

Comparaison de différentes méthodes d'enregistrement de la nasalité: données fibroscopiques et signal acoustique nasal

Julie Montagu et Angélique Amelot

Laboratoire de Phonétique et Phonologie
UMR 7018, CNRS/ Sorbonne Nouvelle
19, rue des Bernardins 75005 Paris
Tel.: ++33 (0)1 43 26 37 80 – Fax: ++33 (0)1 44 32 05 78
julie_montagu@yahoo.com

ABSTRACT

In this study, we compare the acoustic nasal signal (SAN) and the interpolated curve of the height of the velum. The results show that there is a good correspondance between the velar articulation and the resulting SAN.

1. INTRODUCTION

Vu la complexité du langage, il apparaît nécessaire d'aborder chaque phénomène de parole de plusieurs points de vue, et de les mettre en confrontation. La production de la nasalité fait principalement intervenir les articulateurs suivants : le velum, la langue et les lèvres. Il existe de nombreuses instrumentations pour étudier la nasalité, plus ou moins invasives : méthode EMG, fibroscopique, aérodynamique, acoustique. Aucune de ces méthodes ne permet d'étudier seule les articulateurs dans leur ensemble.

Cet article fait suite à Montagu & Amelot [1], une étude comparative entre les apports de deux méthodes d'enregistrement de la nasalité, le débit d'air nasal (DAN) et le signal acoustique nasal (SAN). Au terme de cette étude préliminaire, nous avons montré que le DAN et le SAN ne sont pas corrélés aux mêmes phénomènes physiologiques, mais qu'ils apportent des informations complémentaires pour expliquer les phénomènes de coarticulation entre voyelles et consonnes. Il semblerait que le SAN serait mieux corrélé aux mouvements du velum que le DAN quant à lui serait mieux corrélé à l'impédance dans le conduit vocal. Nous avons donc mis en place une expérimentation en enregistrant de manière simultanée les mouvements du velum par fibroscopie et le SAN.

Il est important pour mieux comprendre la nasalité vocalique de connaître le rapport entre l'articulatoire et la résultante acoustique.

2. METHODOLOGIE

2.1. Corpus et locuteur

Le corpus se compose de 96 logatomes de structure $C_1V_1CV_2C_1V_1$ répétés deux fois à la suite en débit normal et une fois en débit rapide (non traité ici). Il s'agit du même type corpus utilisé dans la première étude comparative entre les données de DAN et de SAN, auquel

il a été ajouté deux contextes consonantiques et deux contextes vocaliques. $C_1 = [t, d, s, z, l, n]$, $V_1 = [i, \text{œ}, a, u]$ et $V_2 = [\text{ē}, \text{ā}, \text{ō}]$ (la cible). Exemple : [tatōta].

Un seul locuteur a subi l'expérimentation. Il s'agit d'un homme de 30 ans, locuteur natif du français et résidant actuellement en région parisienne.

2.2. Acquisition des données

Données fibroscopiques

Les données vidéo fibroscopiques ont été prises à l'Hôpital Européen G. Pompidou (par le Dr. Lise Crevier-Buchman) à l'aide de la station ATMOS, à la vitesse de 25 images par seconde. Chaque image correspond donc à une période de 40 millisecondes (ms). L'endoscope flexible a été inséré par la narine gauche sans anesthésie. Celui est posé sur la face supérieure du velum au bord du port vélo-pharyngé, pour plus de détails voir Amelot [2]. Le signal acoustique a été enregistré à partir d'un micro-cravate de type Sony fixé sur le fibroscope.

Données acoustiques

Le signal acoustique oral (SAO) a été enregistré avec un micro cravate placé sur le locuteur à environ 15 cm de sa bouche. Le signal acoustique nasal (SAN) a été enregistré avec un micro nasal (Montagu [3]) placé à l'aide d'une olive nasale dans la narine droite du locuteur.

Ces deux signaux acoustiques ont été enregistrés simultanément en stéréo sur un minidisque.

