

Étude de la contribution acoustique de la structure formantique à l'identification du ton chuchoté

Xuelu Zhang¹, Rudolph Sock^{1,2}

- (1) U.R.1339 LILPA/Equipe de Recherche Parole et Cognition & Institut de Phonétique de Strasbourg (IPS), Université de Strasbourg (UNISTRA), 22, rue Descartes, 67084 Strasbourg, France
- (2) Language, Information and Communication Laboratory (LICOLAB), Faculté des Lettres, Université Pavla Jozefa Safarika, Šrobárova 2, 040 59 Košice, Slovaquie

zhang_xuelu@yahoo.fr

RESUME

Cette étude examine la contribution de la structure formantique du segment vocalique à l'identification du ton que ce segment porte, et cela en voix chuchotée. Le mandarin a été choisi en tant que langue cible parce que les traits tonals (*tone features*) en mandarin s'appuient acoustiquement sur deux dimensions : le registre et le contour. Nous supposons qu'en l'absence de F0, la structure formantique subirait néanmoins une modification, en fonction du ton et fournirait des indices acoustiques des traits tonals à l'auditeur. Nous nous intéressons aux rapports entre les deux dimensions de traits tonals et à la modification de la structure formantique. À travers l'analyse des données acoustiques issues de 13 sujets locutrices, nous avons observé une divergence d'importance dans les intervalles F2-F3 et F3-F4, en fonction du ton. Cette divergence semble liée aux contrastes tonals en registre et non aux contours mélodiques. Cette distinction semble dépendre d'ailleurs de la nature de voyelle.

ABSTRACT

A study of the acoustic contribution of formant structure to tone identification in whispered speech

This study explores the contribution of vowel formant structure to tone identification, particularly in whispered speech. Mandarin was chosen for this study as the target language, as its tone features are acoustically built up on two dimensions : register and contour. We posit that in case of absence of F0, the formant structure would always undergo modification associated to tone, and hence provide acoustical tone cues to the listener. Also, we are interested in relations between the two dimensions of tone features and modifications of formant structures. By analysing the acoustic data collected from 13 female speakers, we found remarkable differences between F2-F3 and F3-F4 values, with regards to a specific tone. Such differences seem to be linked to tone register contrasts and not to melodic contours. Furthermore, this divergence in formant pattern intervals seems to depend on the nature of the vowel.

MOTS-CLES : Identification du ton, mandarin, traits tonals, structure formantique, sensation psychoacoustique de la hauteur de la voix, voix chuchotée.

KEYWORDS : Tone identification, Mandarin, tone features, formant structure, psycho-acoustical sensation of pitch, whispered voice.

1 Introduction

Le ton lexical est un phénomène bien étudié dans le domaine de la phonologie. En tant que composante de la structure lexicale, le ton permet de réaliser une distinction au niveau sémantique à travers des patterns figés de modification mélodique.

Ce phénomène suprasegmental varie d'une langue à l'autre. Il est pourtant possible de décomposer tous types de tons en traits tonals (*tone features*). Selon le pattern de modification du pitch, un système tonal pouvait être catégorisé en tant que système de registre (*Level-Pitch Register System*), système de contour (*Gliding-Pitch Contour System*) ou système qui combine ces deux premiers (*Register-Contour Combination*). Depuis le développement de la phonologie autosegmentale, la notion du système de contour est remise en question : on considère que le contour tonal peut être subdivisé davantage. Goldsmith considérait le contour tonal comme une séquence de tons de registre, et l'a décrit en employant des traits binaires. Jusqu'aujourd'hui, beaucoup d'études sur les tons dans des langues asiatiques emploient cette description à traits binaires, mais aussi un système de traits tonals à quatre registres que Yip a proposé en 1980. Dans les études sur le mandarin et sur d'autres variantes du chinois, le système tonal de Chao Y.R. à cinq partitions (soit cinq registres) est toujours d'actualité. Dans toutes ces descriptions, le registre reste la base des traits tonals : le ton est défini par des registres cibles à atteindre.

