

Xuelu Zhang

*Institut de Phonétique de Strasbourg,
EA1339- E.R. Parole et Cognition, LiLPa,
Maison Interuniversitaire des Sciences
de l'Homme -Alsace, Université de Strasbourg*

Rudolph Sock

*Institut de Phonétique de Strasbourg,
EA1339- E.R. Parole et Cognition, LiLPa,
Maison Interuniversitaire des Sciences
de l'Homme -Alsace, Université de Strasbourg,
Faculté des Lettres, Université Pavol Jozef Šafárik
Košice, Slovaquie*

Indices acoustiques des tons du chinois mandarin en voix normale et en voix chuchotée

Abstract

This paper presents an acoustic study which contributes to depicting the acoustic cues of Mandarin Chinese tones in whispered speech. By analysing modifications of the acoustic properties of the vowel /a/, we have found a correlation between the duration of vocalic segments and tones, in both normal and whispered voice. The durational differences among vowels which carry different tones are more noticable in whispered voice than in normal voice. In addition, while observing intensity and formants trajectories, we have uncovered a curve pattern which highlights the relationship between tones, intensity and formants trajectories in whispered speech. This pattern suggests that vowel intensity and formants may carry tonal information in a coordinated manner.

Keywords

Lexical tone, perception, whispered voice, Mandarin Chinese, fundamental frequency, intensity, formants.

1. Introduction

Le chinois mandarin est une langue tonale, parlée actuellement par plus d'un milliard de locuteurs, dont la majorité se trouve en Chine continentale, à Taiwan et au Singapour.

Il est admis universellement que le mandarin possède quatre tons lexicaux. L'une des raisons qui justifie l'emploi de ces tons se trouve dans le fait que le mandarin repose sur un système monosyllabique : chaque syllabe correspond à plusieurs morphèmes ayant des signifiés différents, ex. : la syllabe /pa/ est potentiellement le signifiant phonémique de multiples morphèmes, et dans un énoncé, le ton qu'elle porte réduit le choix du signifié concerné.

À l'oral, la compréhension du mandarin s'appuie énormément sur les indices fournis par des structures syntaxiques et des paramètres segmentaux et suprasegmentaux¹. Le ton est l'un des indices suprasegmentaux à fonction distinctive lexicale permettant de produire le sens lexical. Prenons de nouveau pour exemple la syllabe /pa/ : au ton 1, /pā/ signifie « huit » ; au ton 2, /pá/ signifie « retirer » ; au ton 3, /pǎ/ signifie « tenir » ; et au ton 4, /pà/ signifie « papa ».

2. Système tonal du mandarin

Un ton se produit et se fait percevoir principalement à travers la hauteur de la voix. Du point de vue phonétique, il existe deux aspects acoustiques principaux dans la réalisation des tons. Le premier aspect est le registre de la hauteur de la voix : en thaï, par exemple, les trois tons réalisés respectivement par des contours mélodiques plats dans un registre haut, dans un registre moyen et dans un registre bas permettent de distinguer des mots à travers différents registres musicaux de leurs référents sonores. Le deuxième aspect est le parcours mélodique, qui implique donc la variation (ou la maintenance) de la hauteur de la voix à l'intérieur d'une syllabe. En japonais, un contour mélodique montant et un contour mélodique descendant peuvent mettre en évidence la distinction des signifiés lexicaux qui partagent les mêmes signifiants phonémiques. Ces deux aspects ne sont pas exclusifs l'un de l'autre. Dans la perception de nombreuses langues, il convient de tenir compte de ces deux aspects. Tel est le cas du chinois mandarin.

Dans le système phonologique du mandarin, les voyelles apparaissent comme les « noyaux » des structures syllabiques, et sont porteuses de la tonalité. Ainsi, dans cette recherche, les propriétés acoustiques des segments vocaliques seront nos cibles analytiques.

Le mandarin connaît quatre tonèmes définis par quatre types de contours mélodiques couvrant cinq registres, comme dans la figure 1 (Chao, 1930). Cette méthode descriptive divise le cadre sonore de la hauteur de la voix dans la réalisation des tons, en cinq parties numérotées de 1 (la plus basse) à 5 (la plus haute). À l'intérieur de ce cadre :

- le premier tonème se présente par une voix plate, maintenue à la hauteur 5,
- le deuxième tonème se présente par une mélodie montante de 3 à 5,

¹ Est suprasegmental un trait qui ne peut pas être identifié par opposition paradigmatique mais seulement par contraste syntagmatique, c'est-à-dire par comparaison avec un autre qui précède ou qui suit dans la séquence : « [...] les traits suprasegmentaux sont des traits dont la fonction contrastive dans la séquence temporelle n'est pas fondée sur la qualité phonétique de segments individuels. (t.p.n.) » (Lehiste, 1970 : 1—3).

- le troisième tonème se présente par une mélodie qui descend de 2 à 1, puis remonte à 4,
- le quatrième tonème se présente par une chute de la hauteur de la voix, de 5 jusqu'à 1.

Dans la figure 1 se trouve aussi un tableau de référence de ces registres par rapport aux notes de musique.

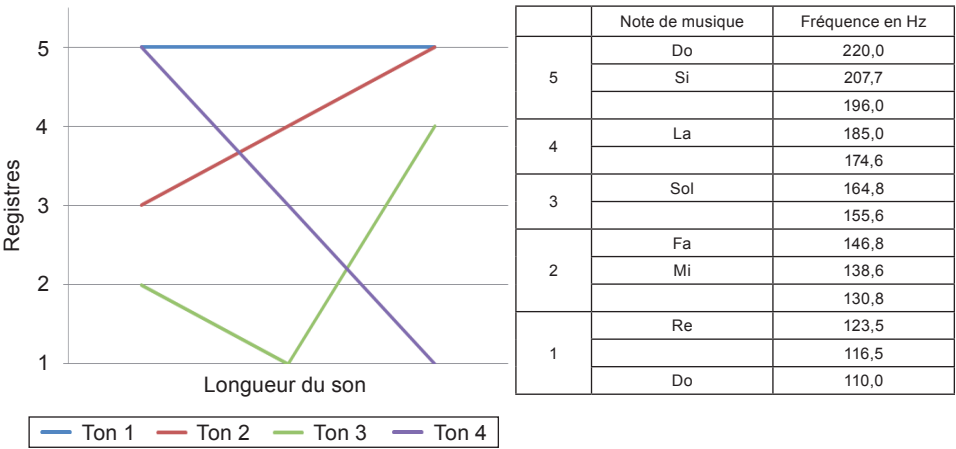


Figure 1. Les contours mélodiques et registres des tonèmes du mandarin (à gauche) et les cinq partitions du système tonal du mandarin (à droite) de Chao (1930)

Pour réaliser un tonème du mandarin, le parcours de la hauteur de la voix s'étend d'un registre à l'autre (sauf au premier ton) et établit ainsi un contraste entre les registres en respectant leurs frontières. Les registres sont également en contraste au niveau inter-tonal. Les cadres musicaux qu'ils définissent ne sont pas relativisés par l'existence des parcours présumés de la voix, mais par la différence physiologique inter-individuelle.