2.3. Traitement des données

Synchronisation :

Les deux types de données (fibroscopique & acoustique nasal) ont été alignés à partir des signaux acoustiques (oraux) captés respectivement par les deux micro-cravate. Un clac de départ a servi de référence à la synchronisation.

Nous avons ensuite remplacé le signal acoustique des données fibroscopiques par le signal acoustique simultané de SAO et SAN, sur le logiciel SoundForge 5, à la même fréquence d'échantillonnage de 48000 Hz .

Traitement des fichiers vidéo :

La sélection des séquences vidéo au format .avi est rigoureusement fonction des 40 millisecondes (ms)

auxquelles correspondent 1 image et n'admet aucun décalage. La longueur des fichiers vidéo a été standardisée à 840 ms, soit 22 images pour les logatomes les plus longs (contexte [d] dV₁dV₂dV₁). Nous avons ensuite extrait les 22 images des fichiers vidéo (.avi) sur le logiciel Saint Paint.

La hauteur du velum a été mesurée image par image, en prenant comme point de référence une image où le velum est maximalelement ouvert (phase de respiration). Il s'agit d'une mesure relative, en pixels. Deux programmes conçus par Shinji Maeda sur Matlab 7 nous ont permis d'obtenir une courbe interpolée au format .text puis .wav, à partir des points déterminés de la hauteur relative du velum toutes les 40 ms.

Analyse acoustique :

Nous avons procédé à une analyse acoustique comparative fine entre les signaux acoustiques SAO et SAN sur SoundForge 5. Nous avons ainsi segmenté et mesuré les décalages temporels aux frontières des voyelles nasales cibles à la ms. Les phénomènes d'anticipation et de persévération de la nasalité ont seulement été observés sur le [l], la consonne du corpus la plus perméable à la nasalisation selon l'échelle de nasalisation fournie par Clements [4]. Pour les contextes occlusif et fricatif, nous avons observé et mesuré deux types de phénomènes de coarticulation qui ont une incidence acoustique calculable grâce à ce type d'enregistrement acoustique (SAO/SAN). En effet, on observe des « fuites nasales » qui correspondent à du SAN relevé avant le début de la voyelle nasale, pendant le relâchement après l'explosion pour les occlusives ou à la fin de la friction pour les fricatives. Le second phénomène de coarticulation est la « queue nasale » à la fin des voyelles nasales suivies de ces consonnes.

Comparaison :

Nous avons enfin reporté la segmentation temporelle précise sur les courbes de hauteur relative du velum ; afin d'étudier dans quelle mesure notre hypothèse de départ, à savoir la corrélation entre le SAN et les mouvements du velum, pourrait être établie avec ce type de données.

3. RESULTATS ET DISCUSSION

3.1. Mesures temporelles

Dans l'étude comparative entre les mesures de débit d'air nasal (DAN) et de SAN (Montagu & Amelot [1]), les phénomènes de coarticulation mesurés temporellement n'apparaissent que dans le contexte [titṽti]. Ici, nous observons des décalages entre le SAO et le SAN aux frontières des voyelles nasales dans pratiquement tous les contextes consonantiques et vocaliques, pour la production de ce locuteur.

Décalage au début des voyelles nasales

Nous avons mesuré le décalage temporel (en ms) entre le début du SAN et le début du SAO. Cependant, les traces de faible amplitude relevées sur le SAN à la fin du [s] souvent de manière discontinue (partie 2 sur la figure 1) seront considérées comme négligeables et dépendantes de ce locuteur (variation individuelle) et /ou des conditions d'enregistrement.