Cependant, d'autres phonologues insistent sur une description des tons à contours avec la direction du mouvement du pitch en tant que trait tonal. Yip (2001) a révisé le système de traits tonals à registres tout en prenant en compte la réalisation phonétique de différents types de tons, et elle est arrivée à la conclusion que la description des tons à contours, selon un tel système, n'est pas satisfaisante. Phonétiquement, un ton à contour consiste plutôt en une cible à atteindre et un éloignement de la cible, alors qu'un ton à registre se centre sur les cibles et non sur le mouvement.

Nous nous intéressons donc à cette problématique de représentation des traits tonals, et nous considérons qu'il serait probablement nécessaire de repenser la définition des traits tonals suivant l'analyse des réalisations phonétiques du ton. Dans cette étude, nous avons choisi le mandarin en tant que langue cible, parce que dans cette langue monosyllabique, le ton se réalise dans le même cadre temporel que la syllabe, ce qui facilite l'observation et l'analyse. De plus, le système tonal du mandarin présente une combinaison de tons à registres et de tons à contours, et étant donné ce fait complexe, il connaît effectivement des désaccords dans sa description.

Le mandarin possède quatre tons lexicaux. Les études phonétiques et phonologiques existantes ont montré que ces tons se produisent dans deux partitions de toute l'étendue du pitch tonal (*pitch range*). Le 1er ton et le 4ème ton se réalisent, durant toute syllabe, dans une partition plus haute alors que le 2ème et le 3ème ton se réalisent dans une partition plus basse. Cette division est décrite comme [+/-Upper] (+/-U). Les registres cibles des deux tons dans chaque partition sont en contraste,

et leurs contrastes peuvent être transcrits par [high/low]¹ (h/l). Les quatre tons du mandarin peuvent être transcrits comme suit :

- Ton 1(T1) : [+U, h]
- Ton 2(T2) : [-U, h]
- Ton 3(T3) : [-U, l]
- Ton 4(T4) : [+U, l]

L'analyse dans cette recherche reposera principalement sur cette description des tons du mandarin. Et nous tenterons de remettre ces traits tonals en question en l'absence de F0 (le paramètre acoustique principal de la perception du ton). Autrement-dit, le but de cette recherche est de trouver des indices acoustiques des traits tonals, lorsque ces derniers ne peuvent plus se réaliser à travers F0. Afin d'observer des stratégies de compensation spontanée déployées par le locuteur, nous avons exigé des productions en voix chuchotée, lors du collectage des données acoustiques. Il s'avère que lorsque des natifs qui pratiquent une langue tonale communiquent en chuchotant, ils arrivent à se comprendre avec beaucoup d'aisance, malgré la perte d'informations tonales provoquée par l'absence de F0. Ils doivent donc avoir trouvé un autre moyen de coder-décoder les informations relatives aux tons. Et cela se ferait, d'après nous, au niveau articulatoire-acoustique grâce à un processus de réafférences sensorielles.

2 Compensation de la perte de F0 en voix chuchotée

La littérature phonétique, dans le domaine du discours chuchoté, en liaison avec les tons, est assez limitée et relativement récente, même si elle est pourtant fascinante. Les chercheurs qui ont étudié la prosodie dans le discours chuchoté ont découvert, en examinant divers paramètres acoustiques, des indices liés aux tons. Selon Segerbäck (1965), l'activité vibratoire des plis vocaux ne serait même pas indispensable pour toute intercompréhension dans une langue tonale. L'auditeur sinophone utilise des indices secondaires (étant donné que F0 fournit des indices primaires) de manière flexible dans le jugement du ton (Liu et al., 2004). La durée du segment vocalique, par exemple, varie en fonction du ton et cette variation pourrait être significative pour la perception du ton (Blicher et al., 1990). Un autre exemple est le contour de l'intensité qui est corrélé avec F0. Il fournirait également des informations tonales (Whalen et al., 1992 ; Chang et al., 2007).

