge agreement for identification of /ɦ/ occurrences was 100%.

Average intra-judge errors for measurement of F1 and F2 onset frequency onset were 9 Hz, and 20 Hz, respectively. The Pearson correlation coefficients for F1 and F2 between the first and the second measurements were 0.79 and 0.72 ($p < 0.01$), respectively. Two *t*-tests were performed to determine whether differences existed in the measurement of F1 and F2 onset, respectively. The results were non-significant for both F1 onset frequency ($p = 0.51$) and F2 onset frequency ($p = 0.41$). The intra-judge measurement error for F1 and F2 midpoint frequency were 7.6 Hz and 16.5 Hz, respectively. The corresponding Pearson correlation coefficients for measurement of F1 and F2 midpoint frequency were 0.83 and 0.79 ($p < 0.01$), respectively. The results of *t*-testing indicated no significant difference in the re-measurement of F1 midpoint frequency ($p = 0.43$) and F2 midpoint frequency ($p = 0.37$).

Average inter-judge measurement error for F1 and F2 onset frequency was 8 Hz and 19 Hz, respectively. The Pearson correlation coefficients for measurement of F1 and F2 frequency between the two judges were 0.81 and 0.78 ($p < 0.01$), respectively. Two *t*-tests were performed to determine whether differences existed in the measurement of F1 and F2 onset, respectively. The results were non-significant for both F1 onset frequency ($p = 0.44$) and F2 onset frequency ($p = 0.40$). The inter-judge measurement error for F1 and F2 midpoint frequency were 6 Hz and 14 Hz, respectively. The corresponding Pearson correlation coefficients for measurement of F1 and F2 midpoint frequency were 0.84 and 0.81 ($p < 0.01$), respectively. The results of *t*-testing indicated no significant difference in the inter-judge measurement of F1 midpoint frequency ($p = 0.39$) and F2 midpoint frequency ($p = 0.36$).

Results

Duration of /h/

The average durations of /h/, as produced by the AE and MAE speaker groups, are listed in Table 1. A one-way analysis of variance (*ANOVA*) test was performed across the

4 speaker groups (AE females, AE males, MAE females, MAE males). The test was significant [$F(3,2213) = 17.87$, $p<0.001$]. Post-hoc testing using a Scheffe *t*-test with Bonferroni adjustment identified significantly shorter /h/ duration for the AE females compared to MAE females ($p<0.001$). The AE males also produced significantly shorter /h/ durations compared to the MAE males ($p<0.001$). The AE females did not differ significantly from the AE males ($p = 0.847$). The MAE females did not differ from the MAE males ($p = .874$).

Duration ratio

The average duration ratios calculated for the AE and MAE speaker groups are listed in Table 1. Results of *ANOVA* testing across the four speaker groups were significant [$F(3,2213) = 15.71$, $p<0.001$]. Follow-up Scheffe *t*-testing identified a significantly smaller duration ratio for the AE males compared to MAE males ($p<0.001$), indicating that /h/ comprised a smaller portion of the overall */hVd/* syllable produced by AE males. The duration ratios did not differ significantly between AE females and MAE females ($p = 0.883$), or between AE males and AE females ($p = 0.239$). MAE males had a significantly larger duration ratio compared to MAE females ($p<0.001$), indicating that /h/ comprised a larger portion of the overall */hVd/* syllable when spoken by MAE males compared to MAE females.

Occurrence of /ɦ/

The number of /ɦ/ occurrences found in the productions of the AE and MAE speaker groups is listed in Table 2. A total of 279 /ɦ/ productions were identified across 21 AE speakers (8 female, 13 male). A total of 140 /ɦ/ productions were identified across

10 MAE speakers (3 female, 7 male). Using a chi-square test of proportions, the combined AE groups produced significantly more /ɦ/ tokens compared to the combined MAE group (p <.005). Within each language group, a significantly higher occurrence of /ɦ/ productions were found among males compared to females (p <0.05).

