JARO 15: 823–837 (2014)
DOI: 10.1007/s10162-014-0475-7

JARO
Journal of the Association for Research in Otolaryngology

Research Article

Computational Model Predictions of Cues for Concurrent Vowel Identification

ANANTHAKRISHNA CHINTANPALLI,[1] JAYNE B. AHLSTROM,[1] AND JUDY R. DUBNO[1]

[1]*Department of Otolaryngology-Head and Neck Surgery, Medical University of South Carolina, 135 Rutledge Avenue, MSC 550, Charleston, SC 29425-5500, USA*

Received: 20 August 2013; Accepted: 3 June 2014; Online publication: 8 July 2014

ABSTRACT

Although differences in fundamental frequencies (F0s) between vowels are beneficial for their segregation and identification, listeners can still segregate and identify simultaneous vowels that have identical F0s, suggesting that additional cues are contributing, including formant frequency differences. The current perception and computational modeling study was designed to assess the contribution of F0 and formant difference cues for concurrent vowel identification. Younger adults with normal hearing listened to concurrent vowels over a wide range of levels (25–85 dB SPL) for conditions in which F0 was the same or different between vowel pairs. Vowel identification scores were poorer at the lowest and highest levels for each F0 condition, and F0 benefit was reduced at the lowest level as compared to higher levels. To understand the neural correlates underlying level-dependent changes in vowel identification, a computational auditory-nerve model was used to estimate formant and F0 difference cues under the same listening conditions. Template contrast and average localized synchronized rate predicted level-dependent changes in the strength of phase locking to F0s and formants of concurrent vowels, respectively. At lower levels, poorer F0 benefit may be attributed to poorer phase locking to both F0s, which resulted from lower firing rates of auditory-nerve fibers. At higher levels, poorer identification scores may relate to poorer phase locking to the second formant, due to synchrony capture by lower formants. These findings suggest that concurrent vowel identification may be partly influenced by level-dependent changes in phase locking of auditory-nerve fibers to F0s and formants of both vowels.

Correspondence to: Judy R. Dubno · Department of Otolaryngology-Head and Neck Surgery · Medical University of South Carolina · 135 Rutledge Avenue, MSC 550, Charleston, SC 29425-5500, USA. Telephone: +1-843-7927978; fax: +1-843-7927736; email: dubnojr@musc.edu

Keywords: concurrent vowels, vowel identification, fundamental frequency cue, formant frequency cue, auditory-nerve fibers, phase locking

INTRODUCTION

Listeners with normal hearing have the ability to extract a target talker from the speech of multiple talkers. Differences in fundamental frequency (F0) among talkers are one important cue for talker segregation. For example, when listening to two vowels simultaneously, identification of both vowels improves as the F0 difference between the two vowels increases (e.g., Scheffers 1983; Assmann and Summerfield 1990; Arehart et al. 1997; Summers and Leek 1998; Vongpaisal and Pichora-Fuller 2007; Chintanpalli and Heinz 2013). Although vowel segregation clearly benefits from F0 differences, the ability to segregate and identify two simultaneous vowels with identical F0s suggests that listeners use multiple cues for identification, including differences in formants between vowels (see Micheyl and Oxenham 2010 for review).

Physiological and computational modeling studies of auditory-nerve (AN) fibers show that phase locking of a fiber with a characteristic frequency (CF) near a vowel formant follows the closest harmonic. Phase locking to the first formant (F1) becomes stronger with increasing vowel level (Young and Sachs 1979), whereas phase locking to the second and third formants (F2 and F3) strengthens from low- to mid-levels (Young and Sachs 1979; Miller et al. 1997; Zilany and Bruce 2007) and then declines at higher levels (Young and Sachs 1979). This decline could be related to broader auditory filters, which reduce phase locking of AN fibers to F2 and F3 due to the presence of lower formants (usually F1) in the same filter. This results in synchrony capture by a lower formant at a higher formant location (Young and Sachs 1979; Palmer 1990). Limited physiological data are available to address level-dependent changes in the strength of phase locking to formants of concurrent vowels. Phase locking to F2 of /i/ was degraded at higher levels when presented simultaneously with /ɑ/ due to synchrony capture by the lower formants of /ɑ/ (Palmer 1990). However, the extent to which synchrony capture by lower formants is influenced by the spacing of formants between vowel pairs is not known. The above-mentioned studies suggest that phase locking to formants varies with vowel level (Young and Sachs 1979; Palmer 1990; Miller et al. 1997; Zilany and Bruce 2007). When vowels have different F0s, phase locking of AN fibers to F0s has been shown to be useful for vowel segregation at moderate levels (Assmann and Summerfield 1990; Palmer 1990; Meddis and Hewitt 1992; Palmer 1992; Keilson et al. 1997; Chintanpalli and Heinz 2013), but the role of F0 differences for segregating vowels at lower and higher levels has yet to be investigated.

The extent to which level-dependent changes in the strength of phase locking to F0s and formant frequencies affect concurrent vowel identification is addressed in the current study by relating vowel identification scores from younger adults with normal hearing to predictions from a physiologically inspired AN model (Zilany et al. 2009). Using tonal signals and vowels, this model has been successfully tested against neurophysiological data in cats and is an extension of previous models (Carney 1993; Zhang et al. 2001; Heinz et al. 2001; Bruce et al. 2003; Tan and Carney 2005; Zilany and Bruce 2006; 2007). Relevant to the current study, the model also captures level-dependent changes in compression, suppression, bandwidth, and shift in the center frequency of auditory filters. Thus, these factors are reflected in the model's predictions of phase locking of AN fibers at each vowel level.

In the current perceptual and modeling study, younger adults with normal hearing listened to two simultaneous vowels with either the same or different F0 over a wide range of vowel levels. Following assessment of vowel identification, computational modeling was used to determine the extent to which changes in phase locking of AN fibers to vowel formants and F0s could account for changes in vowel identification as a function of vowel level and F0 difference. We hypothesized that (a) poorer benefit of F0 cues at lower levels results from poorer phase locking to F0s by AN fibers, due to their lower mean firing rates, and (b) reduction in identification score at higher levels results from degradation in phase locking of AN fibers to second (and higher) formants due to broader auditory filters. The extent of this degradation may be influenced by the role of synchrony capture and formant spacing between vowels.

CONCURRENT VOWEL IDENTIFICATION

Subjects

Twelve adults (mean age=23.5 years; range 20–26 years; eight males) participated in the experiment. Each subject's test ear had audiometric thresholds ≤20 dB HL between 0.25 and 8.0 kHz and normal immittance. Prior to participation, subjects provided informed consent, in accordance with the Institutional Review Board of the Medical University of South Carolina. Subjects were paid for their participation, which required six sessions (approximately 2 h per session).

Stimuli and apparatus

Five vowels (/i/, /ɑ/, /u/, /æ/, /ɝ/) were generated using a MATLAB implementation of a cascade formant synthesizer (Klatt 1980). Vowel duration was 400 ms, including 15-ms raised-cosine rise and fall ramps. Table 1 includes the formant frequencies and bandwidths for each vowel, which were the same as those used in earlier studies of concurrent vowel identification (e.g., Assmann and Summerfield 1994; Summers and Leek 1998; Chintanpalli and Heinz 2013). Figure 1 shows the envelope spectrum for each vowel, computed using linear predictive coding. The formant frequencies (or local maxima in Fig. 1) and their amplitudes differed for each vowel. Note that /ɑ/ and /æ/ have the same F1 (750 Hz), as do /i/ and /u/ (250 Hz).

Subjects listened to pairs of vowels presented simultaneously (concurrent vowels). The vowels in each pair may be the same (e.g., /ɑ/ and /ɑ/) or different (e.g., /ɑ/ and /u/), and each pair had either the same or different F0. There were 25 vowel pairs with different F0s (five identical vowels+ten different vowel pairs, in which one vowel had F0=100 Hz and the other had F0=126 Hz+the same ten vowel pairs with the order of F0 reversed). To be consistent with this condition, the same F0 condition also had 25 vowel pairs (five identical vowels+ten different vowel pairs presented twice, in which both vowels in the pair had F0=100 Hz). Individual vowels were presented at 25, 35, 50, 65, 75, and 85 dB SPL.

TABLE 1

Formant frequencies (in Hz) for five vowels

Vowel	/i/ "beet"	/ɑ/ "father"	/u/ "food"	/æ/ "bat"	/ɝ/ "bird"
F1 (90)	250	750	250	750	450
F2 (110)	2250	1050	850	1450	1150
F3 (170)	3050	2950	2250	2450	1250
F4 (250)	3350	3350	3350	3350	3350
F5 (300)	3850	3850	3850	3850	3850

Values in parentheses in the first column are bandwidths around each formant (in Hz)

The overall level of the vowel pair was ~3 dB higher than the level of each individual vowel. The same and different F0 conditions were equivalent to the 0- and 4-semitone F0 difference used in earlier studies of concurrent vowel identification (e.g., Chintanpalli and Heinz 2013).

Pairs of vowels were converted from digital to analog form using a Tucker–Davis Technologies (TDT) RX6 array processor (sampling frequency= 48.8 kHz), and passed through separate programmable attenuators (TDT PA4) and a mixer (TDT SM3). The vowel pair was passed through a headphone buffer (TDT HB6) and delivered to the subject's test ear through Sennheiser HDA 200 headphones. The right ear was selected as the test ear if the thresholds for both ears were similar; otherwise, the ear that had a better average threshold across frequency was selected as the test ear.

