



HAL
open science

Étude acoustique et statistique de l'influence des consonnes sur la justesse des voyelles subséquentes en voix chantée

Nicole Scotto Di Carlo, André Raphaël

► **To cite this version:**

Nicole Scotto Di Carlo, André Raphaël. Étude acoustique et statistique de l'influence des consonnes sur la justesse des voyelles subséquentes en voix chantée. Travaux interdisciplinaires du Laboratoire Parole et Langage, 1977, 4, pp.237-260. hal-00285548

HAL Id: hal-00285548

<https://hal.science/hal-00285548>

Submitted on 5 Jun 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ÉTUDE ACOUSTIQUE ET STATISTIQUE DE L'INFLUENCE DES CONSONNES SUR LA JUSTESSE DES VOYELLES SUBSÉQUENTES EN VOIX CHANTÉE*

Nicole SCOTTO DI CARLO
André RAPHAEL

I. INTRODUCTION

L'influence des consonnes sur la fréquence fondamentale des voyelles adjacentes passionne depuis longtemps les spécialistes de la voix parlée qui ont consacré de nombreuses études à ce problème.

Comme le fait remarquer A. Di Cristo [3] dans sa thèse de doctorat deux faits importants se dégagent de leurs travaux :

- «d'une part, tous les auteurs constatent que la Fo d'une voyelle est influencée globalement par la consonne précédente : les obstruantes non voisées ayant pour effet d'élever sensiblement leur Fo moyenne tandis que les obstruantes voisées produisent l'effet inverse.

- d'autre part, le mode voisé/non voisé d'une consonne prévoicative influe fortement sur la réalisation du contour de la voyelle subséquente, surtout en ce qui concerne la partie initiale de ce dernier :

- On observe en effet, qu'après une obstruante non voisée, la Fo de la voyelle débute à une valeur relativement élevée et décroît ensuite de façon abrupte jusqu'à la valeur-cible de la voyelle où elle tend à se stabiliser.

- Au contraire, après une obstruante voisée, la Fo de la voyelle commence à un niveau assez bas et s'élève progressivement jusqu'à la partie stable de la voyelle ou parfois jusqu'au terme de cette dernière».

* Cet article a fait l'objet d'une communication au II^o Congrès International de Recherches sur le chant, Paris, 11-13 Juillet 1977.

Il nous a paru intéressant de savoir si ces phénomènes microprosodiques qui affectent non seulement l'attaque mais aussi la tenue de la voyelle post-consonantique, étaient identiques dans le chant ou si leurs effets se trouvaient neutralisés par les contraintes de justesse auxquelles sont soumis les chanteurs.

II. PROTOCOLE EXPERIMENTAL

Les seize consonnes du français, présentées dans un contexte CV où V est constante, ont été émises en chambre anéchoïque par un locuteur masculin et un locuteur féminin, dans les registres grave, médium et aigu.

Le locuteur masculin est un sujet doué d'une excellente oreille musicale, capable de chanter parfaitement juste mais sans aucune technique vocale. Pour la commodité de l'exposé, nous l'appellerons LM/VI (Locuteur Masculin, Voix Inculte) Quant à la locutrice, il s'agit d'une chanteuse professionnelle de réputation internationale que l'on désignera désormais par l'abréviation LF/VC (Locuteur Féminin, Voix Cultivée). Pour éviter que d'éventuelles variations dans la hauteur de l'attaque soient imputables à des difficultés d'ajustement, les fréquences représentatives des registres de chaque locuteur* ont été diffusées par casque Socapex à partir d'un générateur de fréquences Brüel & Kjær, sous contrôle d'un fréquencemètre digital Racal SA 520, à une intensité relativement faible pour ne pas gêner les sujets mais néanmoins suffisante pour leur permettre d'atteindre sans tâtonnements la fréquence-cible. Nous avons demandé aux chanteurs d'émettre à trois reprises la série des consonnes françaises suivies de la voyelle [a] sur la note diffusée par le casque, en essayant de conserver la même force d'articulation tout au long de l'expérience.

Les enregistrements ont été réalisés à la vitesse 38 cm/sec sur magnétophone professionnel monopiste Nagra III, équipé d'un microphone électrostatique omnidirectionnel Philips, type LBB 9060, placé à un mètre du locuteur.