3. La perception de la tonalité en voix normale et en voix chuchotée

Lorsqu'on produit une voyelle en voix normale, les plis vocaux situés au niveau du larynx sont en contact complet l'un avec l'autre et vibrent de manière périodique sous la pression du flux d'air d'origine sous-glottique. L'onde acoustique, résultante de cette vibration glottique, parcourt un trajet dans l'espace supraglottique, dont la forme et le volume sont définis par l'état physique des organes articulatoires. Le segment vocalique que l'on perçoit contient donc plusieurs zones de fréquences, où

les harmoniques sont les plus intenses, soit celles qu'on appelle « les formants ». Parmi ces formants, ce qui est souvent le plus intense, contenant les fréquences les plus basses, s'appelle « la fréquence fondamentale » (F0). Elle reflète principalement l'activité vibratoire des plis vocaux. D'autres formants contenant des fréquences plus hautes sont liés non seulement à la vibration glottique mais aussi à l'état de la cavité supraglottique.

En voix normale, bien que la F0 soit la référence principale dans la perception de la tonalité (Howie, 1976), il est prouvé que d'autres paramètres acoustiques portent également des informations perceptibles sur la tonalité concernée, à tel point que l'activité vibratoire des plis vocaux n'est pas indispensable pour la compréhension même dans une langue tonale (Segerbäck, 1965). La durée du segment, par exemple, varie selon la tonalité dans certains contextes et cette variation est significative pour la perception de la tonalité (Blicher *et al.*, 1990). De plus, la trajectoire de l'intensité est corrélée à la F0 et fournit également des informations sur la tonalité (Whalen *et al.*, 1992). En revanche, la qualité vocalique est censée ne pas influencer de manière systématique le jugement de la tonalité perçue (Tseng *et al.*, 1986 ; Massaro *et al.*, 1983).

En voix chuchotée, nous constatons plusieurs faits : tout d'abord, la glotte n'est pas complètement fermée, les plis vocaux sont proches l'un de l'autre dans la partie antérieure et sont écartées dans la partie postérieure. Quant au flux d'air sous-glottique, il traverse la partie ouverte et provoque un son bruité : l'aspiration. L'onde acoustique résultant d'une voix chuchotée perd la F0 et quelque part sa périodicité vibratoire, tandis que d'autres formants persistent malgré les fréquences déconcentrées à l'intérieur de chaque formant.

Pour des causes aérodynamiques, l'onde acoustique en voix chuchotée a une intensité plus faible que celle en voix normale, avec un décalage d'environ 20 dB. Malgré ce fait, la trajectoire de l'intensité est quand même considérée comme importante dans la perception du pitch en voix chuchotée, à tel point qu'il remplace le contour mélodique (Pike, 1950). De plus, une différence physique entre voix chuchotée et voix normale existe au niveau supraglottique. La forme de l'espace supraglottique est modifiée en voix chuchotée par rapport à celle en voix normale. Cette modification spatiale a lieu non seulement sur le segment potentiellement voisé mais aussi sur toute la réalisation du logatome (Crevier-Buchman *et al.*, 2008). Il est logique d'en déduire que chez le même locuteur, la structure formantique de la même voyelle est modifiée en voix chuchotée (Higashikawa *et al.*, 1996 ; Tartter, 1991) par rapport à la voix normale. Quant à la contribution à la perception de cette modification spectrale, le déplacement des formants à des hautes fréquences vers le haut, tels le F3 et le F4 (Meyer-Eppler, 1957 ; Fónagy, 1969), est considéré comme indiquant une courbe mélodique montante.

4. Hypothèse

En voix chuchotée, vu l'absence des contours mélodiques réalisés par la variation de F0, nous supposons qu'il y a deux types de stratégies dans les modifications de propriétés acoustiques des segments vocaliques pour rappeler la tonalité : le segment peut conserver la même qualité en voix chuchotée par rapport à la voix normale (sauf la F0), à condition qu'en voix normale, d'autres paramètres acoustiques suggèrent la tonalité concernée de manière indépendante de la trajectoire de la F0. Au cas où les corrélations entre la F0 et d'autres paramètres acoustiques en voix normale ne sont pas en fonction de tonèmes, on suppose que la qualité du segment vocalique peut être modifiée sans référence obligatoire à celle de la voix normale, mais toujours en fonction de la tonalité.

Nous tenterons, à travers cette recherche, de découvrir sur quel type de stratégie s'appuieront les indices acoustiques de la tonalité en voix chuchotée. Toutefois, nous n'excluons pas la possibilité que ces méthodologies soient toutes deux employées pour maximaliser l'effet compensatoire.

5. Protocole expérimental

Cette recherche repose sur l'analyse de données acoustiques recueillies dans une expérience de production lancée en mars 2013 à Beijing, en Chine. 13 locutrices d'origine chinoise qui maîtrisent parfaitement le mandarin ont pris part à l'expérience. Nous avons pris un corpus contenant 4 phrases en mandarin, chacune portant la syllabe /pa/ avec l'un des quatre tons dans un contexte phonétique identique. Les contextes syntaxiques et sémantiques ne donnaient aucun indice sur le sens des mots auxquels réfèrent les syllabes cibles. Chaque locutrice a produit 10 répétitions de ce corpus, 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 le logiciel PRAAT.

6. Résultats

Dans le but d'étudier les rapports entre les évolutions des paramètres acoustiques, nous avons pris des mesures de la F0, de l'intensité ainsi que des premiers

quatre formants du début, du milieu et de la fin de chaque segment vocalique en question. Nous avons également pris en compte le paramètre temporel.

6.1. Les trajectoires de la F0

Nous fondons toute l'analyse comparative des données sur les rapports entre la hauteur de la voix (F0) et d'autres paramètres acoustiques. La figure 2 montre l'évolution de F0 des quatre tons à l'intérieur de chaque segment vocalique. En comparant les trajectoires de la F0, nous considérons que les réalisations des tonèmes ont eu lieu dans deux zones de registres : dans la zone de registres hauts se produisent le ton 1 et le ton 3, tandis que dans la zone de registres bas se produisent le ton 2 et le ton 3. Et nous supposons qu'en voix chuchotée, une telle séparation des deux zones se trouvera aux trajectoires des autres paramètres acoustiques afin de rappeler acoustiquement les contours mélodiques des tonèmes.