Procedures

As subjects had no previous experience with the speech materials or tasks in this experiment, a reference sheet of orthographic examples for each vowel ("beet," "father," "food," "bat," "bird") was provided as part of the familiarization procedures and was available to subjects throughout the experiment. Each subject was tested in the following three phases: screening, two stages of practice, and test. During the screening phase, subjects listened to 10 single vowels at 65 dB SPL with F0=100 Hz or F0= 126 Hz (five vowels x two F0 conditions); order of F0 was randomized. These 10 single vowels were repeated five times for a total of 50 single vowels (five vowels x two F0 conditions x five repetitions) per block. Subjects responded by selecting the vowel from a row of five vowel symbols (/i/, /ɑ/, /u/, /æ/, /ɝ/) displayed on a touch-screen monitor. Correct-answer feedback was provided for these conditions only. Subjects were required to achieve ≥90 % correct identification for single vowels to proceed further in the experiment. Although the protocol called for three blocks to achieve the criterion score, each subject achieved ≥90 % during the first block.

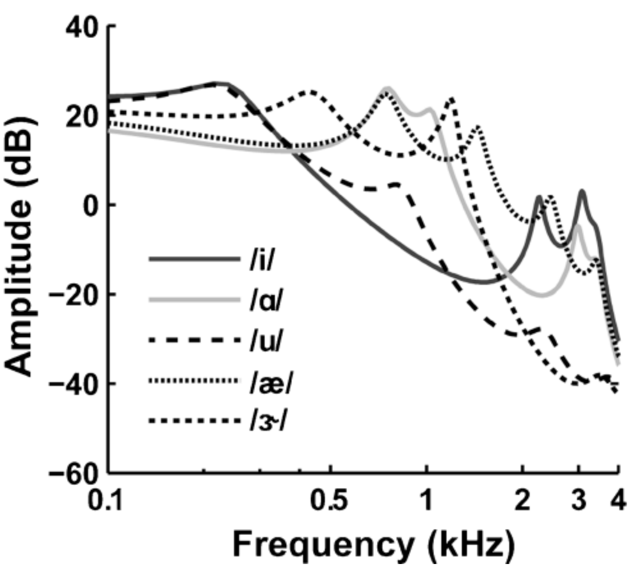

FIG. 1. Envelope spectrum for each of the five vowels presented at 65 dB SPL, computed using linear predictive coding. The local maxima correspond to the formant frequencies of each vowel.

During the first practice stage, each of the 25 vowel pairs with different F0 (described earlier) was presented at 65 dB SPL in random order. Subjects were instructed to identify both vowels in each pair. Two rows of five vowel symbols were displayed on the touch-screen monitor and subjects responded by selecting one vowel from each row; order of selection was ignored. This procedure was then repeated for the 25 vowel pairs with the same F0. In the second stage of practice, 25 vowel pairs with different F0 were presented three times in random order at each of the six vowel levels (65, 75, 85, 50, 35, and 25 dB SPL in a fixed order), for a total of 450 vowel pairs (25 vowel pairs x three repetitions x six levels). Upon completion, another set of 450 vowel pairs with the same F0 was presented in the same vowel level order. Thus, 900 response trials were contained in the second stage of practice. The test phase was similar to the second stage of practice except there were four repetitions of the block, and F0 condition and level within each block were randomized. Overall, 3600 response trials were included during the test phase (900×4).

Percent correct identification scores for both vowels were computed at each level for vowel pairs with both same and different F0. The response was considered correct only if both vowels were identified correctly. Identification scores were rationalized arcsine-transformed to stabilize the variance across conditions (Studebaker 1985). Effects of F0 difference and vowel level on identification scores were assessed with repeated-measures ANOVA and two-tailed *t* tests using the Statistical Package for the Social Sciences (SPSS) software (version 19). Post hoc comparisons were adjusted using Bonferroni corrections. Differences were considered significant with $p<0.05$.

Results

A three-way repeated measures ANOVA, with F0 condition (same vs. different F0), vowel level (six levels), and block (four blocks) as repeated measures, was performed on the rationalized arcsine-transformed identification scores (Studebaker 1985). The main effects of F0 condition [$F_{(1,\ 11)}=169.63$, $p<0.005$], vowel level [$F_{(5,\ 55)}=79.95$, $p<0.005$] and block [$F_{(3,\ 33)}=7.82$, $p<0.005$] were significant. Post hoc comparisons, using Bonferroni correction, showed that the score for the fourth block was significantly higher ($p<0.04$) than the other three blocks. Interactions between F0 condition and block [$F_{(3,\ 33)}=0.90$, $p=0.45$], vowel level and block [$F_{(15,\ 165)}=0.69$, $p=0.80$], and the three-way interaction [$F_{(15,\ 165)}=0.96$, $p=0.50$] were not significant. Because there were no significant interactions with block, identification scores were averaged across the four blocks for each subject; these mean scores were used for subsequent data analyses.

Figure 2A shows mean (±1 SEM) identification scores of both vowels with the same or different F0 as a function of vowel level. Apart from the significant main effects of F0 condition and vowel level (mentioned above), the interaction between F0 and vowel level was also significant [$F_{(5,\ 55)}=21.81$, $p<0.005$]. Post hoc comparisons showed that scores for both F0 conditions improved significantly ($p<0.005$) as vowel level increased from low- to mid-levels and declined significantly at higher levels (asterisks in Fig. 2A). The difference in mean identification scores between different and same F0 (i.e., F0 benefit) was significantly smaller at 25 dB SPL than at higher levels ($p<0.005$).

For identification of identical vowel pairs (Fig. 2B), only a significant effect of vowel level [$F_{(5,\ 55)}=14.14$, $p<0.005$] was observed; F0 condition [$F_{(1,\ 11)}=4.12$, $p=0.067$] and the interaction between F0 and vowel level [$F_{(5,\ 55)}=1.97$, $p=0.098$] were not significant. For identification of different vowel pairs (Fig. 2C), significant effects of F0 [$F_{(1,\ 11)}=156.37$, $p<0.005$], vowel level [$F_{(5,\ 55)}=79.24$, $p<0.005$], and the interaction between F0 and vowel level [$F_{(5,\ 55)}=14.13$, $p<0.005$] were observed. Post hoc comparisons also showed that scores for same and different vowel pairs improved significantly ($p<0.005$) as vowel level increased from low- to mid-levels and declined significantly at higher levels for both F0 conditions (asterisks in Fig. 2B, C). A non-significant effect of F0 condition for identical vowel pairs suggests that vowel segregation using F0 difference is only beneficial for different vowel pairs (i.e., those that include formant differences).

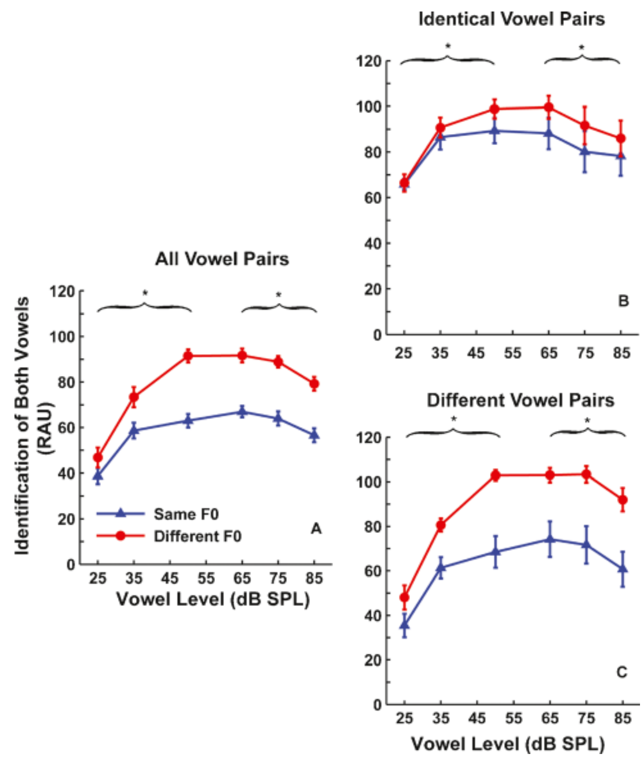

FIG. 2. Identification scores (rau) of both vowels for vowel pairs with the same F0 (*blue triangles*) or different F0 (*red circles*) as a function of vowel level. **A** Identification scores for all vowel pairs (25 pairs). **B** Identification scores for identical vowel pairs (5 pairs). **C** Identification scores for different vowel pairs (20 pairs). For each panel, *error bars* indicate ±1 SEM and *asterisks* indicate significant changes in scores ($p<0.05$) with vowel level for both F0 conditions. Note that scores in panel **A** are not the average of the scores in panels **B** and **C** because the numbers of stimuli are different.

COMPUTATIONAL MODELING: PHASE LOCKING OF AUDITORY-NERVE FIBERS TO VOWEL FORMANTS AND F0s

Motivation and rationale for the modeling approach

The results suggested that vowel identification and F0 benefit were reduced at the lowest level and that vowel identification was reduced at higher levels for both F0 conditions. Computational modeling was used to investigate the extent to which level-dependent changes in phase locking of AN fibers to vowel formants and F0s may explain these results.

Using an AN model (Zilany et al. 2009), formant difference cues available for identification were estimated by quantifying phase locking of AN fibers to formants (predicted by average localized synchronized rate, ALSR; Young and Sachs, 1979) when vowel pairs had the same F0. When vowel pairs had different F0s, formant and F0 difference cues available for identification were estimated by quantifying phase locking to formants (predicted by ALSR) and F0s (predicted by template contrast; Larsen et al. 2008). The model predictions for same and different F0 conditions as a function of level were then compared to identification scores obtained from our normal-hearing human listeners under the same level and F0 conditions.