Nous portons un intérêt particulier aux rapports entre la structure formantique du segment vocalique et le ton que ce dernier porte. Malgré le fait que la structure formantique n'influence pas de manière systématique le jugement du ton perçu (Tseng et al., 1986), une modification du spectre existe réellement lorsque le ton varie. Cela implique une modification de la forme du conduit vocal au niveau supraglottique qui, bien-entendu, peut être issue du mouvement vertical du larynx en voix modale. Cependant, l'amplitude des mouvements du larynx est remarquablement réduite lorsqu'on chuchote. Ainsi, l'emploi de la voix chuchotée peut nous fournir de bonnes conditions pour l'observation des gestes liés au ton. Nous supposons que si la structure formantique connaissait une modification en fonction du ton, même en voix chuchotée, cela apporterait davantage d'arguments à l'existence de gestes articulatoires au niveau supraglottique spécifiques, corrélés au ton, ou bien, aux traits tonals.

De plus, du point de vue psychoacoustique, la structure formantique a le potentiel de transmettre des

¹ Si on considère le contour du pitch comme trait principal ici, on peut le décrire par [fall/rise] (Wang, 1967, cité par Zhang, 2014).

informations tonales par la reproduction d'une «impression» musicale. La hauteur du son de la parole que l'humain perçoit est le résultat d'un traitement du son complexe dans le système auditif périphérique et central. On sait que lorsqu'un locuteur emploie une voix modale, le signal de la parole comporte des propriétés, certes hautement complexes, mais périodiques (y compris sinusoïdales), et la hauteur du son perçue chez l'auditeur dépend principalement (mais pas exclusivement) de la réponse du nerf auditif à la stimulation de F0. Et lorsque le locuteur chuchote, le signal de la parole ne contient plus ces propriétés acoustiques périodiques (dont les composantes sinusoïdales), mais plutôt un bruit de bande, modulé en amplitude (de manière irrégulière) et en fréquence. La perception d'un tel bruit chez l'auditeur ressemble à celle d'un son sinusoïdal, à condition que la bande soit assez étroite. Et pourtant, à cause de l'absence de F0, il n'y a plus de réponse du nerf auditif corrélée directement à la hauteur musicale. La hauteur perçue dans ce cas dépend des réponses de plusieurs fibres nerveuses au stimulus acoustique. Plus précisément, les impulsions sur les fibres nerveuses répondent aux fréquences centrales des bandes du bruit. Ces premières diffèrent en durée du cycle. Ainsi dans le système auditif, «...quand on module en fréquence des bruits de ce type (bruit de bande), on fait varier leur fréquence limite et cette modification peut être perçue.» (Zwicker et al., 1981). Cette conclusion a été soutenue par des recherches ultérieures sur la contribution du TFS (*temporal fine structure*) à la perception du pitch (à titre d'exemple, voir Kong et al., 2006).

Ce bruit de bande que la voix chuchotée produit est issu de l'effet de filtrage que le conduit vocal a sur le bruit blanc produit au niveau du larynx. Par rapport à la voix modale, le spectre de la voix chuchotée est différent selon les points suivants : 1) les valeurs formantiques sont plus élevées ; 2) l'énergie se diffuse davantage à des hautes fréquences ; 3) l'intensité acoustique est d'environ 20 dB au-dessous de celle de la voix modale. Ces différences impliquent que la redistribution de l'énergie sonore dans le spectre est à prendre en compte dans l'analyse d'une voix chuchotée. Et on se demande ainsi, quels sont les formants les plus affectés par le ton. Sur cette dernière problématique, les études existantes se contredisent. Higashikawa et al. (1999) supposent que la perception du ton chuchoté est influencée par une variation simultanée du F1 et du F2, alors que Meyer-Eppler (1957) et Fónagy (1969) considéraient que le déplacement des formants à des hautes fréquences vers le haut, tels le F3 et le F4, indique une courbe mélodique montante. Leurs données avaient montré, de manière générale, que le contour mélodique était le trait tonal le plus associé à la modification de la structure formantique.

