

L'influence de la coarticulation et du contexte sur les paramètres acoustiques des consonnes nasales n (ن) et m (م).

The influence of coarticulation and context on the acoustic parameters of nasal consonants n (ن) and m (م).

DAOUD Messaoud^{1*}

¹ Université d'Alger 2, messaoud.daoud@univ-alger2.dz

Date de réception: 17/01/2023 Date de revision: 10/07/2024 Date d'acceptation:01/12/2024

Résumé

La prononciation est un outil important pour une expression claire et une compréhension humaine dans la communication orale, mais il existe plusieurs phénomènes dans différents contextes responsables de la complexité du traitement sur les signaux vocaux.

Dans ce travail, nous avons réalisé une étude acoustique arabe sur les consonnes nasales /m/(م) et /n/(ن) en utilisant PRAAT comme outil d'analyse des effets contextuels sur les paramètres consonnes des consonnes nasales.

Mots clés : la langue arabe, les consonnes nasales, paramètres acoustiques, le contexte.

Abstract

Pronunciation is an important tool for clear expression and human understanding in oral communication, but there are several phenomena in different contexts responsible for the complexity of processing on speech signals.

In this work, we performed an Arabic acoustic study on the nasal consonants /m/(م) and /n/(ن) using PRAAT as a tool for analyzing the contextual effects on the consonant parameters of nasal consonants.

Keywords : Arabic language, nasal consonants, acoustic parameters, context.

* DAOUD Messaoud, Email: messaoud.daoud@univ-alger2.dz

Introduction :

L'homme se distingue des autres êtres vivants par la production de la parole, qui est une activité très complexe qui nécessite le contrôle des organes impliqués, et les sons de la parole sont coarticulés ce qui permet de produire des phrases et des mots.

Les mouvements articulatoires peuvent en effet être modifiés de façon à minimiser l'effort à produire pour les réaliser à partir d'une position articulatoire donnée ou pour anticiper une position à venir, ces effets sont connus sous le nom de réduction, d'assimilation et de coarticulation, ces phénomènes sont en grande partie responsables de la complexité des traitements réalisés sur les signaux de parole.

La langue arabe fait partie des langues du monde qui possèdent les consonnes nasales /m/(م) et /n/(ن), la coarticulation et leur effet sont étudiées dans plusieurs langues et dans différents contextes.

Cependant, des recherches ont été consacrées à l'étude de la coarticulation des consonnes nasales en arabe, ce qui nous a incité à mener des recherches dans ce domaine, y compris cette recherche. En fait, ce travail est une étude acoustique en arabe qui utilise PRAAT comme outil d'analyse.

Dans la présente étude, nous chercherons à répondre à la question suivante: Quel est l'effet des contextes sur les paramètres acoustiques des consonnes nasales?

1. Les phonèmes de la langue Arabe :

La langue arabe est une langue consonantique qui s'écrit de droite à gauche. Elle se compose d'un certain nombre de symboles ou graphismes que nous résumons comme suit:

- 28 consonnes ;
- trois voyelles brèves [a], [i] et [u] ;
- trois voyelles longues [a], [i] et [u] ;
- le [sukun] ;
- les caractéristiques propres à la langue (emphase, gémation, ...).

La particularité de la langue Arabe se trouve dans la présence des phonèmes arrières glottales [ʔ] (ء) et [h] (ه), pharyngales [ħ] (ح) et [ʕ] (ع), vélaire [q] (ق), de l'affriquée [g] (ج) et du phonème emphatique [d] (ض) (BAHI, Fethi. 2012 : 25).