Formant frequencies

The difference values in F1 and F2 frequency measured at onset and midpoint locations for each vowel type are listed in Table 3.[2] Among the AE speakers, F1 difference values across the various vowels ranged from 34-108 Hz for AE females and 31-119 Hz for AE males. The F1 difference values ranged from 40-121 Hz for MAE females and 35-113 for MAE males. A series of alpha-adjusted two-tailed t-tests were performed to evaluate whether the overall (collective) difference values in F1 frequency differed between speaker groups. No attempt was made to evaluate individual vowel patterns. Rather, the analysis was based on determining whether absolute differences in vowel formants were apparent between onset and midpoint locations. Assuming vowel formant patterns were relatively stable between onset and midpoint locations, the test would yield a non-significant result. Results obtained for the F1 values indicated no significant difference in onset and midpoint values for each of the speaker groups (p >.05).

The difference values in F2 frequency ranged from 65-226 Hz for the AE females and 59-208 Hz for the AE males. The F2 difference values ranged from 88-200 Hz for MAE females and 79-194 for MAE males. A series of alpha-adjusted two-tailed t-tests were performed to evaluate whether the overall difference values in F2 frequency differed between speaker groups. Results obtained for the F2 values indicated no significant difference in onset and midpoint values for each of the speaker groups (p >.05).

Discussion

Use of /h/ in the phrase, "*Say /hVC/ again"* has been tacitly assumed to provide a neutral phonetic context in which to study the articulatory characteristics of speech either preceding or following /h/ articulation. Yet, the stability or neutrality of /h/ has been drawn into question by recent research suggesting a sex difference in various aspects of /h/ articulation (Koenig, 2000; Koenig et al., 2005). We sought to further explore the variable nature of /h/ articulation from the standpoint of accented and non-accented English, as well as to directly examine the influence of /h/ on subsequent vowel articulation. The present results appear to support the suggestion that /h/ is not exclusively sex-linked and may indeed vary according to non-biological factors. In spite of these variations, /h/ articulation does not appear to have a strong influence on neighboring vowel articulation.

Duration of /h/

The duration of /h/ produced by men and women within each language group was remarkably similar. This finding supports the first hypothesis posed in the present study and is in agreement with the results of Koenig (2000), who also found no differences in /h/ duration between men and women. The lack of difference in /h/ duration would seem to confirm that this particular feature of /h/ articulation is not influenced by anatomical and physiological differences between sexes. There were however, significant differences in /h/ duration between AE and MAE groups, which serve to reject the second hypothesis. Both AE males and females produced /h/ with a shorter duration compared to the MAE groups. So, although there were no apparent sex differences within language groups, there were clear differences between language groups. The most obvious explanation for these group differences would be to attribute them to language-based influences. All of the MAE participants were classified as *consecutive* bilinguals. As such, it is conceivable that their articulation of /h/ was more reflective of the Mandarin language rather than the English language. It is not unusual for languages to share the same IPA symbol for various phonemes, yet not show identical articulatory patterns (Flege, 1987). This is likely to be the case for /h/. The present study was not designed to examine acoustic features of the Mandarin language; however we would predict that the duration of /h/ in MAE more closely matches that of native Mandarin than English.

Additional insight to possible language and sex differences in /h/ articulation can be gained by considering the results from the duration ratios. Recall, a duration ratio was calculated to provide an estimate of the contribution of /h/ to the overall word duration. The present group of AE male and female speakers showed no sex difference in duration ratio. In other words, AE speakers produced /h/ durations, as well as overall */hVd/* durations of similar length. A markedly different pattern emerged for the MAE speakers. A sex difference was identified for duration ratio with MAE males tending to produce */hVd/* syllables with a larger proportion of /h/ compared to MAE females. The duration ratio for MAE males was also significantly different compared to the AE males. A conclusion to be drawn from the duration ratio data is that articulation of /h/ had a differential effect on the overall duration of the */hVd/* syllable. Among the male and female AE speakers, both /h/ duration and */hVd/* duration were produced with consistent, stable durations. The same could not be said for MAE speakers.

