# A PRELIMINARY INVESTIGATION OF THE EFFECT OF HIGH-PASS FILTERING ON THE PERCEPTION OF VOCALIC NASALITY

Richard W. Lungstrum

University of Pennsylvania

## Introduction

This paper describes a preliminary investigation of some of the acoustic cues available for the perception of vocalic nasality in naturally produced speech. Specifically, it suggests and examines the hypothesis that acoustic cues at frequencies above 1,200 Hz are sufficient for discrimination between lexical items in a language, whose difference in production is as nearly as possible exclusively the coupling or non-coupling of the nasal tract during vowel production.

The articulatory and acoustic aspects of vocalic nasality have been fairly well understood for some years. The phenomenon was defined and discussed theoretically by House and Fairbanks (1956), and quite similarly by Fant (1960), in section 2.4, "Nasal Sounds and Nasalization" (pp. 139-161), and was examined in Hattori, Yamamoto and Fujimura (1958), and notably in Delattre (1954). Recently (1985) Hawkins and Stevens investigated the effect of low-frequency cues (in the region of F1) in the perception of nasal quality in synthesized speech data, (following e.g. Fujimura (1960, 1961)). Without entering on a detailed review of their article it is pertinent to mention here that Hawkins and Stevens (and Stevens, Fant and Hawkins (1986)) suggest that the most salient acoustic cue to vocal nasality is the presence of a resonance-antiresonance pair in the vicinity of the first formant, and that the introduction of a pole-zero pair into the vocal tract transfer function in the vicinity of the first formant produced stimuli which were identified reliably as nasalized monophthong vowels.

This work establishes that low-frequency cues are sufficient to the perception of vocalic nasality. But some of Hawkins and Stevens' stimulus types are very sensitive to manipulation of their low-frequency acoustic structure, while other types are quite insensitive to the same manipulation (certain vowel stimuli were particularly sensitive to the location of their first formant and the nasal pole-zero with respect to their reliable identification or "natural" character), suggesting that other acoustic information may provide important concurrent cues for the perception of nasality, at least for some vowels. The usual understanding of the Motor Theory of Speech Perception maintains that speech perception is mediated through processes which control the production of speech. All acoustic effects (perhaps not only acoustic) of the articulations of speech sounds are available as cues to their perception. Thus it is likely that there may be several sufficient cues for a speech sound, all of which work together to form a coherent total impression, which is interpreted in terms of the articulation which could produce it. In such a situation, it is unlikely that any one acoustic cue is absolutely necessary to the perception of a particular speech sound. Hawkins and Stevens provide no evidence either way as to the necessity of the low-frequency acoustic cues they investigate.

The present investigation examines the hypothesis that higher-frequency acoustic information provides sufficient cues for the reliable identification of vocalic nasality. This hypothesis is supported in Delattre (1954) and somewhat ambiguously in Fant (1960). These sources describe a nasal formant near 2,000 Hz, or in the latter case, a zero at 1,800 Hz and an extra peak in the F3 region. Delattre mentions in connection with this formant, "A cette fréquence, l'audibilité est assez forte, malgré la faible intensité." (p. 104) Data which support this hypothesis also stand to refute another: that low-frequency cues (specifically the pole-zero pair in the vicinity of F1) are necessary to perception of vocalic nasality.

Phonologists have given much attention to the genesis of nasal vowels (e.g. Ferguson, (1975)). The common if not exclusive phonetic explanation for the advent of significant nasality in vowels is as a result of anticipatory coarticulation before nasal consonants. Subsequently the nasal consonants are redundant as information markers, and are eliminated from the string, with vowel nasality carrying the significant distinctive information. Tape-splicing tests by Ali, et al. (1971) and Bond (1976) investigate the coarticulatory nasalization through which such a linguistic change purportedly is effected. See also Hattori, et al. (1958). Malécot (1960a,b) recognizes just such a situation emergent in English forms on the canon

XC VN ([+C -vd -cont] (^ C)) #

=> XC [V +nas] ([+C -vd -cont] (^ C)) #

This results in minimal pairs of English words distinguished by the presence or absence of vocalic nasality.

