ISCA Archive
http://www.isca-speech.org/archive

3rd International Conference on
Spoken Language Processing (ICSLP 94)
Yokohama, Japan
September 18-22, 1994

ANALYSIS OF PITCH DEPENDENCE OF PHARYNGEAL, FAUCAL, AND LARYNX-HEIGHT VOICE QUALITY SETTINGS

John H. Esling, Lynn Marie Heap, Roy C. Snell, and B. Craig Dickson

Department of Linguistics, University of Victoria,
P.O. Box 3045, Victoria, B.C. V8W 3P4 Canada

Speech Technology Research, Ltd., 1623-B McKenzie Avenue,
Victoria, B.C. V8N 1A6 Canada

ABSTRACT

To examine the possibility that pharyngeal voice quality settings are not independent of pitch differences or of their effect on each other, pharyngealized voice, faucalized voice, raised-larynx voice and lowered-larynx voice were produced under controlled phonetic conditions at eight separate, incremental pitch intervals and analyzed using auditory, spectrographic, and vowel-formant tracking. Results suggest an interdependent relationship between raised-larynx voice and pharyngealization, and between lowered-larynx voice and faucalization. These relationships are realized at several pitch increments, and may reflect some pitch dependence; that is, pharyngealized voice may mask raised-larynx voice at low frequencies, and faucalization may mask lowered-larynx voice auditorily at high frequencies. At the highest frequencies (falsetto range), pharyngealized and faucalized voices, and to some extent lowered-larynx voice, cannot be distinguished. A close front vowel is differentiated the most consistently, while a close back vowel is most often undifferentiated for pharyngeal quality. To explain these relationships, it is posited that for raised-larynx voice at certain (low) frequencies, the vocal tract is configured in a manner that is the same as for pharyngealization; and that for faucalization at certain (low) frequencies, the vocal tract is configured in a manner that is the same as for lowered-larynx voice.

I. INTRODUCTION

It has sometimes been implied in describing long-term articulatory configurations, or voice quality settings, that the various postures in the taxonomy are independent of pitch and amplitude variation, and that they contribute independently to the sum total of combined settings that make up the description of a given individual's voice [1]. In fact, it has remained unclear how many of these various settings interact with each other, and whether some of them may only occur under certain pitch or amplitude conditions. This paper investigates some of the least well-defined settings in voice quality description: the larynx-height and pharyngeal settings, that is, raised-larynx voice, lowered-larynx voice, pharyngealization, and faucalization. The aim of the research is to determine whether these settings are independent of pitch changes, and whether they behave independently of each other.

Raised-larynx voice and lowered-larynx voice are described as vertical movements of the larynx which result in raising of all formants and lowering of all formants, respectively [2]. Pharyngealization is taken to be the retraction of the tongue and the constriction of the muscles of the pharynx walls at the level of the epiglottis [2, 3], causing a rise in F1 and a fall in F2. Faucalization is only tentatively described physiologically as an approximation of the faucal pillars (palatoglossal and palatopharyngeal muscles), narrowing the upper pharynx, raising F1 and lowering F2 [2, 4]. It is the interaction between the musculature for raising or lowering the larynx, constricting the pharynx, and raising or lowering pitch which prompted the use of a variety of pitch levels in this study, since changes in laryngeal posture to produce different F0s may influence the range of gestures by the supralaryngeal vocal tract.

II. DATA ACQUISITION

The methodology relies on the controlled phonetic production by a trained phonetician of: (1) raised-larynx voice, (2) lowered-larynx voice, (3) pharyngealized voice, (4) faucalized voice, and (5) modal voice as a control. Each setting was produced on three peripheral sustained vowels, [i, ɑ, u], at each of eight pitch increments, recorded on two channels with an AKG C1000 condenser microphone and a laryngograph on a DAT recorder, and captured at 10K samples/second on the CSL 4300B (Computerized Speech Lab) system in anticipation of LPC treatment. The same regular increments along the musical scale were used as in research comparing phonation types [5]. Target frequencies span over two octaves: F (87.31 Hz), A (110 Hz), C# (138.59 Hz), F (174.61 Hz), A (220 Hz), C# (277.18 Hz), F (349.23 Hz), A (440 Hz). First, formant measurements of all tokens were taken from spectrograms. These measurements were then checked against a formant-extraction routine designed within the CSL environment [6]. Using the Sona-Match program [7] to plot and display the captured data, the first two vowel formants were differentiated using cluster analysis. Auditory evaluations of the set of three vowels at each pitch were performed separately, with each token played back in the order it was produced. A second auditory evaluation was performed at a later time on tokens of [ɑ] only, for comparison with the acoustic analysis.

