VOCAL TRACT VOCALIC NOMOGRAMS : ACOUSTIC CONSIDERATIONS.

A Crucial Problem : Formant Convergence.

Pierre BADIN   Louis-Jean BOE

Institut de la Communication Parlée (U.A. CNRS 368)

Laboratoire de la Communication Parlée
E.N.S.E.R.G. - I.N.P.G.
46, av. Félix Viallet - 38031 GRENOBLE Cédex

Institut de Phonétique de Grenoble
Université des Langues et des Lettres.
B.P. 25 - 38040 GRENOBLE Cédex, FRANCE

ABSTRACT

Presented by FANT in 1960, nomograms have not been thoroughly exploited for studying vocalic productions. It took a long time before we noticed the utilisation of this kind of tool with STEVENS' quantal theory (/7/) and, more recently, with the attempts by LADEFOGED and BLADON to reproduce FANT's nomograms (/5/). In spite of its simplicity, the four-tube model can be used to interpret the articulatori-acoustic relations (formants-cavities affiliations) and to describe the main vocalic types.

We first make a revisit of FANT's explanations for the "affiliations" between formants and cavities, and we try to refine his findings, in relation with losses in the vocal tract (especially at the glottis). We study the phenomenon of formant convergence ("focalization"), and more precisely in the case of |i|. Finally, we compare the results of our simulations with natural speech.

INTRODUCTION

Presented by FANT in 1960, nomograms have not been thoroughly exploited for studying vocalic productions. It took a long time before we noticed the utilisation of this kind of tool with STEVENS' quantal theory (/7/) and, more recently, with the attempts by LADEFOGED and BLADON to reproduce FANT's nomograms (/5/). We think that nomograms are still very powerful tools in the field of articulatori-acoustic relations for vowel production, for vocalic system interpretation and prediction, as well as for formant measurement up to F5 (e.g. F'2 estimation).

In the first section, we give a brief description of FANT's nomograms, and we recall FANT's explanations about the "affiliation" phenomenon. Then, we study the effect of losses, especially at the convergence point between F2 and F3, in the case of an intermediate lip opening.

1. FANT'S VOCALIC NOMOGRAMS : A REVISIT

1.1 The Four Tubes Model : Basic Resonances

In order to mimic in a simple way the acoustical behavior of the vocal tract when the constriction is moving along the midline from glottis to lips, FANT defined a four-tube model (/4/, p.71-79). Here, we retain the following configurations for the 4 sections :

Pharynx cavity : $A_4 = 8.\text{cm}^2$, L_4 varying ;
Tongue constriction : $A_3 = 0.65\text{cm}^2$, $L_3 = 5.\text{cm}$;
Mouth cavity : $A_2 = 8.\text{cm}^2$, L_2 varying ;
Lips : $A_1=0.16$, $4.\text{cm}^2$ or no lips, $L_1=1.\text{cm}$.

The constriction center coordinate X_c (measured from the glottis position) can vary from -2.5cm to 17.5cm, keeping : $L_2 + L_3 + L_4 = 15\text{cm}$.

In order to understand the affiliation phenomenon, we have made nomograms for the different cavities alone. We have supposed very low heat and viscosity losses, no wall vibration, no lip radiation load, and we have used the acoustic model developped by BADIN & FANT (1984).

Fig.1a shows the resonances of the back cavity alone : integer multiples of the half wavelength resonance $c/2L_4$ (solid lines). The dashed line at the bottom corresponds to a "HELMHOLTZ resonance" between the back cavity and the constriction neck.

Fig.1b shows the resonances of the "mouth + lips" front cavity alone. In the case of no lip section, the front cavity produces resonances at odd integer multiples of the quarter wavelength resonance $c/4L_2$ (dotted lines). In the case of a very small lip opening ($A_1 = 0.16\text{cm}^2$), we obtain the same behavior as for the back cavity : resonances at integer multiples of the half wavelength resonance $c/2L_2$ (upper dashed lines), plus a "HELMHOLTZ resonance" between the mouth cavity and the lip constriction (bottom dashed line). For the intermediate case ($A_1 = 4.\text{cm}^2$), the resonances have an intermediate behavior (solid lines).