3 Réflexions et hypothèses

Comme il est mentionné ci-dessus, nous portons un intérêt particulier aux rapports entre la modification de la structure formantique et les traits tonals. Cet intérêt est dû principalement au fonctionnement du système auditif dans la perception du bruit de bande et à la distribution de l'énergie sonore dans le spectre de la voix chuchotée. Vu que la hauteur mélodique du bruit de bande ne se réalise pas à travers une seule bande formantique mais plusieurs, l'un des objectifs de cette étude est de mettre au jour les formants qui seraient les plus concernés par un ton chuchoté. L'autre objectif est de mettre en lumière le trait tonal/les traits tonals suivant lequel/lesquels s'oriente la modification de la structure formantique.

Concernant les formants les plus liés au ton, et malgré la position de Higashikawa et al. (1999), nous sommes assez en ligne avec les conclusions de Meyer-Eppler et celles de Fónagy, pour les raisons suivantes : 1) Dans le spectre du signal acoustique du chuchotement, nous avons observé plus d'énergie présente au niveau du F3, du F4 et du F5, alors que peu d'énergie se présente au niveau du F1 ; 2) la variation du F1 et celle du F2 subissent des contraintes que les frontières des

phonèmes vocaliques imposent au niveau articulatoire, ainsi ces formants peuvent être peu efficaces pour indiquer les contrastes tonals, surtout lorsqu'il s'agit d'une langue tonale à vocalisme riche. Etant donné que le « pitch fantôme » de la voix chuchotée, qui n'existe pas dans le spectre mais est reconstitué dans le système perceptif, semble être lié à la modulation en fréquence dans le bruit, nous supposons que la hauteur et la variation de ce « pitch fantôme » se réaliseraient à travers les intervalles entre les fréquences centrales des formants voisins, probablement celui entre les formants à hautes fréquences.

En ce qui concerne le(s) trait(s) tonal(s) qui se présente(nt) en voix chuchotée, nous supposons que la modification de la structure formantique est capable de réaliser, non seulement le contraste en contour, mais aussi celui en registre. Une langue tonale dont les tons sont en contraste uniquement au niveau du registre n'empêche pas les utilisateurs de cette langue de chuchoter ces tons, alors les informations des registres ont la possibilité de se diffuser à travers la structure formantique².

4 Protocole expérimental

Les résultats de cette étude s'appuient sur l'analyse de données acoustiques recueillies dans le cadre d'une expérience de production orale lancée en mars 2013 à Beijing en Chine. 13 locutrices d'origine chinoise, âgées entre 18 et 21 ans, ont pris part à l'expérience. Elles maîtrisent parfaitement le mandarin. Pour l'expérience, nous avons établi un corpus contenant 12 phrases en mandarin, chacune portant respectivement les syllabes /pa//pi//pu/, avec l'un des quatre tons dans le même contexte phonétique, soit /pV_tsɿ/. Aucun sandhi tonal ne s'est produit dans cette expérience.

Chaque sujet locutrice a produit 10 répétitions de ce corpus, respectivement en voix normale et en voix chuchotée. L'enregistrement des voix des locutrices a été réalisé dans un environnement silencieux, avec un enregistreur Marantz Professional© PMD661 et un microphone Sennheiser e845 S®. Le recueil des données a été réalisé sur PRAAT et le traitement statistique des données a été réalisé sur Excel 2010 et sur GraphPad Prism 5.

5 Résultats expérimentaux

5.1 Analyse des fréquences centrales des bandes formantiques

Afin de suivre la variation spectrale à l'intérieur du segment vocalique, nous avons pris les valeurs des fréquences centrales des bandes formantiques aux points 0%, 20%, 40%, 60%, 80% et 100% de la durée totale du segment et ainsi défini les courbes de variation spectrale en temps normalisé. Des analyses ANOVA à deux facteurs ont été conduites en considérant la moyenne des 10 répétitions de chaque sujet locutrice comme une répétition non appariée, et comme deux facteurs analysés, le *ton* et le *temps normalisé*. Pour les analyses ANOVA, nous donnons les données de variance avec le *ratio F* correspondant à la variabilité entre les sujets et la valeur de *p*. Seuls les résultats avec une probabilité de moins de 5% ($p < 0,05$) sont considérés comme significatifs.