To compute vowel formants, pitch-epoch separations were calculated for each token to allow pitch-synchronous LPC analysis. This step in the procedure also ensures through pitch extraction that the voices were produced at the desired pitch. For vowels produced at frequencies above 277 Hz, not using pitch-synchronous analysis pro-

Produced VQ: at Pitch	Modal	Raised Larynx	Lowered Larynx	Pharyngealized	Faucalized
87 Hz	Modal	RLx/Phar	LLx	Phar	LLx/Phar
110 Hz	Modal	RLx+Phar	LLx+Fauc	Phar	LLx/Fauc
138 Hz	Modal	RLx	LLx+Fauc	Phar/RLx	LLx/Fauc
174 Hz	Modal	RLx	Fauc+LLx	Phar/RLx	Fauc
220 Hz	Modal	RLx	Fauc+LLx	RLx/Phar	Fauc
277 Hz	Modal	RLx	LLx/Fauc	RLx/Phar	Fauc
349 Hz	Falsetto	RLx/Fals	LLx/Fals	RLx+Phar	Fauc
440 Hz	Falsetto	RLx/Fals	LLx/Fals/Fauc	Phar/Fauc/ RLx/Fals	Phar/RLx/ Fauc/Fals

Table 1. Auditory assessment of original production of [i, ɑ, u] in sequence with modal, raised-larynx, lowered-larynx, pharyngealized, and faucalized voice quality settings at 8 pitches.

Produced VQ: at Pitch	Modal	Raised Larynx	Lowered Larynx	Pharyngealized	Faucalized
87 Hz	Modal/LLx	RLx	LLx/Phar	LLx/Phar	LLx/Phar
110 Hz	Modal	RLx/Phar	LLx/Fauc	Phar/Fauc	LLx/Fauc
138 Hz	Modal	RLx/Phar	LLx	Phar/LLx	LLx
174 Hz	Modal	RLx	Fauc/LLx	RLx	Fauc/LLx
220 Hz	Modal/Fauc	RLx/Fauc	Fauc/LLx	RLx/Fauc	Fauc/LLx
277 Hz	Modal/Fauc	RLx/Fauc	LLx/Fauc	RLx/Fauc	Fauc/LLx
349 Hz	Falsetto	RLx/Fals	Fauc/Fals	Fauc/Fals	Fauc/Fals
440 Hz	Falsetto	RLx/Fals	Fals/Fauc	Fauc/RLx/Fals	Fauc/Fals

Table 2. Auditory assessment of original production of the vowel [ɑ] alone with modal, raised-larynx, lowered-larynx, pharyngealized, and faucalized voice quality settings at 8 pitches.

duced better results. Spectrograms were obtained with the following parameters: frame length 75 points, pre-emphasis 0.8, and a Blackman window. The centre of each band was measured by eye; then the LPC formant history was overlaid on the spectrogram, and the numerical results of LPC were saved to disk for statistical measurement. LPC formant history was obtained using pitch-synchronous analysis with a 12th-order filter and pre-emphasis set to 0.9. The values from LPC formant history were then used to compute formant means and z-scores. Using an algorithm in the Sona-Match program, F1 and F2 were plotted within 80% confidence ellipses to show trends and to guide statistical comparisons.

III. RESULTS

The results of the auditory assessment of the original tokens are summarized in Table 1. There is an apparent relationship between lowered larynx and faucalization, as both qualities are perceived in the samples of lowered larynx produced above the lowest pitch. Raised larynx appears to have a similar relationship with pharyngealization, as it is difficult to distinguish them above low pitch. It appears that at lower pitches, pharyngealization is the dominant perceptual quality when both raised-larynx and pharyngealized settings are the intended target; and that lowered larynx is the dominant quality at low pitch when either lowered-larynx or faucalized settings are the target.