Fig.1c shows the behavior of the constriction resonances : when the constriction tube is open at both ends, it produces the half wavelength resonance $c/2L_3$ (middle and right part of the figure), whereas when it is closed at the glottis, it produces resonances at odd integer multiples of the quarter wavelength resonance $c/4L_3$ (left part).

1.2 Affiliation and Coupling : focal points

Fig.2 (a, b, c) represent, for the same simplified boundary conditions as in 1.1, the resonances of the complete four-tube system (solid lines) and the contributions from the different cavities (dashed lines) deduced from Fig.1. This nomogram demonstrates clearly the affiliation phenomenon : whenever a solid line is very close to a dashed line the resonance of the whole system depends mainly on the cavity corresponding to the dashed line ; on the opposite, when the solid line departs from the dashed line, there is a little affiliation and coupling phenomena occur.

When two dashed lines cross each other, i.e. a resonance associated with the front cavity and a resonance associated with the back cavity have close frequencies for a given position of the tongue constriction, we observe a focalization of the formants : we call **"focal point"** (/2/) this convergence region where the affiliation of the lower and upper resonances switches from one cavity to the other when the tongue constriction is shifted.

Coupling between two resonant systems is a classical phenomenon : it always spreads apart the natural frequencies of the two systems. More precisely, the greater the coupling, the larger the frequency spreading, and conversely, the smaller the coupling (i.e. the constriction cross area), the more prominent the formant convergence. The case of a small lip opening (|u|, Fig.2c) shows a good example of large coupling : the HELMHOLTZ resonances associated with the "back cavity + tongue constriction" resonator and with the "mouth cavity + lip constriction" are spread into F1 and F2 when coupled through the constriction tube. The case of a middle lip opening (|i|, Fig.2b) shows a good example of F2/F3 convergence : on the glottis side of the convergence point, F2 is clearly associated with the front cavity, and F3 with the back cavity, whereas it is the opposite on the lip side. We call an |i| configuration close to the convergence point "focal" |i|, and configurations on the glottis side and on the lip side respectively "prefocal" and "postfocal" |i|'s.

By generalization of the above examples, it is possible to define other formant convergences and other vocal types. The vowel |a| with open lips shows a F1/F2 convergence between the HELMHOLTZ back cavity resonance and the front cavity first resonance (Fig.2b). The vowel |y| has a F2/F3 convergence corresponding to the HELMHOLTZ "mouth cavity + lips" resonance and the back cavity half wavelength resonance in the case of constricted lips (Fig.2c).

The constriction presents an interesting behavior in the central region of the nomogram corresponding to |u| (Fig.2c). By chance (because L2 is exactly one third of the total length of the configuration without lips), there is a point where three dashed lines cross each other : the half wavelength resonances of the front and back cavities, and of the constriction itself. The constriction resonance is not modified, and the resonances of the front and back cavities are spread apart. This focal point is one of the two convergence points for |u|, but we should mention that, owing to the lip constriction, the amplitude for F3, F4 and F5 is rather low. For more open lips, the situation is different : this triple point does not exist, and the resonance frequency of the constriction is much modified by the coupling. We can induce that the formants F3 and F4 must be rather sensitive to the location and to the size of the constriction.

1.3 Losses effects

To have a better insight into the coupling phenomena we have neglected in section 1.2 the losses and the boundary effects such as wall vibration or lip radiation. We now include these effects in the simulation. The contribution of different types of losses to formants and bandwidths has already been discussed somewhere else (/4/, /1/). We just recall that most of the losses are due to the radiation at the lips and to the glottis resistance. The effect of the losses due to the lip radiation or to the glottis is to decrease the amplitudes and to broaden the bandwidths of the associated resonances. The losses at the lips increase with lip opening, and the losses at the glottis increase with glottis opening.

Because of affiliation, the lip and glottis opening conditions have a selective influence upon the resonances of the associated cavities. It is interesting to analyze the nomogram in a region of formant convergence. The selective effects of the glottis losses are depicted on Fig.3 : a 3-D representation of the nomogram, for a middle lip opening ($A_1=4.\text{cm}^2$), is given for a "small" glottis opening (a) (glottis area $A_g=10.\text{mm}^2$, resistance $R_g=43\Omega$) and for an "moderate" glottis opening (b) ($A_g=4.6\text{mm}^2$, $R_g=100\Omega$) . We clearly see that when the glottis opens, the amplitude of the formant related to the back cavity decreases and the bandwidth increases, due to the increase of glottis losses. The same phenomenon happens for a variation of the lip opening (which would be associated with a displacement of the convergence point, because of the correlated variation of the length end correction at the mouth, /4/, p.36).

