Simulation numérique des apériodicités vocales dues aux fluctuations de la tension musculaire

Jean Schoentgen[1, 2] Dhouha Rezgui[2] Francis Grenez[2]
(1) Fonds de la Recherche Scientifique, Rue d'Egmont 5, B - 1000 Bruxelles, Belgique
(2) B.E.A.M.S., Université Libre de Bruxelles, CP165/51, 50, Av. F.-D. Roosevelt, B-1050 Bruxelles, Belgique
jschoent@ulb.ac.be, rezgui_dhouha@yahoo.com,fgrenez@ulb.ac.be

RÉSUMÉ

L'objectif est le dévelopement et l'étude d'un modèle du jitter vocal destiné à la sythèse numérique de la qualité de voix. Le modèle est numériquement stable et compact et évite les hypothèses *ad hoc* d'un modèle existant. Le jitter vocal est obtenu à partir des fluctuations de la tension musculaire causées par les contractions élémentaires superposées du muscle thyro-aryténoïdien impliquant l'activité concomitante de plusieurs unités motrices. Les paramètres de contrôle sont le nombre d'unités motrices, le temps mort et le taux d'émission des neurones moteurs, ainsi que le temps latent et le temps de montée des contractions musculaires élémentaires. La présentation inclut une comparaison avec un modèle existant ainsi qu'une étude de l'influence des paramètres physiologiques sur un indice connu du jitter de la fréquence vocale.

ABSTRACT

Numerical simulation of vocal aperiodicities owing to muscle tension fluctuations.

The presentation is devoted to a model of vocal jitter, which is inspired by physiology. The simulation is numerically stable and compact so that it may be used in speech synthesis. Synthetic vocal jitter is obtained via simulated muscle tension fluctuations, which are the outcome of the superposition of the TA muscle twitches that involve the activity of several motor units. The control parameters are the number of active motor units, the dead time and firing rate of the motor neurones, as well as the rise time and latency of the muscle twitches. The presentation includes a comparison with an existing model as well as an analysis of the dependence of a popular cue of vocal jitter on physiological parameters.

MOTS-CLÉS : Jitter vocal, fluctuations de la tension musculaire, unités motrices, neurones moteurs.

KEYWORDS: Vocal frequency jitter, muscle tension fluctuations, motor units, motor neurons.

1 Introduction

L'objectif est la présentation d'un modèle des perturbations de la fréquence vocale qui est inspiré de la physiologie du muscle thyro-arythénoïdien, la comparaison à un modèle existant (Titze, 1991) et l'étude de l'influence des paramètres physiologiques sur des indices connus des perturbations de la fréquence vocale Fo. Le modèle est numériquement compact et exempt de régimes physiquement ou physiologiquement impossibles. Le temps de calcul permet son utilisation pour la synthèse vocale.

10.21437/JEP.2018-31

Kreiman a publié une énumération des causes possibles du jitter vocal, dans le sens large du terme, qui renvoie à une vaste gamme de vibrations irrégulières dont la majorité est mieux connue sous d'autres noms (Kreiman & Sidtis, 2011).
Le modèle qui est discuté ici simule les causes du jitter et du scintillement de la fréquence vocale lorsque les vrais plis vocaux vibrent pseudo-périodiquement et monodiquement dans un même mécanisme phonatoire. Lors de l'émission d'une voyelle soutenue, les perturbations de la périodicité stricte, des plus lentes aux plus rapides, sont les suivantes.

— La dérive ou la déclinaison de Fo.
— Le pleurage vocal qui désigne le tremblement physiologique causé par le flux sanguin pulsé et la respiration.
— Le tremblement vocal d'origine neurologique qui est dû à des taux d'émission non stationnaires des potentiels d'action des neurones moteurs.
— Le scintillement et le jitter vocal qui sont dus à la tension fluctuante des muscles laryngés.