Thus recognizing that minimal pairs of words exist in (some forms of) English between which discrimination relies on the recognition of vocalic nasality, it is straightforward to design a test around these words for the sufficiency of higher-frequency acoustic cues for vowel nasality. This can be done by isolating the higher-frequencies by high-pass filtering the signal to remove all lower-frequency information, and then presenting the filtered tokens as stimuli to listeners for their judgement, to determine if the remaining higher-frequency acoustic information contains sufficient cues for reliable discrimination based on nasality differences.

In such a test, care must be taken to assure that the steady-state vowels contain the only cues for discrimination between pairs of words. Two other kinds of cues particularly could also be present, and must not be allowed in the test stimuli lest they interfere with the investigation, namely:

1) the presence of a nasal consonant following the vowel (particularly the acoustic manifestation of nasal murmur), and

2) significant suprasegmental coarticulatory cues to nasal presence, particularly significant differences of vowel duration (see e.g., House and Fairbanks (1953))

Both of these factors are identifiable using evidence from spectrography and waveform records.

## Preparation of the test materials

Data were generated to represent the naturally occurring oppositions in American English of the general type:

(consonant,) consonant, vowel{0 ~ nasal}, voiceless obstruent
or: (C)CV(0 ~ N)T.

The core data are (near) minimal pairs whose discrimination crucially depends upon differentiation of a nasal quality. The core data consist of these minimal pairs: individual data of the general class having a nasal presence (henceforth, "nasal data") were not admitted unless the corresponding data without the nasal presence (hereafter, "oral data") also occur in English, and vice-versa. One pair was chosen for each combination of vowel quality and final consonant place of articulation. The only exception allowed was for the vowel quality /au/, the only patent diphthong in the test data, for which the oral data contain a final fricative (mouth). No nearer minimal pair was found, and it was desired to include this vowel quality in the test data so that each of the vowel qualities occurring in the target canon in American would be included in the investigation. (Stimuli with back low unrounded and back mid-low round vowels were grouped together in the responses. Note that in the dialects of some of the respondents the vowel phonemes represented in these types are merged.) A few other, non-paired data were admitted to the corpus. These consisted of oral types, differing in each case from one of the oral members of the core data pairs only in perceived vowel height (e.g., the words *tint*, *tit*, and *Tet*). This measure allowed for the possibility of altered perceived vowel height as an effect of nasalization of vowels in the test producing confusion with data containing oral vowels of adjacent heights. See J. Ohala (1975), Abramson, *et al.* (1981), Beddor, *et al.* (1986), Wright, (1975, 1986), Krakow, *et al.* (1987). No significant confusion did result, and no further discussion of this aspect is necessary here.)

The test words were:

| | | | | |
|---|---|---|---|---|
| pip | pimp | pep | pop | pomp |
| tit | tint | Tet | fought | font |
| pitch | pinch | | hock | honk |
| pick | pink | peck | pup | pump |
| he'p | hemp | | putt | punt |
| pet | pent | | much | munch |
| wretch | wrench | | truck | trunk |
| tap | tamp | | mouth | mount |
| pat | pant | | | |
| catch | can't'y' | | | |
| tack | tank | | | |

Tokens of each word were produced in such a manner that the nasality present in the nasal data was manifested in vocal nasalization, with no overt nasal consonants. The velar port opened during the vowel articulation, coupling the nasal tract, and phonation ceased in glottal stop before oral obstruction was effected, but after some portion of the formant transitions had been produced indicating the place of articulation of the lost stop. The oral data similarly end in glottal stop before oral closure, but after more or less of the formant transitions had transpired.

Three tokens of each word were recorded in a frame (say the word ____ one more time) in a low-noise environment, on unspliced new high quality magnetic tape, using a Crown model 600 reel-to-reel recorder and Electro-Voice 631 B microphone. The data were digitized at a sampling rate of 22,000 Hz using a MacNifty Audio Digitizer from MacNifty Central of Minneapolis, MN and recorded in digital form using the SoundWave software program from Impulse, Inc. of Minneapolis, MN., and an enhanced Apple

Macintosh SE computer.

I reviewed the tape to assure that the recorded tokens were recognized reliably as those intended. The auditory impression of these data was that of words excerpted from a context of continuous speech, not like citation forms, but not unnatural or peculiar. (words with final -ch sounded the most colloquial, specifically, urban-north-eastern).

