Nasalité consonantique et coarticulation : étude perceptive

Tiphaine Ouvaroff, Solange Rossato

Institut de la Communication Parlée
Université Stendhal - 1180, Avenue Centrale, BP 25, 38040 GRENOBLE CEDEX 9
Tél. +33 (0)4 76 82 43 37
Tiphaine.Ouvaroff@icp.inpg.fr, Solange.Rossato@icp.inpg.fr

ABSTRACT

This paper investigates the coarticulation of consonantal nasality from a perceptive point of view. The aim of this study is to determine in CV utterances to which extent the consonant is perceived in the following vowel and compare these perceptual boundaries between oral and nasal consonants. Results show that, although the vowel is actually nasalized (low velum and consistent nasal air flow), the listeners don't attribute the nasalization of the vowel to the presence of a nasal consonant. After the release, the nasal feature of the consonant is lost, only the place of articulation is perceived until 60 ms after the release.

1. INTRODUCTION

De nombreuses études se sont intéressées au phénomène de coarticulation de la nasalité (entre autres Krakow & Beddor [4], Krakow et al. [5], Abramson et al. [1]) : il est avéré que les voyelles en contexte de consonne nasale sont nasalisées par effet de propagation du trait nasal de la consonne. Des enregistrements réalisés au sein de l'ICP nous ont permis d'observer le phénomène de coarticulation de la nasalité d'un point de vue articulatoire et aérodynamique, et cette étude perceptive se place dans la continuité de ces travaux.

Des mesures EMA (Articulographie Electro-Magnétique) réalisées en 2002 nous ont permis d'évaluer d'un point de vue articulatoire le phénomène de coarticulation consonantique de la nasalité, grâce à des données quantitatives quant à l'abaissement vélaire (Rossato et al. [8]). Dans des séquences de type NV, nous avons pu observer une coarticulation progressive de la consonne nasale sur la voyelle orale. La figure 1 montre la trajectoire vélaire durant un segment VNV : le velum reste abaissé au même niveau que la consonne nasale durant toute la durée de la voyelle, cette hauteur étant dépendante de la hauteur de la voyelle.

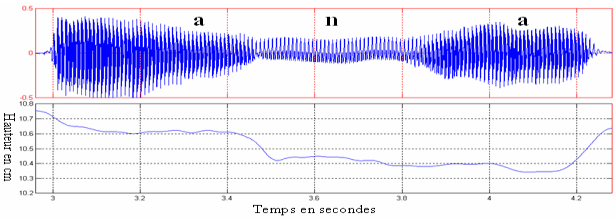

Figure 1 : Trajectoire du velum dans l'axe des Y (en bas) synchronisée au signal acoustique pour un segment [ana].

Des enregistrements aérodynamiques EVA réalisés en 2003 sur le même corpus et pour le même locuteur, ont permis l'observation d'un débit nasal important dans la voyelle suivant la consonne nasale (Ouvaroff [6]). Ce débit nasal n'est pas réservé exclusivement au [a] : les voyelles fermées [i], [y] et [u] attestent un débit nasal élevé (Figure 2), d'ailleurs supérieur à celui des voyelles nasales.

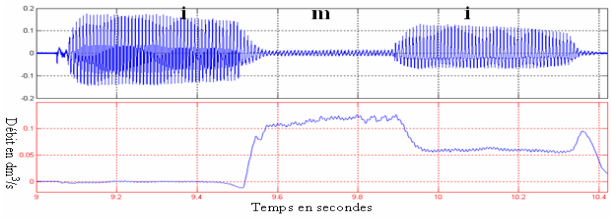

Figure 2 : Débit nasal en dm^3/s (en bas) synchronisé au signal acoustique pour un segment [imi].

Plusieurs travaux ont étudié d'un point de vue perceptif le rapport entre la consonne nasale et la voyelle en contexte. Krakow & Beddor [4] observent que la voyelle orale [ɛ], nasalisée car placée en contexte consonantique nasal [mɛn], est perçue nasale même placée en contexte CVC ou #V#. Cette voyelle est même perçue plus nasalisée qu'en contexte approprié. La nasalisation de la voyelle modifie la perception de sa hauteur : les voyelles hautes sont perçues plus basses tandis que les voyelles basses sont perçues plus hautes (Wright [11]). Krakow et al. [5] ont montré pour des locuteurs américains que ce couplage nasal n'amène pas nécessairement à une mauvaise perception de la hauteur de la voyelle quand la nasalisation de cette voyelle provient d'un phénomène de coarticulation. D'autre part, la hauteur intrinsèque du velum dépend de la hauteur de la voyelle. Abramson et al. [1] a montré que cette hauteur vélaire de la voyelle affecte les frontières perceptives de la nasalité dans un continuum acoustique [d-n]. Il existe un rapport étroit en terme de perception entre la consonne et la voyelle, et c'est sur ce rapport que s'appuie notre test de perception de la nasalité consonantique.