Occurrence of /ɦ/

A significantly higher occurrence of /ɦ/ productions was found among males compared to females. This pattern was apparent for both AE and MAE speaker groups. The more frequent occurrence of /ɦ/ among males supports the third hypothesis posed in the present study. Koenig (2000) and others (Lindqvist, 1972; Sawashima, 1968) have also found males to produce /ɦ/ more frequently than females. The laryngeal anatomy of males contributes to less overall glottal area during running speech compared to females (Bless & Abbs, 1983; Titze, 1989), which is likely to facilitate the occurrence of /ɦ/ articulations. Even though males in both groups were found to produce /ɦ/ more often than their female counterparts, the overall occurrences of /ɦ/ productions differed between language groups. Collectively, the combined AE speakers produced twice as many /ɦ/ articulations compared to MAE speakers. The variable pattern of /ɦ/ production across the AE and MAE groups serves to reject the fourth hypothesis posed in the present study. If /ɦ/ articulations were due exclusively to sex differences in laryngeal anatomy, the pattern displayed by the MAE participants should have closely mirrored the AE participants. The absence of such a pattern further underscores the probable influence of language accent on /*ɦ*/ articulation. Furthermore, /ɦ/ occurs most often in the production of connected speech, so it seems likely that the style of speaking employed by an individual in their use of connected speech might also contribute to its variable occurrence.

Formant Frequencies

The coarticulatory effects of /h/ on subsequent vowel production were estimated by examining the absolute difference in F1 and F2 frequency at vowel onset compared to a vowel midpoint location. Across AE and MAE speakers, the absolute difference values for F1 frequency were typically less than 100 Hz, and the difference values for F2 frequency were less than 200 Hz. The absolute values reflected the difference (in Hz) in estimating the center frequency of the overall formant bandwidth. As such, the observed differences in F1 and F2 frequencies at onset and midpoint frequencies were judged to be within acceptable limits for estimation of vowel formants. That is, the vocal tract posture for vowel production did not markedly differ across the earliest portions of the vowel immediately following /h/ articulation. This conclusion was also supported by the non-significant differences between onset and midpoint formant values for each speaker group. The minimal change in formant values provides compelling evidence of /h/ having little coarticulatory influence on vowel production, regardless of sex or speaker accent. It would be worthwhile to explore whether minimal coarticulatory influences are apparent in speakers of accented English whose first language does not contain /h/. For example, the Russian language does not contain /h/ and results in a substitution of either [h → g] or [h → x] in accented English (Ladefoged & Maddieson, 1996; Thompson, 1991). The influence of substituted phonemes on subsequent vowel production would serve to confirm whether /h/ is indeed unique in regard to its phonetic neutrality.

Say /hVC/ again

So what can be concluded in regard to use of the phrase, "*Say /hVC/ again"* in speech and voice research? Our findings generally support the results of past research indicating no sex difference in the duration of /h/, with sex differences apparent for /ɦ/ (Koenig, 2000; Lindqvist, 1972). Alone, these results suggest a direct link to basic biological influences on of speech and voice production. However, when examining /h/ duration and /ɦ/ articulation in the context of accented and non-accented English, it would appear that non-biological factors can also influence the production of /h/. Perhaps the variability observed in /h/ is not that surprising when considering that articulation of /h/ is primarily regulated at the level of the larynx. Koenig (2000) observed that articulation of /h/ can be accomplished in any number of ways. The only required articulation for production of this phoneme is related to approximation of the vocal folds and this gesture can be accomplished by variations in vocal fold abduction, vocal fold tension, and transglottal airflow. Furthermore, individual speaker differences in /h/ articulation are commonly observed in running speech (Koenig et al., 2005). The present study was confined to acoustic measurement of /h/ articulation but the results of our analysis tend to confirm that /h/ articulation is not static between sexes or between language varieties.

In spite of the variable nature of /h/, there appears to be little spill-over to the articulation of neighboring vowels, as inferred by measurement of formant frequencies. Ladefoged (2005) suggested that the sound source of /h/ is technically not air being forced through the glottis but rather turbulent air that moves across the edges of the open vocal folds and other surfaces of the vocal tract. Because the origin of the sound is deep within the vocal tract, the resonances of the whole vocal tract are more prominent and the sound is articulated like a noisy vowel. The absence of ‘upper’ vocal tract constriction is likely to facilitate the production of vowels with minimal, if any, influence from /h/. The present study found vowel formant patterns to remain quite stable immediately following articulation of /h/.

It is important to recognize that this study centered on features of normal speech articulation among speakers of AE and MAE. However, the */hVC/* context has been used to examine vowel production in various speech disordered conditions, including alaryngeal speech (Cervera et al., 2001), apraxia (Katz, Bharadwaj, & Stettler, 2006), deaf speech (McCaffrey & Sussman, 1994), dysarthria (Goberman & Elber, 2005), foreign accent syndrome (Kurowski et al., 1996), stuttering (Prosek, Montgomery, Walden, & Hawkins, 1987), as well as dysphagia (DiMaggio & Bell-Berti, 1998). Determining the presence and severity of a speech disorder is typically guided by comparing the features of the speech disorder to a model of normal speech production. Presumably, the model of normal /hVC/ articulation examined in the present study provides a basis for examining disordered vowel production, assuming a similar methodology is employed.