Procedures

Figure 3 shows the steps involved in computing ALSR and template contrast using the AN model. The input to the AN model (step 1 of Fig. 3) was concurrent vowels with the same or different F0 and the output was the peri-stimulus time histograms (PSTH) from an AN fiber at a single characteristic frequency (CF). The PSTHs were predicted from fibers at 100 different CFs (CF_1, CF_2 … CF_{100}, step 1 of Fig. 3), which were spaced logarithmically between 100 and 4000 Hz. The first 10 ms of each PSTH was excluded to avoid onset effects. High spontaneous rate (SR) fibers were used for the modeling-based phase locking predictions for the following reasons: (a) high SR fibers constitute the majority of AN fibers (Liberman 1978) and thus provide a good approximation for the population of AN fibers, (b) phase locking predictions from high SR fibers would not substantially change with the addition of medium-and-low SR fibers, as phase locking for pure tones, a reasonable approximation of synthetic vowels, does not depend largely on SR (Johnson 1980; Louage et al. 2004), and (c) phase locking predictions from high SR fibers are robust for understanding level effects even though their firing rates are saturated at mid-high levels (Young and Sachs 1979).

Average localized synchronized rate: phase locking to vowel formants. ALSR was computed from the model-generated PSTHs to quantify the strength of phase locking to formants of concurrent vowels having the same or different F0s (step 2 of Fig. 3). ALSR shows a peak at the harmonic nearest to each of the vowel formants; thus, the frequency spectrum of the vowel is reflected in AN fiber responses. The ALSR for a given harmonic was the average value of the synchronized rates from all fibers whose CFs were within ±0.25 octave of that harmonic. The synchronized rate for a fiber at CF was obtained by multiplying the PSTH with a 39-ms Hamming window and then computing the Fourier transform. This rate was normalized by dividing the RMS of the Hamming window (see Miller et al. 1997; Zilany and Bruce 2007). In the current study, the ALSR was computed at harmonics of 100 Hz when F0s were the same and at harmonics of 100 and 126 Hz when F0s were different. The maximum ALSR from the two harmonics (with the same F0) on either side of the formant value was taken as a metric to quantify the strength of phase locking to that formant. Because F1 and F2 are generally most important for vowel identification (Peterson and Barney 1952), the current study focused on the ALSR for F1 and F2 of both vowels.

Template contrast: phase locking to vowel F0. Template contrast was also computed from the model generated PSTHs to quantify the strength of phase locking to F0s of the concurrent vowels. As shown in step 3 of Figure 3, for each fiber at a certain CF (e.g., CF_1), periodicity was extracted by computing the auto-correlation function (ACF) of the PSTH. Each ACF was then multiplied by an exponential function. The time constant (tau) of this function was dependent on the exact value of CF (Cariani 2004). The ACF computed from a lower value CF (e.g., CF_1) had a slower time constant, whereas the ACF computed from a higher value CF (e.g., CF_{100}) had a faster time constant. The value of tau (ms) was varied per CF based on Cariani (2004; tau=30 ms for CF<100 Hz; tau=16 ms for 100≤CF<440 Hz; tau=12 ms for 440≤CF<880 Hz; tau=10 ms for 880≤CF<1320 Hz; tau=9 ms for CF≥1320 Hz). The CF-dependent time constants of the exponential functions were designed to account for the lower F0 limit (i.e., ~30 Hz; Pressnitzer et al. 2001) and the effect of peripheral filtering on pitch perception (Bernstein and Oxenham 2005). The pooled ACF was then computed by summing across these multiplied ACFs to obtain periodicity information across 100 fibers. A similar computation of the pooled ACF has been used previously in many pitch-related studies (e.g., Licklider 1951; Meddis and Hewitt 1991; Cariani and Delgutte, 1996; Meddis and O'Mard 1997; de Cheveigné 2005). A periodic sieve (or template) was then defined by placing narrow bins (with bin-width=50 microseconds) at every integer multiple of the template period in the pooled ACF. Template contrast at a given template period (1/template F0) was the ratio of the mean value in the periodic sieve to the overall mean of the pooled ACF (Larsen et al. 2008). Template contrast was then computed for various template periods ranging from 80 to 1,000 Hz. Template contrast >1 suggests good phase locking at each template F0 whereas template contrast ≤1 suggests

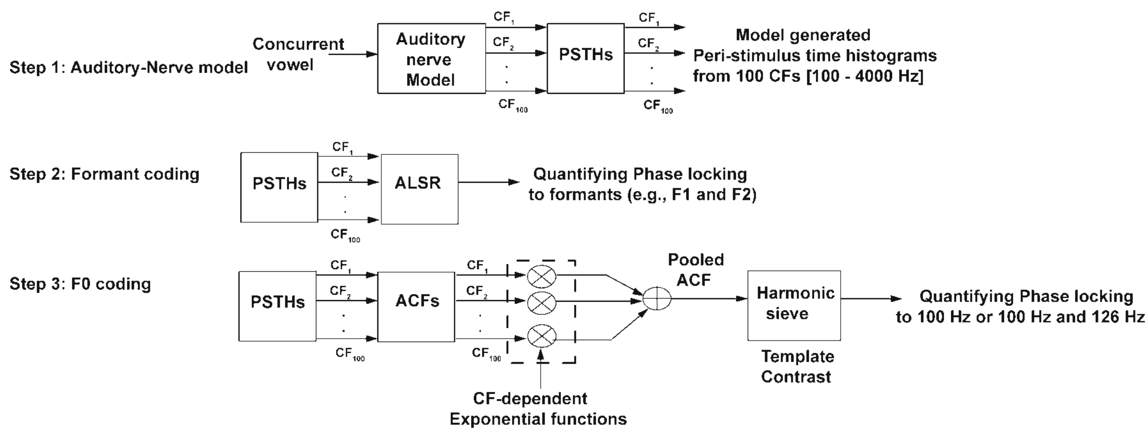

FIG. 3. Block diagram describing procedure to quantify the strength of phase locking of AN fibers to vowel formants and F0s using the auditory-nerve model. *Step 1* shows peri-stimulus time histograms (PSTHs) predicted from the model. CF_i corresponds to the *i*th CF of the AN fiber. The value of CFs ranged logarithmically between 100 and 4000 Hz, where CF_1=100 Hz and CF_{100}=4000 Hz. *Step 2* shows average localized synchronized rate (ALSR) for quantifying phase locking to vowel formants (i.e., to F1 and F2 of each vowel in the pair). *Step 3* shows template contrast for quantifying phase locking to F0 of each vowel in the pair. The auto-correlation function (ACF) is computed from PSTH and then multiplied by a CF-dependent exponential function to estimate F0 information at each fiber. The multiplied ACFs across many CFs are summed to obtain the pooled ACF (i.e., estimating F0 information across a population of fibers). Template contrast is computed from the pooled ACF. Template contrast >1 indicates good phase locking to F0. See text for additional details.

poor phase locking. A higher template contrast value suggests stronger phase locking. For the current study, phase locking of AN fibers to F0s was quantified by template contrast at 100 Hz when F0 was the same and at 100 and 126 Hz when F0 was different. Because phase locking to 100 Hz (using template contrast) when F0 is the same for the two vowels cannot contribute to vowel segregation, template contrast for this condition was not considered in the current analyses. When F0 is different for the two vowels, strong phase locking to at least one of the two F0s (100 and 126 Hz) is required for vowel segregation using the F0 difference cue.

Results and discussion

Effect of formant separation on low-frequency formants for synchrony suppression at the F2 location at higher levels. Palmer (1990) showed that a fiber near F2 of the vowel /i/ phase locked to F1 of /ɑ/ for the vowel pair /i, ɑ/ at relatively high levels. One explanation is that formant frequency spacing between two vowels may affect the strength of synchrony capture by F1 at high levels and may explain differences in identification across vowel pairs as a function of level. The following metric was developed to quantify the minimum formant distance (in octaves) between two vowels of the vowel pair: $\log_2$ [minimum F2 (Hz)/maximum F1 (Hz)]. Lower values for the minimum formant distance indicate that formants between vowels are closely spaced and may result in synchrony capture by the lower formant due to broader auditory filters at higher than lower levels. To test this assumption, two vowel pairs /ɑ, æ/ and /i, æ/ were selected for modeling based on the spacing of formants between the two vowels (see Table 1). The minimum formant distance for /ɑ, æ/ is $\log_2$ (1050/750)=0.48 octave and for /i, æ/ is $\log_2$ (1450/750)=0.98 octave. As minimum formant distance was lower for /ɑ, æ/, we expected the strength of synchrony capture by F1 would be stronger than for /i, æ/.

Figure 4 shows the effect of vowel level on synchronized rate (response from a simulated fiber at a single CF) for /ɑ, æ/ with the same F0. The CF value for each fiber was selected so that it was close to the F2 of each vowel. The simulated fiber shown in the left column (CF=1015 Hz; closer to F2 of /ɑ/) phase locked to F2 of /ɑ/ at 25 dB SPL (shown by the arrow in Fig. 4A). The same fiber phase locked to F2 of /ɑ/ at 50 dB SPL (Fig. 4B) but also responded strongly to F1 of /ɑ/ (or /æ/). At 75 and 85 dB SPL (Fig. 4C, D), the fiber did not respond to F2 of /ɑ/, but instead phase locked to F1 of /ɑ/ (/æ/) possibly due to broader auditory filters. This resulted in synchrony capture by F1 at the F2 place of /ɑ/. Another simulated fiber, shown in the right column (CF=1426 Hz; closer to F2 of /æ/), phase locked to F2 of /æ/ at 25 and 50 dB SPL (Fig. 4E, F) but evidence of synchrony capture by F1 of /æ/ (or /ɑ/) at the F2 place of /æ/ was observed at higher levels (Fig. 4G, H).