R. Cook écrit que les régions dans le spectre du contour intonatif en-dessous et au-dessus du pic de vibrato sont appelées les régions du pleurage et scintillement vocal respectivement (Cook, 1999). Titze (1994) précise que le pleurage et le scintillement sont des modulations dans l'intervalle de 1 à 2 Hz et de 10 à 12 Hz respectivement. Le scintillement vocal n'est que rarement discuté dans la littérature, mais on sait qu'il influence les indices de perturbations courants (Alzamendi & Schlotthauer, 2017). Il apparaît ensemble avec le jitter dans des simulations inspirées de la physiologie du muscle vocal, ce qui suggère qu'ils ont une origine commune. En pratique, le jitter vocal rapporte les perturbations à court terme tandis que le scintillement vocal observé s'étale sur quelques cycles.
Nous utilisons ici le concept de jitter vocal dans l'acception courante du terme qui désigne la variabilité de cycle-à-cycle de la fréquence vocale ou des durées des cycles glottiques. Les indices acoustiques qui rapportent le jitter de la fréquence vocale sont à la fois populaires et critiqués. Ils sont populaires car ils ont la réputation d'être des marqueurs distinctifs de la qualité vocale d'une majorité de locuteurs. Ils sont critiqués car les utilisateurs perdent parfois de vue les conditions sous lesquelles les indices courants peuvent être obtenus fiablement, mais aussi parce que ces conditions suggèrent que le jitter vocal, dans le sens restreint utilisé ici, contribue de façon négligeable aux qualités plus extrêmes ou spectaculaires.
Les fluctuations de la tension du muscle thyro-arythénoïdien (TA) sont à même de causer des perturbations audibles de la fréquence vocale lorsque les vrais plis vocaux vibrent exclusivement, pseudo-périodiquement et monodiquement. On s'attend à ce que les fluctuations de la tension d'autres muscles, le crico-thyroïdien (CT) par exemple, contribuent moins car l'inertie du cartilage thyroïde est susceptible de les lisser.
La tension d'un muscle squelettique est l'effet de l'activité simultanée de plusieurs unités motrices. Une unité motrice est composée d'un neurone moteur qui innerve un groupe de fibres musculaires qui se contractent ensemble à l'arrivée d'une impulsion électrique, aussi appelée potentiel d'action, émise par le neurone. Cette contraction simultanée de plusieurs fibres est appelée une contraction élémentaire (ou secousse). La tension musculaire totale est la résultante de la co-existence de nombreuses secousses qui se chevauchent dans l'espace (suite à l'activité simultanée de plusieurs unités motrices) et dans le temps (suite à la succession rapide des potentiels d'action d'un seul neurone moteur.)

2 Le modèle "$\mathcal{N} - P$" des sources neurologiques des apériodicités vocales

Le modèle est ici appelé "$\mathcal{N} - P$" parce qu'il repose sur une perturbation aléatoire normale "$\mathcal{N}$" des positions "P" dans le temps des contractions musculaires élémentaires. En effet, Titze à proposé en 1991 un modèle des fluctuations de la tension du muscle TA. Les composants en sont (i) un modèle de la contraction élémentaire, (ii) la superposition des contractions élémentaires qui se suivent dans le temps simulant l'activité d'une seule unité motrice et (iii) la superposition asynchrone de plusieurs séquences de contractions simulant l'activité concomitante de plusieurs unités motrices.
Le modèle de la contraction élémentaire comprend un paramètre qui est le temps de monté T_m (Herman, 2007; Titze, 1991). La contraction $s(t)$ proprement dite démarre à l'instant $t = 0$ et son amplitude est égale à l'unité.

$$s(t) = \frac{t}{T_m} e^{1-\frac{t}{T_m}} \tag{1}$$

Les positions t_j des contractions dues aux potentiels d'action d'un seul neurone moteur sont obtenues en centrant une distribution normale $\mathcal{N}$ sur des positions équidistantes d'une durée μ qui est égale à l'inverse du taux d'émission λ en Hz d'un neurone moteur. L'écart-type σ des perturbations est fixée à l'aide du coefficient de variation $\nu = \sigma/\mu$ de la durée moyenne μ entre potentiels d'actions. La durée T_l est le temps de latence qui sépare l'arrivée du potentiel d'action et le début de la contraction élémentaire.

$$t_j = \mathcal{N}(j \times \mu, \nu \times \mu) + T_l \tag{2}$$

Finalement, la tension musculaire τ est obtenue en superposant de façon asynchrone les contractions élémentaires (1) dues à N_u unités motrices. Les positions $t_{j,k}$ des contractions élémentaires de l'unité motrice $k = 1, N_u$ sont les suivantes.

$$t_{j,k} = \mathcal{N}(j \times \mu, \nu \times \mu) + \mathcal{N}_k(0, \mu) + T_{l,k} \tag{3}$$