In conclusion, prior to this study the neutrality of /h/ in the “*Say /hVC/ again*” framework has gone untested. While /h/ is ostensibly no more stable than other consonants articulated between men and women, as well as across language groups, it appears to have a minor influence on subsequent vowel articulation. Use of finer grained analyses of spectral and temporal overlap may prove revealing of influences of /h/ upon neighboring phonemes. Until then, we suggest that /h/ plays a neutral role in subsequent supraglottal articulations.

Footnotes

[1] Simultaneous bilingualism is also referred to as spontaneous bilingualism. In this form of language use, an individual has presumably spoken (or has been spoken to) two or more languages in the home since birth (Shenker, 2004). The second form of bilingual language use is referred to as consecutive bilingualism. In this form of bilingualism, an individual learns one language after already knowing another. This is the situation for individuals who become bilingual as older children and adults (Halsband, 2006). Consecutive bilingualism is by far the most common form of bilingualism.

[2] The average F1 and F2 values for the present group of participants have been previously reported (Chen et al., 2001a). In this earlier study, F1 and F2 frequencies were estimated at the mid-point of the vowel segment using a 50 msec time window.

References

Bless, D., & Abbs, J. (1983). *Vocal fold physiology: Contemporary research and clinical issues*. San Diego: College-Hill Press.

Bohn, O. (2004). How to organize a fairly large vowel inventory: the vowels of Fering (North Frisian). *Journal of the International Phonetic Association, 34*, 161-173.

Borden, G., Harris, K., & Raphael, L. (2003). *Speech science primer: Physiology, acoustics, and perception of speech* (4th ed.). Baltimore: Lippincott, Williams & Wilkins.

Byrd, D. (1994). Relation of sex and dialect to reduction. *Speech Communication, 15*, 39-54.

Cervera, T., Miralles, J., & Gonzalez-Alvarez, J. (2001). Acoustic analysis of Spanish vowels produced by laryngectomized subjects. *Journal of Speech, Language, and Hearing Research, 44*, 988-996.

Chen, Y., Robb, M., Gilbert, H., & Lerman, J. (2001a). Vowel production by Mandarin speakers of English. *Clinical Linguistics & Phonetics, 15*, 427-440.

Chen, Y., Robb, M., Gilbert, H., & Lerman, J. (2001b). A study of sentence stress production in Mandarin speakers of American English. *Journal of the Acoustical Society of America, 109*, 1681-1689.

Cheshire, J. (2002). Sex and gender in variationist research. In J. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change*. Oxford: Blackwell Publishers.

Cox, F. M. (2006). The acoustic characteristics of /hVd/ vowels in the speech of some Australian teenagers. *Australian Journal of Linguistics, 26*, 147-179.

DiMaggio, J., & Bell-Berti, F. (1998). A study of vowel production in persons with dysphagia. *Journal of the Acoustical Society of America, 104*, 1853.

Flege J. (1992). The intelligibility of English vowels spoken by British and Dutch talkers. In R. Kent (Ed.), *Intelligibility in speech disorders: Theory, measurement, and management.* New York: John Benjamins Publishing Company.

Flege, J. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*, 47-65.

Goberman, A., & Elmer, L. (2005). Acoustic analysis of clear versus conversational speech in individuals with Parkinson disease. *Journal of Communication Disorders, 38*, 215-230.

Halsband, U. (2006). Bilingual and multilingual language processing. *Journal of Physiology, Paris 99,* 355-369.

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

Johnson, K. (1997). *Acoustic and auditory phonetics.* Malden, MA: Blackwell Publishers.

Katz, W., & Assmann, P. (2001). Identification of children's and adults' vowels: intrinsic fundamental frequency, fundamental frequency dynamics, and presence of voicing. *Journal of Phonetics, 29*, 23-51.