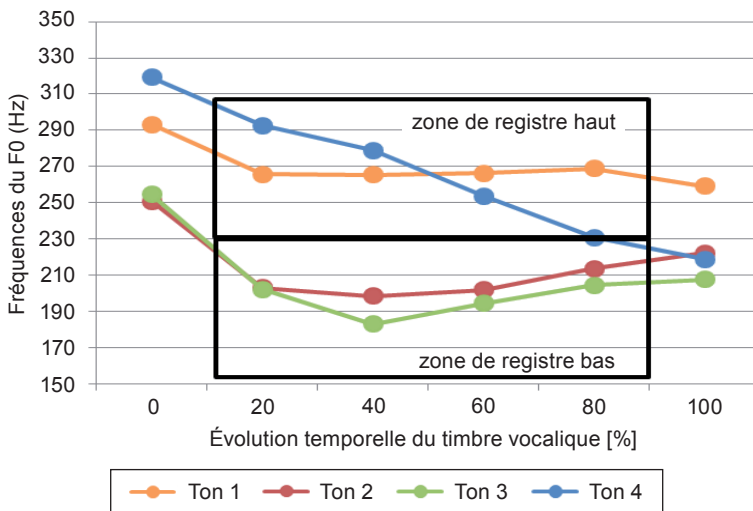


Figure 2. L'évolution de F0 sur l'axe temporel pour les quatre tons

6.2. Les durées vocaliques

Les résultats des durées de /a/ en voix normale et en voix chuchotée sont résumés dans la figure 3. Dans les deux graphiques, les axes verticaux indiquent la longueur du segment vocalique en milliseconde (ms), et sur les axes horizontaux est indiquée la durée en moyenne des quatre tonèmes.

En comparant les durées des deux modes de voix, une tendance générale de variation nous interpelle : en voix normale, les segments du ton 2 sont les plus

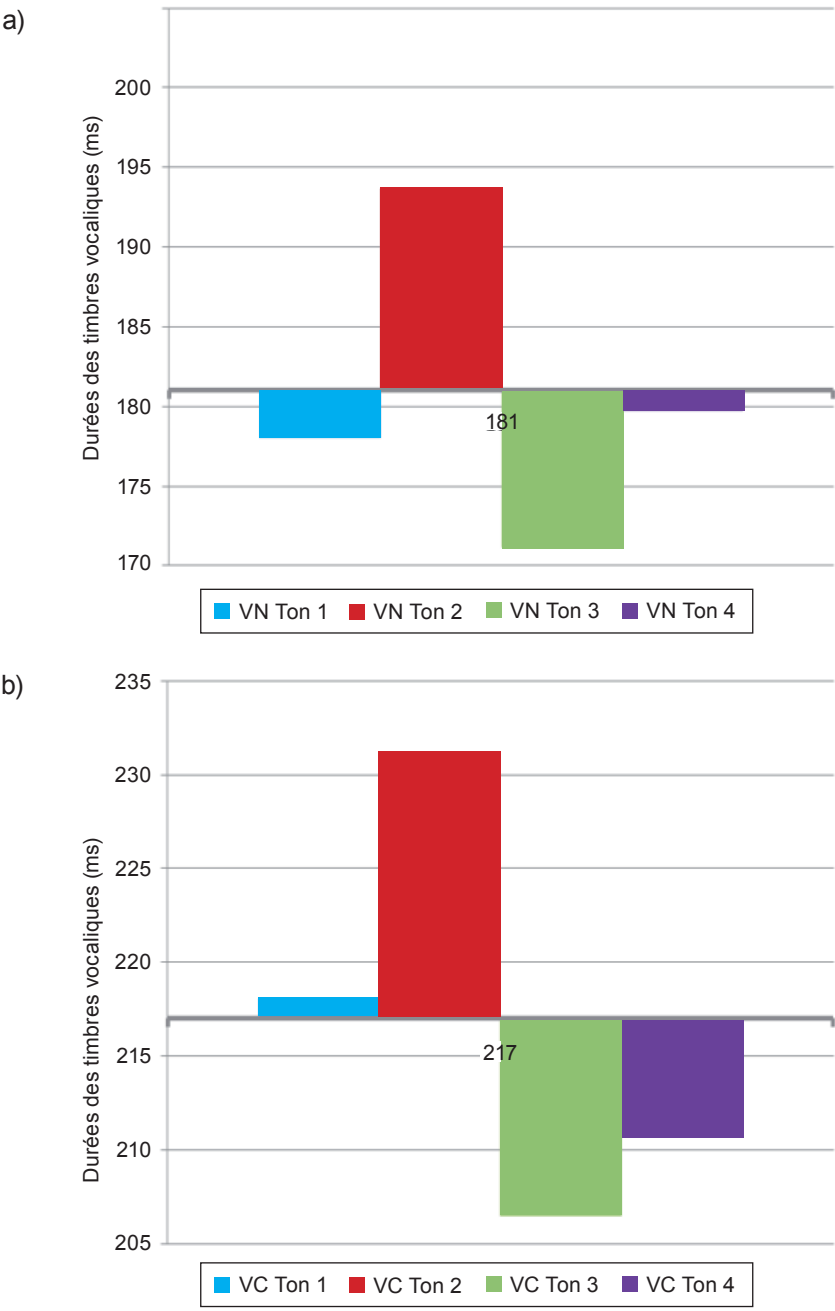


Figure 3. Durées des segments vocaliques portant quatre tons en voix :
a — normale, **b** — chuchotée

prolongés, tandis que ceux du ton 3 sont les plus brefs. En voix chuchotée, les segments vocaliques sont tous plus prolongés par rapport à la voix normale, mais les deux pôles de durée persistent en voix chuchotée.

Les segments du ton 1 et du ton 4 ont des durées moyennes, quel que soit le mode de la voix. Cependant, la différence de durée entre ces deux tons n'est pas significative en voix normale, mais elle devient plus importante en voix chuchotée.