A more detailed analysis of Fig.3 leads to a notion of **"bandwidth inversion"**, especially clear for the small glottis opening : on the glottis side of the focal point, the bandwidth of F2 (which is associated with the front cavity) is greater than that of F3, whereas on the lip side the bandwidth of F3 (which is then associated with the front cavity) is greater : this inversion is the consequence of affiliation. When the glottis opening increases this effect decreases, but for a moderate glottis opening, the inversion effect occurs again, the role of the two cavities being then exchanged (wider bandwidths for the back cavity resonance, Fig.3b).

2. EXPERIMENTAL ILLUSTRATION

In this section, we compare some results of the above simulation study with equivalent situations for real speech.

2.1 F2/F3 convergence for |i|

A recent and interesting attempt to reproduce FANT's nomograms has been the one by LADEFOGED & BLADON (/5/). They present a series of sonagrams corresponding to sounds for which the tongue constriction is progressively shifted from the pharynx to the teeth, every other parameter being supposed constant : they noticed regular shifts of the formant frequencies corresponding to what was expected from FANT's nomograms, but they mentioned that, for a vowel in one of their series, "F3 would appear to have suddenly assumed a value comparable

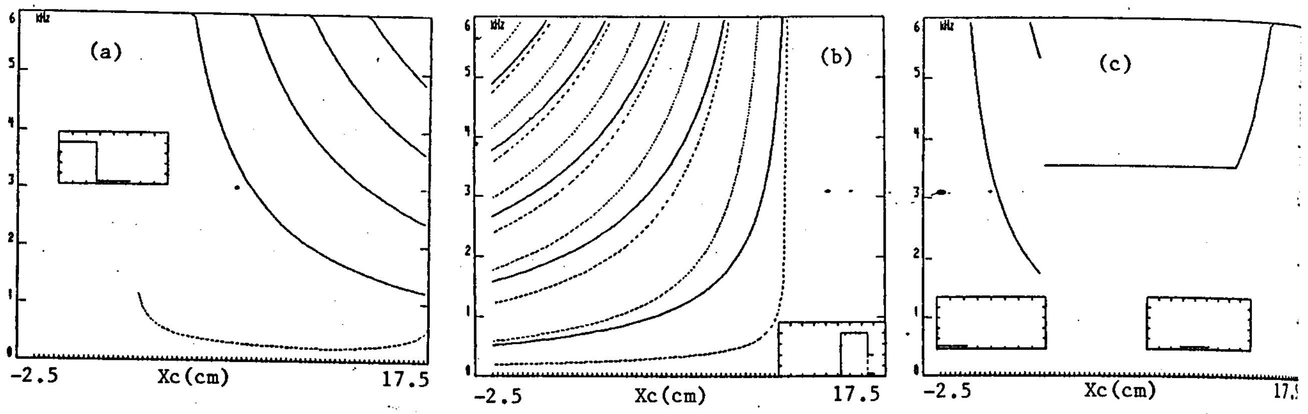

Fig.1 : Resonances of the different cavities vs. tongue constriction location.
(a) back cavity : pharynx + tongue constriction (b) front cavity : mouth + lip constr. (c) tongue constr.