Katz, W., Bharadwaj, S. V., & Stettler, M. P. (2006). Influences of electromagnetic articulography sensors on speech produced by healthy adults and individuals with aphasia and apraxia. *Journal of Speech, Language, and Hearing Research, 49*, 645-659.

Kent, R., & Read, C. (2002). *Acoustic analysis of speech* (2nd ed.). Albany, NY: Delmar.

Koenig, L. (2000). Laryngeal factors in voiceless consonant production in men, women, and 5-year-olds. *Journal of Speech, Language, and Hearing Research, 43*, 1211-1228.

Koenig, L., Mencl, W., Lucero, J. (2005). Multidimensional analyses of voicing offsets and onsets in female speakers. Journ*al of the Acoustical Society of America, 118*, 2535-2550.

Kurowski, K., Blumstein, S., & Alexander, M. (1996). The Foreign Accent Syndrome: reconsideration. *Brain and Language, 54*, 1-25.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell Publishing.

Lane, H., Denny, M., Guenther, F, Matthies, M., Menard, L., Perkell, J., Stockmann, E., Tiede, M., Vick, J., & Zandipour, M. (2005). Effects of bite blocks and hearing status on vowel production. *Journal of the Acoustical Society of America, 118*, 1636-1646.

Lindqvist, J. (1972). Laryngeal articulation studied on Swedish subjects. *Quarterly Status and Progress Report* (Speech Transmission Laboratory, Royal Institute of Technology, Stockholm), *2-3,* 10-27.

Manuel, S., & Stevens, K. (1989). Acoustic properties of /h/. *Journal of the Acoustical Society of America, 86,* S49.

Mattingly, I. (1966). Speaker variation and vocal tract size. *Journal of the Acoustical Society of America, 39*, S1219.

McCaffrey, H. & Sussman, H. (1994). An investigation of vowel organization in speakers with severe and profound hearing. *Journal of Speech and Hearing Research, 37*, 938-951.

Nearey, T. (1997). Speech perception as pattern recognition. *Journal of the Acoustical Society of America, 101,* 3241-3254.

Nishi, K., & Rogers, C. (2002). On the relationship between perception and production of American English vowels by native speakers of Japanese: A pair of case studies. *Journal of the Acoustical Society of America, 112*, S2250.

Peterson, G., & Barney, H. (1952). Control methods used in a study of vowels. *Journal of the Acoustical Society of America, 24,* 175-184.

Peterson, G., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America, 32,* 693-703.

Pols, L.C.W., van der Kamp, L.J., & Plomp, R. (1969), Perceptual and physical space of vowel sounds. *Journal of the Acoustical Society of America, 46*, 458-467.

Prosek, R., Montgomery, A., Walden, B., & Hawkins, D. (1987) Formant frequencies of stuttered and fluent vowels. *Journal of Speech and Hearing Research, 30*, 301-305.

Reuter, M. (1971) Vokalerna in finlandsvenskan: en instrumentell analys och ett försök till systematisering enligt särdrag. *Festskrift till Olav Ahlbeck, 28*, 240-249.

Robb, M., Gilbert, H., & Lerman, J. (2005). Influence of gender and environmental setting on voice onset time. *Folia Phoniatrica et Logopaedica, 57*, 125-133.

Sancier, M. & Fowler, C. (1997). Gestural drift in a blingual speaker of Brazilian Portuguese and English. *Journal of Phonetics, 25*, 421-436.

Sawashima, M. (1968). Movements of the larynx in articulation of Japanese consonants. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics,* (University of Tokyo), *2*, 11-20.

Schiavetti, N., Metz, D., Whitehead, R., Brown, S., Borges, J., Rivera, S., & Schultz, C. (2004). Acoustic and perceptual characteristics of vowels produced during simultaneous communication. *Journal of Communication Disorders, 37*, 275-294.

Shenker, R. C. (2004). Bilingualism in early stuttering: Empirical issues and clinical implications. In A. K. Bothe (Ed.), *Evidence-based treatment of stuttering: Empirical bases and clinical applications*. (pp. 81-96). Mahwah, New Jersey: Lawrence Erlbaum Associates Publishers.

Shriberg, L., & Kent, R. (2003). *Clinical phonetics* (3rd ed.). Boston: Allyn and Bacon.

Steinlen, A. J. (2005). *The influence of consonants on native and non-native vowel production: a cross linguistic study.* Gunter Narr: Verlag.