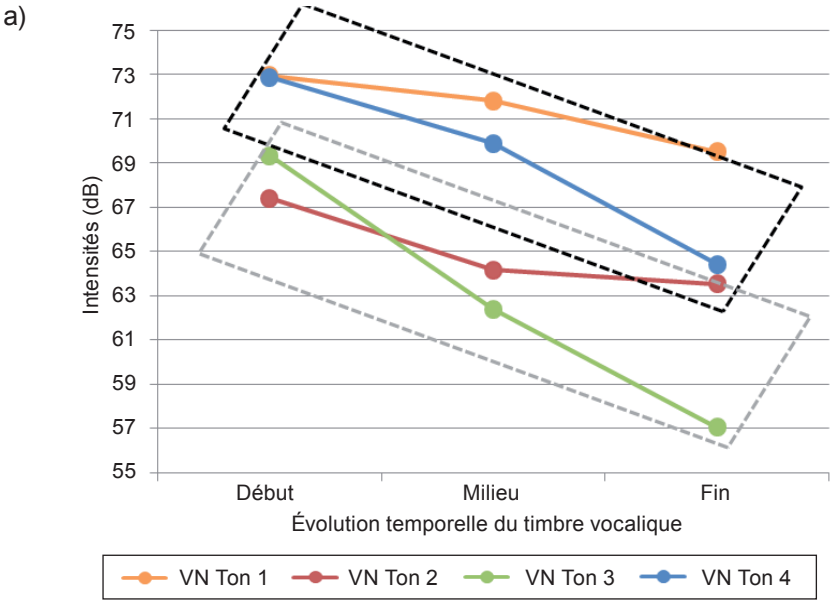
6.3. Les trajectoires de l'intensité

La figure 4 illustre les trajectoires de l'intensité en fonction des tonèmes sur un plan intensité (dB)-temps (ms). Dans le graphique (a), on peut observer une divergence de ces trajectoires dans deux zones d'intensité, l'une plus forte que l'autre, rappelant le cas des trajectoires de la F0.

Les tons hauts, soit le ton 1 et le ton 4, causent des trajectoires de l'intensité dans la zone d'intensité plus forte. Les tons bas, soit le ton 2 et le ton 3, entraînent des trajectoires de l'intensité dans la zone d'intensité plus faible.

À l'intérieur de chaque zone, le décalage entre les deux trajectoires est faible au début du segment. Mais à la fin du segment, il arrive au maximum.

Dans le graphique (b) qui illustre le cas en voix chuchotée, les deux zones d'intensité couvertes par les trajectoires persistent, mais se chevauchent en grande partie. À l'intérieur de chaque zone, la hiérarchie des trajectoires est maintenue, tandis que le décalage entre les deux trajectoires devient moins important.



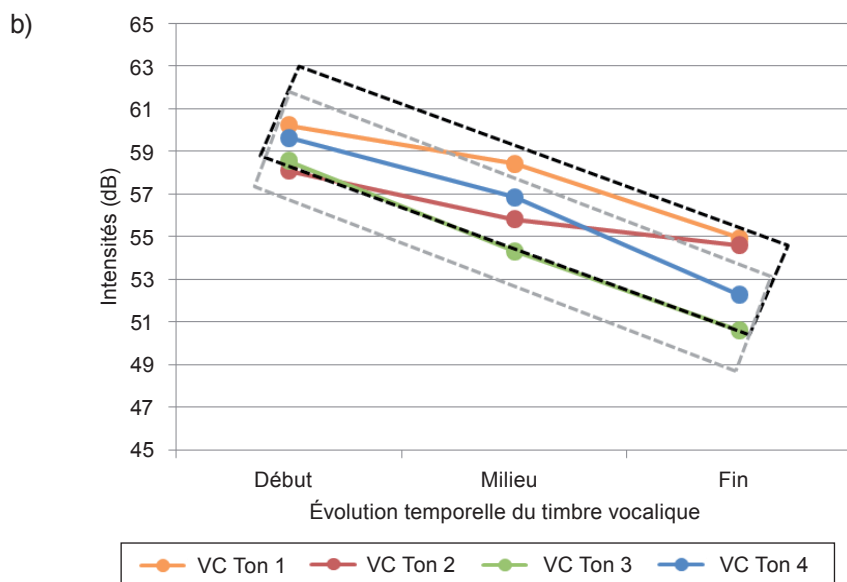


Figure 4. Trajectoires de l'intensité des quatre tons en voix: a — normale, b — chuchotée

6.4. Les spectres formantiques

— F1

En tant que formant le plus intense après la F0, et donc le plus proche de cette dernière, le F1 est souvent considéré comme une référence importante pour juger de la hauteur de la voix. Pourtant, la figure 5 illustre les trajectoires du F1 de /a/ en fonction des tonèmes sur le plan fréquence (Hz)-temps (ms) et montre que ce formant fournit probablement peu d'indices sur la tonalité en voix chuchotée.

Dans le graphique (a), nous observons une hiérarchie sur l'axe fréquentiel entre les trajectoires du F1 pour les quatre tons en voix normale. Malgré les croisements provoqués par un changement de registre qui survient au niveau des trajectoires de la F0 et de l'intensité, les trajectoires du F1 aux quatre tons ne connaissent aucun croisement entre elles.

La trajectoire du F1 du ton 3, qui a la F0 la plus basse, se situe la plus haute dans la hiérarchie fréquentielle, suivie par la trajectoire du F1 du ton 2. Les trajectoires du F1 des tons hauts, soit le ton 1 et le ton 4, se trouvent en bas dans la hiérarchie, avec celle du ton 1 la plus basse. Cette hiérarchie est l'inverse de celle des trajectoires de l'intensité.

Dans le graphique (b), nous observons des chevauchements entre les trajectoires. La hiérarchie fréquentielle n'est plus respectée et les différences fréquentielles entre les trajectoires sont faibles.