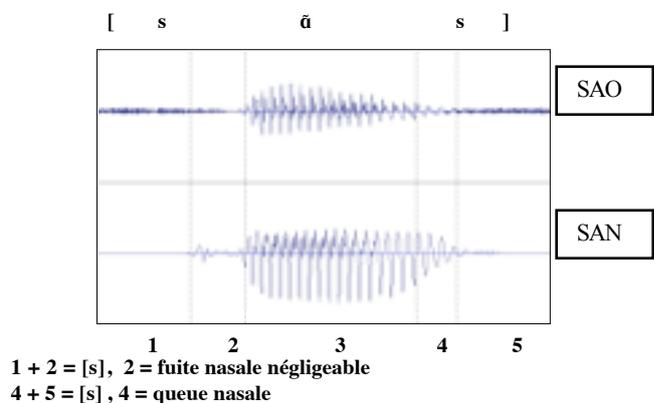
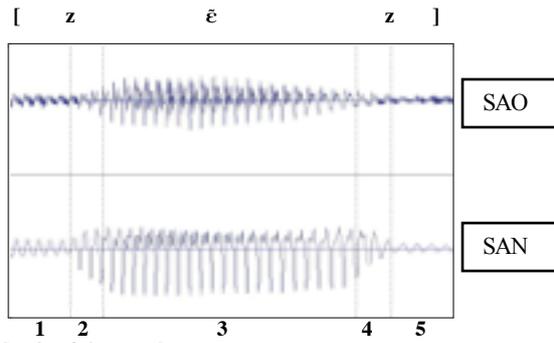


Figure 1 : SAO et SAN de la voyelle nasale [ã] dans le contexte [s] du logatome /sœsãsœ/

Par contre, pour le [t] et les consonnes voisées [d] et [z] précédant les voyelles nasales, il se produit un phénomène de « fuite nasale » pendant le relâchement de l'occlusion pour les deux premières ou à la fin de la friction pour la troisième, phénomène non négligeable qui a une incidence acoustique que l'on observe sur le SAN par un tracé relativement ample (partie 2 sur la Figure 2). Ce phénomène de coarticulation entre la fin de ces consonnes et le début des voyelles nasales sera également considéré comme une variation individuelle.

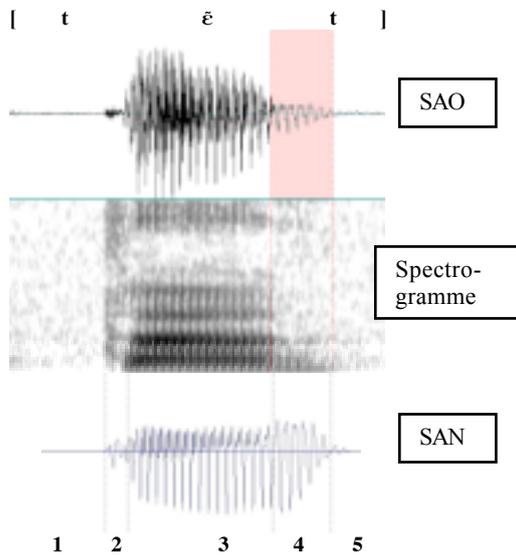


1 + 2 = [z], 2 = fuite nasale
4 + 5 = [z], 4 = queue nasale

Figure 2 : SAO et SAN de la voyelle nasale [ɛ̃] dans le contexte [z] du logatome /zœzɛ̃zœ/

Décalage à la fin des voyelles nasales

Nous avons mesuré le décalage temporel (en ms) entre la fin du SAN et celle du SAO à la fin des voyelles nasales. Le phénomène de queue nasale correspond sur le SAO à des vibrations de faible amplitude à la fin des voyelles nasales, visibles sur le spectrogramme dans les basses fréquences (Figure 3). Sur le SAN, on observe un changement de forme des vibrations souvent accompagnée d'une augmentation de l'amplitude du signal.



1 + 2 = [t], 2 = fuite nasale
4 + 5 = [t], 4 = queue nasale

Figure 3 : SAO, spectrogramme et SAN de la voyelle [ɛ̃] en contexte [t] dans le logatome [tœtɛ̃tœ].

Les résultats présentés en table 1 montrent que le phénomène de queue nasale semble dépendre plus largement du type de consonne qui suit une voyelle nasale. En effet, pour expliquer ce phénomène, il faut trouver ce que ces consonnes occlusives et fricatives ont en commun au niveau articulatoire qui puisse provoquer le même type d'incidence acoustique (et aérodynamique).

Table 1 : Mesures du phénomène de « queue nasale » en ms, correspondant au décalage temporel entre le SAN et le SAO à la fin de chaque voyelle nasale suivie par les consonnes [t], [d], [s] et [z].