La classification des phonèmes de l'Arabe est basée, comme toute autre langue, sur le lieu et le mode d'articulation :

- le lieu d'articulation est la zone du conduit vocal qui participe à la formation du son. Il représente la position de la constriction totale (cas des occlusives) ou partielle (cas des fricatives) d'une zone spécifique du conduit vocal lors du passage de l'air provenant des poumons. La position de la constriction peut être bilabiale, labiodentale, palatale, vélaire, pharyngale, glottale, etc. ;

- le mode d'articulation est lié aux diverses sources d'excitation du conduit vocal, nous pouvons citer à titre d'exemples :

- le mode occlusif ;
- le mode fricatif ;
- la nasalité ;
- l'emphase : les travaux dans le domaine parlent de la présence d'un second point d'articulation pharyngale ;
- le voisement lorsqu'il y a vibrations des cordes vocales lors du passage de l'air, on dit que le son est sonore ou voisé, dans le cas contraire, le son est dit sourd ou non voisé (RAI, Ali. 2013 : 20).

En Arabe, les trois signes qui représentent les voyelles redoublées à la fin des noms indéterminés et les voyelles finales se prononcent comme si elles sont suivies du son nasal [n], on parle de [tanwin].

Exemples : [-un], kitabun; [-an], kitaban; [-in], kitabin (BAHI, Fethi. 2012 : 25).

2. Les consonnes nasales :

En phonétique articulatoire, les consonnes produites en abaissant le voile du palais sont appelées consonnes nasales, il existe aussi des voyelles nasales, les consonnes nasales sont presque toujours des occlusives, mais il existe aussi dans certaines langues telles que le gaélique des fricatives nasales (obtenues toutefois par deux articulations liées par une phase de désocclusion pour produire la fricative).

Ce sont des occlusives car bien que le flux pulmonaire s'écoule par la cavité nasale, il est stoppé dans la bouche.

Les occlusives nasales font partie de la famille des continues, c'est-à-dire qu'elles ne bloquent pas l'air expiré (au contraire des occlusives ou fricatives orales) et sont quasiment toujours voisées, d'un point de vue acoustique, les

occlusives nasales ont un spectre sonore compris entre 200 et 2000 Hz (HEMRI, Fodil. 2012 : 23).

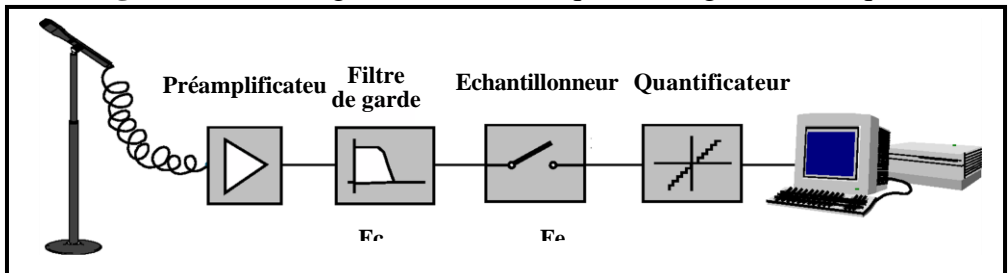
Au système vocalique arabe, le système consonantique contient deux consonnes nasales: /m/ (م) /مفتاح/ miftah/ une clé et /n/ (ن) /نمر/ namir/ un tigre. Ces deux consonnes exigent que la sortie d'air s'opère par les deux cavités - nasale et orale (ELBEREGLI, Othman. 2017 : 132).

3. Le niveau acoustique

La parole apparaît physiquement comme une variation de la pression de l'air causée et émise par le système articulatoire, la phonétique acoustique étudie ce signal en le transformant dans un premier temps en signal électrique grâce au transducteur approprié: le microphone (lui-même associé à un préamplificateur). De nos jours, le signal électrique résultant est le plus souvent numérisé. Il peut alors être soumis à un ensemble de traitements statistiques qui visent à en mettre en évidence les traits acoustiques: sa fréquence fondamentale son énergie et son spectre. Chaque trait acoustique est lui-même intimement lié à une grandeur perceptuelle: pitch, intensité et timbre.

L'opération de numérisation schématisée à la (figure 1) requiert successivement: un filtrage de garde un échantillonnage et une quantification (THIERRY, Dutoit. 2000 : 6).

Figure N° 1: Enregistrement numérique d'un signal acoustique.