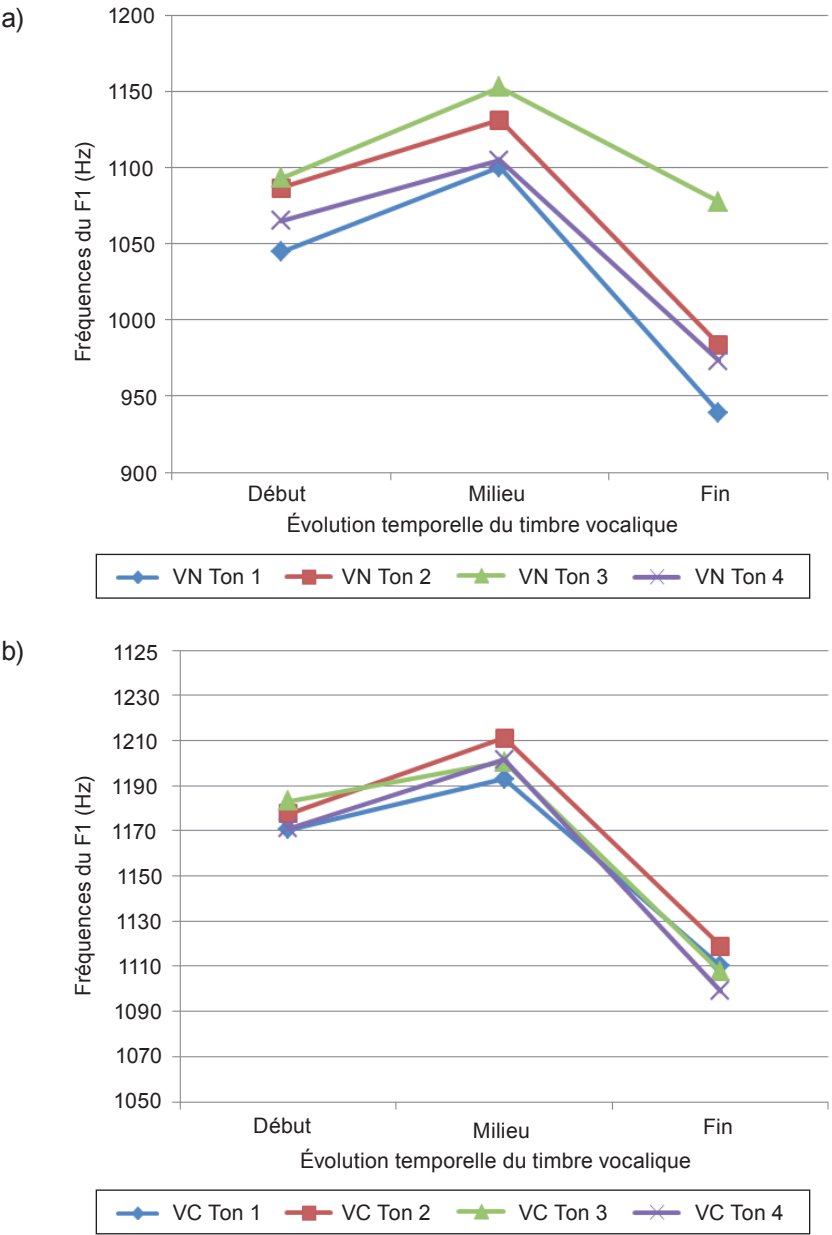


Figure 5. Trajectoires du F1 de /a/ aux quatre tons en voix : **a** — normale, **b** — chuchotée

— F2

Pour ce qui est des valeurs du F2, nous les résumons dans la figure 6. Et nous avons observé les trajectoires aux quatre tons en parallèle en parole normale.

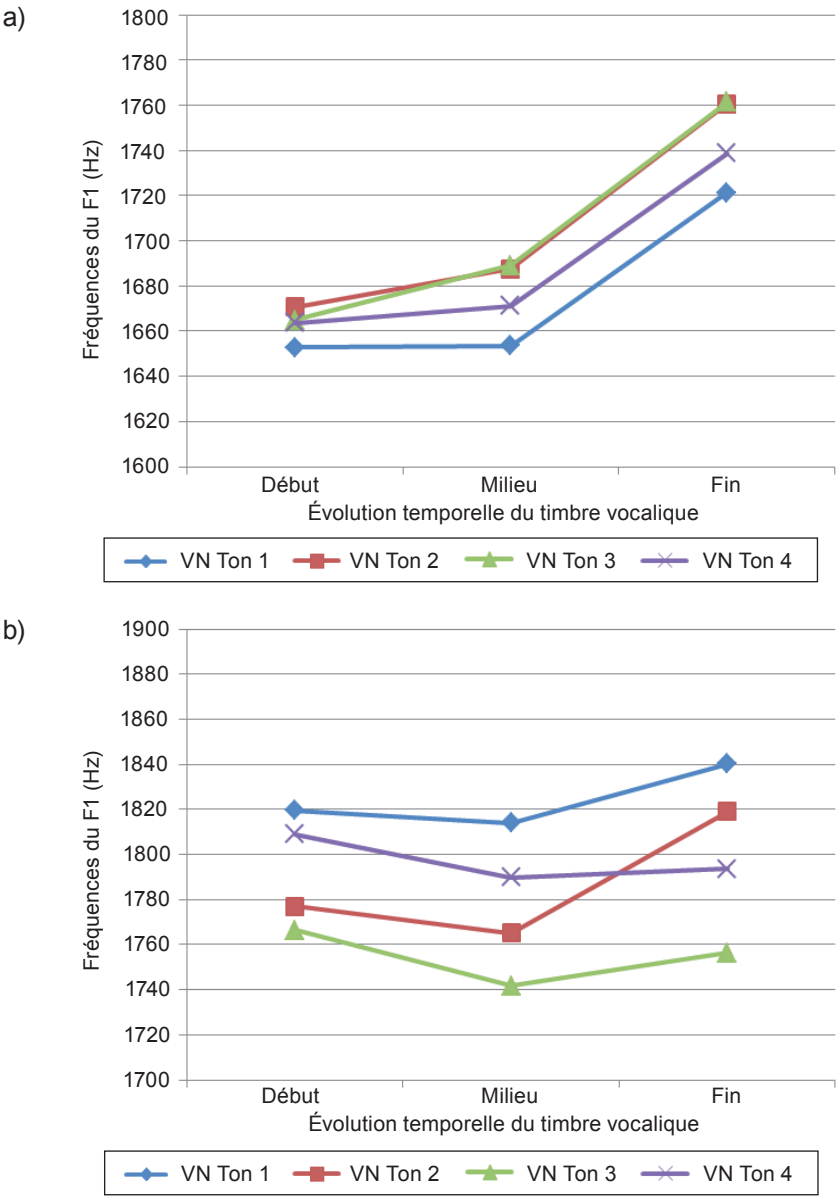


Figure 6. Trajectoires du F2 de /a/ aux quatre tons en voix : **a** — normale, **b** — chuchotée

Les trajectoires du F2 du ton 2 et du ton 3 se chevauchent, se situant plus haut dans la hiérarchie fréquentielle, suivies par celle du ton 4. La trajectoire du F2 du ton 1 se trouve la plus basse sur l'axe fréquentiel. Entre les trajectoires voisines, le décalage est maintenu à 20 Hz.

Dans le graphique (b), contrairement au cas de la voix normale, les trajectoires du ton 2 et du ton 3 en voix chuchotée se limitent dans une zone plus basse (celui du ton 3 au plus bas) sur l'axe fréquentiel, sauf à la fin du segment ; et celles du ton 1 (la plus haute) et du ton 4 se situent dans une zone plus haute. Cette division en zones rappelle les trajectoires de la F0 et de l'intensité.

À l'intérieur de chaque zone, le décalage entre les deux trajectoires est faible (environ 10 Hz) au début du segment, mais il augmente et arrive au sommet à la fin du segment (entre 50—60 Hz).

— F3

La hiérarchie fréquentielle des trajectoires s'opère également au niveau du troisième formant, comme le montrent les graphiques de la figure 7.

En voix normale (voir le graphique (a)), il est toujours possible de séparer deux zones de fréquences contenant respectivement les trajectoires aux tons hauts et les trajectoires aux tons bas.