Il s'agit de définir, dans des séquences de type CV, les 'frontières perceptives' de la coarticulation. Puisque la nasalisation de la voyelle peut être attribuée à un phénomène de coarticulation, cela nous permet-il de retrouver le trait nasal de la consonne précédente ? Pour cela, nous comparons les frontières perceptives de la coarticulation de consonnes qui ne s'opposent que par le trait nasal ([b d] versus [m n]).

2. MÉTHODE

Nous avons donc réalisé un test perceptif comportant des stimuli de type 'backward gating', utilisé précédemment par Smits [9] et Pols & Schouten [7] dans leur étude sur la perception des consonnes.

2.1. Stimuli

Les signaux utilisés proviennent d'enregistrements EMA réalisés en 2002 au sein du laboratoire (dispositif EMA Carstens AG100) pour lesquels nous disposons des trajectoires des articulateurs (y compris le velum). Dans ce corpus, des séquences de la forme CV ont été sélectionnées, où C = [b d m n] et V = [a i u]. La sélection des séquences s'est faite sur la qualité acoustique du signal et sur sa régularité. Les stimuli ont été extraits par technique de gating, procédé consistant à dévoiler le signal par étapes, grâce au logiciel PRAAT. La découpe s'est faite par rapport à l'instant de relâchement acoustique de la consonne (noté '(0)'), et ce par tranches de 20 ms environ (les stimuli sont découpés au passage par zéro le plus proche). 10 stimuli ont été ainsi extraits pour chaque séquence CV : le premier a son onset 60 ms avant le relâchement de la consonne (un pré-test nous a montré que nous avons une bonne identification de al consonne) et le dernier 120 ms après le relâchement (fin des transitions formantiques). Un exemple de ce découpage est donné en Figure 3. L'offset est fixe et se situe à la fin de la partie stable de la voyelle. Nous avons des stimuli de durée variable. Notre test comporte au final (10 tranches x 4 consonnes x 3 voyelles) = 120 stimuli CV. Des stimuli de la forme VC ont également été découpés et utilisés dans le test de perception, mais sont pour l'instant en cours d'analyse.

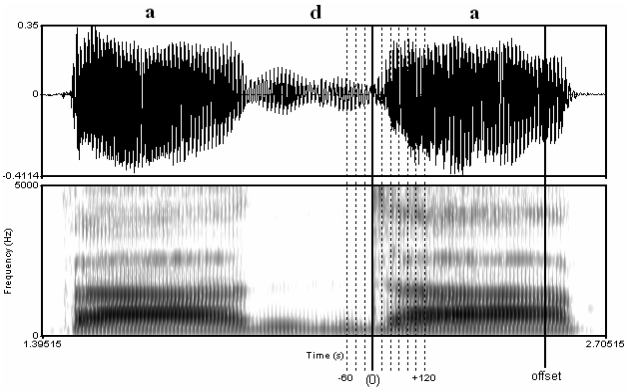

Figure 3 : Exemple de découpe des stimuli d'une séquence [ada] par technique de 'backward gating' : 10 pas de gating à offset fixe et à onset variable, de 60 ms avant jusqu'à 120 ms après le relâchement de la consonne (0).

2.2. Procédure

Le test perceptif a été implémenté grâce au logiciel Dreamcard Revolution et se divise en trois parties correspondant aux trois voyelles [a i u]. La consigne donné au sujet est la suivante : «Déterminer la consonne que vous entendez avant la voyelle, sachant qu'elle a été coupée. » Le sujet a le choix entre 5 réponses fermées : [b] [d] [m] [n] et {aucune consonne} dans le cas où la consonne n'est pas perçue par le sujet.