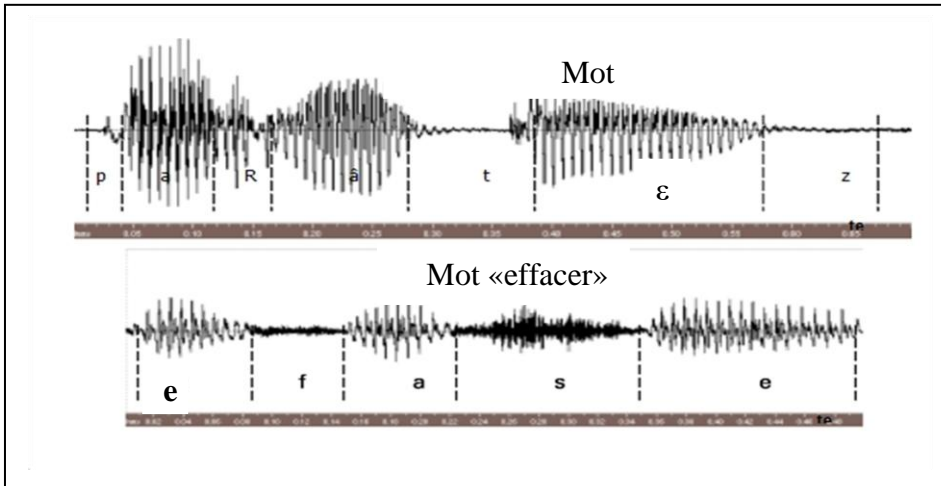
Source : Dutoit, 2000, p6.

3.1. Audiogramme

L'échantillonnage transforme le signal à temps continu $x(t)$ en signal à temps discret $x(nT_e)$ défini aux instants d'échantillonnage multiples entiers

de la période d'échantillonnage T_e . Celle-ci est elle-même l'inverse de la fréquence d'échantillonnage f_e . Pour la téléphonie, on estime que le signal garde une qualité suffisante lorsque son spectre est limité à 3400 Hz et l'on choisit $f_e = 8000$ Hz. Pour les techniques d'analyse, de synthèse ou de reconnaissance de la parole la fréquence peut varier de 6000 à 16000 Hz. Pour les applications multimédia les fréquences sous-multiples de 44.1 kHz sont de plus en plus utilisées : 22.5 kHz, 11.25 kHz (THIERRY, Dutoit. 2000 : 8).

Figure N° 2: Audiogramme des signaux de parole des mots «parenthèse» et «effacer».



Source : Dutoit, 2000 , p8.

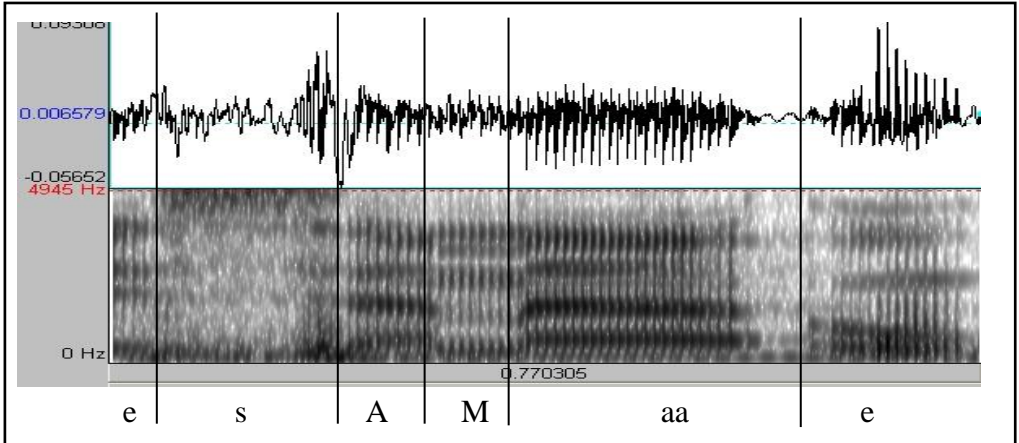
La figure 1 représente l'évolution temporelle ou audiogramme du signal vocal pour les mots «parenthèse» et «effacer», on y constate une alternance de zones assez périodiques et de zones bruitées appelées zones voisées et non-voisées (THIERRY, Dutoit. 2000 : 9).