À l'intérieur de la zone de fréquences plus hautes, la trajectoire du F3 du ton 1 se situe au plus haut, en gardant un décalage d'au moins 50 Hz par rapport à la trajectoire du ton 4. À l'intérieur de la zone de fréquences plus basses, la trajectoire du ton 2 se situe au plus bas, en parallèle avec celle du ton 3, avec un décalage de 20 Hz.

Dans le graphique (b), les trajectoires du F3 de la voix chuchotée se distribuent toujours dans deux zones de fréquences. Comme dans le cas du F2 en voix chuchotée et celui de l'intensité, dans la zone contenant des fréquences plus hautes se trouvent les trajectoires du ton 1 (plus haute) et du ton 2 (plus basse) ; dans la zone de fréquences plus basse se trouvent les trajectoires du ton 2 (plus haute sauf au début du segment) et du ton 3 (plus basse).

À l'intérieur de chaque zone, les deux trajectoires des tons débutent avec peu de différences dans les valeurs fréquentielles et enregistrent un décalage croissant entre eux sur tout le long du segment vocalique.

— F4

Bien que le F4 contienne des fréquences hautes peu perceptibles à l'oreille, nous avons étudié les trajectoires de ce formant en considérant son importance en voix chuchotée.

Le graphique (a) de la figure 8 illustre les trajectoires du F4 en voix normale. Les trajectoires du ton 1 et du ton 2 se trouvent dans une zone plus haute sur l'axe fréquentiel, tandis que celles du ton 3 et du ton 4 se situent dans une zone basse (sauf la trajectoire du ton 3 qui débute dans la zone de fréquences plus hautes). La hiérarchie n'est pas clairement définie, puisque les trajectoires du F4 se croisent à un moment.

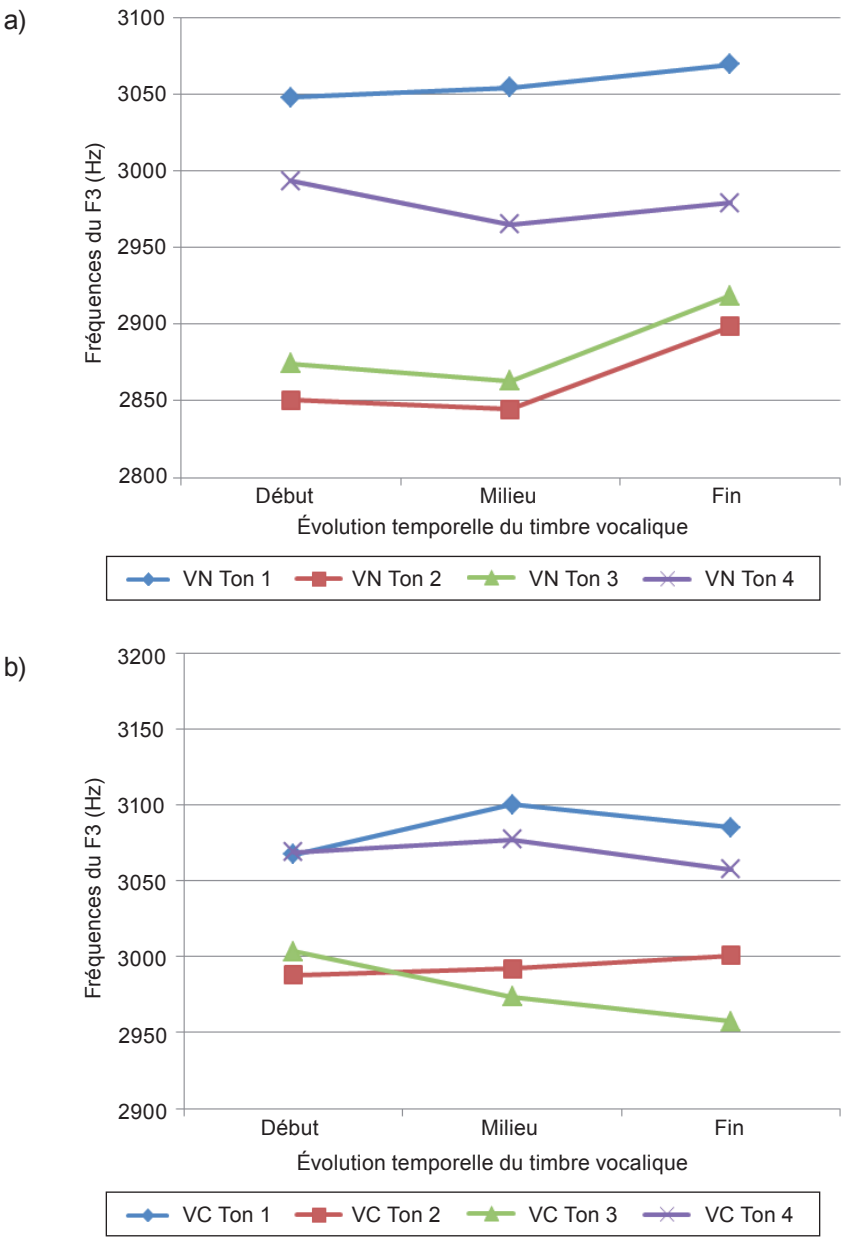


Figure 7. Trajectoires du F3 de /a/ aux quatre tons en voix : **a** — normale, **b** — chuchotée