Wide-band spectrograms were made of the original tape data, and reviewed to assure the absence of nasal consonants in the nasal data. Measurements of the vowel durations were made from the SoundWave waveform records, referring also to the spectrograms. Measurements were made from the first glottal pulse whose waveform was characteristic of the vowel articulation, Corresponding with the first pulse of the spectrographic records exhibiting legible formant bands for at least the first three formants) to the cessation of harmonic production. Values for the three tokens of the same type were averaged, and the averages for the minimal pairs were compared. In nearly all cases the nasal data showed greater vocalic duration than their oral counterparts. This confirmed expectations based on literature on vowel duration in different consonantal environments (cf. House and Fairbanks (1953), Delattre and Monnot (1968)).

Using the editing facility of the SoundWave program, the nasal data were shortened to approximate as nearly as possible the durations of the corresponding oral data. I excised the number of glottal pulses in the vocalic portion of each nasal token whose total duration most nearly equalled the average difference between oral and nasal tokens in the pair. I excised non-adjacent glottal pulses, and removed a similar number of pulses from the earlier (non-nasalized) and later (nasalized) portions of the relatively steady-state vowels. Pulses were identified from an initial increasing intersection of the axis to the following initial increasing intersection of the axis.

The version of SoundWave I used to prepare the data lacks a zero-axis display, and thus required an externally fabricated display axis for editing purposes. As I lacked the requisite programming skills for Macintosh software (and the source code for the program), a full-screenwidth display of silence was produced, and using a straightedge, the resulting zero line was traced lightly on the video display screen in pencil. Care was taken to avoid parallax-induced errors because the surface of the display is a few millimeters in front of the phosphor of the cathode ray tube. In this case, the axis was drawn so it lay correctly in the field of view when my head was in a comfortable position. Records appeared on the screen beginning and ending with periods of silence several milliseconds long, so it was a simple matter to keep my plane of view alligned so the zero line in these silent periods corresponded with the axis.

On playback, the data records which had been edited in this way sounded natural to colleagues I played them for, and on side-by-side comparison with the unedited records, the edited records were nearly indistinguishable. When the editing produced any distracting noise, the record was re-edited. Eventually edited data were produced of which most were of very good quality. In one or two cases I myself was able to hear "something," even after several attempts, but on inquiry, other listeners at the lab were not aware of this presence. At this point, I considered the editing to be good enough. The data records now consisted of three groups:

1) Oral digitized data, individual record of each token, unedited
2) Nasal digitized data, individual records, unedited
3) Nasal digitized data, individual records, edited to normalize their vocalic duration to that of their corresponding oral data by type class (min. pair)

Now test tapes were made. Each tape contained the data from group one together with nasal data, either from group two or group three. The order of the data on the test tapes was randomized, a different random order was used on each of the four tapes. Two of the tapes consisted of data from {1+2} (hereafter called unedited-data tapes), and two were from {1+3} (hereafter, edited-data tapes). Each tape consisted of 120 tokens. Each token was presented twice, with a silent pause of approximately 3/4 second between the repetitions; a pause of approximately 3 seconds was left between consecutive tokens.

One edited-data tape and one unedited-data tape were re-recorded through a high-pass filter, with a low cutoff at 1200 Hz. This produced four tapes used in the listening tests:

1) Unedited-data tape, unfiltered
2) Unedited-data tape, high-pass filtered
3) Edited-data tape, unfiltered
4) Edited-data tape, High-pass filtered

The main test in the investigation used tapes three and four. Spectrograms were made of a sample of the filtered data to check the filter operation..

## Test procedure

Test participants were eight adults, four male and four female. One was a native speaker of new-world Spanish, with extremely good command of American English; seven were native speakers of English, from seven different dialect areas in North America. The test was administered to the participants in three sub-groups at different times. The tapes were played through a loudspeaker at a level comfortable for the listeners, who sat four to ten feet from the speaker in a reasonably isolated and noise-free room. Test procedure and response instructions were given to the participants, along with a response sheet. Participants were told that all the data would represent words of American English, and the words were listed on the instruction sheet in an order unrelated to that of any of the test sequences. If a word was not familiar in all respondants' dialects, the word was defined or explained. Respondants were then asked to turn the instructions face down, and not to use them directly as reference for their answers. Next, the test procedure was discussed. Participants were again instructed to write after each double reproduction of a token, on the numbered line on the answer sheet corresponding to the serial position of the token in the test, the American English word they believed they heard. Tapes three and four were played, with a recess of a few minutes between the tapes.

