

ETUDE DE LA PERCEPTION DES NOTES COURTES CHANTEES EN PRESENCE DE VIBRATO

Christophe d'Alessandro & Michèle Castellengo

LIMSI-CNRS BP133-91403 Orsay Cédex, France.
LAM, URA 868-CNRS, Tour 65-66, Université Paris VI, 4 place Jussieu,
75005 Paris, France.

ABSTRACT

This paper presents some results on perception of short vocal vibrato tones, using a method of adjustment. Means, standard deviations were computed and histograms plotted from 19 sets of responses (11 subjects, 28 types of short tones ranging from 1/2 vibrato cycle up to 2 vibrato cycles in steps of 1/4 cycle, 4 types of initial phase, mean frequency of 440 Hz, vibrato frequency of 6 Hz and vibrato amplitude of 100 cents). The main results are: A) for short tones, the pitch does not correspond to the mean frequency; B) the pitch depends on the shape of end of the tone; C) the pitch converges towards the mean frequency as the duration increases; D) the overall pattern of F0 has an influence on perception, and some simple patterns seem to behave better perceptually.

1 Introduction

L'étude de la perception de hauteur tonale dans la voix chantée en présence de vibrato de la fréquence fondamentale est un domaine qui a été relativement peu étudié. Une étude [5] (par la méthode d'ajustement) sur la perception de notes longues, qui comportent plusieurs cycles de vibrato par note, conclut à la perception d'une hauteur moyenne (arithmétique ou géométrique, ce qui est peu différent) pour des amplitudes et des fréquences de vibrato de l'ordre de grandeur de celle rencontrées dans le chant. Il est notable que cette perception moyenne est contestée par plusieurs auteurs (par exemple [4] p. 46) qui affirment que la moyenne mais aussi les deux hauteurs extrêmes peuvent être entendues, en fonction du contexte. Cependant, les différents auteurs s'accordent pour reconnaître qu'en dehors de consignes ou de contraintes particulières, les sujets auditeurs perçoivent la hauteur moyenne. Les résultats sur les notes longues ne peuvent s'appliquer pour

expliquer comment sont appréciées les notes courtes, qui abondent dans les exécutions musicales. Lors d'une première étude sur les notes courtes [1] nous avons émis quelques hypothèses sur la base d'expériences préliminaires qu'il était nécessaire de reprendre de façon plus systématique. Cette communication présente une étude pour mesurer la hauteur tonale perçue lors de l'émission de notes courtes vibrées, hors de tout contexte musical, donc dans des conditions d'écoute de test psychoacoustique. Malgré la différence notable de cette situation avec celle d'une situation musicale réelle, nous pensons que les résultats peuvent se révéler utiles pour expliquer des phénomènes observés dans la production des chanteurs ou dans la perception des auditeurs.

2 Méthode

Une méthode d'ajustement a été utilisée pour mesurer la hauteur tonale des notes courtes. Les exemples synthétiques étaient produits par un synthétiseur à formants en parallèle. Les formants, maintenus fixes avec des valeurs ($F1 = 650 Hz$, $F2 = 1100 Hz$, $F3 = 2900 Hz$, $F4 = 3300 Hz$), correspondaient à un /a/. Les expériences ont été menées par 11 sujets, tous pourvus d'une éducation musicale. Certains sujets ont effectué plusieurs fois le test et 19 jeux de réponses ont été utilisés comme données expérimentales. Les stimuli étaient présentés de façon binaurale par un casque Beyer DT48 à 80 dB SPL. Pendant la présentation de la première série d'exemples, certains sujets ont attribué la différence entre les sons à une différence de timbre plutôt qu'à une différence de hauteur. La consigne donnée a donc été de se concentrer sur la différence de hauteur, mais l'influence de petites variations de hauteur sur le timbre mérite probablement une étude spécifique. Pour chaque test, les stimuli sont constitués d'une paire de sons: un son non vibré toujours identique et un son vibré

que le sujet pouvait choisir parmi 12 sons de fréquences différentes ordonnées. Le vibrato suit une loi sinusoïdale autour d'une fréquence moyenne de 440 Hz, loi de fréquence (période $1/6s=167ms$), ce qui correspond à une double croche avec la noire à 88, soit une note musicale courte mais courante dans la pratique musicale. L'amplitude totale du vibrato est de 100 cents (1/2 ton), la fréquence variant entre 428 et 450 Hz. Les fréquences des 12 sons d'appariement s'échelonnent entre ces deux limites, par pas de 2 Hz, les sujets devant répondre par un numéro entre 1 et 12. Le seuil différentiel de fréquence (pour des sons purs, à 440Hz et 80dB SPL) est légèrement supérieur à 2 Hz à cette fréquence. Les deux sons d'un même stimuli sont de durée égale, et séparés par un silence de 300 ms. D'un test à l'autre, la durée des sons varie en fonction du nombre fractionnaire de cycles de vibrato présent. La plus petite durée utilisée dans ces tests correspond à une demi période de vibrato, soit une durée de 82 ms environ, et la plus grande à deux cycles complets, soit 343 ms. Les 7 durées utilisées s'étagent par palier d'un quart de cycle (41 ms). Pour chaque durée, quatre phases initiales ont été présentées, et 28 tests d'ajustement, correspondant à 28 formes ont donc été proposés, comme le résume la figure 1. Une expérience préliminaire, a permis de tester la perception de notes longues de dix cycles complets de vibrato, avec la même méthodologie, et donne des résultats identiques à ceux de rapportés dans [5] pour les mêmes valeurs de vibrato et de fréquences fondamentales: la hauteur perçue est alors la moyenne soit environ 440 Hertz.