Figure 5 shows the effect of synchronized rate for /i, æ/ with the same F0. The simulated fiber shown in

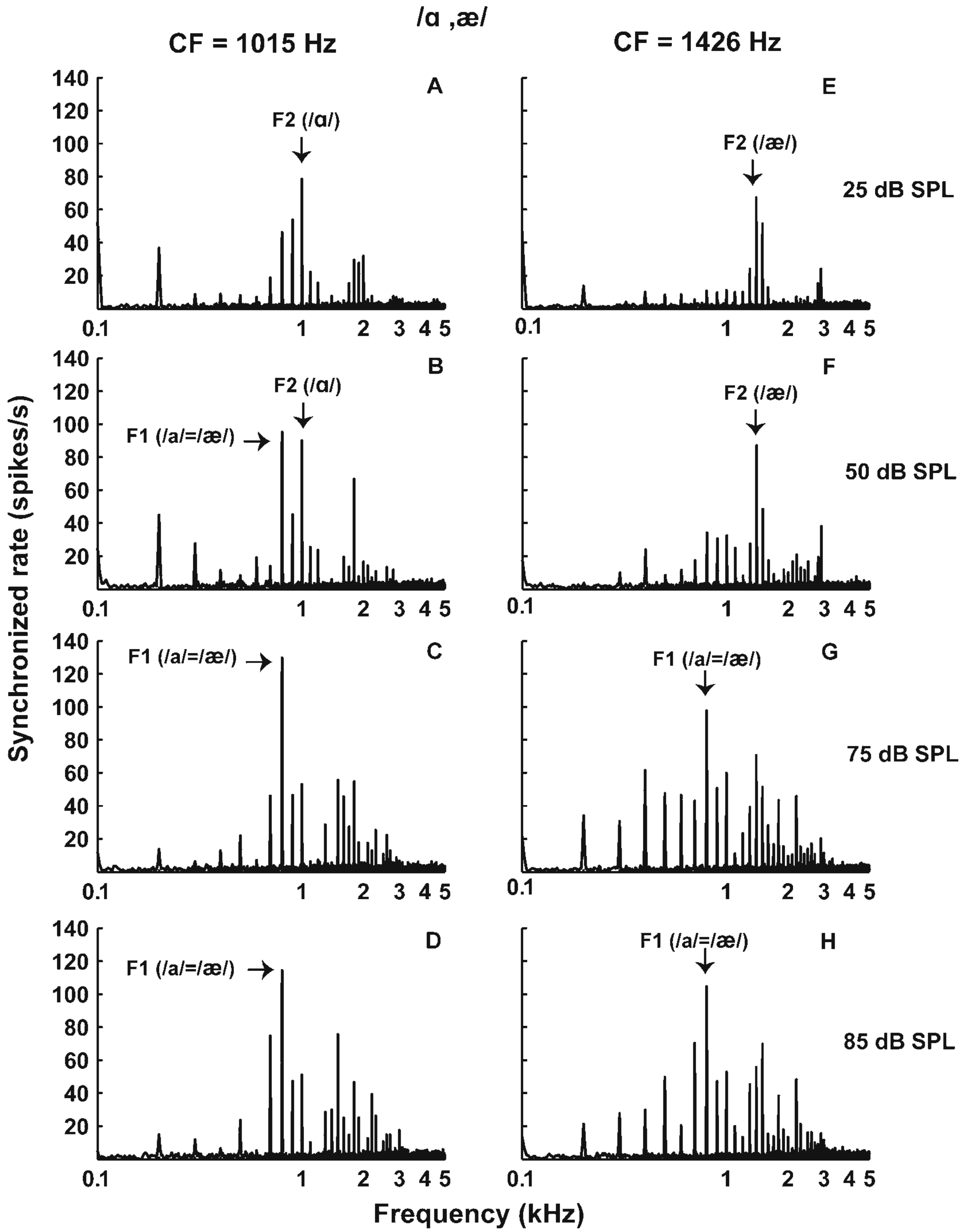

FIG. 4. Synchronized rate from two simulated fibers for the vowel pair /ɑ, æ/ with the same F0. The CF of each fiber was closer to F2 of each vowel. First column (panels **A**–**D**) shows synchronized rates at CF=1015 Hz (near F2 of /ɑ/), whereas the second column (panels **E**–**F**) shows synchronized rates at CF=1426 Hz (near F2 of /æ/). Each *row* corresponds to a different vowel level. Both vowels have the same F1 (750 Hz). The *arrow* in each panel shows phase locking to the vowel formant. See text for additional details.

the left column (CF=2235 Hz; closer to F2 of /i/) phase locked to F2 of /i/ at 25 and 50 dB SPL (Fig. 5A, B). At 75 and 85 dB SPL, the same fiber phase locked to F1 of /æ/ and a prominent peak at 2100 Hz near the F2 place of /i/ was also observed (Fig. 5C, D). The emergence of the peak at 2100 Hz could be due to the effect of broader auditory filter at higher than lower levels, which results in the best frequency (>1500 Hz) shifting to a lower frequency (Carney 1999). Another simulated fiber shown in the right column (CF=1426 Hz; closer to F2 of /æ/) phase locked to F2 of /æ/ at all vowel levels (Fig. 5E–H). Comparing vowel pairs /ɑ, æ/ (Fig. 4) and /i, æ/ (Fig. 5), synchrony capture by F1 was stronger for /ɑ, æ/ and could be related to closer formant spacing between vowel pairs.

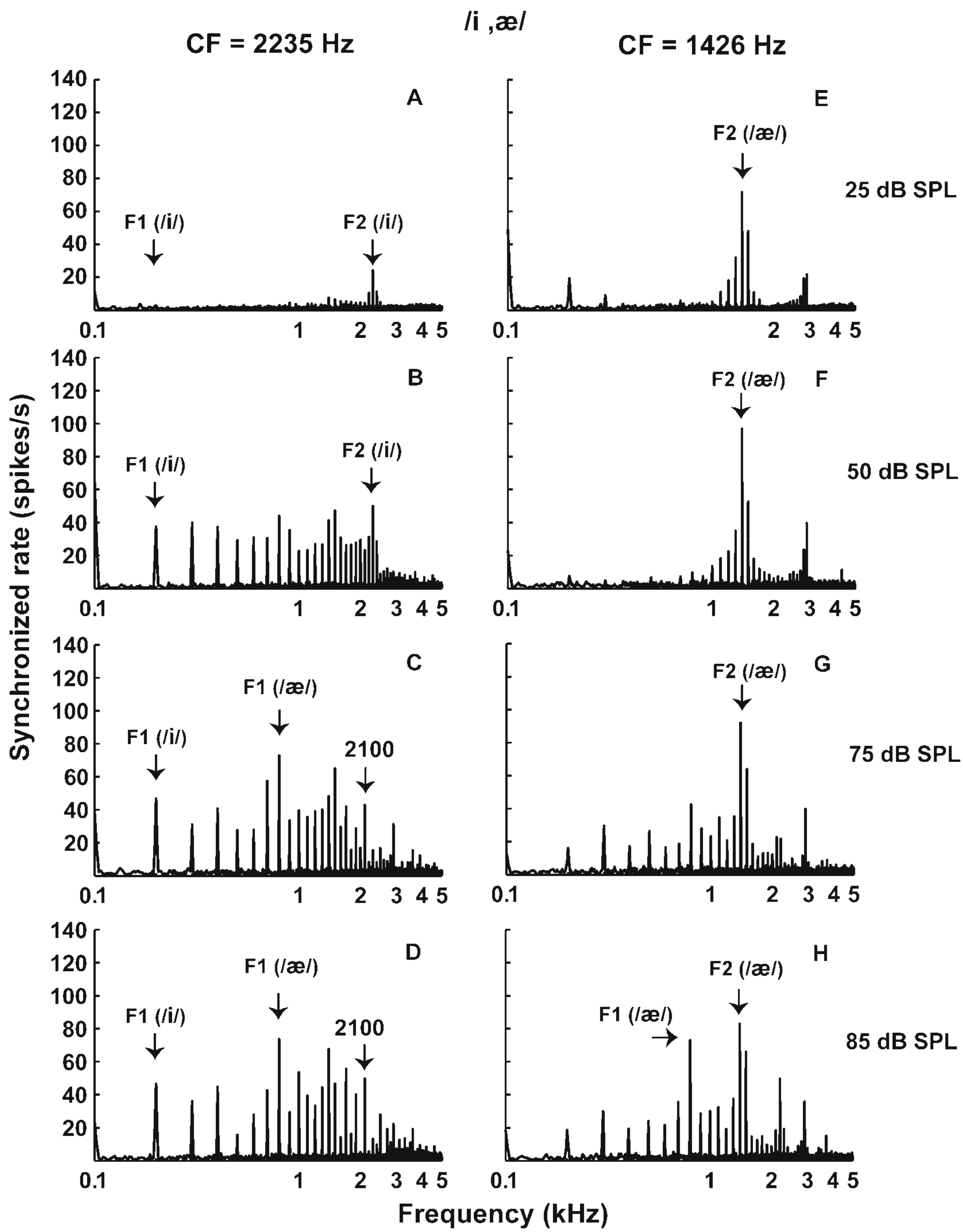

FIG. 5. Similar to Figure 4, except for the vowel pair /i, æ/ with the same F0 and at two different simulated fibers. **A–D** CF=2235 Hz (near F2 of /i/). **E–H** CF=1426 Hz (near F2 of /æ/). The location of F1 of /i/ is also shown by *arrows* in panels **A–D**. Note that there is no synchrony capture by F1 of /i/.

Role of phase locking for formant difference cues. Figure 6A shows the ALSR for the response to the vowel pair /ɑ, æ/ with the same F0 (100 and 100 Hz) at 25, 50, 65, and 85 dB SPL. ALSR peaks at the harmonic of 100 Hz nearest to each of the formants of the two vowels (arrows). Figure 6B shows the maximum ALSR around F1 and F2 of /ɑ/ and /æ/ as a function of level. The reduction in the ALSR of F2 for both vowels may be due to the presence of F1 in the same auditory filter, which results in synchrony capture by F1 at higher but not lower levels (see Fig. 4).