## Results:

Responses were coded by four criteria:

1) whether a response was made
2) whether the presence of nasality of the word in the response agreed with that of the original production of the token
3) whether the vowel quality (aside from nasality) of the word in the response agreed (phonemically) with that of the token word
4) whether the token word was correctly identified in the response (i.e., the response given is the same word as was intended by the speaker who recorded the test tape)

The results were grouped to allow independent analysis of the effect of several variables on the accuracy of responses to the intended stimuli. These included filtration of the stimuli prior to presentation, the place of articulation of final consonant, intended presence or absence of nasality, and vowel quality (phoneme) other than nasality. Additionally, vowels were then grouped into front, back (monophthongs), non-high, and high (i.e., /i/). The rates of response, and of correct response, were tabulated, and these rates were

TABLE 1 difference in rate of correct judgement of intended presence or absence of vocalic nasality (underlining = negative value)

| /r | oral<br>u - f | nasal<br>u - f | o - n | u - f | unf.<br>o - n | fil.<br>o - n |
|---|---|---|---|---|---|---|
| lab | .017 | .018 | .025 | .017 | .025 | .026 |
| alv | .012 | .050 | .061 | .016 | .092 | .030 |
| pa-al | 0 | .021 | .011 | .011 | 0 | .021 |
| vel | .009 | 0 | .004 | .005 | 0 | .009 |
| i | .020 | .053 | .037 | .016 | .073 | 0 |
| e | .011 | .028 | .010 | .006 | .028 | .011 |
| æ | 0 | .010 | .026 | .006 | .021 | .031 |
| ʌ | .011 | .052 | .021 | .032 | 0 | .041 |
| ɑ/o | 0 | 0 | 0 | 0 | 0 | 0 |
| au | .042 | .083 | .146 | .021 | .208 | .083 |
| front | .010 | .023 | .026 | .005 | .043 | .010 |
| back | .006 | .030 | .012 | .018 | 0 | .024 |
| hi(=i) | .020 | .053 | .037 | .016 | .073 | 0 |
| nonhi | .005 | .012 | .015 | .008 | .012 | .019 |
| total | .010 | .007 | .027 | .002 | .036 | .019 |

TABLE 2 difference in rate of correct identification of intended vowel quality apart from nasality

| /r | oral<br>u - f | nasal<br>u - f | o - n | u - f | unf.<br>o - n | fil.<br>o - n |
|---|---|---|---|---|---|---|
| lab | .018 | .022 | .135 | .019 | .133 | .137 |
| alv | .019 | .051 | .087 | .013 | .122 | .052 |
| pa-al | 0 | 0 | .021 | 0 | .021 | .021 |
| vel | .008 | .001 | .093 | .005 | .089 | .096 |
| i | .001 | .039 | .252 | .020 | .231 | .279 |
| e | .017 | .058 | .046 | .050 | .039 | .014 |
| æ | 0 | .156 | .109 | .078 | .187 | .031 |
| ʌ | .021 | .010 | .006 | .016 | .011 | 0 |
| ɑ/o | .016 | .042 | .046 | .029 | .033 | .059 |
| au | 0 | 0 | 0 | 0 | 0 | 0 |
| front | .016 | .026 | .140 | .003 | .161 | .119 |
| back | .005 | .012 | .015 | .004 | .007 | .024 |
| hi(=i) | .001 | .039 | .252 | .020 | .231 | .279 |
| nonhi | .011 | .027 | .048 | .007 | .066 | .028 |
| total | .008 | .012 | .088 | 0 | .098 | .078 |

TABLE 3 difference in rate of correct identification of the American English word intended as stimulus

|  | oral | nasal |  |  | unf. | fil. |
|---|---|---|---|---|---|---|
| /r | u - f | u - f | o - n | u - f | o - n | o - n |
| lab | .224 | .034 | .035 | .094 | .164 | .094 |
| alv | .087 | .038 | .139 | .030 | .201 | .076 |
| pa-al | .031 | 0 | .005 | .016 | .021 | .010 |
| vel | .018 | .019 | .069 | .019 | .068 | .069 |
| i | .093 | .035 | .210 | .029 | .273 | .145 |
| e | .142 | .047 | .055 | .133 | .124 | .029 |
| æ | .083 | .145 | .072 | .032 | .087 | .041 |
| ^ | .004 | .042 | .083 | .018 | .106 | .060 |
| α/o | .087 | .025 | .093 | .058 | .123 | .061 |
| au | .083 | 0 | .125 | .042 | .166 | .083 |
| front | .104 | .052 | .116 | .043 | .204 | .048 |
| back | .035 | .036 | .010 | .035 | .011 | .010 |
| hi(=i) | .093 | .035 | .210 | .029 | .273 | .145 |
| nonhi | .093 | .015 | .030 | .043 | .004 | .024 |
| total | .092 | .018 | .072 | .040 | .127 | .017 |