La distribution normale $\mathcal{N}(j \times \mu, \nu \times \mu)$ positionne les contractions musculaires dues à un neurone moteur et la distribution $\mathcal{N}_k(0, \mu)$ décale aléatoirement les positions des contractions qui sont dues à des unités motrices différentes. Dans le modèle de Titze, $\mathcal{N}$ et $\mathcal{N}_k$ sont les seuls sources de variabilité qui sont toujours présentes. Titze a exploré, en option, des variations inter-unités du taux d'émission λ et de l'amplitude de la contraction élémentaire, sans découvrir des différences qualitatives avec le cas homogène.
La fréquence vocale non perturbée $\bar{F}_o$ est fixée par l'expérimentateur et la tension $\bar{\tau}$ est la moyenne des superpositions spatiales et temporelles des contractions musculaires (1). La fréquence instantanée f_o et la racine carrée de la tension musculaire instantanée $\sqrt{\tau}$ sont supposées être proportionnelles (Titze, 1991). Par conséquent, la relation entre perturbations instantanées $f_o - \bar{F}_o$ et la tension instantanée τ est la suivante.

$$f_o - \bar{F}_o = \bar{F}_o \times (\sqrt{\tau/\bar{\tau}} - 1) \tag{4}$$

(Titze, 1991) utilisait le développement de Taylor de la relation (4). La présence du quotient $\tau/\bar{\tau}$ dans la formule des perturbations instantanées (4) rend superflue la modélisation des amplitudes absolues des contractions (1).
Le modèle "$\mathcal{N} - P$" suscite des questions théoriques et numériques qui sont discutées ici. Les réponses possibles sont à la base du modèle "$\mathcal{G} - I$" qui est expliqué ci-après.

(i) Le modèle "$\mathcal{N} - P$" est encombrant d'un point de vue numérique. Il exige le placement, la copie et l'addition de milliers d'exemplaires de la contraction (1) pour une durée de simulation d'une seconde. Aussi, la simulation des potentiels d'action exige une modélisation numérique distincte de celle des secousses. Les temps $t_{j,k} - T_l$ dans (3) fixent alors les positions des impulsions unitaires qui simulent les potentiels d'action.

(ii) La relation (2) montre que le modèle est non-physique en principe et on constate qu'il l'est en pratique dès que le coefficient de variation $\nu = \sigma/\mu$ des durées des intervalles Δt_{iii} entre secousses est > 0.2. A partir de cette valeur, on observe des durées négatives, c.-à-d. on constate que la $(n+1)_{ieme}$ secousse musculaire peut précèder la n_{ieme}. Ces erreurs ne peuvent pas être interceptées car la modélisation repose sur les positions absolues des secousses.

(iii) L'utilisation obligée du coefficient de variation ν comme paramètre de contrôle implique que le taux d'émission λ et la dispersion σ des durées inter-impulsions Δt_{iii} d'un même neurone ne sont pas indépendants. L'un diminue lorsque l'autre augmente.

(iv) La distribution Gaussienne $\mathcal{N}$ dans (2) et (3) est problématique pour une autre raison. Elle tolère des intervalles inter-impulsions Δt_{iii} plus courts que le temps de réfraction T_{refr} qui devrait empêcher un neurone d'émettre deux impulsions successives arbitrairement proches.

(v) (Titze, 1991) compare la contraction simulée (1) à des contractions observées et constate que la durée de la première dépasse de 30% voire 50% la durée des secondes. Cette observation soulève la question à savoir si la forme de la contraction influence les perturbations de Fo observées ainsi que le problème de la prolongation non-nécessaire du temps de calcul suite à l'utilisation de contractions simulées extra-longues.

(vi) L'obtention des valeurs des indices de perturbations de la fréquence vocale F_o est obscure (Titze, 1991). Il semble qu'elle implique une moyenne des fréquences instantanées f_o sur une durée fixe.

3 Le modèle "$\mathcal{G} - I$" des sources neurologiques des apériodicités vocales

Le modèle "$\mathcal{G} - I$" est inspiré du modèle "$\mathcal{N} - P$", mais évite les hypothèses *ad hoc* de celui-ci. Il est appelé "$\mathcal{G} - I$" ici parce qu'il repose sur le tirage au sort à l'aide d'une distribution $Gamma$ $(\mathcal{G})$ des durées des intervalles (I) entre potentiels d'actions simulés.