Pour une raison de commodité, nous prenons ici la voyelle /a/ chuchotée à titre illustratif (les

² Dans la littérature (ex. Whalen et al., 1992 ; Blicher et al., 1990), la divergence d'autres paramètres acoustiques, tels que l'intensité et la durée tonale, nous semble associée davantage au contraste en contour tonal et moins au contraste en registre.

résultats de l'analyse de /i/ et de /u/ sont proches de ceux de /a/, quel que soit le sujet). En outre, nous renvoyons le lecteur à une de nos études connexes qui présente des données pour des productions en voix normale (Zhang et al., 2015). Cette étude montre des variations des paramètres acoustiques, en fonction du ton, aussi bien en voix normale qu'en voix chuchotée.

La figure 1 présente les courbes de variation des premiers quatre formants de /a/ aux quatre tons du mandarin. En observant les courbes de variation spectrale en fonction du ton sur chaque graphique, nous concluons que le ton a effectivement tendance à influencer la structure formantique. Pourtant les rapports entre les traits tonals et la variation formantique ne sont pas linéaires. Le trait [+U], qui sépare T1 et T4 de T2 et de T3, semble avoir fait augmenter les valeurs formantiques du F2, F3 et F4. Cette augmentation est relativement faible par rapport à la valeur absolue du formant en question, surtout par rapport à celles des formants à hautes fréquences, et cela s'est effectivement montré non-significatif ($F(1,21)=0,0015$, $p=0,9691$ pour T1 et T4, et $F(1,21)=0,2316$, $p=0,2316$ pour T2 et T3). Le trait [h/l], qui sépare T1 de T4 et T2 de T3, semble avoir légèrement plus d'effet en séparant T2 de T3, et moins d'effet en séparant T1 de T4. D'ailleurs il est difficile de préciser si l'effet du contraste [h/l] sur la variation formantique se manifeste tout au long du segment ou seulement à la fin du segment. Le résultat des analyses statistiques a montré que cet effet n'est pas significatif ($F(1,20)=0,5120$, $p=0,4758$ pour T1 et T4, et $F(1,22)=0,9885$, $p=0,3221$ pour T2 et T3).

L'observation de la variation des formants ne nous révèle pas de résultats significatifs relatifs aux rapports entre les traits tonals et les valeurs absolues des formants. Nous allons donc nous tourner vers une observation des intervalles entre les fréquences centrales des formants.

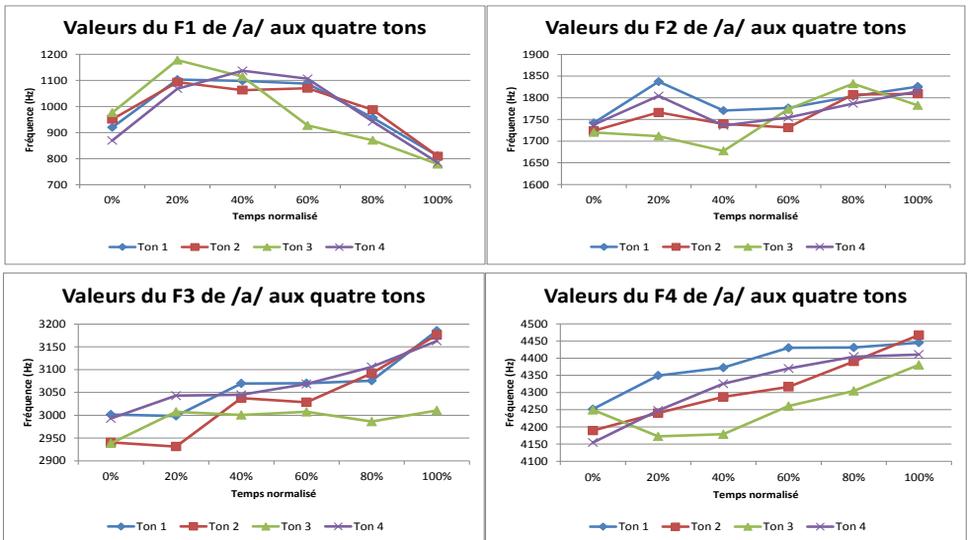


FIGURE 1 : Divergence des quatre premiers formants de /a/ portant quatre tons (cas du sujet FSY).