Afin d'évaluer la validité du corpus ainsi obtenu, nous l'avons soumis à un double test, concernant d'une part la justesse des sons émis et d'autre part la reconnaissance des phonèmes perçus. Tous

* Les fréquences sélectionnées étaient respectivement dans le grave, le médium et l'aigu : 125, 153, 245 Hz pour le locuteur masculin et 262, 523, 1046 Hz pour le locuteur féminin.

les items étant rigoureusement justes, il nous a suffi d'éliminer les phonèmes dont l'identification prêtait à confusion.

Les éléments retenus ont ensuite fait l'objet d'une analyse spectrographique au Sonograph Kay Electronic Vibralyser 7031A (analyse en filtre étroit de 40 à 4000 Hz et de 80 à 8000 Hz, largeur de bande : 45 Hz).

III. RESULTATS

1. Résultats de l'analyse acoustique :

a) Incidences sur la tenue :

Pour étudier l'influence des consonnes sur la justesse de la voyelle subséquente, on a relevé la valeur moyenne de chaque tenue vocalique afin d'établir la moyenne des fréquences fondamentales des consonnes sourdes et sonores par registres. On a ensuite calculé pour le grave, le médium et l'aigu, la justesse relative des voyelles suivies de consonnes voisées et non voisées, par référence à la fréquence-cible.

Les résultats confirment ceux obtenus pour la voix parlée : dans les trois registres, la valeur moyenne de la fréquence fondamentale est plus élevée après les sourdes qu'après les sonores. (Tableau I)

Par conséquent, dans la voix chantée aussi, les consonnes non voisées ont tendance à élever la hauteur moyenne de la voyelle subséquente et les consonnes voisées à l'abaisser. Néanmoins, ces perturbations qui sont de l'ordre de 5 à 8 Hz dans la parole [4] dépassent rarement 1 Hz dans le chant, et sont de ce fait inférieures au seuil d'intégration auditive.

- On remarque en outre, que, dans les trois registres, mais plus particulièrement dans le grave, ce sont les consonnes sonores qui perturbent davantage la justesse des voyelles subséquentes. A l'intérieur de la catégorie voisé/ non voisé, les différences n'excèdent toutefois jamais 1/16 de ton et ne sont donc pas perçues.

b) Incidences sur l'attaque :

L'étude de l'influence des consonnes sur l'attaque des voyelles subséquentes a été réalisée à partir du calcul de l'intervalle musical compris entre les attaques et les tenues vocaliques. Cet intervalle, exprimé en savarts, est affecté d'un signe positif lorsqu'il s'agit d'une attaque haute ou en-dessus et d'un signe négatif

lorsqu'il s'agit d'une attaque basse ou en-dessous, les attaques au niveau étant évidemment égales à zéro.

Les résultats, schématisés par les figures 1 et 2, font apparaître des perturbations moins importantes pour le LM/VI que pour le LF/VC, et pour les deux locuteurs, des perturbations plus importantes dans l'aigu que dans le grave et le médium. Il faut signaler, en outre, que chez les deux locuteurs, dans les registres grave et médium, les transitions des consonnes dévoisées sont positives (configuration concave) et celles des consonnes voisées négatives (configuration convexe) alors que dans l'aigu les transitions des consonnes voisées et non voisées sont toujours négatives, ce qui signifie que, contrairement à la voix parlée où seules, les consonnes sonores ont un pouvoir aggravant, alors que les sourdes correspondantes ont tendance au contraire à élever la fréquence fondamentale à l'attaque, dans le chant, ceci n'est exact que dans l'infra-grave, le grave, le bas médium et une partie du médium. A partir du haut médium, toutes les consonnes aggravent l'attaque de la voyelle subséquente. L'effet perturbateur des consonnes augmente donc avec la fréquence. C'est ainsi par exemple que dans l'aigu, pour la consonne [d], on enregistre chez la locutrice une attaque basse de 406 savarts, soit plus de huit tons. Autrement dit, il y a un intervalle de dixième entre l'attaque et la tenue de la voyelle. Mais, ainsi que nous le faisons remarquer dans notre communication au Congrès de Phonétique de Leeds [10], la durée de ces attaques n'est généralement pas suffisante pour permettre à l'oreille de les intégrer comme des attaques fausses. Elles sont seulement ressenties d'une façon désagréable par les auditeurs-naïfs. Quant aux oreilles exercées, elles perçoivent en même temps que l'attaque un élément parasite qu'elles sont toutefois incapables d'identifier.