(i) Modélisation de la contraction élémentaire musculaire.
Nous avons utilisé le modèle (1) que nous comparons à une contraction triangulaire dont le temps de descente du maximum à zéro est le double du temps de monté de zéro au maximum. A temps de montée égal, la durée $(T_{1\%})$ du modèle exponentiel (1) est typiquement $2.5\times$la durée du modèle triangulaire qui épouse mieux les formes observées et qui permet de réduire le temps de calcul.

(ii) Superposition versus convolution.
Le placement à l'instant t_i de la contraction revient à décaler $s(t)$ en la recopiant. Une alternative est la convolution de $s(t)$ avec une impulsion unitaire $\delta(t - t_i)$. La superposition des contractions élémentaires est numériquement équivalente à la simulation des potentiels d'action par une superposition d'impulsions unitaires suivie d'une convolution avec la contraction $s(t)$.

(iii) Simulation des durées des intervalles inter-impulsions.
La simulation des potentiels d'action d'un neurone moteur repose sur le tirage au sort des durées des intervalles inter-impulsions Δt_{iii}. Deger (2012) montre que la distribution $\mathcal{G}(k, b)$ imite le mieux la distribution observée des Δt_{iii} d'un neurone. Les deux paramètres k et b de la distribution sont liés à la moyenne $\mu = 1/\lambda$ et au coefficient de variation ν des durées Δt_{iii} de la manière suivante.

$$k = 1/\nu^2, b = 1/\mu\nu^2 \tag{5}$$

La contrainte physique $\Delta t_{iii} \geq 0$ est satisfaite quelque soit la valeur des paramètres. La contrainte physiologique $\Delta t_{iii} > T_{refr}$ implique la condition $\nu < 1$ qui est nécessaire mais pas suffisante. C'est pourquoi, les durées Δt_{iii} tirées au hasard et $< T_{refr}$ sont omises.

(iv) Calcul des durées de cycles.
La fréquence instantanée f_o est obtenue via la relation (4). Ensuite, les durées de cycle T_o sont obtenues en cumulant le temps qui est nécessaire à la phase ϕ de croître de 2π sur base de la relation canonique $d\phi = 2\pi \times f_o \times dt$. Nous utilisons cette relation pour obtenir les durées de cycles simulées pour les modèles "$\mathcal{N} - P$" et "$\mathcal{G} - I$".

4 Méthodes

(i) Nous avons calculé l'indice J_{ppq5} des perturbations des durées de cycles afin de faciliter la comparaison avec des données publiées (Boersma & Weeninck, 2014). Nous utilisons une variante de J_{ppq5} qui remplace la moyenne courante sur 5 cycles par une moyenne courante sur un nombre de cycles qui est équivalent à une durée de $50ms$ (c.-à-d. 5 cycles lorsque $Fo = 100Hz$). L'avantage en est que la durée moyenne des cycles n'est pas impliquée explicitement dans la définition du jitter vocal.

(ii) La valeur de indice J_{ppq5} ou d'autres qui lui sont similaires n'est pas suffisante afin de comparer les perturbations simulées et naturelles. En effet, ces indices rapportent des perturbations à court terme dont les valeurs peuvent êtres très semblables pour des scintillements lents et larges ou rapides et petits et il n'existe pas de données publiées qui permettraient de trancher.

(iii) Nous avons estimé le scintillement vocal à l'aide de indice S_{10} qui rapporte les perturbations par rapport une moyenne courante de $100ms$ des durées de cycles desquelles le $jitter$ a été soustrait. La moyenne courante sur $100ms$ est conforme à la définition du scintillement vocal (Titze, 1994). On observe que le quotient J_{ppq5}/S_{10} découvre des différences entre options de modélisation qui échappent aux indices de $jitter$ conventionnels (cf. Tableau 1, colonne 9 comparée aux colonnes 4 et 7).

(iv) Afin d'offrir un point d'ancrage à la discussion des résultats de simulations, nous avons mesuré l'indice J_{ppq5} du jitter et l'indice S_{10} du scintillement des longueurs de cycles pour un corpus de voyelles [a] soutenues par 35 et 59 locuteurs dont la voix a été libellée respectivement $G = 0$ ou $R = 0$ sur les échelles $GRBAS$ (U.MA., 2018). L'analyse porte sur un intervalle de $1sec$ placé $1sec$ après l'attaque de la voyelle échantillonnée à $44kHz$. Le Tableau 1 rapporte les quartiles des indices.

