REVIEW OF FRENCH WORK ON VOCAL SOURCE - VOCAL TRACT INTERACTION

R. CARRE

Laboratoire de la Communication Parlée, ICP Unité Associée au CNRS,
INPG-ENSERG, 46 Avenue Félix Viallet, 38031 GRENOBLE CEDEX

ABSTRACT

The French research work on vocal source - vocal tract interaction was mainly developped in Grenoble and began as early as 1974 with various studies on vocal tract modeling and on source modeling. We have studied both detailed and simplified models of source and tract in order to assess the interaction effects. Modeling of the subglottal cavities has been also carried out. Today, various spectral analysis allow us to complete this approach, to have new data and to formulate new interpretations.

MODELING OF SOURCE - TRACT COUPLING

Vocal tract impedance loading the vocal source.

For a long time, the vocal tract impedance was considered to be low compared with the source impedance. Thus, time-varying vocal tract was considered having no influence on the source.

In reality, several studies do not support these simplification (FANT - 1960, HOLMES - 1976). The coupling phenomenon can be studied by computing the vocal tract input impedance seen from the vocal source. With a model of the vocal tract including all the different losses (wall vibrations, viscosity, heat and radiation), we may suppose that this model gives a good approximation of the vocal tract impedance. The coupling effect must be maximum for high values of the vocal tract impedance. This impedance was calculated for the first three formants and for different French vowels (MRAYATI, GUERIN, BOE - 1976). The resistive part of the input impedance of the vocal tract is given in table 1 for the first two formants.

Judging from the resistive part of the input impedance at resonance, the coupling would, at first, affect the first formant of /a/, /ɔ/, /ɑ/, and the second formant of /ɑ/, /i/.

Modeling is a useful tool for studying the coupling phenomenon ; we therefore developed a simplified model of the vocal tract load impedance, suited for simulations in the time domain of the glottal flow shape. Foster type circuits were introduced, each representing the properties of a formant (GUERIN, MRAYATI, CARRE - 1976). Such simulations with one or two formant circuits have generally been used in these studies.

A simplified representation of the vocal tract impedance is given in figure 1. The elements (L1, R1, C1, L2, R2, C2) of the two Foster circuits are related with the formant frequencies F1 and F2.

This impedance was coupled with a two mass model of the vocal cords (ISHIZAKA, FLANAGAN - 1972), and the effect of the coupling was studied on different parameters :

- the fundamental frequency of the vowel (GUERIN, DELOS, MRAYATI - 1976, GUERIN, BOE - 1980) : the coupling cannot take into account the intrinsic pitch of the vowels, its effect is opposite ;
- the intensity of the vowels (GUERIN, BOE - 1977) : the coupling effect is negligeable ;
- the formant frequencies and the bandwidths (CARRE, GUERIN - 1980) : when coupling is strong, the formant frequency increases due to the glottal inductance effect ; on the other hand, the bandwidth increases due decreasing source resistance.

The fundamental frequency variations are given in figure 2 for different vowels.

Modeling the subglottal cavities (AL ANSARI, GUERIN - 1981) has small effects on the fundamental frequency, but greater effects on the dissymmetry factor and the opening factor of the glottal flow shape.

Finally, the dynamic impedance of the two mass model was measured (CARRE, GUERIN - 1980) : the impedance value corresponding to an average glottal area is not representative of the dynamic effect. The dynamic impedance of the source was found to be small enough to permit a strong coupling effect in the case of high load impedance. This effect was noticed on formant frequencies.

Figure 3 shows the evolution of the harmonic frequency number 6 of the source spectrum for a F0 increasing in relation with the variation of the vocal cords tension in two situations : two mass model (1) short circuited, (2) loaded by the vocal tract. In the case of the situation 1, harmonic amplitude evolution is linear. In the situation 2, the harmonic number 6 is attenuated at the place of the formant. This attenuation can be used to calculate the impedance of the source for this frequency value. The formant frequency increase, due to the coupling effect, can also be measured.

Moreover, a synchronisation effect on the highest harmonic around the formant can be observed. The two maxima around the formant correspond to a negative resistance and thus to an amplification. Is it a means for adjusting the formant on the most important harmonic ?

A simulation of a two beam model of the source does not improve very much the coupling modeling (PERRIER, GUERIN, AULOGE - 1981).

As a main result, the two mass model is always the best simple representation of the vocal cord behavior.

STUDY OF THE GLOTTAL FLOW SHAPE WITH COUPLING, MALE AND FEMALE VOICE

By inverse filtering, the opening factor of the glottal flow was studied and was considered stable whatever the coupling may be (CHENG, GUERIN - 1985). On the other hand, the dissymmetry factor is a decreasing non linear function of the first formant frequency.