Table 2 summarizes auditory judgements specifically of the quality of the [ɑ] vowel, which is known to have pharyngeal colouring and is even described as approaching epiglottal place of articulation [3]. For [ɑ], raised larynx and pharyngealization are not heard as primarily pharyngealized at 87 Hz as in Table 1, and all other tokens at 87 Hz are heard as lowered larynx. At higher pitches, the pattern is similar to Table 1 in that faucalization is realized in productions of lowered larynx, and raised larynx is realized in productions of pharyngealization.

Formant analysis helps to clarify the nature of the similarity between raised-larynx and pharyngealized settings, and between lowered-larynx and faucalized settings. Both raised larynx and pharyngealization show an increase in F1 relative to modal voice, with pharyngealization showing the lower F2. Both lowered larynx and faucalization show a

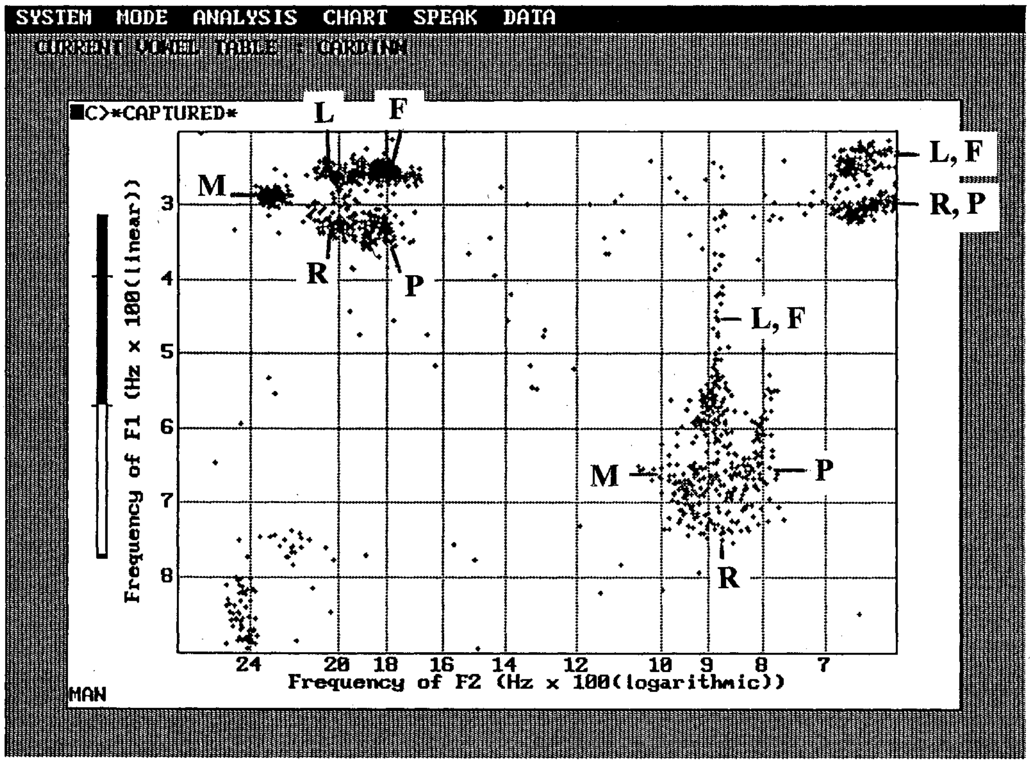

Fig. 1. Sona-Match output showing the groupings of 5 settings on sustained vowels [i, ɑ, u] at 110 Hz.