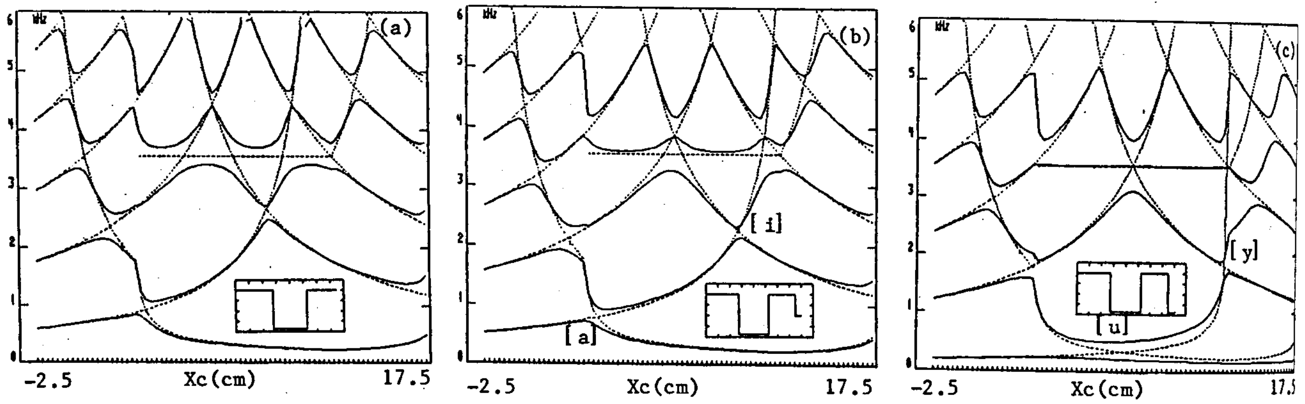

Fig.2 : Resonances of the whole four tubes system vs. constriction location.
(a) maximum lip opening (L_1=0.) ; (b) middle lip opening (A_1=4.cm^2) ; (c) small lip opening (A_1=0.16cm^2).

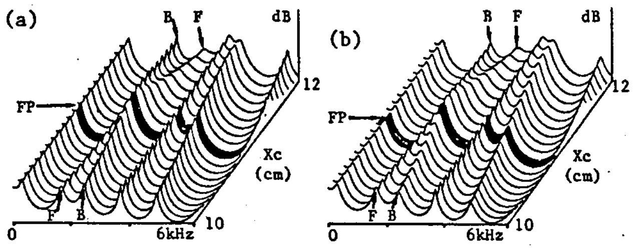

Fig.3 : 3-D Nomogram (transfer functions) around the F2/F3 convergence for |i| (A_1=4.cm^2).
(a) small glottis opening (b) moderate glottis opening.
F : front cavity B : back cavity FP : focal point.

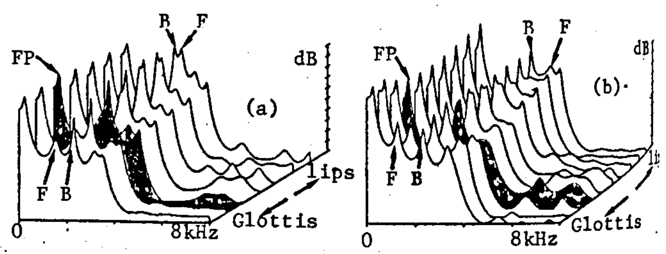

Fig.5 : 3-D spectral representation of focal |i|s
(a) in |tiwit| (b) in |ziwiz|.

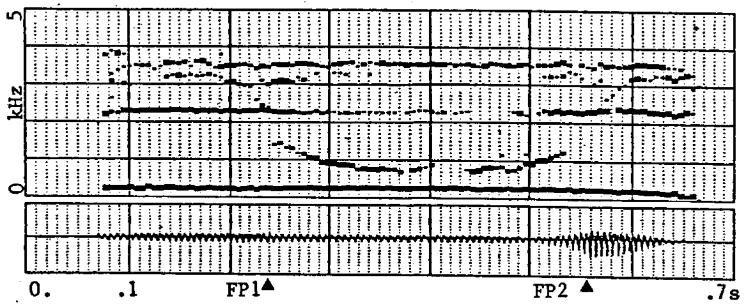

Fig.4 : Formants tracked for the natural sound [iwit].

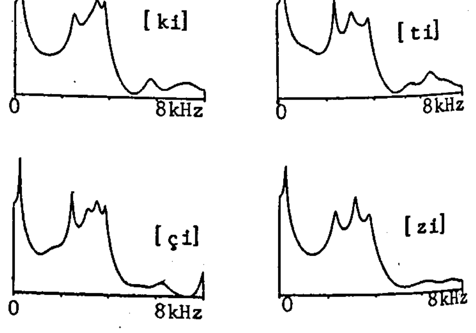

Fig.6 : spectra of |i| in different contexts.

to that of F4 in the previous vowel". By reference to the above simulations, we know that F2 and F3 can merge into a focal point for a given position of the tongue constriction : we believe that an answer to LADEFOGED & BLADON's difficulties is that F2 and F3 are actually merged into a single formant.