TABLE 4 difference in the rate of response to stimuli

|  | oral | nasal |  |  | unf. | fil. |
|---|---|---|---|---|---|---|
| /t | u - f | u - f | o - n | u - f | o - n | o - n |
| lab | .008 | .025 | .009 | .017 | .017 | 0 |
| alv | .018 | .014 | .013 | .003 | .003 | .029 |
| pa-al | 0 | 0 | 0 | 0 | 0 | 0 |
| vel | .008 | .011 | .005 | 0 | .004 | .015 |
| i | .010 | .010 | 0 | .010 | 0 | 0 |
| e | .017 | .028 | .028 | .021 | .033 | .022 |
| æ | 0 | 0 | 0 | 0 | 0 | 0 |
| ^ | .032 | 0 | .026 | .016 | .010 | .042 |
| α/o | .014 | .014 | .028 | 0 | .042 | .014 |
| au | 0 | 0 | 0 | 0 | 0 | 0 |
| front | .003 | .004 | .012 | .004 | .012 | .011 |
| back | .024 | .006 | .003 | .009 | .012 | .018 |
| hi(=i) | .010 | .010 | 0 | .010 | 0 | 0 |
| nonhi | .015 | .003 | .010 | .010 | .004 | .016 |
| total | .010 | 0 | .008 | .006 | .003 | .013 |

compared, yielding the figures in the tables. Underlined figures are negative values. The rows represent the groupings just named. The columns give the difference of response for intended oral vowels, unfiltered minus filtered; nasal vowels, unfiltered minus filtered; all oral vowels minus all nasal vowels; all unfiltered stimuli minus filtered; unfiltered oral vowel stimuli minus nasal vowel stimuli, and filtered oral vowel stimuli minus nasal vowel stimuli. The tables present values for:

1: rate of judgement in agreement with the intended nasality or orality of the vowel in the stimulus;
2: rate of correct identification of the vowel phoneme except for judgement of nasality;
3: rate of correct absolute identification of intended word;
4: rate of response overall.

In addition, the figures of table 1, column [u - f] are presented graphically, following the divisions of the table, showing the correlation of effect of high-pass filtering on perception of vocalic nasality with (graph 1) following consonantal environment, (graph 2) vowel phoneme environment, and (graph 3) vowel phonological class environment

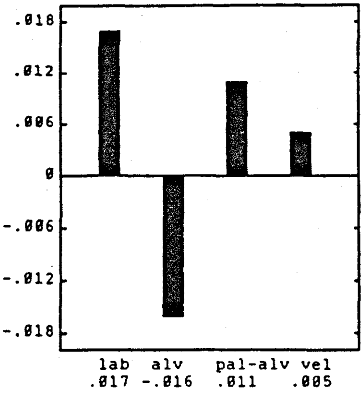
## Discussion of results

As evident in table 1 in column [u-f], the effects of filtering on the perception of the nasal-oral vowel distinction were roundly minute. Indeed, for several subcategories in the data, the listeners' judgements of the filtered stimuli agreed more successfully with the intended produced nasality value than do those of unfiltered stimuli. It does not appear that these effects are significantly dependent on the place of articulation of the latent final consonant, nor on the other qualities of the vowel itself. All the figures lie in a small region near zero (positive figures indicate judgements were more successful for unfiltered stimuli than filtered ones, negative figures indicate the reverse).