Sun, C. (2006). *Chinese: A linguistic introduction.* Cambridge: Cambridge University Press.

Svirsky, M. A., & Tobey, E. A. (1991). Effect of different types of auditory stimulation on vowel formantfrequencies in multichannel cochlear implant users. *Journal of the Acoustical Society of America*, *89,* 2895-2904.

Swartz, B. (1992). Gender differences in voice onset time. *Perceptual and Motor Skills, 75*, 983-992.

Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning, 41*, 177–204.

Titze, I. (1989). Physiologic and acoustic differences between male and female voices. *Journal of the Acoustical Society of America, 85*, 1699-1707.

van Wieringen, A., & Wouters, J. (1999). Natural vowel and consonant recognition by cochlear implantees. *Ear & Hearing. 20*, 89-103.

Whiteside, S., & Marshall, J. (2001). Developmental trends in voice onset time: Some evidence for sex differences. *Phonetica, 58,* 196-210.

Whiteside, S., Henry, L., & Dobbin, R. (2004). Sex differences in voice onset time: A developmental study of phonetic contexts effects in British English. *Journal of the Acoustical Society of America, 116,* 1179-1183.

Xue, S., & Hao, G. (2003). Changes in the human vocal tract due to aging and the acoustic correlates of speech production: A pilot study. *Journal of Speech, Language, and Hearing Research, 46*, 689-701.

Yoshioka, H. (1981). Laryngeal adjustments in the production of the fricative consonants and devoiced vowels in Japanese. *Phonetica, 38*, 236-251.

Table 1. Average duration (msec) of /h/ and the ratio of /h/ duration to overall word duration produced by male and female speakers of American English (AE) and speakers of Mandarin accented English (MAE). Standard deviations are shown in parentheses.

	AE			MAE		
Group	tokens	/h/	ratio	tokens	/h/	ratio
Female	562	119 (35)	0.384 (.083)	635	129 (40)	0.388 (.078)
Male	478	117 (34)	0.374 (.088)	542	131 (43)	0.408 (.083)

Table 2. Total number (T #) of occurrences of /ɦ/ produced by male and female speakers of American English (AE) and speakers of Mandarin accented English (MAE). The percentage (%) of the entire speech sample containing /ɦ/, as well as the total number of participants (P#) in each group (n = 20 per group) producing /ɦ/ is also listed.

	AE			MAE		
Group	T #	%	P #	T #	%	P #
Female	99	15%	8/20	24	4%	3/20
Male	180	27%	13/20	116	18%	7/20

Table 3. Mean absolute difference in F1 (ΔF1) and F2 (ΔF2) frequency values (Hz) between vowel onset and mid-point for 11 vowels produced by Mandarin and American male and female speakers in a */hVd/* context. The corresponding standard deviation is shown in parentheses

	Male				Female			
	Mandarin (N = 20)		American (N = 20)		Mandarin (N = 20)		American (N = 20)	
Vowel	ΔF1	ΔF2	ΔF1	ΔF2	ΔF1	ΔF2	ΔF1	ΔF2
/i/	35 (29)	79 (64)	31 (28)	59 (42)	40 (39)	88 (99)	34 (26)	65 (101)
/e/	73 (52)	194 (168)	55 (47)	126 (116)	111 (96)	131 (81)	70 (50)	118 (76)
/u/	47 (59)	126 (103)	50 (48)	149 (133)	74 (62)	129 (95)	47 (34)	106 (110)
/o/	113 (112)	185 (209)	119 (98)	208 (210)	106 (102)	200 (199)	56 (47)	151 (134)
/ɑ/	76 (94)	95 (105)	100 (94)	99 (96)	120 (101)	108 (99)	66 (65)	120 (116)
/ɪ/	65 (65)	117 (119)	81 (53)	182 (136)	78 (66)	126 (134)	83 (66)	158 (92)
/ɛ/	67 (58)	109 (95)	74 (64)	117 (100)	82 (85)	158 (152)	90 (66)	129 (85)
/æ/	61 (62)	111(109)	84 (72)	144 (103)	79 (79)	145 (116)	81 (62)	226 (167)
/ʌ/	80 (69)	91 (87)	92 (87)	126 (120)	86 (82)	103 (98)	78 (57)	137 (98)
/ʊ/	46 (41)	137 (123)	88 (80)	177 (103)	76 (54)	163 (263)	98 (66)	172 (99)
/ɔ/	81 (81)	157 (144)	106 (101)	160 (154)	121 (86)	151 (129)	108 (93)	140 (94)

Figure Legends

Figure 1. Example of a typical acoustic display used for measuring the duration of /h/ and */hVd/* duration. The particular phrase is, "Say heed again."