Contexte consonantique	Contexte vocalique	Voyelles nasales \tilde{v}		
		[ɛ̃]	[ɑ̃]	[õ]
[t]	[tat \tilde{v} ta]	41	41	44
	[tœt \tilde{v} tœ]	40	42	42
	[tit \tilde{v} ti]	42	41	44
	[tut \tilde{v} tu]	42	42	41
	<i>moyenne</i>	<i>41,25</i>	<i>41,5</i>	<i>42,75</i>
[d]	[dad \tilde{v} da]	53	56	64
	[dœd \tilde{v} dœ]	55	57	62
	[did \tilde{v} di]	56	57	53
	[dud \tilde{v} du]	56	57	56
	<i>moyenne</i>	<i>55</i>	<i>56,75</i>	<i>58,75</i>
[s]	[sas \tilde{v} sa]	21	26	34
	[sœs \tilde{v} sœ]	22	25	29
	[sis \tilde{v} si]	35	25	34
	[sus \tilde{v} su]	21	29	26
	<i>moyenne</i>	<i>24,75</i>	<i>26,25</i>	<i>30,75</i>
[z]	[zaz \tilde{v} za]	20	28	33
	[zœz \tilde{v} zœ]	20	24	25
	[ziz \tilde{v} zi]	22	22	27
	<i>moyenne</i>	<i>20,66</i>	<i>24,66</i>	<i>28,33</i>

Il s'agit de leur regroupement dans la classe des « constrictives » (classement articulatoire). C'est dans la configuration suivante que se produit le phénomène de queue nasale : lorsqu'une voyelle nasale est suivie par une consonne constrictive, la constriction (fermeture totale ou resserrement) dans la cavité buccale se produit avant que le port vélo-pharyngé ne soit remonté ce qui provoque l'échappement de l'air phonatoire par la cavité nasale, dont les résonances naturelles dans les basses fréquences sont captées par le SAN.

Sur la table 1, on peut remarquer d'une part qu'entre une voyelle nasale et une consonne occlusive, le décalage de SAN par rapport au SAO à la fin de la voyelle nasale est plus important que dans le cas de la coarticulation avec une consonne fricative. D'autre part, dans la paire des occlusives, le voisement du [d] semble d'autant plus favoriser ce phénomène alors que pour les fricatives, les mesures de décalage de SAN pour [z] sont légèrement moins importantes que pour [s].

En regardant les moyennes par contexte mais surtout par voyelle, il apparaît enfin assez clairement que ce

phénomène de queue nasale dépend à mesure égale de l'articulation de la voyelle nasale, et plus précisément de son ouverture. On note une ordonnance régulière entre les mesures de décalage de SAN par rapport à la fin du SAO pour les trois voyelles nasales dans tous les contextes. Plus la voyelle nasale est fermée, plus le phénomène de queue nasale est long dans la coarticulation avec les occlusives et les fricatives.

3.3. Comparaison de la hauteur relative du velum et du SAN

Nous avons comparé les données du SAN avec les hauteurs relatives du velum et plus précisément nous avons reporté la segmentation effectuée sur chaque logatome en comparant le SAO et le SAN. Nous observons une excellente corrélation entre le SAN et les courbes de hauteur relative du velum pour les logatomes avec un contexte occlusif ou fricatif entourant la voyelle nasale (Figure 4).