No great differences between maximum and minimum values are found among the responses categorized for either kind of environment, consonantal or vocalic. The categorizations based on consonantal environment (graph 1) show a variation of 0.033, between a highest value for pre-labial environments (+0.017) and a lowest value for pre-alveolar environments (-0.016). Among the environmental groupings the widest variations are shown among vocalic environments. There is variation of 0.053 between the highest value, for data containing the vowel [ʌ] (+0.032), and the lowest for those with the vowel [i]

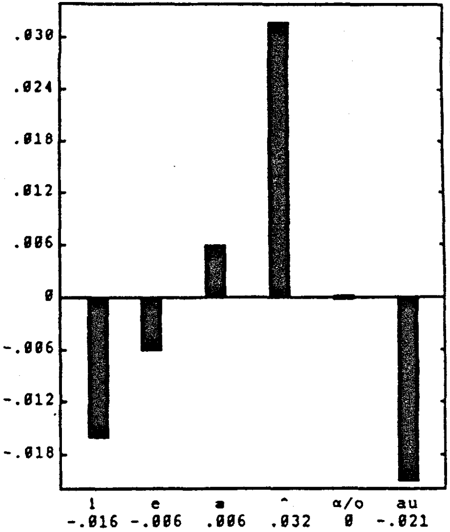
GRAPH 1

CORRELATION OF FOLLOWING CONSONANTAL ENVIRONMENT WITH EFFECT OF HIGH-PASS FILTERING AT 1,200 C/S ON PERCEPTION OF VOCALIC NASALITY



GRAPH 2

CORRELATION OF VOWEL PHONEME ENVIRONMENT WITH EFFECT OF HIGH-PASS FILTERING AT 1,200 C/S ON PERCEPTION OF VOCALIC NASALITY



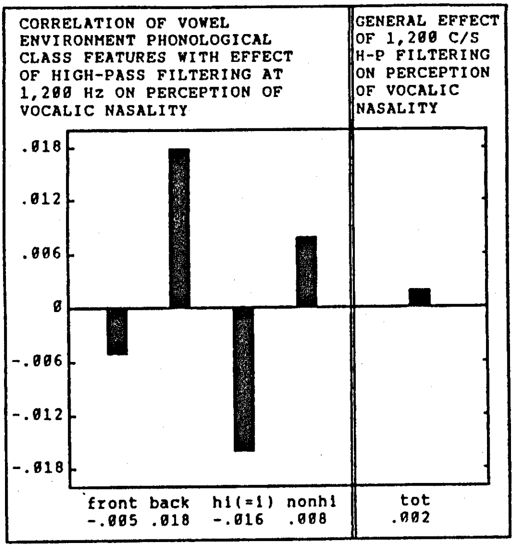
(-0.021). Tiny as these differences are, graph 2 makes clear that their distribution does have symmetry. From [i] through [e], [æ] and [ʌ], each step shows a larger effect of filtering, and each step is associated with a lower frequency F2 than the step before it. However the next step in this line, to [ɑ/o], does not yield another increase in effect, but a fall back to 0 effect (each subgroup [ɑ] and [o] shows the identical 0 value for this variable). One may note that the F2 peaks for [ɑ] and [o] are near or just below the filter cutoff of 1,200 Hz, while that of [ʌ] is slightly above it, near 1,500 Hz. [i] and [e] have higher frequency F2, and their negative effects from filtering may reflect an interaction of F2 with the expected approx. 2,000 Hz formant and/or 1,800 Hz antiformant. But with effects generally so small, I don't want to try to make any more of this.

Similarly, differences were not great across some common phonological oppositions. Scores for front and back vowels were -0.005 and +0.018 respectively, or a difference of 0.023. The only high monophthong environment in the test, [i], shows -0.016, while the value for the other monophthongs together is +0.008, for a difference of 0.024. Small as these differences are, it is worth noting that they exist, and that in each case one of the pair of environments shows a negative value. And it is particularly interesting that [i] stimuli are better identified after filtering than before (see graph 3). (One might also look at oppositions between consonantal environments characterized as grave and nongrave, or apical and dorsal)

## General observations

In general the responses indicate that high-pass filtering at 1,200 Hz has negligible effect on the perception of nasality in the test stimuli. While there is some variation in the degree of this effect among various groups of stimulus type, and more variation among the subcategorizations recognized in the response data, nowhere is any significant confusion between the nasal and oral stimuli introduced by the filtration. Thus, relatively low-frequency cues are not necessary to the perception of vocalic nasality. Experimentation with synthetic speech (Hawkins & Stevens, (1985), following on Fujimura, (1960, 1961)) has shown that these low frequency cues are sufficient for fairly reliable perception of nasality. Fant (1960) recognized effects at various frequency ranges resulting from coupling the nasal tract in producing vocalic sound, however no investigation as yet has been undertaken to compare the relative salience of cues or combinations of cues at different regions in the spectrum. The present data indicate that higher-frequency cues are sufficient to the perception of vocalic nasality in the absence of acoustic information at lower frequencies. Experiments with synthetic speech could be undertaken