3.2. Spectrogramme :

Un spectrogramme est une représentation visuelle de la parole. Il permet de voir l'évolution de l'énergie dans l'échelle des fréquences en fonction du temps. L'amplitude du spectre (l'énergie) y apparaît sous la forme de niveaux de gris dans un diagramme en deux axes : le temps (de gauche à droite) et la fréquence (de basse fréquence (son grave) à haute fréquence (son aigu)). Ils mettent en évidence l'enveloppe spectrale du signal et permettent par conséquent de visualiser l'évolution temporelle des formants. Les périodes

voisées y apparaissent sous la forme de bandes verticales plus sombres (LE MANH, Tuan. 2007 : 20).

Figure N° 3: spectrogramme de signal de parole de mot "esamaae (السماء)



3. 3. Le formant :

L'appareil phonatoire étant constitué de différentes cavités. Lors du passage de l'air à travers ces cavités il est amplifié et subit différentes transformations dues aux degrés d'ouverture et de fermeture au niveau de chaque cavité à la position de la langue, des lèvres, etc. Ces cavités possèdent des fréquences de résonance qui renforcent certaines régions du spectre des sources excitatrices. Les maxima de la courbe de réponse en fréquence du conduit vocal sont appelés «Formant». Chaque son a ses formants caractéristiques. Sur un spectrogramme les formants sont représentés par le des bandes noires (le degré de noirceur correspondant à l'énergie).

Les formants nous permettent de décrire aussi bien les cibles vocaliques correspondant aux transitions (passage entre deux sons consécutifs). Ce qui montre leur très grande importance pour l'analyse acoustique en phonétique. Au moins trois formants sont exigés pour produire les différentes voyelles. Généralement on peut aller jusqu'à cinq formants pour produire une parole de haute qualité.

Les valeurs des formants sont très influencées par le lieu d'articulation des phonèmes. Elles donnent une image de la configuration articuloire du

conduit vocal car elles correspondent aux fréquences de résonance du conduit vocal.

Les expériences ont montré que la position fréquentielle des trois premiers formants caractérisait le timbre vocalique. L'explication qu'on donne aux trois formants notés F1, F2 et F3 est la suivante (DJOUADI, Mahieddine. 1991 : 35):

- F1 naît dans la cavité résonnante comprise entre le larynx et le dos de la langue.
- F2 naît dans la cavité résonnante située entre le dos de la langue et les lèvres.
- F3 dépend de l'arrondissement des lèvres.

Ajouter à cela le fait que plus la cavité de résonance est large plus la fréquence correspondante est basse. Réciproquement plus cette cavité est petite, plus la fréquence est haute.

L'ensemble de ces notions citées ci-dessus nous permet d'interpréter aisément la majorité des résultats dans l'étude acoustique des phonèmes.

3.4. Paramètres prosodiques d'un signal de parole

Du point de vue acoustique la prosodie désigne les phénomènes liés à la variation dans le temps des paramètres de hauteur, d'intensité et de durée. La hauteur est essentiellement liée à la fréquence fondamentale f_0 qui correspond au niveau physiologique de la production de la parole à la fréquence de vibrations des cordes vocales. L'intensité est essentiellement liée à l'amplitude et à l'énergie du son mais dépend aussi partiellement de sa durée (BAHI, Fethi. 2012 : 13).

3.4.1. Fréquence fondamentale

Une analyse d'un signal de parole n'est pas complète tant qu'on n'a pas mesuré l'évolution temporelle de la fréquence fondamentale ou pitch. Elle s'étend approximativement de 70 à 250 Hz chez les hommes, de 150 à 400 Hz chez les femmes, et de 200 à 600 Hz chez les enfants (THIERRY, Dutoit. 2000 : 12).