A strategy was developped to calculate the fundamental frequency, the asymmetry factor and the opening factor of the glottal flow, from the coupling, for male and female voices. Three independant parameters are used (subglottic pressure, vocal cord tension and first formant frequency). Three dependant parameters are then obtained (fundamental frequency, opening factor and asymetry factor). High quality synthesis was obtained by this means.

STUDY OF THE VOWEL SPECTRUM FOR INCREASING F0

Stable vocal tract shape

The conditions of the coupling simulation for increasing F0 (two mass model and vocal tract loading) was reproduced for natural vowel production (CARRE - 1981a, 1981b). The vocal tract was kept as stable as possible by the speaker during the F0 variations. The stability was controled by closed glottis analysis. As in the simulation, we observed an increase of the formant frequencies and bandwidths for a strong coupling when comparing the transfer function of the vocal tract obtained for closed glottis conditions and the envelope of the harmonic components. Moreover, in some cases, the harmonic evolutions do not exactly follow the formant envelope when coupling is important. For example, for the vowel /a/, male voice, a deep of about 10 dB appears near the first formant for the harmonic number 6 (figure 4). Does this correspond to the synchronisation effect observed in the simulation ? The formant frequency could be synchronized on the highest harmonic. Formant measurements by autocorrelation LPC techniques have shown such an effect. It could be also due to the analysis method.

Natural vocal tract shape

When the larynx is moving freely during the F0 increase, the larynx generally rises up and an increase of the first formant frequency is observed when the source tract coupling is known to be important. In this case, the male vocal tract changes into a female tract size by reduction of the length on the source side (CARRE, LANCIA, WAJSKOP - 1968). A clear correlation between the impedance of the vocal tract load and the male/female formant frequency factors exists.

CONSEQUENCES IN PRODUCTION, ANALYSIS AND SYNTHESIS OF VOWELS

The results reported by SUNBERG (1982) on singing show that, for a better energy transmission, some adjustments (larynx displacements for an important source tract coupling, modifications of the tract shape...) are carried out to equate the formant and the biggest harmonic. This is specially true for female voice, for high F0. CHENG (1986) obtained the same results for synthetic vowels in the case of high fundamental frequencies : listeners prefer high F0 synthetic vowels for strongest harmonic closed to formant. With all the possibilities of adjustment, the synchronization effect described above could be added.

The formant analysis of the vowels produced by female speakers could have to be weighted by the fundamental frequency measurement. Such an interaction has to be taken into account for speech recognition. In the case of speech synthesis, the acoustic and physiological coupling effects cannot be ignored for a better synthesis quality. CHENG (1986) used special formant circuits, the damping of which being controled by the shape of the source signal. With such circuits, the synthesis quality is improved.

Diphone synthesis which is the result of sound concatenation obtained for a specific F0, is not well adapted to take into account the modifications due to various coupling.

CONCLUSIONS

We have now a better knowledge about the source-tract coupling : variations of vowel characteristics have been studied (intrinsic pitch and intensity, formant and bandwidth variations) and the two mass model is always the best simple representation of the vocal tract behavior.

Coupling characteristics (due to F0 variations) have to be taken into account for analysis and recognition. The perceptual importance of coupling characteristics have also to be tested for synthesis.

BIBLIOGRAPHY

AL ANSARI A., GUERIN B. (1981)
Effet du couplage source - conduit vocal sur les caractéristiques de l'onde de débit.
Actes des 12èmes Journées d'Etude sur la Parole, MONTREAL.

CARRE R. (1980)
Modèles auditifs et modèles de production de parole.
10th International Congress on Acoustics, SYDNEY.

CARRE R. (1981)
Vocal source - vocal tract coupling. Effects on the vowel spectrum.
Proceeding of the IV th. FASE Symposium, VENEZIA.

CARRE R. (1981)
Couplage conduit vocal - source vocale. Conséquences en analyse de parole.
Actes des 12èmes Journées d'Etude sur la Parole, MONTREAL.

CHENG Y.M., GUERIN B. (1985)
Introduction du couplage source - conduit vocal dans un synthétiseur à formants.
Actes des 14èmes Journées d'Etude sur la Parole, PARIS.

CHENG Y.M., GUERIN B. (1985)
Control Parameters in male and female glottal sources.
Fourth International Vocal Fold Physiology Conference, NEW HAVEN.