GRAPH 3


CORRELATION OF VOWEL ENVIRONMENT PHONOLOGICAL CLASS FEATURES WITH EFFECT OF HIGH-PASS FILTERING AT 1,200 Hz ON PERCEPTION OF VOCALIC NASALITY
GENERAL EFFECT OF 1,200 C/S H-P FILTERING ON PERCEPTION OF VOCALIC NASALITY
.018
.012
.006
0
-.006
-.012
-.018
front -.005
back .018
hi(=i) -.016
nonhi .008
tot .002

to determine the effect of higher-frequency cues when lower-frequency information is still present which does not contain the familiar pole-zero pair in the vicinity of the first formant. (Delattre (1954) abstractly suggested something similar, p.105) It appears that there are at least two groups of cues sufficient to recognition of vowel nasality under certain conditions, and it should be investigated in what combinations these cues are effective.

## REFERENCES

Abramson, A.S., P.W. Nye, J.B. Henderson, C.W. Marshall. 1981. Vowel height and the perception of consonantal nasality. Journal of the Acoustical Society of America (JASA) 70.329-39.

Ali, L., T. Gallagher, J. Goldstein, R. Daniloff. 1971. Perception of coarticulated nasality. JASA 49.538-40.

Beddor, Patrice S., R. A. Krakow, L. M. Goldstein. 1986. Perceptual constraints and phonological change: A study of nasal vowel height. Haskins Laboratories Status Report on Speech Research SR-85

Channon, R. and L. Shockley. 1986. In Honor of Ilse Lehiste. Dordrecht: Foris.

Bond, Z.S. 1976. Identification of vowels exerpted from neutral and nasal contexts. JASA 59.1229-32.

Delattre, P. 1954. Les attributs acoustiques de la nasalité vocalique et consonantique. Studia Linguistica 8.103-109.

_____, and Monnot, M. 1968. The Role of Duration in the identification of French nasal vowels. International Review of Applied Linguistics 6.267-288

Fant, C.G.M. [1960] Acoustic Theory of Speech Production. The Hague: Mouton

Ferguson, C.A. 1975. Universal Tendencies and 'Normal' Nasality. In Ferguson et al. 175-212.

_____, Hyman, L.M. and Ohala, J.J. 1975. Nasálfest (papers from a symposium on nasals and nasalization) Stanford: Language Universals Project, Stanford University.

Fujimura, O. 1960. Spectra of Nasalized Vowels. Quarterly Progress Report MIT Research Laboratory in Electronics. 58.214-8.

____ 1961. Analysis of Nasalized Vowels. Quarterly Progress Report MIT Res. Lab. in Elec. 62.191-2.

Hattori, S., K. Yamamoto, and O. Fujimura. 1958. Nasalization of vowels in relation to nasals. JASA 30.267-74.

Hawkins, S., and K.N. Stevens. 1985. Acoustic and perceptual correlates of the non-nasal--nasal distinction for vowels. JASA 77.1560-75.

House, A.S., and G. Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. JASA 25.105-13.

Krakow, R.A., P.S. Beddor, L.M. Goldstein, and C.A. Fowler. 1987. Coarticulatory Influences on the Perceived Height of Nasal Vowels. Haskins Laboratories Status Report on Speech Research SR-92 31-53.

Malécot, A. 1960a. Vowel nasality as a distinctive feature in American English. Language 36.222-9.

____ 1960b. Nasal Syllabics in American English. Journal of Speech and Hearing Research 3.268-74.

Ohala, J.J. 1975. Phonetic explanations for nasal sound patterns In Ferguson et al. 289-316.

____ and J.J. Jaeger eds. 1986. Experimental Phonology. New York: Academic

Stevens, K.N., Fant, G. and Hawkins, S. 1986. Some acoustical and perceptual correlates of nasal vowels. In: Channon and Shockley.

Wright, J.T. 1975. Effects of Vowel Nasalization on the Perception of Vowel Height. In Ferguson et al. 373-387.

____ 1986. The behavior of nasalized vowels in the perceptual vowel space. In Ohala and Jaeger. 45-68.