2. Résultats du traitement statistique :

a) Tenue :

Les données relatives à l'influence des consonnes prévocales sur la justesse de la tenue nous ont parues assez cohérentes et explicites pour ne pas nécessiter un traitement statistique. Ce qui n'était pas le cas pour les données concernant l'influence des consonnes sur l'attaque des voyelles subséquentes.

b) Attaques :

Une première analyse de variance réalisée sur ordinateur T 1600 avec le programme NAVA mis au point par J. Pynte, attaché de recherches au C.N.R.S. fait apparaître un F hautement significatif (22,71) pour le facteur "Registres"; les facteurs "Sexe",

"Catégories" ainsi que les autres interactions ne présentant pas un F significatif. (Tableau II). Ces résultats se trouvent confirmés par deux analyses de variance avec décomposition orthogonale de la variance inter, la première portant sur les six groupes constitués par les trois registres associés aux deux locuteurs, la seconde sur les seize groupes constitués par les seize consonnes du français.

Dans la première analyse (Tableau III), une différence très significative ($F = 22,96$) apparaît entre les registres grave et médium pris ensemble, et le registre aigu. On constate également un F significatif (4,08) entre le groupe des mesures relevées dans le registre aigu du locuteur et celui des mesures relevées dans le registre aigu de la locutrice.

Le locuteur ayant, comme on l'a vu, une voix inculte et la locutrice une voix cultivée, il était nécessaire de savoir si c'était le facteur "sexe" ou le facteur "culture vocale" qui était pertinent. Pour cela, nous avons procédé à une vérification portant sur le même corpus émis dans les mêmes conditions par une voix féminine inculte. Les mesures obtenues se rapprochant davantage de celles du LM/VI que de celles relevées pour le LF/VC, nous pensons pouvoir affirmer que c'est la technique vocale plutôt que la catégorie vocale qui est significative dans l'aigu.

Dans la deuxième analyse (Tableau IV), c'est au niveau de l'opposition voisé/non voisé, pour les occlusives et les constrictives qu'apparaissent deux F très significatifs (respectivement 13,82 et 14,94), les autres comparaisons se révélant toutes non significatives. Il est intéressant de comparer cette dernière analyse à celle que nous avons effectuée sur un corpus en phonation continue avec l'opposition Voix Chantée Articulée / Voix Chantée Esthétique [10] et qui faisait apparaître une différence significative entre les occlusives d'une part et les autres consonnes prises ensemble d'autre part ($F = 12,14$) ainsi qu'entre les constrictives d'une part et les nasales regroupées avec les liquides d'autre part ($F = 4,29$). (Tableau V)

Il semble donc que le facteur "mode d'articulation" joue un rôle décisif dans un corpus en phonation continue, alors que dans un corpus de logatomes, c'est le facteur "voisement" qui se révèle déterminant.

IV. CONCLUSION

Afin de pouvoir vérifier si ces conclusions s'appliquent à toutes les conditions d'expérimentation et déterminer