(v) Les colonnes à l'extrême droite du tableau 1 rapportent les valeurs des mêmes indices pour une série de 1000 simulations des perturbations par du bruit blanc Gaussien des durées de cycles avec des valeurs aléatoires pour la fréquence vocale $\bar{Fo}$ $(100\ Hz\ - 200\ Hz)$ et du coefficient de variation ν_{Fo} $(0.1\% - 0.5\%)$.
Un test non-paramétrique de Kolmogorov-Smirnov indique que les indices pour les corpus "$R = 0$" et "$G = 0$" ne diffèrent pas statistiquement. Par contre, les quotients diffèrent statistiquement significativement entre les colonnes 4 et 9 ainsi que 7 et 9 ($p < 10^{-3}$).

	Corpus "G=0"			Corpus "R=0"			Bruit blanc	
Q	S_{10}	J_{ppq5}	J_{ppq5}/S_{10}	S_{10}	J_{ppq5}	J_{ppq5}/S_{10}	J_{ppq5}	J_{ppq5}/S_{10}
05%	0.08	0.12	0.76	0.08	0.12	0.81	0.09	2.58
25%	0.12	0.16	1.06	0.11	0.19	1.23	0.15	3.30
50%	0.15	0.20	1.26	0.15	0.24	1.53	0.22	3.92
75%	0.19	0.26	1.72	0.20	0.29	2.25	0.29	4.42
95%	0.30	0.39	2.47	0.34	0.48	3.25	0.36	5.17

Tableau 1 : Quantiles des indices du jitter (J_{ppq5} en %) et scintillement (S_{10} en %) vocal pour deux corpus de voyelles [a]. Colonnes à l'extrême droite : simulation des perturbations par du bruit blanc Gaussien des durées de cycles.

(vi) Deux séries de 1000 simulations chacune ont été réalisées afin d'étudier les modèles "$\mathcal{G} - I$" et "$\mathcal{N} - P$" et de les comparer. Les modèles sont implémentés dans $Python$. La fréquence d'échantillonnage est à 200 kHz, la durée d'une simulation est une sec et la fréquence vocale moyenne $\bar{F_o}$ est à 100 Hz recto tono. Au début de chaque simulation, les valeurs des paramètres de contrôle sont choisies au hasard dans un intervalle dont les étendues relatives sont identiques afin de faciliter la comparaison entre paramètres (Tableau 2).

	Paramètre	Valeur typique	Intervalle	Références
Temps mort	T_{refr}	2.5 ms	1.25-3.75 ms	(Roark *et al.*, 2002)
Taux d'émission	λ	30 Hz	15-45 Hz	(Roark *et al.*, 2002)
Ecart-type$\times\lambda$	ν	0.1	0.05-0.15	(Titze, 1991)
Nombre U.M.	N_u	100	50-150	(Titze, 1991)
Latence	T_l	15 ms	7.5-22.5 ms	(Titze, 1991)

Tableau 2 : Valeurs typiques et intervalles des paramètres de modélisation.

En outre, la contraction triangulaire ou la contraction exponentielle (1) a été choisie au hasard pour chaque simulation ($p = 1/2$). Tous les paramètres sont identiques pour toutes les unités motrices qui sont co-actives lors d'une simulation. Le temps de montée T_m est fixé à 0.02 sec pour toutes les unités motrices et toutes les simulations. Aucun effort n'a été fait pour ajuster les modèles aux indices observés pour des locuteurs.

(vii) La contribution relative de chaque paramètre est quantifiée à l'aide d'une analyse par régression linéaire multiple des variables z-normalisées. L'inverse S_{10}^{-1} et J_{ppq5}^{-1} des indices est utilisé comme variable dépendante afin de mieux linéariser le lien entre variables. Tous les tests statistiques sont réalisés à l'aide du logiciel R.