Figure 6C shows mean (±1 SEM) identification scores for the vowel pair /ɑ, æ/ with the same F0 as a function of vowel level. Post hoc comparisons showed that the scores improved significantly at lower levels and declined significantly at higher levels. The significant increase in identification scores at lower levels (Fig. 6C) may be related to enhanced phase

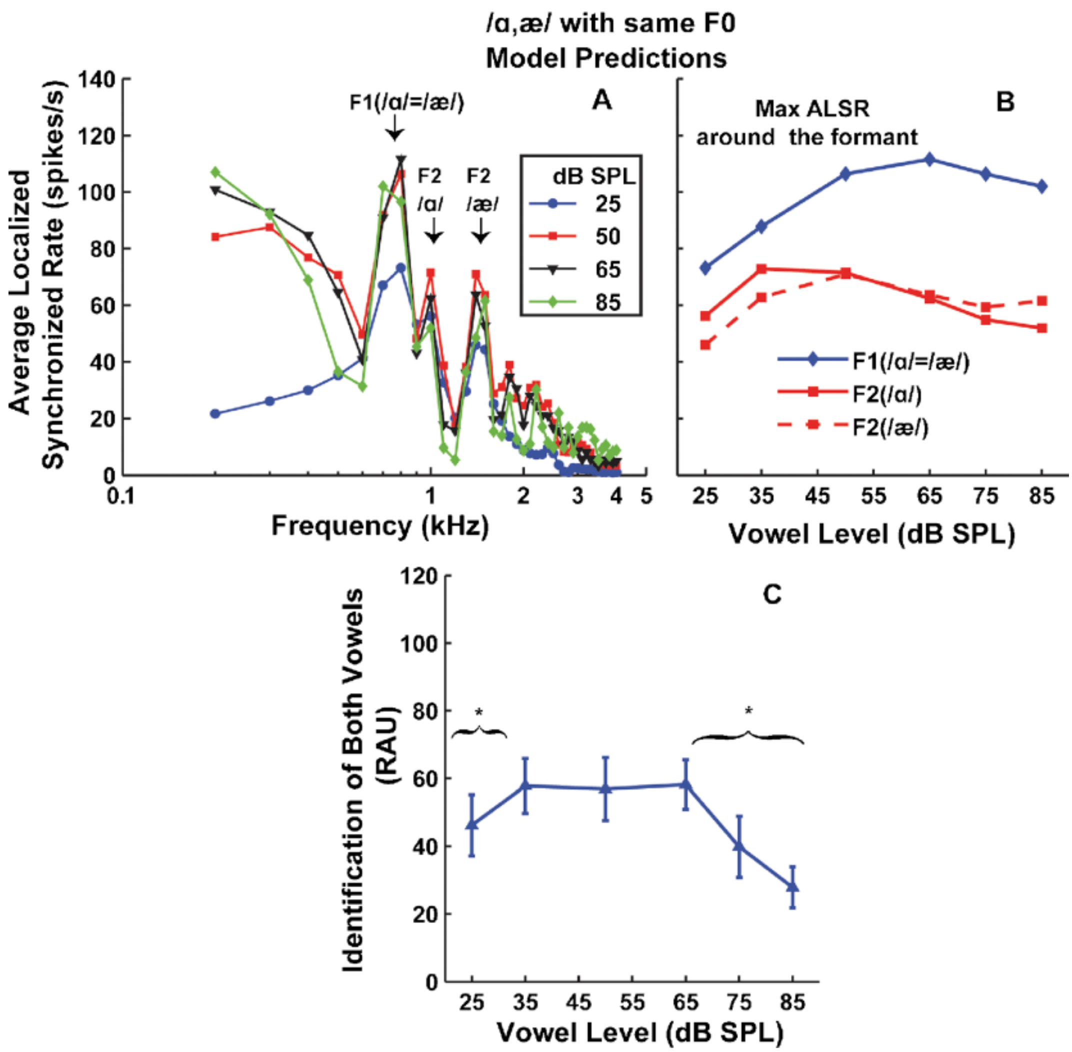

FIG. 6. Predicted phase locking of AN fibers to formants and identification scores for the vowel pair /ɑ, æ/ with the same F0. **A** Average localized synchronized rate (ALSR) for /ɑ, æ/ at 25, 50, 65, and 85 dB SPL. A peak (indicated by the *arrow*) occurs at the harmonic of 100 Hz nearest to each of the formant frequencies (F1, F2) of /ɑ/ and /æ/. **B** ALSR for F1 and F2 of /ɑ/ and /æ/ as a function of vowel level, obtained from **A**. The ALSR for F1 and F2 of the two vowels are shown by *blue diamonds* and *red squares*, respectively. The *solid blue line* shows F1 of /ɑ/ (or /æ/). The *solid red line* indicates F2 of /ɑ/, whereas the *dotted red line* indicates F2 of /æ/. **C** Identification scores of the vowel pair /ɑ, æ/ with the same F0 as a function of vowel level. *Error bars* indicate ±1 SEM, and *asterisks* indicate significant changes in scores ($p<0.05$) with increasing vowel level.

locking to F1 and F2 of the two vowels (Fig. 6B), whereas the significant decline at higher levels (Fig. 6C) may be related to degraded phase locking to F2 of both vowels (Fig. 6B).

With another vowel pair /i, ae/, ALSR peaks as expected at the harmonic of 100 Hz nearest to F1 and F2 of /i/ and /ae/ (Fig. 7A). However, there was a change in the ALSR peak for F2 of /i/ from 2300 to 2100 Hz at higher levels (see arrow at 2100 Hz in Fig. 7A and also see Fig. 5G, H). Phase locking to F1 and F2 of the two vowels improved with increasing level and then asymptoted at higher levels (Fig. 7B). There was no reduction in the ALSR of F2 for either vowel, as synchrony capture by F1 was weaker, possibly due to wider formant spacing for this vowel pair (Fig. 5).

The significant increase in identification scores for the vowel pair /i, æ/ when F0 was the same (Fig. 7C) could be partly attributed to enhanced phase locking to F1 and F2 (Fig. 7B). Identification scores were not significantly different at the two highest levels (Fig. 7C), which could be due to similar phase locking to F1 and F2 of /i/ and /æ/ (Fig. 7B).

Role of phase locking for formant and fundamental frequency difference cues. Figure 8A, B is similar to Fig. 6A, B except that F0 was different for the vowel pair /ɑ, æ/ (100 and 126 Hz, respectively). Patterns of phase locking to F2 of both vowels with

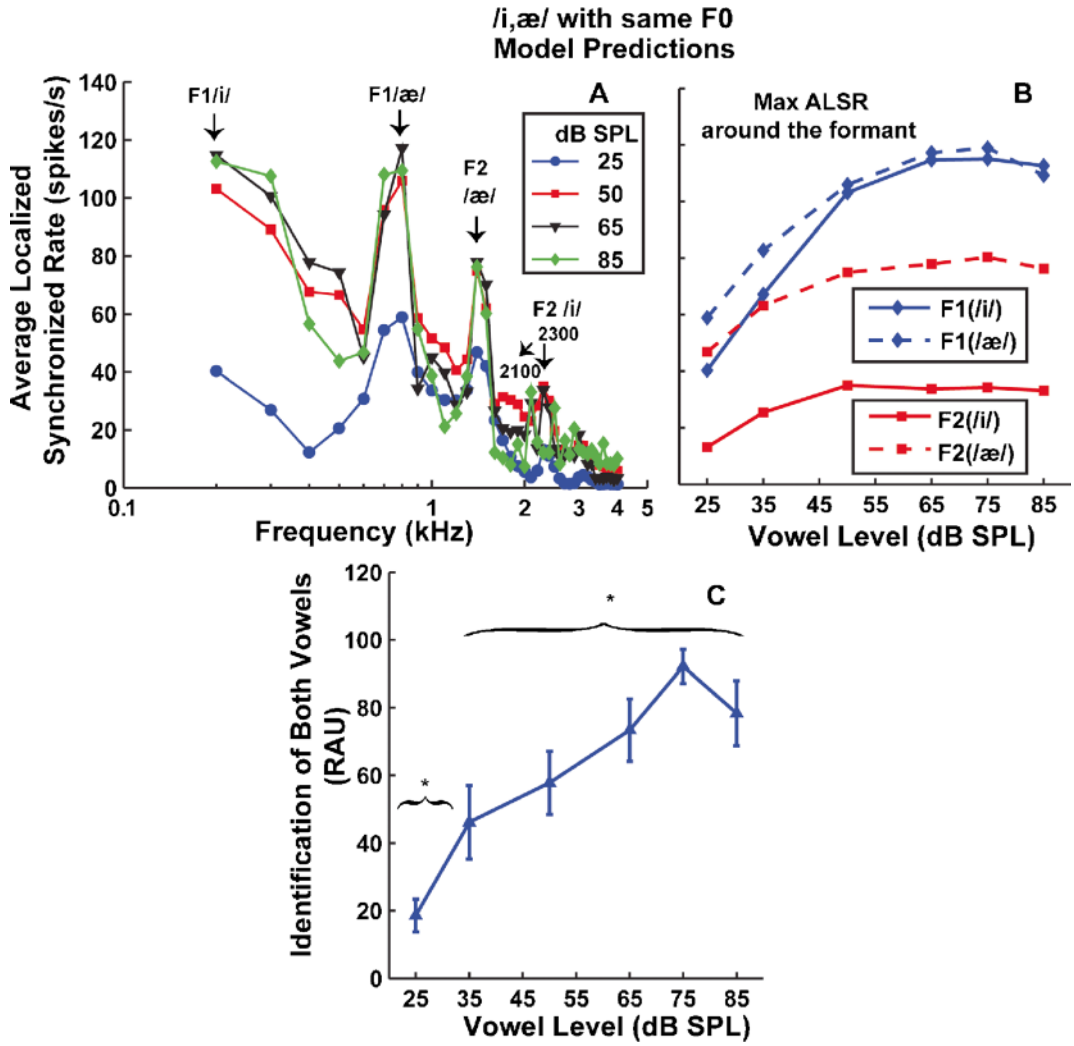

FIG. 7. Similar to Figure 6, except for the /i, æ/ vowel pair. The harmonic shift at 2100 Hz for F2 of /i/ is shown by the *arrow* for 85 dB SPL in panel **A**. Note that the maximum ALSR value around F1 of /i/ occurs at 200 Hz across vowel levels.

increasing level (red, Fig. 8B) were similar to those when F0 was the same (red, Fig. 6B). Regardless of F0 conditions, phase locking to F2 of both vowels was reduced due to synchrony capture by F1 at higher but not lower levels (compare Figs. 8B and 6B), suggesting that formant difference cues were less salient at higher levels.