2.3. Sujets

16 sujets ont participé au test perceptif, mené au sein du laboratoire de l'ICP. Tous les sujets sont des français natifs, exceptés deux, de nationalité américaine et syrienne : leurs résultats ont été exploités étant donné que leur langue comporte les oppositions consonantiques labiale/dentale et orale/nasale.

3. ANALYSE DES RÉSULTATS

Nous avons observé les résultats d'identification correcte des consonnes orales et nasales [b d m n], et les éventuelles différences selon le mode et le lieu de la consonne. Les 'frontières perceptives' (logit) de la consonne, ont été extraites à l'aide d'un test de régression logistique binaire avec le logiciel SPSS qui nous fournit également les différences significatives.

2.1. Identification correcte de la consonne dans la voyelle : mode et lieu

Les taux d'identification correcte des consonnes en fonction du pas de gating sont représentés Figure 4. Les résultats en termes de frontières de perception de la consonne sont différents selon le mode (oral/nasal) et le lieu (labial/coronal) de la consonne.

Contraste de mode

Nous observons une différence significative entre les consonnes orales et nasales ($p<0,001$). Ces résultats montrent que la nasalisation de la voyelle ne suffit pas à inférer une coarticulation nasale : les frontières perceptives des consonnes nasales se situent peu après le burst (+4,8 ms pour [m], +2,2 ms pour [n]). Nous observons même que le sujet, dans 40% des cas, ne perçoit déjà plus la consonne nasale 20 ms avant même le relâchement de la consonne. En revanche les consonnes orales sont identifiées correctement jusqu'à 42,2 ms après le relâchement pour le [b] et 54,3 ms pour le [d]. Nous n'observons pas d'interaction entre le mode et le pas de gating : les courbes d'identification orale et nasale ont la même pente, mais décalées dans le pas de gating.

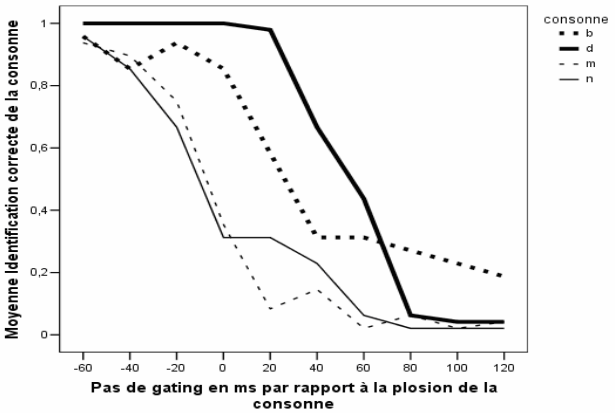

Figure 4 : Pourcentage d'identification correcte des consonnes [b d m n], en fonction du pas de gating (temps en ms par rapport au relâchement de la consonne (0)).

Nous pouvons donc dire au vu de nos résultats que les effets de coarticulation, en terme d'abaissement du velum et de débit nasal, ne permettent pas aux sujets d'attribuer la nasalisation de la voyelle à la présence d'une consonne nasale en contexte gauche.

Contraste de lieu

Nous observons des frontières perceptives significativement différentes selon le lieu d'articulation de la consonne. Nous avons une interaction entre le mode oral/nasal et le lieu labial/coronal (interaction significative, $p<0,001$) : c'est seulement en contexte oral que le lieu a une influence sur les frontières de perception. En effet, cette différence n'est observée qu'en contexte de consonne orale [b d] ($p<0,001$), et non en contexte nasal [m n]. En contexte oral, la consonne coronale [d] est perçue environ 12,1 ms de plus dans la voyelle que la consonne labiale [b] : les frontières perceptives des consonnes nasales ne présentent qu'environ 2,6 ms de différence. Nous observons également pour les consonnes orales une interaction entre le lieu et le pas de gating, en d'autres termes la frontière de perception des coronales est plus raide que celles des labiales. La consonne coronale est en effet perceptible plus longtemps dans la voyelle car l'effet de coarticulation augmente avec l'amplitude de mouvement de la langue. Nos résultats perceptifs indiquent une interaction entre le mode et le lieu de la consonne : selon que la consonne est orale ou nasale, les frontières de perception correcte du lieu changent.