3 Résultats

Tous les sujets ont évoqué des difficultés pour juger la hauteur dans certains cas, a cause du mouvement du fondamental. Certains sujets ont même évoqué la possibilité d'entendre deux hauteurs, et la consigne a été d'ajuster à la meilleure hauteur possible (un tel phénomène est décrit dans [2] pour des glissandi de sons purs). Les résultats sont très cohérents malgré ces difficultés. La figure 1 résume les différentes formes (notées de 1 à 28) ainsi que les moyennes et écarts types obtenus. Les écarts types sont dans l'ensemble plutôt faibles, de l'ordre du seuil différentiel. Ils ne varient pas de façon significative d'une forme à l'autre.

3.1 Moyennes

La figure 2 reporte la moyenne des fréquences perçues en fonction de la forme (notée de 1 à 28, dans l'ordre). Les 7 courbes comportent 4 points chacune, pour les stimuli de même durée: l'axe des numéros de formes est aussi l'axe des durées croissantes, avec quatre phases initiales pour chaque durée. Pour une même durée les hauteurs perçues peuvent être très différentes: les formes de plus courte durée peuvent, par exemple, suivant la phase initiale du vibrato, être perçues de 437 à 447 Hertz. En observant la courbe des 4 points d'une même durée, on remarque que ces formes se reproduisent à intervalle de 4: la courbe 1 ressemble à la 5, la 2 à la 6, la 3 à la 7. Les courbes de plus longues durées ont un ambitus plus petit. Cette "pseudo-périodicité" perceptive semble liée à la périodicité du vibrato: la forme 5 (resp. 6, 7) reproduit la forme 1 (resp. 2, 3) à une période de vibrato près, au début. La fin des formes semble donc prédominer perceptivement. Cette figure montre une convergence des hauteurs perçues, lorsque la durée augmente, vers la valeur moyenne perçue pour les notes longues. La figure 3 se déduit de la figure 2 en joignant les points, selon les durées croissantes, qui possèdent une même phase initiale (soit les points 1,5,9,13,17,21,25 pour la phase 1). Les 4 courbes obtenues oscillent, en convergeant vers la moyenne. Ces oscillations sont liées à la variation finale de la forme finale, et il semble que ce n'est pas la phase initiale qui domine perceptivement. La figure 4 se déduit de la figure 2 en joignant les points qui possèdent une demi-période finale commune. Les 4 courbes convergent de façon asymptotique, sans oscillations, vers la moyenne des notes longues. Il semble que la forme finale du son gouverne la perception de hauteur, avec une pondération due à la durée. Un tel phénomène a été observé [3] [2] pour différents glissandi de fréquence.

3.2 Histogrammes

Les histogrammes représentant le nombre de réponses obtenues pour une fréquence donnée et une forme fixée indiquent la dispersion des réponses, donc leur certitude. La figure 5 montre les 4 histogrammes obtenus pour une durée d'1/2 cycle. Deux situations apparaissent: les phases 1 et 3 donnent des pics marqués, indiquant que les sujets ont perçu un peu de fréquences différentes; les phases 2 et 4 montrent un étalement des réponses, qui implique une certitude plus faible. L'examen des histogrammes, qui ne sont pas tous reproduits ici par manque de