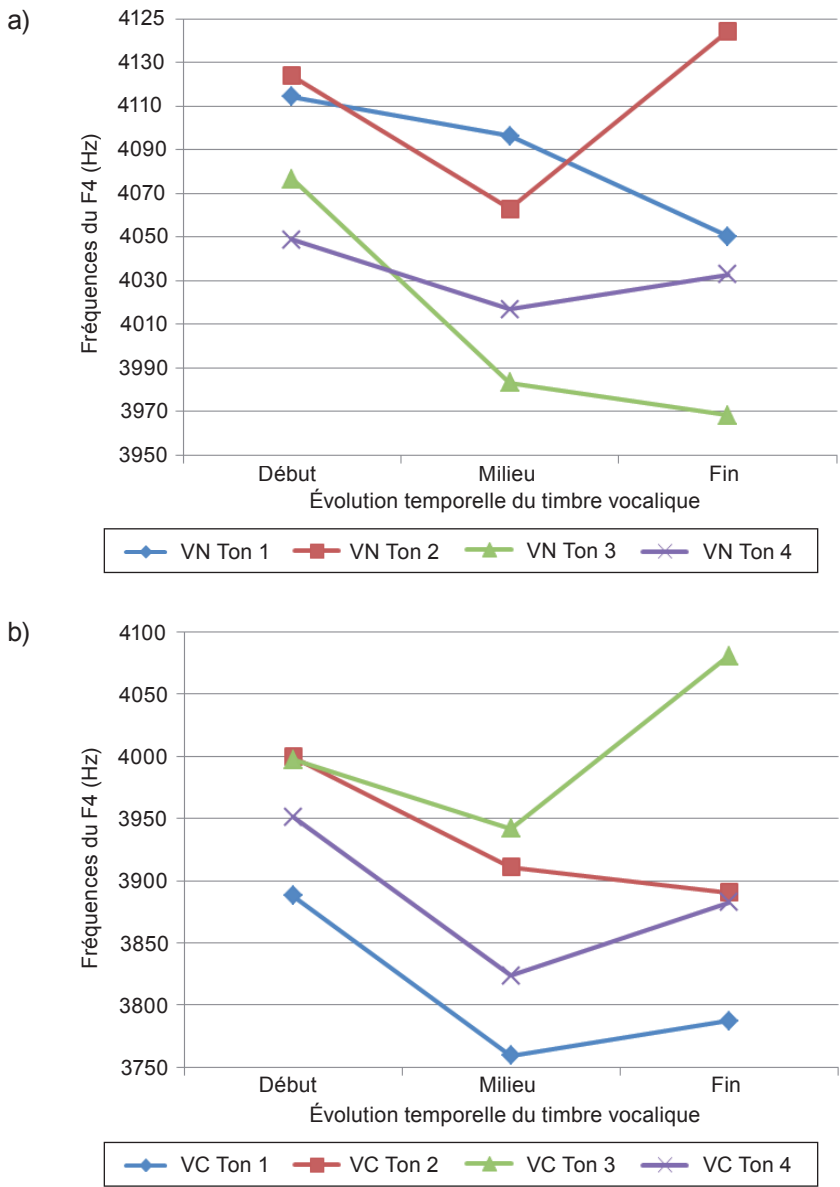


Figure 8. Trajectoires du F4 de /a/ aux quatre tons en voix : **a** — normale, **b** — chuchotée

Au début du segment vocalique, la trajectoire du ton 1 et celle du ton 2 sont proches, à des valeurs fréquentielles plus hautes, suivies par la trajectoire du ton 3. Cependant, à la fin du segment, nous observons de nouveau un décalage élargi (par rapport à ceux dans d'autres phases temporelles) entre les trajectoires des deux tons dans la même zone, 100 Hz entre le ton 1 et le ton 2, et 60 Hz entre le ton 3 et le ton 4.

Dans le graphique (b), nous observons que la zone de fréquences plus hautes est occupée de nouveau par les trajectoires du ton 2 (plus basse) et du ton 3 (plus haute), et que la zone de fréquences plus basses est occupée par les trajectoires du ton 1 (plus haute) et du ton 4 (plus basse); excepté la trajectoire du ton 4 qui débute dans la zone plus haute. Le décalage entre les deux trajectoires à l'intérieur de la zone s'accroît graduellement tout le long du segment vocalique. À la fin du segment, le décalage au niveau du F4 entre le ton 2 et le ton 3 arrive à 200 Hz et celui entre le ton 1 et le ton 4 arrive à 100 Hz.

7. Discussion générale

Dans nos analyses, nous avons étudié les trajectoires de la F0 en voix normale en tant que repère principal. Et pour mettre au jour des indices, nous avons choisi les durées des segments vocaliques, les trajectoires de l'intensité et des quatre premiers formants en tant que paramètres acoustiques cibles.

Dans l'observation des trajectoires de la F0 des quatre tonèmes, nous avons défini deux zones de registres couvertes par la F0. Dans la zone de fréquences hautes se situent les trajectoires du ton 1 et du ton 4, tandis que dans la zone de fréquences basses se situent les trajectoires du ton 2 et du ton 3. Malgré les différents types de contours mélodiques en fonction des tonèmes, les trajectoires de la F0 ne connaissent pas de croissance entre les deux zones.

Nous avons comparé également les durées vocaliques en voix normale et en voix chuchotée. La tonalité exerce effectivement une influence sur la durée du segment vocalique, quel que soit le mode de la voix.

En voix normale, les segments du ton 2 sont les plus prolongés, tandis que les segments du ton 3 sont les plus brefs. Les segments du ton 1 et du ton 4 ont des durées proches du niveau moyen.

En voix chuchotée, cette tendance a été maintenue avec tous les segments vocaliques qui sont plus prolongés par rapport à la voix normale. Pourtant, les différences de durée provoquées par les tonèmes sont devenues plus importantes. Nous supposons ainsi qu'en voix chuchotée, la durée du segment vocalique occupe une place significative dans la perception de la tonalité. Lorsque la F0 est absente, c'est la durée qui prendra le relais; elle devient de ce fait plus perceptible.

Dans l'observation des trajectoires de l'intensité des segments vocaliques en voix normale, au lieu de trouver des tracés similaires à ceux de la F0, nous avons reconnu la séparation des deux zones sur l'axe de l'intensité acoustique. Les deux zones d'intensité ont été contrastées tout le long du segment vocalique. Dans la zone d'intensité la plus forte se situent les trajectoires du ton 1 et du ton 4, tandis que dans la zone d'intensité la plus faible se situent les trajectoires du ton 2 et du

ton 3. À l'intérieur de chaque zone, le décalage entre les deux trajectoires augmente par rapport à l'évolution temporelle du segment vocalique.

Pourtant, en voix chuchotée, l'intensité du segment devient faible par rapport à la voix normale. Les deux zones couvertes par les trajectoires se croisent en grande partie, et à l'intérieur de chaque zone, le décalage entre les trajectoires est devenu moins important.

En conclusion, nous supposons qu'au niveau de l'intensité, la trajectoire, ainsi que la zone d'intensité qu'elle occupe, portent des informations sur la tonalité. Cependant, en voix chuchotée, leur efficacité en tant qu'indicateur de la tonalité est réduite. En prenant en compte les modifications de durées vocaliques mentionnées ci-dessus, cet affaiblissement de l'efficacité de l'intensité est probablement à un tel point que la référence à d'autres paramètres acoustiques devient nécessaire pour la distinction de la tonalité.

Nous avons également tenté de trouver des indices sur la tonalité dans les spectres formantiques. En voix normale, nous avons observé des hiérarchies fréquentielles des trajectoires formantiques. Quant aux trajectoires du F1 et du F2, celles du ton 2 et du ton 3 se situent plus haut que celles du ton 1 et du ton 4 sur l'échelle fréquentielle. Concernant les trajectoires du F3, le cas est inversé. Et pour les trajectoires du F4, celles du ton 1 et du ton 2 se situent plus haut que celles du ton 3 et du ton 4. Nous pouvons en conclure qu'en voix normale, la séparation des deux zones existe dans les hiérarchies des trajectoires formantiques, mais à l'intérieur des zones, la divergence des trajectoires est moins régulière et difficile à généraliser.