5.2 Analyse des intervalles entre les fréquences centrales des formants

Dans l'objectif de quantifier l'impression psychoacoustique que la structure formantique pourrait laisser au système auditif, l'analyse des intervalles entre les fréquences centrales des formants repose sur l'emploi de l'échelle *mel*. La conversion de Hertz en *mel* se fait avec la formule bien connue suivante :

$$m = 2595 * \log_{10} \left(1 + \frac{f}{700} \right)$$

À titre illustratif des résultats de cette analyse, nous prenons encore le cas de la voyelle /a/. La figure 2 illustre la variation des intervalles entre les fréquences centrales des bandes formantiques voisines de cette voyelle, en fonction du ton.

En observant les courbes sur chaque graphique, nous constatons que les courbes correspondant aux différents tons se distinguent davantage en largeur de l'intervalle (mesuré en *mel*) et peu en contour. Les différences entre elles sont maximales au milieu du segment vocalique. De plus, en comparant les courbes des trois graphiques, nous avons observé que les intervalles F1-F2 et F2-F3 de /a/ aux tons au trait [+U] sont plus importants que ceux aux tons au trait [-U] (sauf pour ce qui concerne la similarité entre l'intervalle F2-F3 de T2 et celui de T4). Cependant, l'intervalle F3-F4 entre les tons est à l'inverse de cet ordre ; les intervalles F1-F2 et F2-F3 de /a/ aux tons au trait [+U] sont, cette fois-ci, moins importants que ceux aux tons au trait [-U]. Le trait [h/l] semble avoir fait séparer davantage les courbes, et selon ces graphiques cette séparation se manifeste davantage au milieu du segment. Une remarque importante s'impose ici. Par commodité, les résultats de /i/ et de /u/ ne sont pas donnés ici. Toutefois, signalons que l'effet du trait [+/-U] sur les intervalles F3-F4 de ces deux voyelles est différent de celui sur /a/ : T1 et T4 provoquent des intervalles plus importants entre le F3 et le F4, par rapport à T2 et T3.

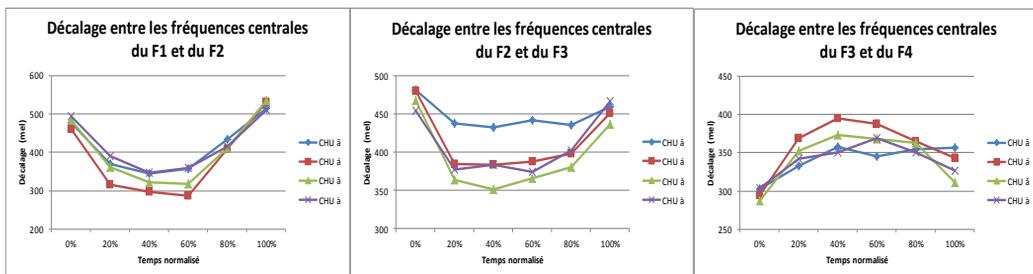


FIGURE 2 : Les intervalles entre les fréquences centrales des bandes formantiques voisines de /a/ en fonction du ton (cas du sujet ZBJ) : /ā/ =/a/+T1, /á/ =/a/+T2, /ǎ/ =/a/+T3, /à/ =/a/+T4.

Pour connaître davantage l'influence du ton sur la structure formantique, nous avons effectué une analyse statistique de tous ces intervalles entre les formants en fonction du ton. Le résultat des analyses statistiques ($p=ns$) montre que la variation du ton ne provoque pas systématiquement de différences significatives sur les intervalles entre les formants, et pourtant ces intervalles ont des sensibilités différentes à la variation tonale. Le tableau 1 illustre le nombre de sujets (sur la base de 13 sujets) qui effectuent des différences significatives lorsque le ton varie de l'un (marqué en rang, transcrit en traits tonals) à l'autre (marqué en colonne, transcrit en traits tonals).