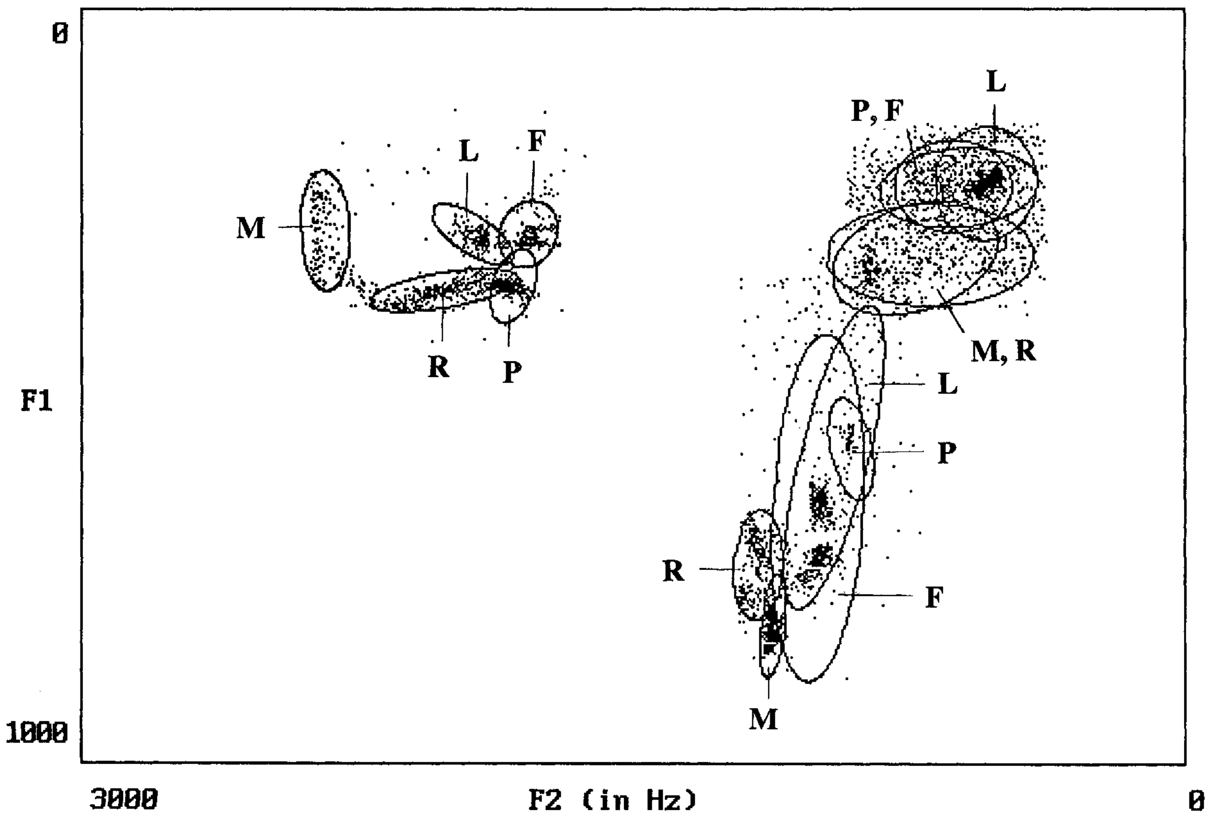

Fig. 2. Clusters (80% confidence ellipses) of 5 settings on sustained vowels [i, ɑ, u] at 174 Hz.

decrease in F1 relative to modal voice, with faucalization showing the lower F2. As illustrated in Figure 1 for vowels at 110 Hz and in Figure 2 for vowels at 174 Hz, this relationship is strongest for [i], varies mainly in the F1 dimension for [ɑ], and often appears masked in the case of [u]. Statistical evaluation (Table 3) confirms that these related pairs show significant (50 Hz) differentiation in F1 of [ɑ] but not in F2. For [i], these pairs show significant (100 Hz) differentiation in F2 at most pitches. Differences from modal voice of the order found for [i] in the illustration are significant, while for [u] they are not. It should be noted that values for [u] are also confounded by the effect of lip rounding.

	Raised Larynx ~				Lowered Larynx ~			
i	Phar		F2	F3	Fauc	F1	F2	F3
ɑ	87	F1		F3	87	F1		F3
u				F3				
i	Phar		F2	F3	Fauc		F2	F3
ɑ	110				110	F1		
u								
i	Phar		F2	F3	Fauc			F3
ɑ	138				138	F1	F2	F3
u			F2	F3				F3
i	Phar		F2	F3	Fauc			
ɑ	174		F2		174	F1		F3
u			F2	F3				
i	Phar		F2		Fauc	F1	F2	F3
ɑ	220	F1			220			F3
u								F3
i	Phar				Fauc		F2	F3
ɑ	277	F1	F2		277	F1		F3
u		F1	F2	F3			F2	
i	Phar			F3	Fauc		F2	
ɑ	349			F3	349	F1		F3
u				F3				
i	Phar		F2	F3	Fauc		F2	
ɑ	440				440		F2	F3
u				F3			F2	

Table 3. Statistical differentiation of vowel formant means; Fn = significant at 50 Hz for F1, 100 Hz for F2 and F3.