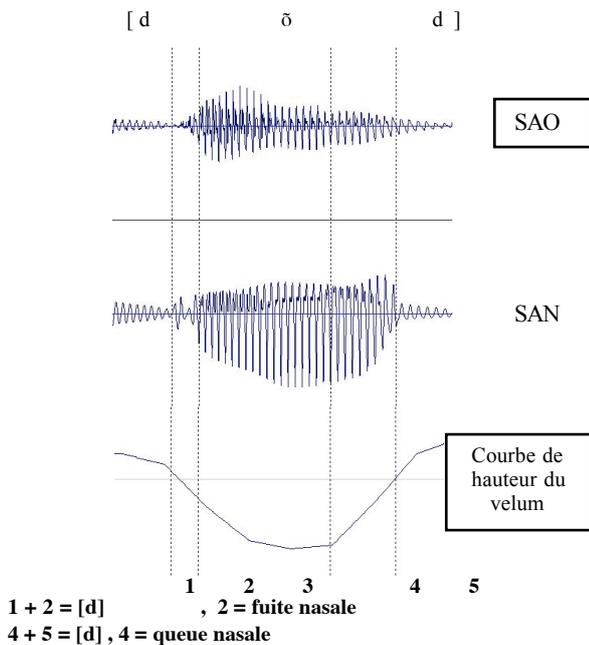


Figure 4 : SAO, SAN et courbe du velum pour la voyelle [õ] en contexte [d] dans le logatome [dœdõdœ].

4. CONCLUSION ET PERSPECTIVE

Nous pouvons remarquer que la corrélation pourrait être encore plus précise avec des données fibroscopiques prises avec une caméra plus rapide, c'est-à-dire avec le double ou le triple d'images. Cependant que la courbe de hauteur relative du velum ne permet pas une segmentation précise entre consonnes et voyelles nasales, la position du velum est basse et varie peu (Figure 5), ni de voir les phénomènes de coarticulation aux frontières des voyelles nasales tels que les fuites ou les queues nasales. Il faut également souligner le traitement long de ce type de données, qui ne peuvent être prises que dans certaines conditions (présence et contrôle d'un médecin). Les informations contenues dans le signal acoustique nasal

résultant possèdent donc plusieurs avantages. Premièrement, la précision des informations que l'on obtient dans la comparaison des deux signaux acoustiques permet de rendre compte très localement de phénomènes de coarticulation entre consonnes et voyelles nasales. Deuxièmement, les mouvements articulaires du velum et la méthode d'enregistrement acoustique (SAO/SAN) est facilement reproductible et le traitement plus rapide.

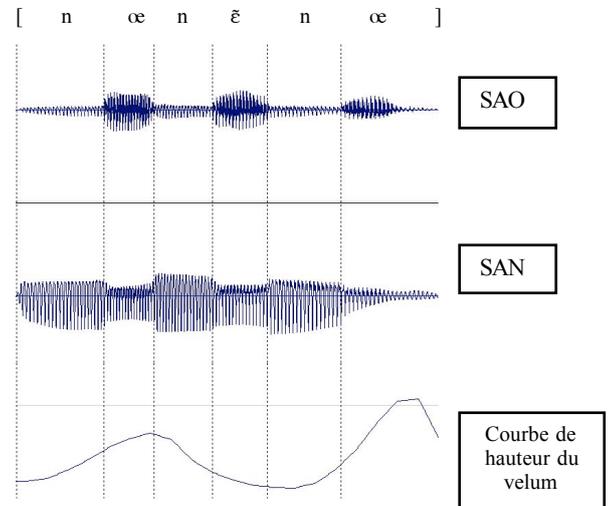


Figure 5 : SAO, SAN et courbe du velum pour la voyelle [ẽ] en contexte [n] dans le logatome [nœnõnœ].

BIBLIOGRAPHIE

- [1] Montagu & Amelot, Comparaison des apports de différentes méthodes d'enregistrement de la nasalité : étude préliminaire. *Rencontre Jeunes Chercheurs en Parole, Toulouse, 2005.*
- [2] Amelot, A. *Etude aérodynamique, fibroscopique, acoustique et perceptive des voyelles nasales du français.* Thèse de Phonétique, Université Paris 3, 203 pp, 2004.
- [3] J. Montagu. Les sons sous-jacents aux voyelles nasales en français parisien : indices perceptifs des changements. In *XXVèmes Journées d'Étude sur la Parole*, pages 385-388, Fès, Maroc, 24-27 Avril 2004.
- [4] Clements, G.N. and Sylvester, O. Patterns of nasality and obstruction in Ikwere, an African language with nasal harmony, *Atas do 1º Encontro internacional do GTLI da ANPOLL/UFPA. Linguas Indigenas Brasileiras: Fonologia, gramática e história, Belem, 2001.*