CHENG Y.M. (1986)
Etude du concept source - filtre interactif pour la synthèse de la parole. Analyse des voyelles nasales.
Thèse de l'Institut National Polytechnique, GRENOBLE.

DELOS M., GUERIN B., MRAYATI M., CARRE R. (1976)
Study of intrinsic pitch of vowels.
91th ASA Meeting, WASHINGTON.

FANT G. (1960)
Acoustic theory of speech production.
Mouton. THE HAGUE.

GUERIN B., BOE L.J. (1977)
Etude des caractéristiques acoustiques intrinsèques des voyelles françaises.
9th International Congress on Acoustics, MADRID.

HOLMES J.N. (1976)
Formant excitation before and after glottal source.
IEEE Conf. on Acoustics, Speech and Signal Processing, PHILADELPHIA.

ISHISAKA K., FLANAGAN J.L. (1972)
Synthesis of voiced sounds from a two - mass model of the vocal cords.
Bell System Tech. Journal, 51, 1233-1268.

MRAYATI M., GUERIN B., BOE L.J. (1976)
Etude de l'impédance d'entrée du conduit vocal. Couplage source - conduit vocal.
Acustica, 35, 330-340.

PERRIER P., GUERIN B., AULOGE J.Y. (1982)
Simulation d'un modèle continu de la source vocale.
Actes du Symposium FASE/DAGA, GOTTINGEN.

Vowels	/u/	/o/	/ɔ/	/ɑ/	/a/	/ɛ/	/e/	/i/	/y/
R-F1(Ohms)	50	128	275	330	265	102	95	38	50
R-F2	175	37	51	350	125	107	93	348	62

TABLE 1. Input impedance of the vocal tract at F1 and F2, for 9 different French vowels.

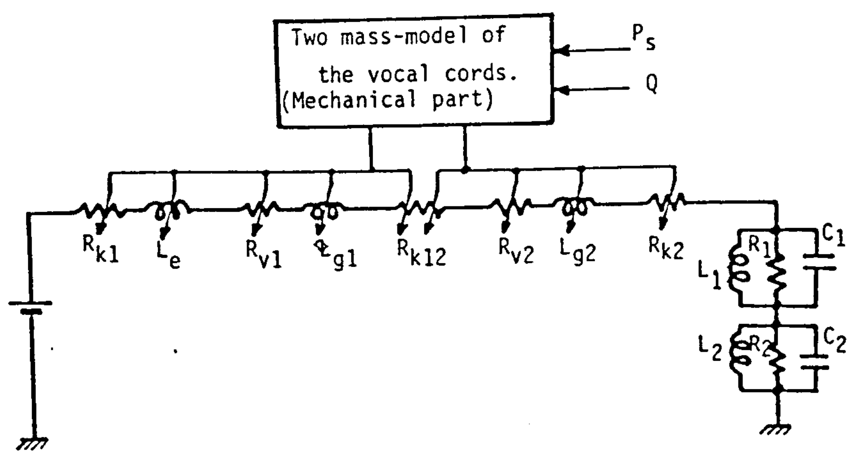

Figure 1. Simulation of the tract load impedance and a two mass model.

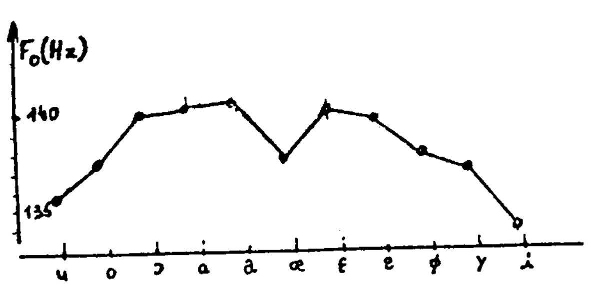

Figure 2. F0 variations for different vowels.

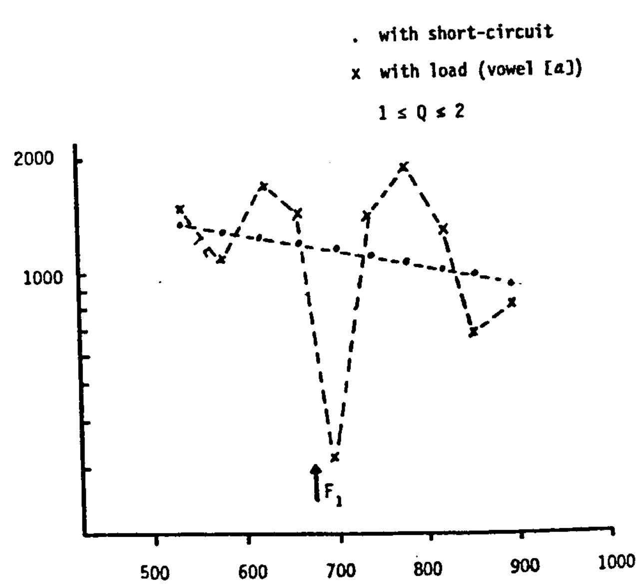

Figure 3. Variations of the harmonic amplitude for a two mass-model : when short circuited, when loaded by a vocal tract (vowel /a/).