5 Simulations et discussion

(i) Une série préliminaire de 100 simulations avait comme objectif la comparaison de deux versions du modèle "$\mathcal{N} - P$". La version originale somme les contractions élémentaires directement et la version alternative convolue les potentiels d'action simulés avec la contraction $s(t)$. Les tensions musculaires calculées sont identiques pour les deux versions (corrélation $\equiv$ 1), mais les temps de calcul diffèrent approximativement d'un ordre de grandeur. C'est pourquoi, toutes les simulations qui sont rapportées ci-après font appel à la convolution. Le gain en temps de calcul découle de l'implémention efficace de la convolution dans la bibliothèque *numpy* et ne peut pas être généralisé à d'autres logiciels.

(ii) Une série de 1000 simulations chacune a été réalisée avec les modèles "$\mathcal{G} - I$" et "$\mathcal{N} - P$". Le Tableau 3 montre les quartiles des indices. Un test de Kolmogorov-Smirnov indique que le quotient J_{ppq5}/S_{10} diffère statistiquement significativement entre "$\mathcal{N} - P$" et "$\mathcal{G} - I$" ($p < 10^{-3}$).

	"$\mathcal{G} - I$"			"$\mathcal{N} - P$"		
Q	$S_{10}(\%)$	$J_{ppq5}(\%)$	J_{ppq5}/S_{10}	$S_{10}(\%)$	$J_{ppq5}(\%)$	J_{ppq5}/S_{10}
05%	0.13	0.16	0.74	0.10	0.17	1.07
25%	0.21	0.29	1.23	0.18	0.31	1.38
50%	0.29	0.52	1.79	0.28	0.53	1.75
75%	0.40	0.84	2.44	0.45	0.87	2.36
95%	0.75	1.75	3.76	0.89	1.67	3.71

Tableau 3 : Quantiles des indices obtenus pour deux modèles du jitter vocal.

(iii) Le Tableau 4 rapporte les coefficients de la régression des indices S_{10}^{-1}, J_{ppq5}^{-1} et J_{ppq5}/S_{10} sur les paramètres des modèles. Le paramètre $\mathbb{C}$ réfère au type de la contraction (triangulaire = "1", exponentielle = "0"). La dernière ligne rapporte la qualité du modèle de régression. Les coefficients de régression dont seulement l'ordre de grandeur est donné sont statistiquement non-significatifs.

	"$\mathcal{G} - I$"			"$\mathcal{N} - P$"		
Paramètres	S_{10}^{-1}	J_{ppq5}^{-1}	J_{ppq5}/S_{10}	S_{10}^{-1}	J_{ppq5}^{-1}	J_{ppq5}/S_{10}
T_l	*10^{-2}*	*10^{-2}*	*10^{-2}*	*10^{-2}*	*10^{-2}*	*10^{-3}*
T_{refr}	*10^{-3}*	*10^{-2}*	*10^{-3}*	*10^{-3}*	*10^{-3}*	*10^{-3}*
λ	$+0.52$	$+0.81$	-0.55	$+0.67$	$+0.70$	-0.15
N_u	$+0.31$	$+0.19$	*10^{-3}*	$+0.24$	$+0.23$	-0.06
ν	-0.57	-0.64	-0.50	-0.40	-0.15	-0.51
$\mathbb{C}$	-0.28	-0.27	-0.13	-0.35	-0.33	-0.06
R^2 ajusté	*0.77*	*0.75*	*0.56*	*0.82*	*0.69*	*0.29*

Tableau 4 : Coefficients de régression linéaire multiples des paramètres z-normalisées.

(iv) L'utilisation du coefficient de variation ν qui dépend du taux d'émission λ via $\nu = \sigma \times \lambda$ facilite le contrôle de la causalité du modèle "$\mathcal{N} - P$" et a été retenue pour le modèle "$\mathcal{G} - I$" à des fins de

comparaison.
En inspectant les coefficients de régression (Tableau 4), on constate que les indices de perturbations S_{10} et J_{ppq5} augmentent toujours avec la dispersion σ des durées inter-impulsions et diminuent avec le taux d'émission λ à condition que la dispersion est faible.
Les indices de perturbation augmentent également lorsque le nombre d'unités motrices diminue et lorsque la secousse triangulaire remplace la secousse exponentielle (1). Une explication possible est que la première est plus courte et moins lisse que la deuxième.