3.4.2. Energie

L'énergie (intensité) E d'un son correspond à l'amplitude de la vibration acoustique. En d'autres termes, elle caractérise le volume sonore qui nous permet de distinguer un son fort d'un son faible. C'est le paramètre prosodique le plus facile à culer, il peut être exprimé pour un signal échantillonné

$$E = \frac{1}{T} \sum_{n=1}^T X_n^2 \quad (1)$$

Etant donné sa dynamique et pour respecter l'échelle perceptive, elle est généralement exprimée en décibels (dB) (BAHI, Fethi. 2012 : 14):

$$E_{db} = 10 \times \log_{10} \left(\frac{1}{T} \sum_{n=1}^T X_n^2 \right) \quad (2)$$

3.4.3. Durée

On appelle durée l'étalement du mouvement oscillatoire. Plus ce mouvement est étalé et plus le son est perçu comme long (RAI, Ali. 2013 : 10).

La durée d'un signal correspond à son temps d'émission sa durée acoustique. C'est le paramètre le plus difficile à préciser car rien n'indique comment le système de contrôle, de production ou de perception de parole mesure le temps. Les indices de durée classiques supposent généralement la donnée d'une segmentation des frontières des unités dont on désire mesurer la durée: la durée d'une unité est alors mesurée par le nombre de trames qui séparent ses frontières de début et de fin. La plupart des systèmes utilisent une segmentation basée sur le phonème (BAHI, Fethi. 2012 : 15).

4. Elaboration du corpus

La première étape à effectuer avant d'entamer les analyses, c'est l'élaboration du corpus de consonnes, le choix de ce dernier n'est pas arbitraire, en ce qui concerne notre travail nous avons pris des phrases contenant les consonnes nasales pris dans différentes positions (Initiale, Médiane et Finale), nous justifions le choix de ce type de corpus par le fait qu'il est préférable de reconnaître les consonnes dans des contextes réels.

Table N° 1 : Les consonnes choisies pour notre étude.

Modes	Types des consonnes	Consonnes arabes	Transcription	Lieu d'articulation
nasale	Voisées	م	[m]	Labiale
nasale	Voisées	ن	[n]	dental

5. Acquisition des données

L'acquisition des données consiste à enregistrer les phrases du corpus choisis, en utilisant un matériel spécifique au traitement du signal vocal. Les phrases du corpus ont été enregistrées par un locuteur arabophone, la fréquence d'échantillonnage choisie est de 11025 Hz.

6. Représentation spectrographique des prononciations des phrases enregistrées:

Le spectrogramme est un moyen de voir le déplacement formantique et dire si la consonne a été bon prononcé ou non, ceci se remarque par une analyse visuelle comparative des fréquences propres à chaque consonne, comme illustré dans la figure 4.

Figure N° 4 : spectrogramme de signal parole de la phrase "min bakh si ni ? mati elahi dafnuhaa" (من بخس نعمة الله دفتها)