En comparant les cas des trois voyelles dans ce tableau, nous avons remarqué, chez les 13 sujets, que la variation tonale provoque plus de différences significatives ($p<0,05$), dans l'intervalle entre les formants dans la réalisation de /a/, et moins dans la réalisation de /i/ (par ex, des différences significatives ont été trouvées dans l'intervalle F3-F4 chez 7 sujets sur 13, lorsqu'elles réalisent /a/, respectivement à T1 et à T3, alors 4 sujets sur 13 ont fait des différences significatives dans l'intervalle F3-F4 pour distinguer /i/ à T1 de /i/ à T3). En comparant les intervalles entre différentes paires de formants, nous avons remarqué que l'intervalle F3-F4 est le plus sensible au contraste tonal, alors que l'intervalle F1-F2 est le moins sensible (ex. le contraste entre T1 et T3 sur /u/

suscite des différences significatives dans l'intervalle F3-F4 chez 5 sujets, dans l'intervalle F2-F3 chez 1 sujet, et dans l'intervalle F1-F2 chez aucun sujet).

Voyelle /a/	Tons	[+U,h]	[-U,h]	[-U,l]	[+U,l]	Voyelle /i/	Tons	[+U,h]	[-U,h]	[-U,l]	[+U,l]	Voyelle /u/	Tons	[+U,h]	[-U,h]	[-U,l]	[+U,l]		
F3-F4	[+U,h]		5	7	5	F3-F4	[+U,h]		3	4	1	F3-F4	[+U,h]		6	5	6		
	[-U,h]			7	5		[-U,h]			1	2		[-U,h]				6	2	
	[-U,l]				5		[-U,l]				2		[-U,l]						5
	[+U,l]						[+U,l]						[+U,l]						
F2-F3	[+U,h]		3	6	5	F2-F3	[+U,h]		3	0	2	F2-F3	[+U,h]		3	1	5		
	[-U,h]			5	3		[-U,h]			2	4		[-U,h]				1	0	
	[-U,l]				2		[-U,l]				1		[-U,l]						3
	[+U,l]						[+U,l]						[+U,l]						
F1-F2	[+U,h]		3	3	2	F1-F2	[+U,h]		2	0	2	F1-F2	[+U,h]		2	0	3		
	[-U,h]			2	2		[-U,h]				2		[-U,h]				0	1	
	[-U,l]				2		[-U,l]				1		[-U,l]						1
	[+U,l]						[+U,l]						[+U,l]						

TABLEAU 1 : Les nombres de sujets qui font de différences significatives (sur la base de 13 sujets locutrices) en intervalles entre les formants, lorsque le ton varie (les tons concernés par la variation, marqués en rang et en colonne, sont transcrits en traits tonals) : les cas de /a/ /i/ /u/.

6 Conclusion et perspective

À travers l'analyse des données acoustiques, nous avons trouvé qu'en mandarin, le ton du segment vocalique chuchoté n'influence pas les fréquences centrales des bandes formantiques, de manière linéaire. Toutefois, ce ton a effectivement une influence sur la structure formantique, et cela se manifeste plutôt dans les intervalles entre les formants voisins qui sont, selon la littérature en psychoacoustique, associés à la sensation de la hauteur du son. Les analyses statistiques révèlent que l'intervalle entre le F3 et le F4, et celui entre le F2 et le F3 connaissent le plus souvent une modification significative ou, au moins, une tendance remarquable lorsque le ton varie. La différence maximale que le ton peut provoquer en intervalle entre les formants, calculée en *mel*, se réalise au milieu du segment, et non au début ni à la fin du segment. Ainsi, la modification de la structure formantique nous semble davantage s'orienter vers les traits tonals du type registre et non vers les traits tonals du type contour. Ce résultat corrobore, jusqu'à un certain degré, la description phonologique des tons à base de registre. Bien entendu, il est probable que le contour tonal se manifeste au moyen d'autres paramètres acoustiques, tel que l'intensité.