IV. DISCUSSION

The findings support the presence of pharyngealization in the raised-larynx setting, and suggest that it also contributes to the lowered-larynx and faucalized settings. Laver and Nolan [2, 8] find that synthesizing raised larynx without raising F0 results in a pharyngeal quality, which suggests that pharyngealization and raised larynx share the same articulatory configuration if F0 is not taken into account. Nevertheless, the acoustic evidence examined here compels us to consider that four distinct qualities are distinguishable, since their clusters are differentiated at a number of pitches for more than one vowel.

Pharyngealization and faucalization are observed to behave in a similar way: in the auditory impressions they convey, in the way they are interpreted at different pitches, and in the way their formants cluster, particularly F2. Taking into account (1) the acoustic predictions for pharyngealization and faucalization, (2) evidence about their production from previous experiments [2, 8], (3) their complementary distribution across the pitch range (pharyngealization -- low pitch; faucalization -- high pitch), and (4) their reduced F2 relative to raised larynx and lowered larynx, it is posited that for raised-larynx voice at certain (low) frequencies, the vocal tract is configured in a manner that is the same as for pharyngealization; and that for faucalization at certain (low) frequencies, the vocal tract is configured in a manner that is the same as for lowered-larynx voice. The acoustic evidence suggests that pharyngealization and faucalization are the result of pharyngeal constriction, so it is also posited that pharyngealization is the constriction which accompanies raised larynx from the lowest pitch, and that faucalization is the constriction that accompanies lowered larynx at higher pitches. Thus, for raised larynx, with the pharynx constricted and as pitch increases to mid range, raising the larynx begins to raise F2 until the distinction is lost again at high pitch; while for lowered larynx, constricting the pharynx as pitch increases produces the effect of faucalization in the mid-pitch range. So with the larynx lowered, at low pitch, the addition of pharyngeal constriction is perceived as pharyngealization; while at higher pitch it is perceived as faucalization.

At the very least, the taxonomy of voice quality assessment should not be taken to imply the independence of pharyngeal, faucal and larynx-height settings. It is doubtful that what are identified auditorily as raised-larynx or faucalized settings are distinctive at every pitch on all vowels. Even where they occur on an open vowel, they may be realized as pharyngealization or lowered larynx, respectively, on a close vowel. The role of F3 and F4 also needs to be fully considered when evaluating pharyngeal quality. To further investigate the possible physiological relationships, existing data for the five settings will be compared with visual laryngoscopic observations.

REFERENCES

[1] J. H. Esling. "Voice Quality." In R. E. Asher et al. (eds.), *The Encyclopedia of Language and Linguistics*, pp. 4950-4953. Oxford: Pergamon Press. 1994.

[2] J. Laver. *The Phonetic Description of Voice Quality.* Cambridge: Cambridge University Press. 1980.

[3] A. Laufer, and I. D. Condax. "The Function of the Epiglottis in Speech." *Language and Speech*, vol. 24, pp. 39-62. 1981.

[4] G. Fant. *Acoustic Theory of Speech Production.* The Hague: Mouton. 1960.

[5] J. H. Esling, B. C. Dickson, and R. C. Snell, "Analysis of Phonation Type Using Laryngographic Techniques." *ICSLP 92 Proceedings*, pp. 1107-1110. Edmonton: University of Alberta. 1992.

[6] B. C. Dickson, et al. *CSL User's Manual.* Pine Brook, NJ: Kay Elemetrics Corporation. 1990.

[7] J. Clayards, P. Ollek, and J. R. Woolsey. *Preliminary Documentation for the Sona-Match Program.* Victoria: STR Ltd. 1992.

[8] F. Nolan. *The Phonetic Bases of Speaker Recognition.* Cambridge: Cambridge University Press. 1983.