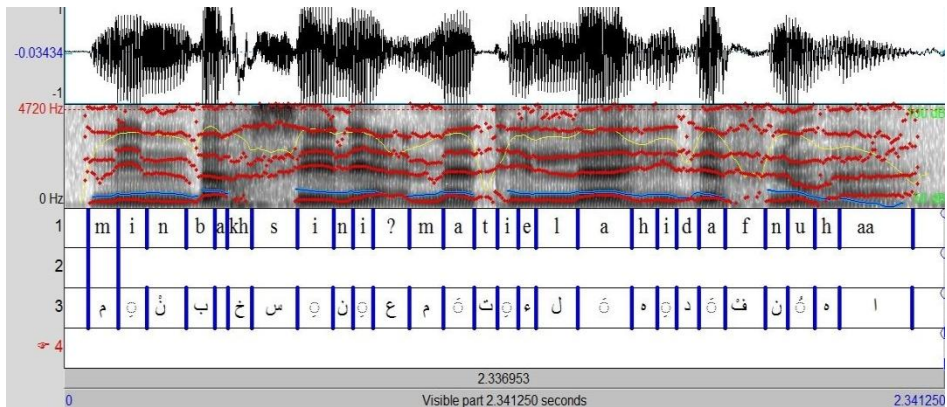
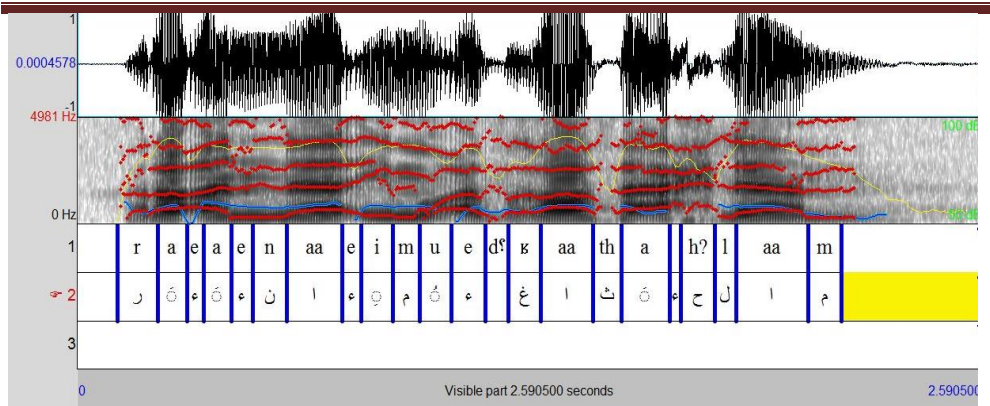


Figure N° 5 : spectrogramme de signal parole de la phrase "raea enaeimu ad'raatha ah'laam" (رأى النائم أضغاث أحلام)



Nous remarquons qu'il est difficile de décider sur caractéristiques visuellement, pour cela l'utilisation de méthodes utilisant des techniques d'analyse plus avancées pour l'extraction des caractéristiques acoustiques est impérative.

7. les caractéristiques acoustiques des consonnes nasales :

Table N° 2 : paramètres acoustiques de consonne n (ن).

	Formants				F0(Hz)	Intensité (dB)	Durée (S)
	F1(Hz)	F2(Hz)	F3(Hz)	F4(Hz)			
Na	291.68	1751.35	2506.21	3876.21	125.70	85.97	0.099
Nu	300.08	1527.21	2385.55	3606.79	124.66	84.63	0.060
Ni	293.34	1891.87	2460.46	3005.00	122.64	86.24	0.052
N	296.58	1599.72	2266.51	3543.68	120.59	84.91	0.106

Table N° 3 : paramètres acoustiques de consonne m (م).

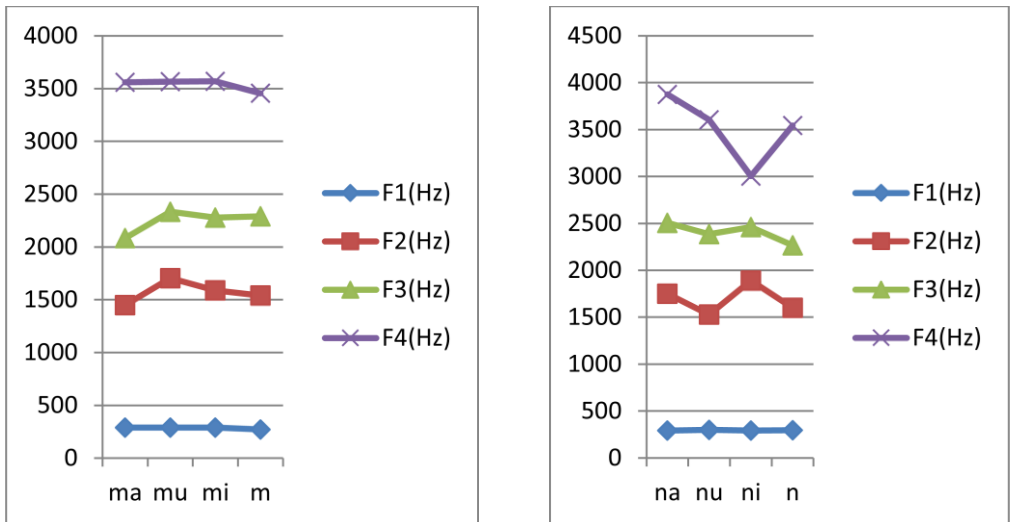
	Formants				F0 (Hz)	Intensité (dB)	Durée (S)
	F1(Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)			
Ma	288.61	1449.02	2083.68	3559.94	117.88	83.10	0.089
Mu	287.87	1704.82	2333.97	3564.84	128.29	86.99	0.077
Mi	289.57	1589.39	2277.94	3569.68	112.80	83.51	0.064
M	271.54	1540.14	2289.80	3455.30	104.80	81.07	0.095