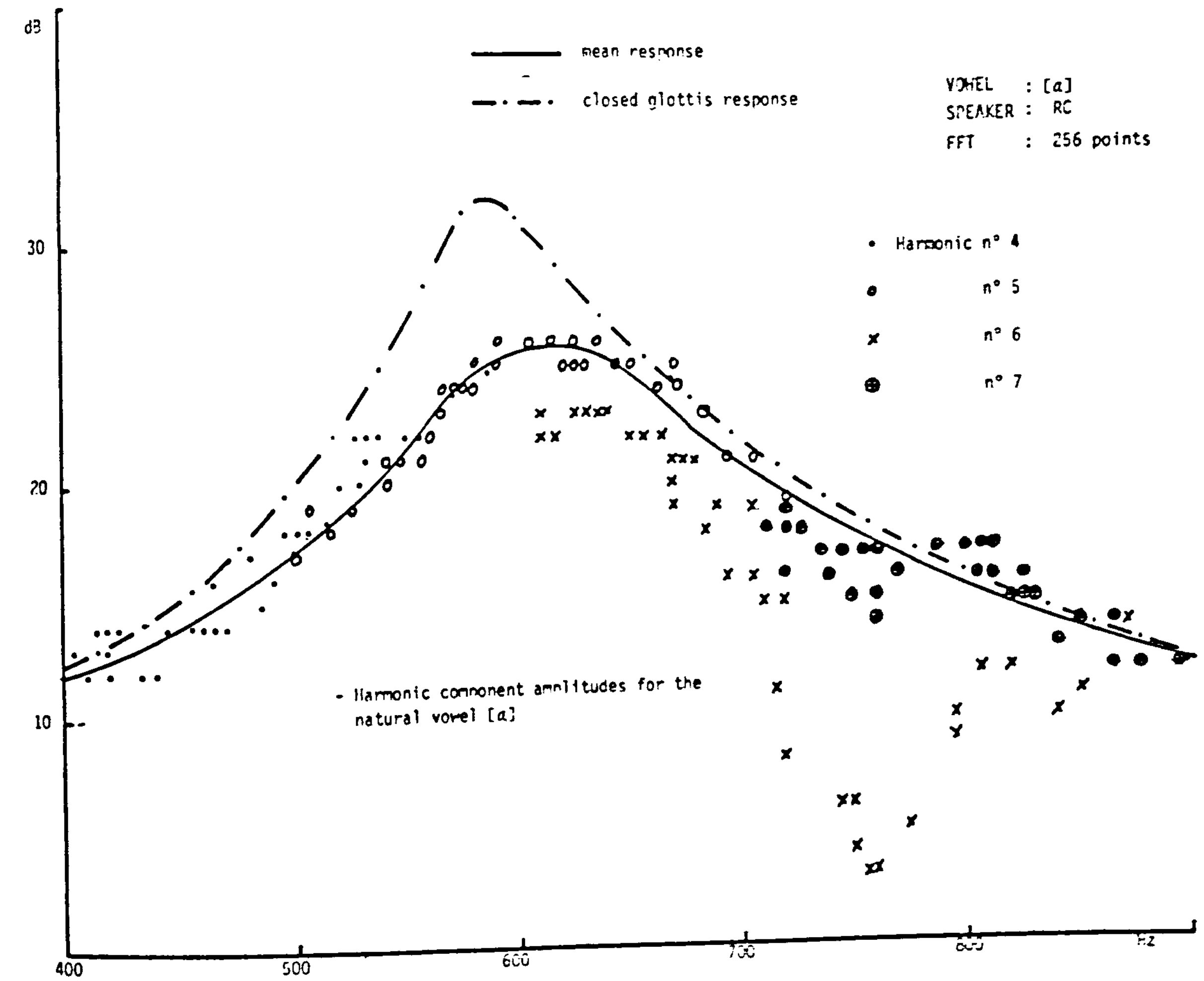

Figure 4. Harmonic component amplitudes for the natural vowel /a/ and F0 increase.