2.2. Perception d'une consonne quelconque : les confusions en termes de traits d'articulation

En plus de l'identification correcte de la consonne, nous nous sommes intéressés à la perception d'une consonne quelle qu'elle soit (toutes réponses sauf [aucune consonne]). Il y a un laps de temps d'environ 30 ms où une consonne est encore perçue, mais pour laquelle les traits de mode ou de lieu d'articulation ne sont plus correctement identifiés.

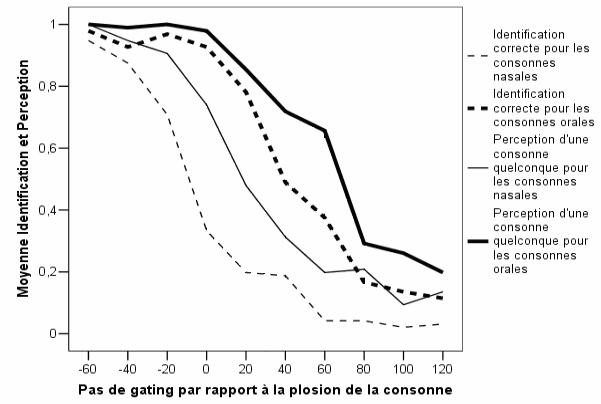

Figure 5 : Représentation des frontières d'identification correcte de la consonne et de perception d'une consonne quelconque pour les consonnes orales (trait épais) et nasales (trait fin).

Ce laps de temps est différent selon la consonne et peut être exprimé par la différence entre la frontière de perception d'une consonne quelconque et la frontière d'identification correcte de la consonne : ces frontières sont résumées Table 1. Ce laps de temps varie dans sa durée (de 20,7ms pour le [b] à 32,9ms pour le [n]). Les confusions pour les consonnes nasales [m n] se font environ dans les 30 premières millisecondes de la voyelle, et entre 40 et 80ms environ après le burst pour les consonnes orales (entre 40 et 60ms pour le [b], entre 60 et 80ms pour le [d]).À ce stade où seules les transitions formantiques sont présentes (après le burst), quelles sont les confusions en terme de perception des traits de la consonne ?

Table 1 : Frontières d'identification correcte et de perception d'une consonne quelconque pour les consonnes [b d m n] (par rapport au burst).

	[b]	[d]	[m]	[n]
Identification correcte de la consonne	42,2 ms	54,3 ms	4,8 ms	2,2 ms
Perception d'une consonne quelconque	62,9 ms	77,9 ms	26,4 ms	35,1 ms

Le mode des consonnes orales est perçu en majorité comme oral mais le lieu est moins bien identifié, notamment pour la consonne [d]. À l'inverse, nous observons pour les consonnes nasales que le lieu est correctement identifié, les principales erreurs concernant le mode : la consonne [m] a tendance à être perçue [b], le [n] perçu [d].

Les résultats en terme de reconnaissance du lieu sont explicables simplement par le fait que l'intervalle de confusion pour les consonnes nasales se situe en début de la transition formantique (Figure 6) : comme l'ont démontré entre autres Stevens & Blumstein[10], les premières 26ms de la transition formantique permettent d'identifier le lieu d'articulation de la consonne. L'intervalle de confusion des consonnes orales se situe trop loin dans la transition pour que le lieu soit encore clairement identifiable, notamment pour le [d].

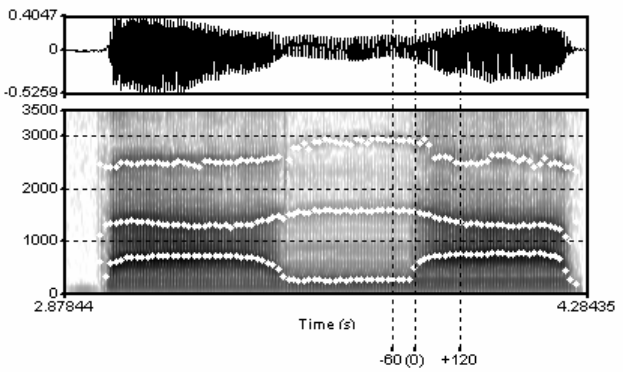

Figure 6 : Spectrogramme et contours formantiques.de la séquence [ana] avec les pas de gating de -60ms et +120ms par rapport à la plosion du [n] (0).