Figure 2. Example of a typical acoustic display demonstrating /ɦ/, in the phrase, "Say heed again." The voicing of /h/ is indicated by the continuous trace of harmonics in the circled area.

Figure 1

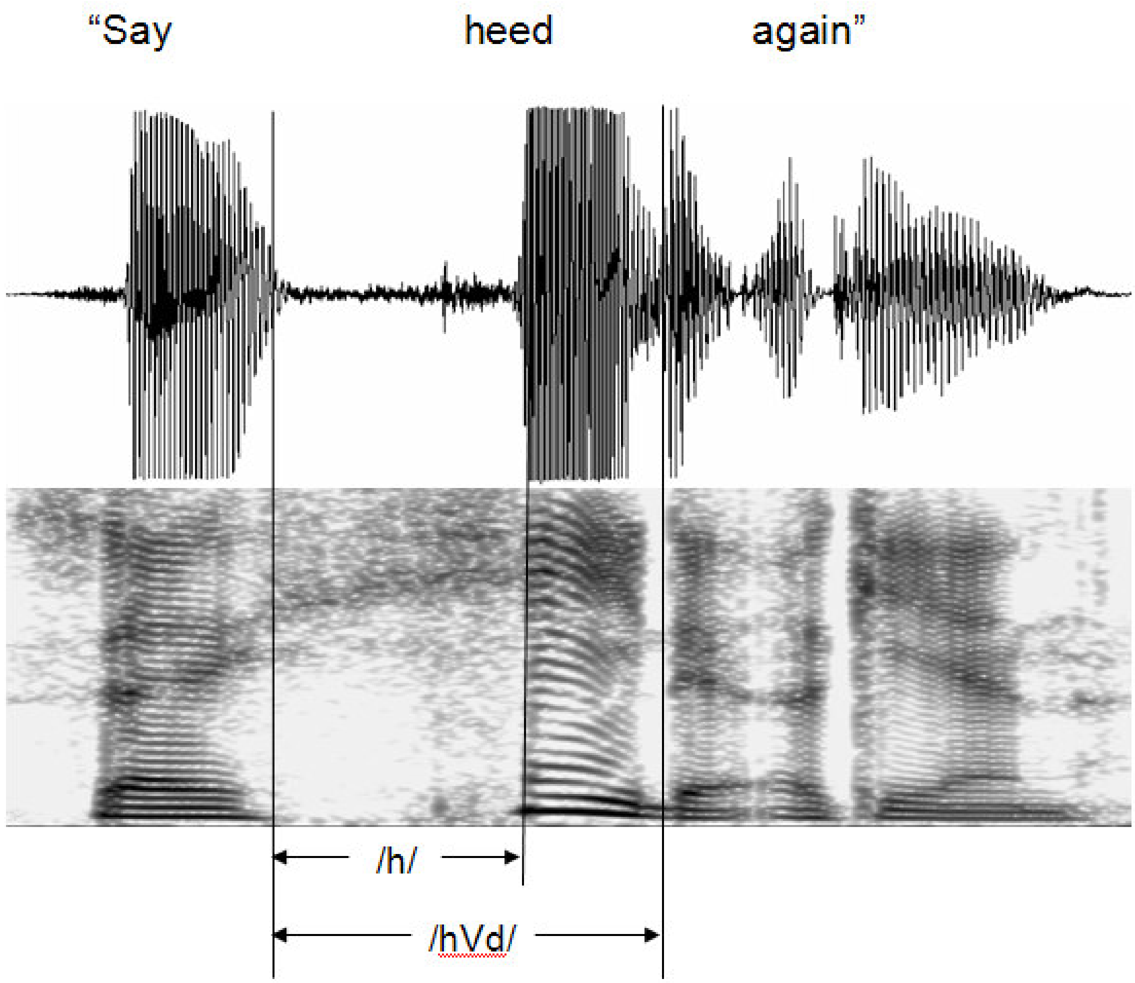

Figure 2