Phase locking of AN fibers to 100 and 126 Hz was computed at each level to determine the extent to which vowel segregation could benefit from the F0 difference cue. Template contrast for 100 Hz of /ɑ/ was <1, suggesting generally poor phase locking at each vowel level (gray, Fig. 8C). Template contrast for 126 Hz of /æ/ was consistently >1, suggesting good phase locking that also varied with vowel level (black, Fig. 8C).

To benefit from the difference in F0, strong phase locking of AN fibers to only one of two F0s is required (Meddis and Hewitt 1992). Hence, in the current study, segregation of the two vowels in the /ɑ, æ/ vowel pair can be achieved using phase locking to 126 Hz (black, Fig. 8C). The strength of this phase locking was poorer at the lowest vowel level, which may have lead to reduced segregation ability and thus may explain the smaller F0 benefit at the lowest level (5.4 %) than at higher levels (mean F0 benefit ranged from 13.6 % to 35.0 % as shown in Fig. 8D).

The significant increase in identification scores for the different-F0 vowel pair /ɑ, æ/ (Fig. 8D) at lower levels may be related to enhanced phase locking to 126 Hz (black, Fig. 8C) and to F1 and F2 (Fig. 8B). Although phase locking to 126 Hz was strong between 65 and 85 dB SPL (black, Fig. 8C), degraded phase locking to F2 of both vowels (red, Fig. 8B) might have negated this effect and contributed to the significant decline in identification scores from 65 to 85 dB SPL. Vowel identification scores also support this interpretation because the decline in identification scores from 65 to 85 dB SPL did not differ significantly [$t(11)$=

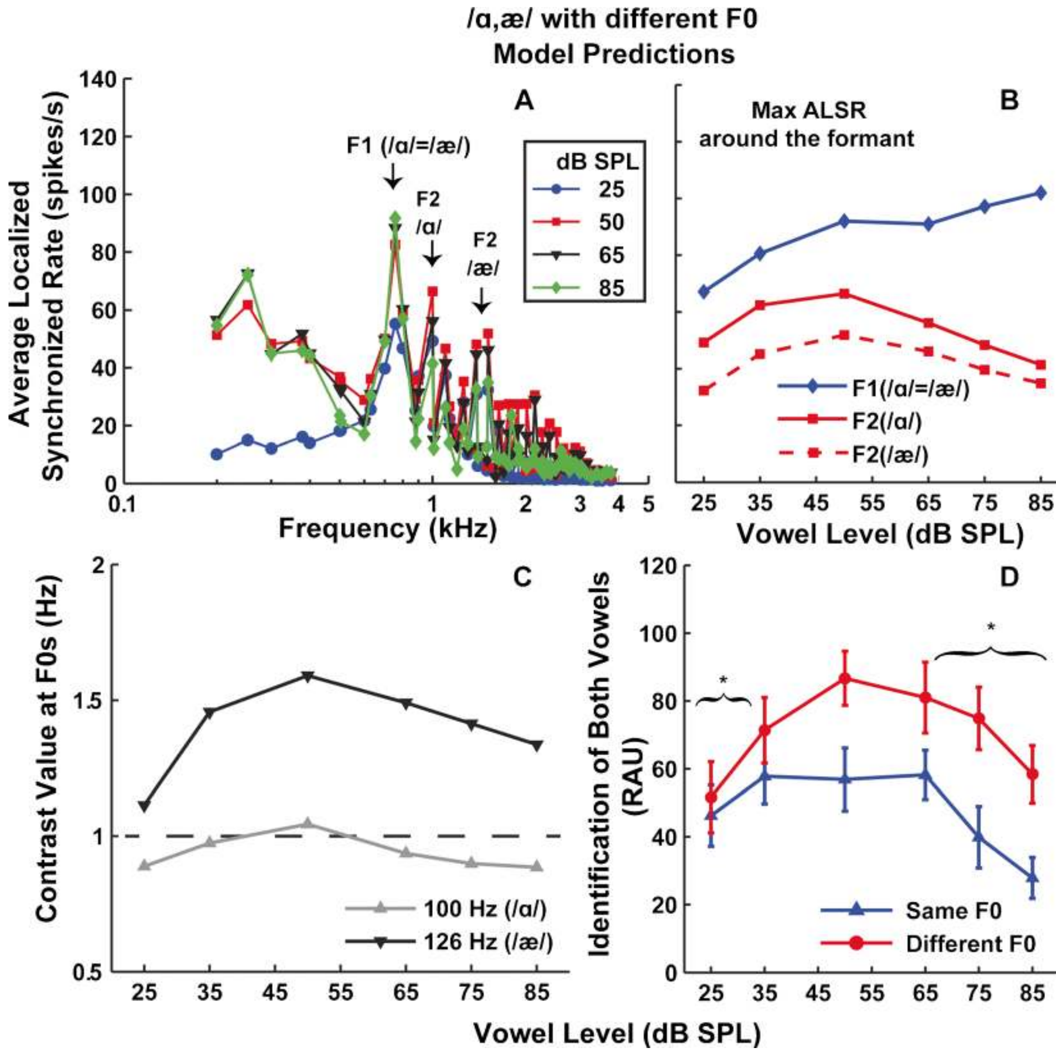

FIG. 8. Predicted phase locking of AN fibers to formants and F0s and identification scores for the vowel pair /ɑ, æ/ with different F0. **A** Average localized synchronized rate (ALSR) for /ɑ, æ/ at 25, 50, 65, and 85 dB SPL. A peak (indicated by the *arrow*) occurs at the harmonic of 100 Hz nearest to each of the formant frequencies of /ɑ/ and at the harmonic of 126 Hz nearest to each of the formant frequencies of /æ/. **B** ALSR for F1 and F2 of /ɑ/ and /æ/ as a function of vowel level. Legends are the same as in Figure 6B. **C** Template contrast for 100 Hz of /ɑ/ (*gray upward triangles*) and 126 Hz of /æ/ (*black downward triangles*) as a function of vowel level. Template contrast >1 indicates good phase locking to F0 whereas ≤1 indicates poor phase locking. **D** Identification scores of the vowel pair /ɑ, æ/ with different F0 (*red circles*) as a function of vowel level. *Error bars* indicate ±1 SEM and *asterisks* indicate significant changes in scores with increasing vowel level ($p<0.05$). Identification scores for same F0 (*blue triangles*) are re-plotted from Fig. 6C.

0.834, $p=0.422$] between same and different F0s (Fig. 8D).

Figure 9A, B is similar to Fig. 8A, B except for the vowel pair /i, æ/ (100 and 126 Hz, respectively). A marginal decline in phase locking to F2 of the two vowels was predicted at higher levels (Fig. 9B). Template contrasts for 100 Hz of /i/ (except at 25 dB SPL) and 126 Hz of /æ/ were >1, suggesting generally good phase locking to both F0s that also varied with vowel level (Fig. 9C).

The significant increase in identification scores for the different-F0 vowel pair /i, æ/ (Fig. 9D) could be partly attributed to enhanced phase locking to 100 or 126 Hz (Fig. 9C), and to F1 and F2 of the two vowels (Fig. 9B). Although phase locking to both F0s was strong for moderate and higher levels (Fig. 9C), the non-significant change in identification scores at these levels (Fig. 9D) could indicate that the small decline in phase locking to F2 of both vowels was not detrimental to vowel identification (Fig. 9B).