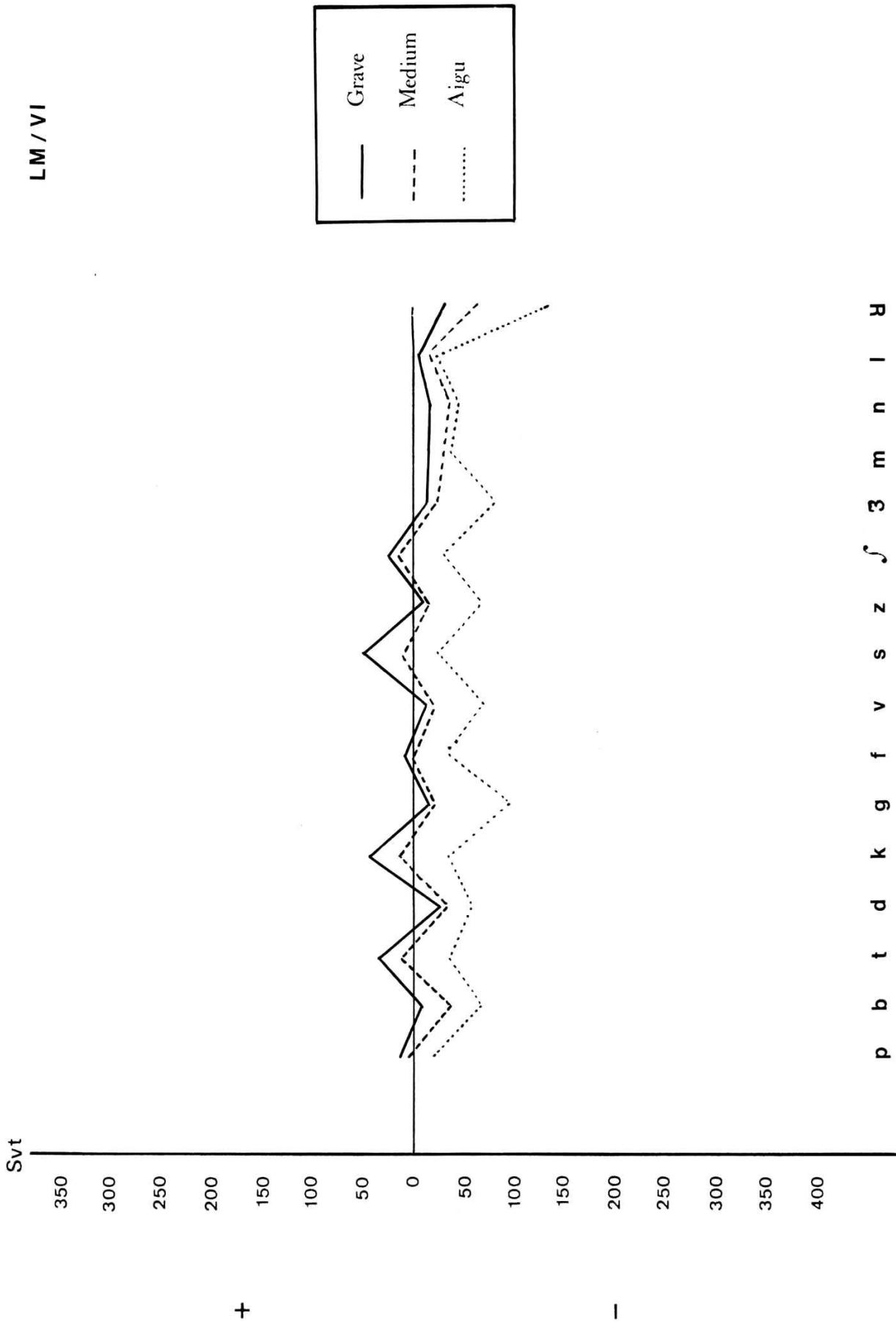
éventuellement la hiérarchie des facteurs significatifs, nous envisageons, dans une étape ultérieure de nos recherches, de poursuivre cette étude sur un nombre plus important de sujets en utilisant les deux types de corpus (corpus en phonation continue et corpus de logatomes chantés) et en épuisant toutes les combinaisons possibles tant au niveau des locuteurs choisis (LM/VC, LM/VI, LF/VC, LF/VI) qu'au niveau des voyelles subséquentes.

BIBLIOGRAPHIE

- 1 - CHAFCOULOFF, M. , DI CRISTO, A. , SEMANDI, L.
Effets de la coarticulation sur les caractéristiques acoustiques des contoïdes fricatives du français.
Travaux de l'Institut de Phonétique d'Aix, 1976, volume 3 : 61-113.
- 2 - CHISTOVICH, L.A.
Variation of the fundamental voice pitch as a discriminatory cue for consonants.
Soviet Physics Acoustics, 1969, 14, 3 : .372-378.
- 3 - DI CRISTO, A.
Contribution à l'étude de la structuration prosodique de la phrase en français moderne.
Thèse de doctorat 3^o cycle. Aix, 19776, 379 p.
- 4 - HOUSE, A.S. & FAIRBANKS, G.
The influence of consonant environment upon the secondary acoustical characteristics of vowels.
Journal of the Acoustical Society of America, 1953, 25, 1: 128-136.
- 5 - KIM, K.
Fo variations according to consonantal environments.
Phonology Laboratory, Berkeley, 1968. (mimeographed).
- 6 - OJAMAA, K., ERICKSON, D. , KUMAR, A.
Coarticulation effects on pitch.
Journal of the Acoustical Society of America, 1970, 48, 1 : A.
- 7 - PYNTE, J.