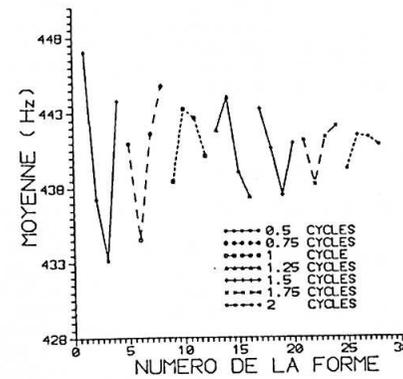


Figure 2

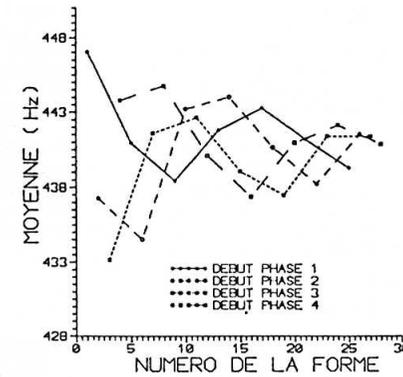


Figure 3

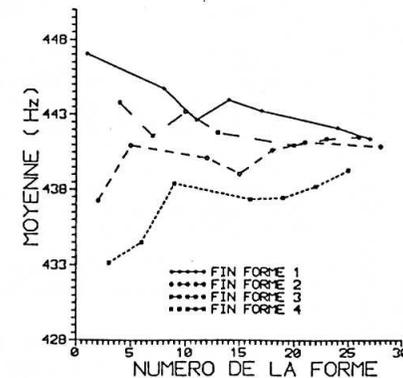


Figure 4

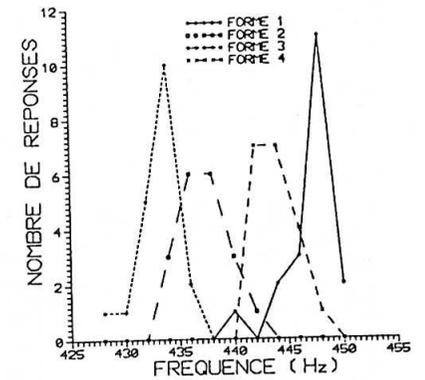


Figure 5

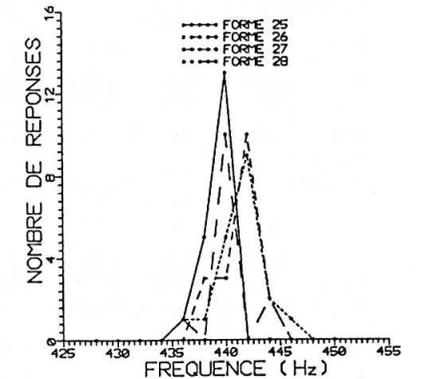


Figure 6

	Phase 1	Phase 2	Phase 3	Phase 4
0.5 cycles	forme 1 	forme 2 	forme 3 	forme 4
Moy.	447.05	437.26	433.16	443.79
E.T.	2.34	2.23	1.92	1.87
0.75 cycles	forme 5 	forme 6 	forme 7 	forme 8
Moy.	440.94	434.53	441.58	444.74
E.T.	1.92	2.19	3.16	1.66
1. cycles	forme 9 	forme 10 	forme 11 	forme 12
Moy.	438.42	443.2	442.63	440.1
E.T.	2.63	2.01	2.5	2.1
1.25 cycles	forme 13 	forme 14 	forme 15 	forme 16
Moy.	441.79	444	439.05	437.37
E.T.	1.99	2	2.78	1.5
1.5 cycles	forme 17 	forme 18 	forme 19 	forme 20
Moy.	443.26	440.63	437.47	440.95
E.T.	1.91	1.77	2.09	1.81
1.75 cycles	forme 21 	forme 22 	forme 23 	forme 24
Moy.	441.16	438.21	441.37	442.10
E.T.	1.54	2.39	1.77	2.35
2. cycles	forme 25 	forme 26 	forme 27 	forme 28
Moy.	439.26	441.47	441.36	440.84
E.T.	1.19	2.09	2.22	1.80

Figure 1

place, donne une indication des formes de variation du fondamental qui semblent perceptivement préférables, surtout pour les sons les plus courts. Les sons plus long possèdent des histogrammes pointus qui indiquent une dispersion faible, comme le montre la figure 6.

4 Conclusions

Les tests d'ajustement pour des sons vocaux vibrés de courte durée indiquent que: A. les sons de courtes durées donnent lieu à des perception de hauteur très différentes de la hauteur moyenne; B. la demi-période finale permet de prévoir 4 sortes d'incidences sur la hauteur perçue: incidence haute pour une 1/2 arche positive (forme 1); incidence moyenne haute pour la 1/2 arche montante (forme 4); incidence moyenne basse pour la demi-arche descendante (forme 2); incidence basse pour la 1/2 arche négative (forme 3); C. avec l'accroissement de durée la hauteur perçue se rapproche de la moyenne; D. la forme globale du son importante, elle se combine avec les formes de fin, et montre une pseudo-périodicité perceptive liée à la périodicité du vibrato; E. les formes courtes qui possèdent un seul maximum (1,3,10,12) sont perçues avec plus de certitude que les autres formes de même durée; F. lorsque la durée s'accroît la certitude des jugement devient égale et haute quelque soit la forme.

Références

- [1] CASTELLENGO M., RICHARD G., d'ALESSANDRO C. (1989). "Study of vocal pitch vibrato perception using synthesis" Proceedings of the 13th Int. Cong. on Acoust. Belgrad, 113-116.
- [2] NABELEK I. V., NABELEK A. K. and HIRSH I. J. (1970). "Pitch of tone bursts of changing frequency" J. Acoust. Soc. Am. 48(2), 536-553.
- [3] ROSSI, M., (1971). "Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole" *Phonetica*, 23, 1-33.
- [4] SEASHORE C. E. (1938). *The psychology of music* Max-Graw Hill, New York.
- [5] SHONLE J. I. and HORAN K. E. (1980). "The pitch of vibrato tones" J. Acoust. Soc. Am. 67(1), 246-252.