(v) Le quotient J_{ppq5}/S_{10} diminue lorsque la dispersion σ ou le taux d'émission λ augmentent. En d'autres mots, l'observation de valeurs > 1 du quotient implique que les rafales d'impulsions des neurones moteurs sont pseudo-périodiques. Pour rappel, la taille du jitter vocal est supérieure à la taille du scintillement vocal pour la parole naturelle (cf. Tableau 1). Titze (1991) ne discute pas le rapport entre jitter et scintillement vocal car l'hypothèse que les potentiels d'action sont pseudo-périodiques est implicite au modèle "$\mathcal{N} - P$". En effet, la condition ν«1 est garante de la causalité du modèle. Le quotient J_{ppq5}/S_{10} augmente aussi lorsque la contraction exponentielle remplace la contraction triangulaire, mais l'impact est faible.

(vi) Le temps de latence de la contraction et le temps de réfraction du neurone moteur n'ont pas d'influence significative. Une explication possible est que le temps de latence est masqué par le terme aléatoire $\mathcal{N}_k$ dans la relation (3) et qu'il est peu probable qu'une durée inter-impulsions Δt_{iii} aussi petite que le temps de réfraction soit observée lorsque le coefficient de variation ν des durées Δt_{iii} est < 0.15.

(vii) Les valeurs des indices S_{10} et J_{ppq5} ont les mêmes ordres de grandeur pour les simulations et les observations sur des locuteurs (tableaux 1 et 3), mais leurs étendues diffèrent. En effet, rien n'a été fait pour ajuster les paramètres de manière à ce que les modèles reproduisent les valeurs observées. Aussi, (Titze, 1991) ne discute pas la possibilité que les perturbations des fréquences instantanées f_o soient modulées à l'intérieur des plis et que les perturbations qui sont simulées par "$N - P$" ou "$G - I$" ne soient pas observables directement. En effet, des conditions laryngées (bénignes) existent qui sont connues influencer le jitter vocal mesuré sans pour cela influencer l'activité des unités motrices. Il se peut, qu'une explication possible est inhérente au modèle $muscle - couverture$ des plis vocaux. Un modèle simple qui n'est pas reproduit ici faute de place, mais qui tient compte des amplitudes de vibration du muscle A_m et de la couverture A_c des plis, suggère que les perturbations instantanées $f_o - \bar{F}_o$ sont multipliées par un terme $A_m/(A_m + A_c)$. Ce terme est égal à l'unité lorsque $A_c = 0$ et égal à zéro lorsque $A_m = 0$. Il suggère donc que les perturbations instantanées sont modulées par la mobilité relative du muscle et de la couverture.

Références

ALZAMENDI G. H. & SCHLOTTHAUER G. (2017). *Describing Voice Period Variability by means of Time Series Structural Analysis*. Proceedings 10th International Workshop : Models and Analysis of Vocal Emissions for Biomedical Applications, Firence, Italy, pages 11-14.

BOERSMA P. & WEENINCK D. (2014). *Praat : doing phonetics by computer [Computer program]*. [Version 5.4.04, retrieved 2014 from http ://www.praat.org].

P. R. COOK, Ed. (1999). *Music, Cognition and Computerized Sound :An Introduction to Psychoacoustics*. Cambridge, Massachusetts : The MIT Press. page 199.

DEGER M., HELIAS M., BOUCSEIN C. & ROTTER S. (2012). Statistical properties of superimposed stationary spike trains. *J. Comput. Neurosci.*, **32**, 443–463.

HERMAN I. P. (2007). *Physics of the Human Body*. Berlin, Heidelberg : Springer. page 281.

KREIMAN J. & SIDTIS D. (2011). *Foundations of Voice Studies : An Interdisciplinary Approach to Voice Production and Perception*. Wiley-Blackwell. page 55.

ROARK R., C.L., LI J., SCHAEFER S., ADAM A. & LUCA C. D. (2002). Multiple motor unit recordings of laryngeal muscles : The technique of vector laryngeal electromyography. *The Laryngoscope*, **112**, 2196–2202.

TITZE I. R. (1991). A model for neurologic sources of aperiodicity in vocal fold vibration. *J. Speech, Hearing Res.*, **34**, 460–472.

TITZE I. R. (1994). *Principles of Voice Production*. Englewood Cliffs, NJ : Prentice Hall. page 332.

U.MA. (2018). *http ://www.atic.uma.es/index_atic.html*. [Online ; accessed 23-January-2018].