To check this hypothesis, we have recorded a series of |CiwiC| sounds, where the transitions |iw| and |wi| correspond roughly to a shift of the tongue constriction, every other parameter being approximately constant (except for lip opening). If C is a dental or post-alveolar consonant, we insure that the tongue constriction shifts from a postfocal |i| (coarticulated with C) to a prefocal |i| (coarticulated with |w|), and thus that the focal point will be gone through. Fig.4 shows the evolution of the formants (tracked from cepstrum) for the sequence |tiwit| : we can easily follow the front cavity resonance going from F3 to F2 and back to F3 through the focal points FP1 and FP2 ; it might be possible to track this resonance even until F4. The spectral representation (obtained by LPC analysis), Fig.5 (a, b), shows a rather striking analogy with the transfer functions from Fig.3 for the behavior of F2 and F3 around the convergence point. This shows that the focal points predicted by our simulations are observable in real speech. It also reconfirms our view on LADEFOGED & BLADON's problem, and provides an explanation for the CHAFCOULOFF & al. (/3/) observations.

This convergence phenomenon may explain a part of the difficulties encountered by phoneticians in measuring F2 and F3 for |i| vowels, and the large dispersion of their data.

2.2 Bandwidth inversion around the F2/F3 focal point

The purpose of this section is to verify on real speech the effect of "bandwidth inversion". Thus we have recorded four sounds with |i| in 4 consonantal contexts, |k|, |ç|, |t| and |z|, corresponding to two articulatory locations (palatal vs. alveolar) and two glottis openings (small vs. large).

According to the transfer functions shown in Fig.3, we could expect the following relations for the bandwidths B2 and B3, and for the amplitudes A2 and A3 of F2 and F3 :

C	glottis opening	articul. location	Bandwidths relation	Amplitudes relation
|k|	small	prefocal	B2 > B3	A2 < A3,
|ç|	large	prefocal	B2 < B3	A2 > A3,
|t|	small	postfocal	B2 > B3	A2 < A3,
|z|	large	postfocal	B2 < B3	A2 > A3.

Fig.6 shows that the spectra of the |i| sounds in the four different coarticulation contexts checks with our expectations. This reconfirms the bandwidth inversion phenomenon around the focal point, and the influence of the glottis losses upon the relative bandwidth values of the formant associated with the back cavity compared to the the one associated with the front cavity.

CONCLUSION

The nomograms have allowed us to interpret the relations between formants and cavities, and to study the influence of the losses. We have illustrated these results with natural speech in a qualitative way, for a focal |i|. In order to obtain quantitative predictions closer to reality, we need to use a more realistic articulatory model : a study is in progress with MAEDA's articulatory model (/6/).

ACKNOWLEDGEMENTS

We are very thankful to C. ABRY for his suggestions at the phonetic level, and to M. CONTINI and C. SCHIRRU as "good" speakers.

REFERENCES

/1/ **BADIN P. & FANT G. (1984)**, "Notes on Vocal Tract Computation", STL-QPSR 2-3/1984, 53-108.
/2/ **BOE L.J. & ABRY C. (1986)**, "Nomogrammes et Systèmes Vocaliques", 15$^{\text{èmes}}$ JEP GALF, 303-306.
/3/ **CHAFCOULOFF M. CHOLLET G. DURAND P. GUIZOL J. & RODET X. (1980)**, "Observation and Modelling of 'Formant' Transitions using ISAAS", IEEE Int. Conf. ASSP, 146-149.
/4/ **FANT G. (1960)**, "Acoustic Theory of Speech Production", Mouton, ('S-Graven Hague).
/5/ **LADEFOGED P. & BLADON A.(1982)**, "Attempts by Human Speakers to Reproduce FANT's Nomograms", Speech Comm. 1, 185-198.
/6/ **MAEDA S. (1979)**, "An Articulatory Model of the Tongue Based on a Statistical Analysis", J. Acoust. Soc. Am. 65, S1, S22 (A).
/7/ **STEVENS K.N. (1972)**, "The Quantal Nature of Speech : Evidence form Articulatory-Acoustic Data", in "Human Communication : a Unified View", 51-66, Ed. By E. DAVID & P. DENES, Mac Graw Hill, New York.