Un fait des plus intéressant est que les consonnes nasales, avant de ne plus être perçues du tout, sont perçues orales dès la disparition du murmure nasal. La nasalisation seule de la voyelle n'est pas suffisante pour identifier la nasalité de la consonne précédente : dès que la consonne est relâchée, la nasalité n'est plus attribuée à la coarticulation. Le même type de résultat a été montré par Beddor & Onsuwan[2] sur la perception des consonnes prénasalisées. Ils observent que dans une séquence [mbV] où le [b] inséré dure environ 27ms, même si la voyelle est nasalisée à 100% les sujets perçoivent en majorité [mb] au lieu de [m] : la nasalisation de la voyelle n'est pas suffisante pour contrer une absence de murmure nasal. Entre 0 et 30ms après le burst, l'indice de nasalité a disparu mais l'indice de présence d'une consonne – matérialisé par les transitions formantiques – indique encore d'après nos résultats le lieu d'articulation de la consonne : seul le trait de lieu subsiste, le [m] est perçu |b] et le [n] perçu [d].

4. CONCLUSION

La nasalisation de la voyelle n'est pas suffisante pour identifier le trait nasal de la consonne précédente une fois le murmure nasal disparu, tandis que le lieu d'articulation, porté par les transitions formantiques, reste perçu. Nous ne pouvons pas dire au vu de ce seul test que la voyelle n'est pas perçue nasalisée, seulement que cette nasalisation n'est pas attribuée à la consonne adjacente. La perception de la nasalisation serait donc à tester indépendamment de la consonne, d'autant plus que la langue française comporte l'opposition voyelle nasalisée/voyelle nasale Il serait intéressant d'extraire la voyelle nasalisée et l'observer dans d'autres contextes, à l'instar de Krakow & Beddor[4] : elles notent, de la même façon que Kawasaki[3], que la perception de la nasalité de la voyelle n'est pas favorisée par la présence adjacente de la consonne nasale. Une prochaine étape sera également d'observer la perception de la nasalité en fonction de la nature de la voyelle.

REMERCIEMENTS

Nous tenons à remercier Pierre Badin, notre locuteur pour les corpus aérodynamique et articulatoire.

BIBLIOGRAPHIE

[1] A. S. Abramson, P.W. Nye, J. Henderson and C.W. Marshall. Vowel height and the perception of consonantal nasality. *Journal of the Acoustical Society of America*, 70:329-339, 1981.

[2] P.S. Beddor and C. Onsuwan. Perception of prenasalized stops. *Proceedings of 15th ICPhS*, Barcelona, Spain, 2003.

[3] H. Kawasaki. Phonetic explanation for phonologicals universals: the case of distinctive vowel nasalization. In *Experimental Phonology*, Edited by J.J Ohala, J.J. Jaeger, New York Academic Press, New York, NY, pages 81-103, 1986.

[4] R.A. Krakow and P.S. Beddor. Coarticulation and the perception of nasality. In *Proceedings of 12th ICPhS*, Aix-en-Provence, France, volume 4, pages 38-41, 1991.

[5] R.A. Krakow, P.S. Beddor, L.M. Goldstein and C. Fowler. Coarticulatory influences on the perceived height of nasal vowels. *Journal of the Acoustical Society of America*, 83:1146-1158, 1988.

[6] T. Ouvaroff. Mesures aérodynamiques de la parole dans le cas des nasales. *Rapport de stage de maîtrise*, 2004.

[7] L.C.W. Pols and M.E.H. Schouten. Identification of deleted consonants. *Journal of the Acoustical Society of America*, 64:1333-1337, 1978.

[8] S. Rossato, P. Badin and F. Bouaouni. Velar movements in French: An articulatory and acoustical analysis of coarticulation. In *Proceedings of 15th ICPhS*, Barcelona, Spain, pages 3141-3144, 2003.

[9] R. Smits. Human consonant recognition for initial and final segments of VCV utterances. In *Speech Hearing and Langage*, Work in process 10, Department of Phonetics and Linguistics, University College London, UK, pages 115-136, 1998.

[10] K.N. Stevens and S.E. Blumstein. Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 64:1358-1368, 1978.

[11] J.T. Wright. Effects of vowel nasalization on the perception of vowel height. In *Nasalfest : Papers from a Symposium on Nasals and Nasalization*. Edited by C.A. Ferguson, L.M. Hyman and J.J. Ohala, Langage Universals Project, Stanford University, Stanford, CA, pages 373-388, 1975.