De plus, en comparant des cas de modification spectrale en fonction du ton de trois voyelles /a i u/, nous avons constaté que les structures formantiques de ces voyelles ont des sensibilités différentes par rapport à la variation du ton, plus précisément, dans l'ordre /a/ > /u/ > /i/. La modification spectrale en fonction du ton dépend également de la voyelle en question. Cela est probablement la conséquence des différences de gestes et de configurations du conduit vocal, donc au niveau supraglottique, dans la compensation de la perte de F0.

Malgré les conclusions présentées ci-dessus, l'étude est loin d'être achevée, car il nous reste deux questions principales à aborder : 1) Si l'on emploie un système tonal à base de registre, les deux couches de traits tonals [+/-U] et [h/l] auront-elles une différence ou un ordre hiérarchique dans leurs influences sur la modification de la structure formantique ? 2) Puisque la modification spectrale en fonction du ton dépend de la nature de la voyelle, quel/quels est/sont le/les aspect/aspects de cette modification qui aura/auront le poids le plus lourd dans la perception de la hauteur de la voix chuchotée ? Nous tenterons de répondre à ces questions ultérieurement avec des analyses plus avancées et des données actuelles dans la direction de la psychoacoustique, et en élaborant de nouvelles expériences qui ciblent davantage ces questions.

Références

- BLICHER D. L., DIEHL R.L., COHEN L.B. (1990). Effects of syllable duration on the perception of the Mandarin Tone 2/Tone 3 distinction : Evidence of auditory enhancement. *Journal of Phonetics* 18(1), 37-49.
- CHANG C., YAO Y. (2007). Tone production in whispered Mandarin. *Proceedings of the 16th ICPH*, 1085-1088. Saarbrücken, Germany.
- FÓNAGY J. (1969). Accent et intonation dans la parole chuchotée. *Phonetica* 20, 177-192.
- HIGA SHIKAWA M., MINIFIE F. D. (1999). Acoustical-perceptual correlates of “whisper pitch” in synthetically generated vowels. *Journal of speech language and hearing research* 42, 583- 591.
- KONG Y.-Y., ZENG F.-G. (2006) Temporal and spectral cues in Mandarin tone recognition. *The Journal of the Acoustical Society of America* 120, 2830-2840.
- LIU S., SAMUEL A. (2004). Perception of Mandarin Lexical Tones when F0 Information is Neutralized. *Language and Speech* 47(2), 109-138.
- MEYER-EPPLER W. (1957). Realization of prosodic features in whispered speech. *Acoustical Society of America* 29, 104-106.
- RYALLS J. H., LIEBERMAN P. (1982). Fundamental frequency and vowel perception. *Journal of the Acoustical Society of America* 72, 1631–1634.
- SEGERBÄCK B. (1965). La réalisation d’une opposition de tonèmes dans des dissyllabes chuchotés. *Studia Linguistica* 19(1-2), 1-54.
- TSENG C. Y., MASSARO D. W., COHEN M. M. (1986). Lexical tone perception in Mandarin Chinese : evaluation and integration of acoustic features. *Linguistics, Psychology and the Chinese Language*. University of Hong Kong Press.
- WHALEN D. H., XU Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica* 49, 25-47.
- YIP M. (2001). Tonal features, tonal inventories and phonetic targets. UCL Working Papers in *Linguistics* 13, 161-188.
- ZHANG J. (2014). Tones, tonal phonology and tone sandhi. *The Handbook of Chinese Linguistics*, Edited by Huang J., Li A., Simpson A. Wiley-Blackwell.
- ZHANG X., SOCK R. (2015). Indices acoustiques des tons du chinois mandarin en voix normale et en voix chuchotée. *Neophilologica* 27, 266-285.
- ZWICKER, E., FELDTKELLER, R., (1981) *Psychoacoustique : l’oreille, récepteur d’information*. Traduit de l’allemand par Christel Sorin, Masson.