Programmation d'une expérience de psycholinguistique.
Cahiers de Psychologie, 1975, vol. 18, 1-3 : 65-74.

- 8 - RICHAUD, M.C. & CORNUT, G. Variations du fonctionnement laryngé en fonction des phonèmes et des transitions phonétiques dans la voix chantée.
Proceedings of the XIIth International Speech and Voice Therapy conference, Padua, 1962 : 166-176.
- 9 - SCOTTO DI CARLO, N.
Le mécanisme des attaques dans le chant.
Hommages à Georges Mounin, C.L.O.S., 1975, 5-6 : 365-371.
- 10 - SCOTTO DI CARLO, N.
Influence de l'articulation sur la musicalité de la phrase chantée. *Travaux de l'Institut de Phonétique d'Aix*, 1976, 3 : 117-146.



- FIGURE 1 -

- TABLEAU I -

REGISTRES	FREQUENCE CIBLE	VALEURS MOYENNES DE FO		JUSTESSE	
		SOURDES	SONORES	SOURDES	SONORES
GRAVE	125 Hz	124,7 Hz	123,8 Hz	- 1,04 Svt	- 4,18 Svt
MEDIUM	153 Hz	151,3 Hz	150,8 Hz	- 4,85 Svt	- 6,29 Svt
AIGU	245 Hz	238,6 Hz	237,6 Hz	- 12,58 Svt	- 13,31 Svt

- TABLEAU II -

ANALYSE DE VARIANCE
(PROGRAMME NAVA DE J. PYNTE)

=====

PLAN : R * S * P < C >

R = Registre, S = Sexe, P = Phonèmes, C = Catégories

S.C.	D.L.	C.M.		F
9862918	2	4931459	R	22,71
870966	1	870966	S	1,54
955762	2	477881	SR	<1
14397066	8	1799633	P	
3474138	16	217133	PR	
4517362	8	564670	PS	
50946392	16	3184149	PSR	
1379784	7	197112	C	<1
3861476	14	275819	CR	1,27
4435612	7	633658	CS	1,12
2525566	14	180397	CSR	<1
<hr/>				
97227060	95			

F (2,16) - Seuil à .05 : 3,63

Seuil à .01 : 6,23

Seuil à .001 : 11,00

- TABLEAU III -

ANALYSE DE VARIANCE I

(DECOMPOSITION ORTHOGONALE DE LA VARIANCE INTER D'A. RAPHAEL)

=====

SIX GROUPES : Grave M, Grave F, Médium M, Médium F, Aigu M, Aigu F.

SOURCE DE VARIATION	S.C.	D.L.	C.M.	F	
1. Grave + Médium / Aigu	98464,08	1	98464,08	22,96	Significatif à .001
2. Grave / Médium	165,12	1	165,12	< 1	
3. Grave M / Grave F	741,13	1	741,13	< 1	
4. Médium M / Médium F	4,2	1	4,2	< 1	
5. Aigu M / Aigu F	17521,91	1	17521,91	4,08	Significatif à .05
VAR. INTER	116896,44	5	23379,28	5,45	Significatif à .001
VAR. INTRA	385899,13	90	4287,76		
VAR. TOTALE	502795,57	95			

F (5,90) - Seuil à .05 : 2,31- F (1,90) Seuil à .05 : 3,95
 Seuil à .01 : 3,22 Seuil à .01 : 6,93
 Seuil à .001 : 4,53 Seuil à .001 : 11,6

- TABLEAU IV -

ANALYSE DE VARIANCE II

(DECOMPOSITION ORTHOGONALE DE LA VARIANCE INTER D'A. RAPHAEL)

=====

SEIZE GROUPES : [p] [t] [k] [b] [d] [g] [f] [s] [ʃ] [v] [z] [ʒ] [m] [n] [l] [ʁ]