Identification scores at higher levels for /ɑ, æ/ and /i, æ/ may be attributed to the strength of synchrony capture by a lower formant, which could be influenced by the interaction between broader auditory filters and formant separation between vowels. To understand whether these findings are consistent across vowel pairs, model predictions were generated (not shown here) for two additional vowel pairs /æ, ɝ/ and /i, ɝ/, which had similar formant distances (0.61 and 1.35

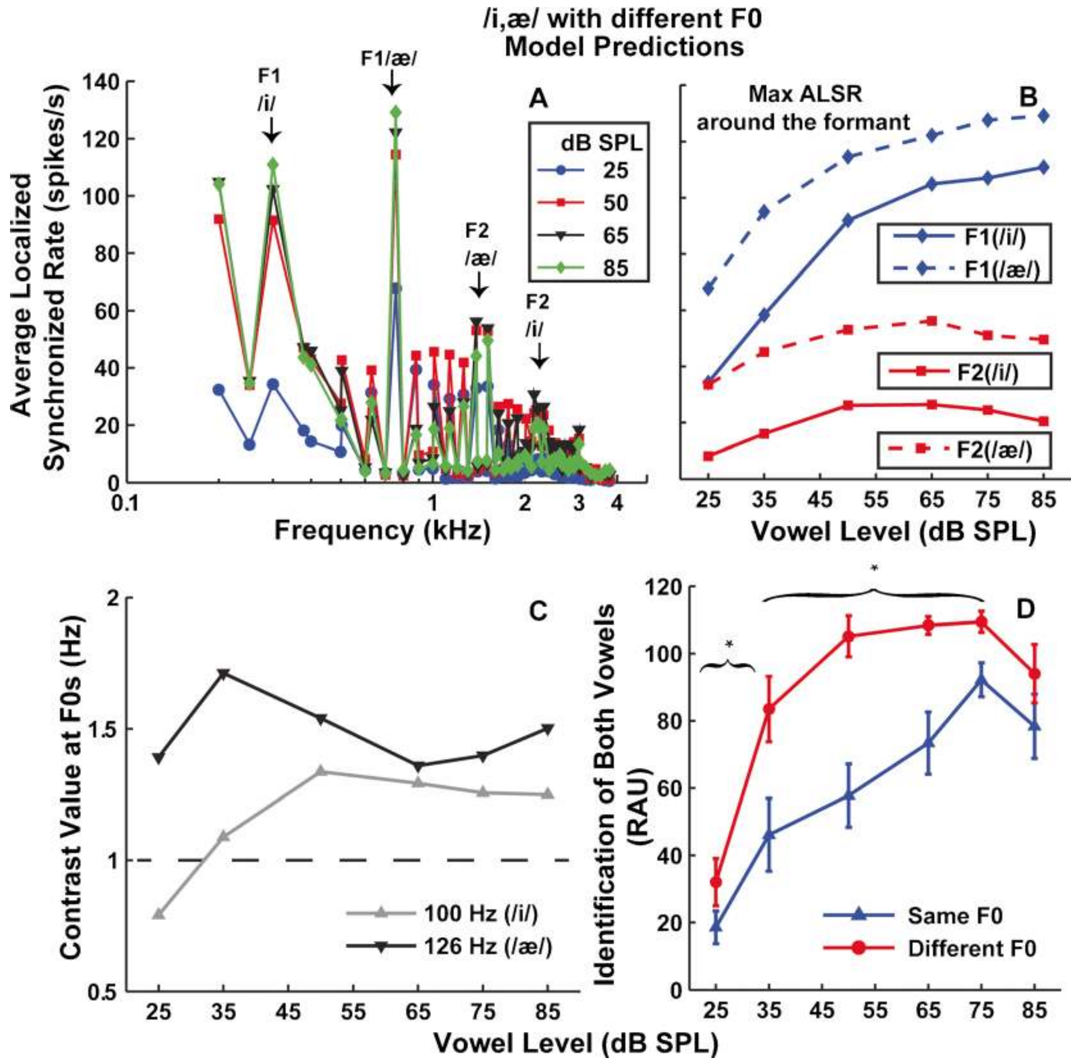

FIG. 9. Similar to Figure 8, except for the /i, æ/ vowel pair. Note that the maximum ALSR value around F1 of /i/ occurs either at 200 or 300 Hz at each vowel level. The *arrow* for F1 of /i/ is shown at 300 Hz for 85 dB SPL.

octaves) and patterns of identification as /ɑ, æ/ and /i, æ/, respectively. Although declines in phase locking at higher but not lower levels were consistent with declines in identification scores (e.g., /ɑ, æ/ and /æ, ɝ/), a modest decline was also observed for the vowel pair (e.g., /i, æ/) whose scores asymptoted as level increased. The predictions from four vowel pairs suggest that formant spacing, synchrony capture and their effects on phase locking of AN fibers could be one of the factors contributing to level-dependent changes in identification scores.

Effect of vowel level on phase locking for fundamental frequency difference cues. As shown by template contrast values in Figs. 8 and 9, patterns of phase locking of AN fibers to F0s with increasing level differed across the two vowels pairs. These differences may reflect an interaction between the spectral characteristics of the vowel pairs (e.g., locations of harmonics, formants, and troughs) and frequency selectivity of auditory filters, which may affect the ability to use F0 differences to segregate vowels. For example, phase locking to 100 and 126 Hz at mid-to-high levels provided cues for segregating the vowel pair /i, æ/ (wider formant spacing), whereas only phase locking to 126 Hz provided cues for segregating the vowel pair /ɑ, æ/ (narrower formant spacing). Moreover, stronger phase locking to 126 Hz at 25 dB SPL for /i, æ/ (Fig. 9C) than for /ɑ, æ/ (Fig. 8C) may have contributed to the larger mean F0 benefit for /i, æ/ (13.5 %; Fig. 9D) than for /ɑ, æ/ (5.4 %; Fig. 8D).

GENERAL DISCUSSION

Significance of the current study

Previous studies of concurrent vowel identification focused mainly on the contribution of F0 difference and formant difference cues at one vowel level (e.g., Summers and Leek, 1998; Chintanpalli and Heinz, 2013). The current study examined the role of F0 and formant frequency difference cues and vowel level (25–85 dB SPL) on the identification of concurrent vowels. The combination of perceptual and computational modeling approaches provides a means to separate the contribution of F0 difference and formant difference cues to vowel identification at various vowel levels.

Identification scores of both vowels improved significantly with increasing level and then declined significantly at higher levels (Fig. 2). Consistent with other studies conducted at one mid-to-high level, listeners took advantage of F0 differences between vowels (i.e., F0 benefit) as revealed by improvement in scores when F0 difference was increased, but this improvement was significantly smaller at the lowest vowel level. Computational model predictions suggest that smaller F0 benefit may be attributed to poorer phase locking to both F0s of AN fibers due to their lower firing rates, and that F0 difference cues could be limited at lower levels. Significant declines in identification scores were observed as level increased regardless of F0 (Fig. 2). However, this decline may be related, in part, to the formant spacing between two vowels. Computational model predictions suggest that the decline at higher levels may be attributed to degradation in phase locking to the second formant of the vowel pairs. Degradation in phase locking was larger for /ɑ, æ/, which had more closely spaced formants between vowels, than for /i, æ/. This suggests that the limited benefit of formant cues at higher levels could depend partially on formant characteristics, and may explain differences in performance across vowel pairs.

Use of modeling to predict level-dependent changes in identification scores of concurrent vowels

The current study suggests that level-dependent changes in phase locking of AN fibers to vowel formants and F0s may underlie level-dependent changes in identification scores. Although the AN model developed by Zilany et al. (2009) was used in the current study, it is expected that other physiologically realistic AN models that capture the basic properties of phase locking of AN fibers (e.g., models of Meddis and colleagues), would produce similar model predictions. As an additional test of the current findings, future computational modeling using phase locking of AN fibers can be developed to predict vowel identification scores. Although model predictions presented here considered F0 difference and formant difference cues separately, previous modeling studies using phase locking of AN fibers (Meddis and Hewitt 1992; Chintanpalli and Heinz 2013) suggest that listeners might utilize differences in vowels' F0 for segregation and then identify the two segregated vowels using formant difference cues. These modeling studies were focused on the effect of F0 difference at a single vowel level. Chintanpalli and Heinz (2013) employed the Meddis-and-Hewitt style F0-segregation algorithm (1992) to predict scores and found that segregation parameters varied with changes in identification scores as a function of F0 difference. As such, different parameter sets might be required at each vowel level to predict identification scores. For future computational models that predict scores, vowel scores obtained here could be useful for selecting the parameters at each vowel levels.

Other physiological explanations of vowel identification

Our model predictions using phase locking of AN fibers may explain some of the level-dependent changes in vowel identification. However, other physiological explanations are also plausible. Based on responses from pure tones (e.g., Louage et al. 2004), the current modeling assumed that phase locking of AN fibers was similar between SR classes for synthetic vowels. However, further neurophysiological study is required to determine if there are any changes in phase locking between SR classes for speech-like stimuli such as vowels. Rate-place cues are available at low-and-mid levels (e.g., Goldstein 1973; Sachs and Young 1979) for estimating formant and F0 difference cues, and it is possible that listeners could utilize these cues (solely or along with phase locking cues) for vowel identification. Lower mean firing rates of AN fibers at low vowel levels could also limit the availability of rate-place cues, as suggested by the smaller F0 benefit. This could be evaluated by expanding the present computational model to include rate-based cues of AN fibers to assess their relation to level-dependent changes in identification scores. In contrast, listeners may not be utilizing rate-place cues for vowel identification at higher levels due to fibers' rate saturation, and may instead rely on AN phase locking mechanisms. Thus, declines in identification scores with increasing level may more likely be attributed to declines in phase locking to formant cues, and synchrony capture by lower formants of the vowel pair. In the same way that formant spacing and auditory filter width interact as vowel level increases, synchrony capture by lower formants may also be determined by the interaction of formant levels and auditory filter width. That is, for a particular F2, synchrony capture by F1 at the F2 location may more likely occur with a higher level F1 than a lower level F1.

The medial olivocochlear (MOC) reflex has also been shown to enhance the response of AN fibers to pure tones in the presence of background noise (i.e., anti-masking effect; see Kawase et al. 1993; Chintanpalli et al. 2012; Smalt et al. 2014). In real-world listening, an anti-masking effect could play a role in understanding target speech in the presence of interfering speech or noise. However, the anti-masking effect of the MOC reflex may not have much influence on perception of concurrent vowels, as this task requires identifying both vowels. Thus, we expect that efferent effects would not alter the model predictions of the current study.