Nous avons mesuré la durée des consonnes nasales n et m de la production de locuteur, par rapport à leur position dans les phrases et les voyelles qui le suit, une différence significative a été observée entre ces différentes positions

au niveau de la durée, nous remarquons une décroissance des valeurs distribuées comme suit : m(0.095s), ma(0.089s), mu(0.077s), mi(0.064s); mais la durée de prononciation ne dépasse pas 0.106 s en général pour le consonne n et 0.095 s pour la consonne m.

Nous avons pu enregistrer qu'il y a une stabilité autour d'une valeur moyenne d'intensité 84.5 dB, et pour les valeurs de fréquence fondamentale F0 nous allons voir la même stabilité autour d'une valeur F0= 120 Hz pour les deux consonnes nasales n et m.

Figure N° 6:valeurs des formants des consonnes nasales n et m en contexte.



Nous avons observé qu'il y a une stabilité des formants autour des valeurs moyennes F1= 284.4 Hz et F4=3537.44 Hz pour le consonant m, et pour le consonant n la stabilité est autour d'une valeur moyenne F1=295.42 Hz, et pour les autres formants, un changement significatif a été observé pour les différentes positions dans le contexte.

Conclusion

Chaque consonne a ses propres caractéristiques acoustiques qui la distinguent des autres. Mais dans la parole continue, ces caractéristiques changent et subissent des modifications qui se traduisent par un changement de la valeur des caractéristiques acoustiques telles que la durée et les valeurs de F2 et F3, et ce phénomène se produit lorsque les consonnes entrent en

contact les unes avec les autres, et ce phénomène est inévitable dans la parole continue.

Au terme de cette recherche, nous avons atteint les constats suivants:

- Il est difficile d'identifier visuellement les caractéristiques des consonnes nasales.
- Une différence significative a été observée entre les positions de consonnes dans le contexte en terme de durée.
- Il existe une stabilité de la valeur d'intensité, la fréquence fondamentale F0 et la valeur de F1 pour les consonnes nasales n et m.
- F2 et F3 sont influencés par le contexte et l'effet de coarticulation avec la voyelle qui le suit.

Bibliographie

- 1- Ali RAI (2013), Apprentissage Des Sons Spécifiques à l'Arabe Standard par Un locuteur Anglais, mémoire de magister, université d'Alger 2.
- 2- Dutoit THIERRY (2000), Introduction au traitement automatique de la parole, Faculté Polytechnique de Mons, Belgium.
- 3- Fethi BAHY (2012), apprentissage des sons spécifiques à l'arabe standard par un locuteur français, mémoire de magister, université d'Alger 2.
- 4- Fodil HEMRI (2012), Les troubles d'Articulé Dentaire chez un arabophone locuteur, mémoire de magister, université d'Alger 2.
- 5- Tuan LE MANH (2007), Analyse acoustique de sons, INRIA de Lorraine, France.
- 6- Mahieddine DJOUDI (1991), Contribution à l'étude et à la reconnaissance automatique de la parole en Arabe standard, Doctorat, Université de Nancy1.
- 7- Othman ELBEREGLI (2017), La nasalisation en langue arabe (vocalique ou consonantique), Norsud, Libye, n 10, p 129-139.