SOURCE DE VARIATION	S.C.	D.L.	C.M.	F	
1. Occl/Constr + Nas + Liq	1154,55	1	1154,55	< 1	
2. Constr / Nas + Liq	2604,99	1	2604,99	< 1	
3. Nas / Liq	1542,41	1	1542,41	< 1	
4. Occl non vois/ Occl vois	59617,35	1	59617,35	13,82	Sign. à .001
5. Constr non vois/ Constr vois	64465,21	1	64465,21	14,94	Sign. à .001
6. p / t + k	282,79	1	282,79	< 1	
7. b / d + g	8433,36	1	8433,36	1,95	
8. f / s + ʃ	63,20	1	63,20	< 1	
9. v / z + ʒ	4482,30	1	4482,30	1,03	
10. t / k	27,91	1	27,91	< 1	
11. d / g	1905,11	1	1905,11	< 1	
12. s / ʃ	20,53	1	20,53	< 1	
13. z / ʒ	115,93	1	115,93	< 1	
14. m / n	3340,00	1	3340,00	< 1	
15. l / ʁ	9712,83	1	9712,83	2,25	
	<hr/>	<hr/>			
VAR. INTER	157768,47	15	10517,89		
VAR. INTRA	345027,10	80	4312,83		
	<hr/>	<hr/>			
VAR. TOTALE	502795,57	95			

F (1,80) - Seuil à .05 : 3,96
 Seuil à .01 : 6,96
 Seuil à .001: 11,7

- TABLEAU V -

ANALYSE DE VARIANCE

(DECOMPOSITION ORTHOGONALE DE LA VARIANCE INTER D'A. RAPHAEL)

=====

28 GROUPES : 14 CONSONNES : [t] [k] [b] [d] [g] [f] [s] [v] [z] [ʒ] [m] [n] [l] [ʁ]
chacune en VCA (Voix Chantée Articulée) et en VCE (Voix Chantée Esthétique)

1. Occl/ autres	21592,98	1	21592,98	12,141	<i>Sign. à .01</i>
2. Constr / Nas + Liq	7637,21	1	7637,21	4,294	<i>Sign. à .05</i>
3. Nas / Liq	42,92	1	42,92	< 1	
4. Occl sde / Occl son.	2024,36	1	2024,36	1,138	
5. Cst. sde / Cst. son.	2114,03	1	2114,03	1,188	
6. b / d + g	24,80	1	24,80	< 1	
7. ʒ / v + z	254,03	1	254,03	< 1	
8. d / g	134,10	1	134,10	< 1	
9. v / z	10,60	1	10,60	< 1	
10. t / k	124,75	1	124,75	< 1	
11. f / s	210,97	1	210,97	< 1	
12. m / n	0,07	1	0,07	< 1	
13. l / ʁ	166,33	1	166,33	< 1	
14. t VCA / t VCE	661,20	1	661,20	< 1	
15. k VCA / k VCE	101,52	1	101,52	< 1	
16. b VCA / b VCE	443,30	1	443,30	< 1	
17. d VCA / d VCE	15927,56	1	15927,56	8,956	<i>Sign. à .01</i>
18. g VCA / g VCE	3224,03	1	3224,03	1,812	
19. f VCA / f VCE	873,20	1	873,20	< 1	
20. s VCA / s VCE	208,90	1	208,90	< 1	
21. v VCA / v VCE	1,15	1	1,15	< 1	
22. z VCA / z VCE	2,04	1	2,04	< 1	
23. ʒ VCA / ʒ VCE	1,05	1	1,05	< 1	
24. m VCA / m VCE	12,49	1	12,49	< 1	
25. n VCA / n VCE	133,17	1	133,17	< 1	
26. l VCA / l VCE	157,77	1	157,77	< 1	
27. ʁ VCA / ʁ VCE	17,74	1	17,74	< 1	

VAR. INTER	56102,27	27	2077,86
VAR. INTRA	64022,52	36	1778,40
VAR. TOTALE	120124,79	63	

F(1,36)- Seuil à .05 : 4,11
Seuil à .01 : 7,39
Seuil à .001 : 13

F(27,36)- Seuil à .05 : 1,80
Seuil à .01 : 2,31
Seuil à .001 : 3,18