Phase locking of AN fibers to amplitude modulated (AM) stimuli has been shown to be converted to rate representations in inferior colliculus (e.g., Krishna and Semple 2000; Nelson and Carney 2004). To the extent that speech-like stimuli, such as vowels, are comparable to AM signals, phase locking of AN fibers to formants and F0s of concurrent vowels might be similarly converted to rate representations; more research is needed to validate this assumption. On a more general level, the role of the central system to enhance the availability of formant difference and F0 difference cues observed at the level of the AN remains to be determined.

Extension of the present work to study the effects of age and hearing loss on vowel identification

The perceptual and modeling framework used in the current study can also be extended to investigate why older adults with normal and impaired hearing have reduced concurrent vowel identification across F0 differences and reduced F0 benefit (e.g., Vongpaisal and Pichora-Fuller 2007; Arehart et al. 1997; Summers and Leek 1998). Physiological and anatomical changes in the cochlea and the decline in number of AN fibers due to increased age and hearing loss (e.g., Makary et al. 2011; Schmiedt et al. 1996) could affect phase locking of AN fibers to formants and F0s of concurrent vowels. Thus, the approach of relating modeling predictions to identification data can provide a basic framework for understanding relative contributions of changes in the cochlea and loss of number of AN fibers due to age and hearing loss.

ACKNOWLEDGMENTS

This work was supported (in part) by research grants R01 DC000184 and P50 DC000422 from NIH/NIDCD and by the South Carolina Clinical and Translational Research (SCTR) Institute, with an academic home at the Medical University of South Carolina, NIH/NCRR Grant number UL1 RR029882. This investigation was conducted in a facility constructed with support from Research Facilities Improvement Program Grant Number C06 RR14516 from the National Center for Research Resources, National Institutes of Health. We thank Ian C. Bruce and Emily Buss for sharing MATLAB code, Michael G. Heinz for his valuable suggestions, Tyler W. Eisenhart for data collection, and Skyler G. Jennings and William J. Bologna for providing valuable comments on a previous version of the manuscript.

REFERENCES

AREHART KH, KING CA, MCLEAN-MUDGETT KS (1997) Role of fundamental frequency differences in the perceptual separation of competing vowel sounds by listeners with normal hearing and listeners with hearing loss. J Speech Lang Hear Res 40:1434–1444

ASSMANN PF, SUMMERFIELD Q (1990) Modeling the perception of concurrent vowels: vowels with different fundamental frequencies. J Acoust Soc Am 88:680–697

ASSMANN PF, SUMMERFIELD Q (1994) The contribution of waveform interactions to the perception of concurrent vowels. J Acoust Soc Am 95:471–484

BERNSTEIN JG, OXENHAM AJ (2005) An autocorrelation model with place dependence to account for the effect of harmonic number on fundamental frequency discrimination. J Acoust Soc Am 117:3816–3831

BRUCE IC, SACHS MB, YOUNG ED (2003) An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses. J Acoust Soc Am 113:369–388

CARIANI P (2004) A temporal model for pitch multiplicity and tonal consonance. In: Lipscomb SD, Ashley R, Gjerdingen RO, Webster P (eds) Proceedings of the 8th International Conference on Music Perception and Cognition. SMPC, Evanston, pp 310–313

CARIANI PA, DELGUTTE B (1996) Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. J Neurophysiol 76:1698–1716

CARNEY LH (1993) A model for the responses of low-frequency auditory-nerve fibers in cat. J Acoust Soc Am 93:401–417

CARNEY LH (1999) Temporal response properties of neurons in the auditory pathway. Curr Opin Neurobiol 9:442–446

CHINTANPALLI A, HEINZ MG (2013) The use of confusion patterns to evaluate the neural basis for concurrent vowel identification. J Acoust Soc Am 134:2988–3000

CHINTANPALLI A, JENNINGS SG, HEINZ MG, STRICKLAND EA (2012) Modeling the anti-masking effects of the olivocochlear reflex in auditory nerve responses to tones in sustained noise. J Assoc Res Otolaryngol 13:219–235

DE CHEVEIGNÉ A (2005) Pitch perception models. In: Plack CJ, Oxenham A, Fay RR, Popper AN (eds) Pitch - Neural coding and perception. Springer, New York, pp 169–223

GOLDSTEIN JL (1973) An optimum processor theory for the central formation of the pitch of complex tones. J Acoust Soc Am 54:1496–1516

HEINZ MG, ZHANG X, BRUCE IC, CARNEY LH (2001) Auditory-nerve model for predicting performance limits of normal and impaired listeners. ARLO 2:91–96

JOHNSON DH (1980) The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. J Acoust Soc Am 68:1115–1122

KAWASE T, DELGUTTE B, LIBERMAN MC (1993) Antimasking effects of the olivocochlear reflex. II. Enhancement of auditory-nerve response to masked tones. J Neurophysiol 70:2533–2549

KEILSON SE, RICHARDS VM, WYMAN BT, YOUNG ED (1997) The representation of concurrent vowels in the cat anesthetized ventral cochlear nucleus: evidence for a periodicity-tagged spectral representation. J Acoust Soc Am 102:1056–1071

KLATT DH (1980) Software for a cascade/parallel formant synthesizer. J Acoust Soc Am 67:971–995

KRISHNA BS, SEMPLE MN (2000) Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. J Neurophysiol 84:255–273

LARSEN E, CEDOLIN L, DELGUTTE B (2008) Pitch representations in the auditory nerve: two concurrent complex tones. J Neurophysiol 100:1301–1319

LIBERMAN MC (1978) Auditory-nerve response from cats raised in a low-noise chamber. J Acoust Soc Am 63:442–455

LICKLIDER JCR (1951) A duplex theory of pitch perception. Experientia 7:128–133

LOUAGE DHG, VAN DER HEIJDEN M, JORIS PX (2004) Temporal properties of responses to broadband noise in the auditory nerve. J Neurophysiol 91:2051–2065

MAKARY CA, SHIN J, KUJAWA SG, LIBERMAN MC, MERCHANT SN (2011) Age-related primary cochlear neuronal degeneration in human temporal bones. J Assoc Res Otolaryngol 12:711–717

MEDDIS R, HEWITT MJ (1991) Virtual pitch and phase-sensitivity studied using a computer model of the auditory periphery. I. Pitch identification. J Acoust Soc Am 89:2883–2894

MEDDIS R, HEWITT MJ (1992) Modeling the identification of concurrent vowels with different fundamental frequencies. J Acoust Soc Am 91:233–245

MEDDIS R, O'MARD L (1997) A unitary model of pitch perception. J Acoust Soc Am 102:1811–1820

MICHEYL C, OXENHAM AJ (2010) Pitch, harmonicity and concurrent sound segregation: psychoacoustical and neurophysiological findings. Hear Res 266:36–51

MILLER RL, SCHILLING JR, FRANCK KR, YOUNG ED (1997) Effects of acoustic trauma on the representation of the vowel/ε/in cat auditory nerve fibers. J Acoust Soc Am 101:3602–3616

NELSON PC, CARNEY LH (2004) A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. J Acoust Soc Am 116:2173–2186

PALMER AR (1990) The representation of the spectra and fundamental frequencies of steady-state single- and double-vowel sounds in the temporal discharge patterns of guinea pig cochlear-nerve fibers. J Acoust Soc Am 88:1412–1426

PALMER AR (1992) Segregation of the responses to paired vowels in the auditory nerve of the guinea pig using autocorrelation. In: Schouten MEH (ed) The Auditory Processing of Speech: From Sounds to Words. Mouton-deGruyter, Berlin, pp 115–124

PETERSON GE, BARNEY HL (1952) Control methods used in the study of vowels. J Acoust Soc Am 24:175–184

PRESSNITZER D, PATTERSON RD, KRUMBHOLZ K (2001) The lower limit of melodic pitch. J Acoust Soc Am 109:2074–2084

SACHS MB, YOUNG ED (1979) Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate. J Acoust Soc Am 66:470–479

SCHEFFERS M (1983) Sifting vowels: auditory pitch analysis and sound segregation. Dissertation, Groningen University

SCHMIEDT RA, MILLS JH, BOETTCHER FA (1996) Age-related loss of activity of auditory-nerve fibers. J Neurophysiol 76:2799–2803

SMALT CJ, HEINZ MG, STRICKLAND EA (2014) Modeling the time-varying and level-dependent effects of the medial olivocochlear reflex in auditory nerve responses. J Assoc Res Otolaryngol 15:159–173

STUDEBAKER GA (1985) A 'rationalized' arcsine transform. J Speech Hear Res 28:455–462

SUMMERS V, LEEK MR (1998) F0 processing and the separation of competing speech signals by listeners with normal hearing and with hearing loss. J Speech Lang Hear Res 41:1294–1306

TAN Q, CARNEY LH (2005) A phenomenological model for the responses of auditory-nerve fibers. II. Nonlinear tuning with a frequency glide. J Acoust Soc Am 114:2007–2020

VONGPAISAL T, PICHORA-FULLER MK (2007) Effect of age on F0 difference limen and concurrent vowel identification. J Speech Lang Hear Res 50:1139–1156

YOUNG ED, SACHS MB (1979) Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. J Acoust Soc Am 66:1381–1403

ZHANG X, HEINZ MG, BRUCE IC, CARNEY LH (2001) A phenomenological model for the responses of auditory-nerve fibers. I. Nonlinear tuning with compression and suppression. J Acoust Soc Am 109:648–670

ZILANY MS, BRUCE IC (2006) Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery. J Acoust Soc Am 120:1446–1466

ZILANY MS, BRUCE IC (2007) Representation of the vowel/ε/in normal and impaired auditory nerve fibers: model predictions of responses in cats. J Acoust Soc Am 122:402–417

ZILANY MS, BRUCE IC, NELSON PC, CARNEY LH (2009) A phenomenological model of the synapse between the inner hair cell and auditory nerve: long-term adaptation with power-law dynamics. J Acoust Soc Am 126:2390–2412