En revanche, en voix chuchotée, les trajectoires du F1 des quatre tons se chevauchent en grande partie. Nous ne considérons donc pas le F1 comme un paramètre important dans la distinction de la tonalité. Quant aux F2 et F3, la séparation des zones apparaît : les trajectoires formantiques du ton 1 et du ton 4 se situent dans une zone plus haute sur l'échelle fréquentielle (le ton 1 plus haut et le ton 4 plus bas), tandis que celles du ton 2 et du ton 3 se trouvent dans une zone plus basse (le ton 2 plus haut et le ton 3 plus bas). Concernant le F4, la séparation des zones persiste, mais cette hiérarchie est inversée. En général, la divergence des trajectoires à l'intérieur de chaque zone est régulière : les deux débutent souvent avec des valeurs similaires et le décalage entre elles augmente par rapport à l'évolution temporelle du segment vocalique.

Malgré la persistance des hiérarchies fréquentielles, les décalages entre les trajectoires formantiques en voix chuchotée sont quantitativement faibles et d'après nous, ne sont pas pertinents en tant qu'indices acoustiques perceptibles sur la tonalité. Pourtant, nous supposons que la séparation des zones parmi les trajectoires, ainsi que la régularité de divergence des trajectoires à l'intérieur de chaque zone, sont liées à des gestes d'ajustements articulatoires en fonction des tonèmes. Ainsi, nous avons déduit un pattern de modifications de paramètres acoustiques en fonction des tonèmes chuchotés, comme dans la figure 9.

Ce pattern inclut la séparation répétitive de deux zones de l'intensité et des formants, ainsi que la divergence des trajectoires de ces paramètres à l'intérieur de chaque zone. Les tons hauts (le ton 1 et le ton 4) se trouvent dans la même zone, tandis que les tons bas (le ton 2 et le ton 3) sont répartis dans l'autre zone. Ce pattern s'adapte à la description des modifications de la majorité des paramètres acoustiques en voix chuchotée (sauf celles du F1, *cf.* le graphique (b) dans la figure 5). En voix normale, ce pattern s'adapte également à la description des variations de l'intensité et à celle du F1. Sur les formants à des fréquences plus hautes, ce pattern s'adapte moins aisément.

Ce pattern est axé sur le contraste des valeurs du paramètre cible au début et à la fin du segment vocalique. Au début des segments, les paramètres acoustiques se distinguent déjà au niveau de l'intensité et des formants, en séparant deux zones en fonction des tons hauts (le ton 1 et le ton 4) des tons bas (le ton 2 et le ton 3). À la fin du segment, le décalage entre les deux trajectoires à l'intérieur de chaque zone se maximise.

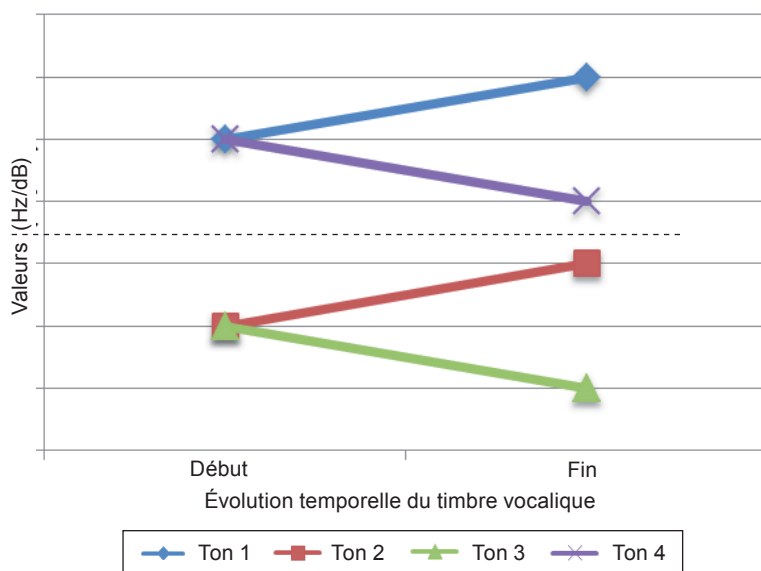


Figure 9. Pattern de l'évolution des formants / de l'intensité

Ces deux phases de distinction reflètent, d'après nous, des modifications articulaires à deux phases. Cette hypothèse reste à vérifier par des expériences articulaires. Mais elle est fondée sur le fait que ce pattern acoustique possède une certaine flexibilité, probablement à cause de son origine articulaire, ex. : en voix chuchotée, l'ordre de la hiérarchie fréquentielle entre les trajectoires du F4 est inversé par rapport à celles des autres paramètres. De plus, en voix chuchotée, nous avons observé que sur l'axe du F2, la trajectoire du ton 2 s'était éloignée autant de la

trajectoire du ton 3 qui partageait la zone de fréquences basses que de la première, et était entrée dans la zone de fréquences plus hautes à la fin du segment vocalique (*cf.* le graphique (b) dans la figure 6). Également, sur l'axe du F4, la trajectoire du ton 4 garde un certain décalage avec celle du ton 1, avec laquelle cette première partage la zone de fréquences basses, et est entrée dans la zone de fréquences hautes au début du segment vocalique (*cf.* le graphique (b) dans la figure 8).

Nous interprétons ces deux cas des travers entre les zones comme des témoins d'une fonction combinatoire des deux phases de distinction : à condition qu'une hiérarchie fréquentielle soit établie entre les trajectoires et sans le souci de croisement qui trouble cette hiérarchie, les trajectoires des paramètres débutent sans respecter strictement les limites des zones, mais doivent les respecter à la fin du segment vocalique, ou ils peuvent débuter dans deux zones différentes, comme convenu, mais terminent sans se limiter à leurs zones. Au niveau acoustique, la distinction de la tonalité prend les informations de ces deux phases en compte, également de celles de la séparation des zones qui concerne tout le segment vocalique.

Pour assurer la pertinence du pattern global, les deux zones doivent maintenir un contraste entre elles. Un cas exceptionnel se trouve dans l'intensité en voix chuchotée, où les trajectoires correspondent à notre pattern global mais où les zones se croisent en grande partie. Nous en concluons ainsi que la séparation des zones dans ce pattern est probablement réalisée par un autre type d'ajustements articulatoires que celui relatif aux deux phases de distinction entre les trajectoires mentionnées ci-dessus. Les deux types d'ajustement articulatoires sont indépendants l'un de l'autre, mais contribuent au même objectif acoustique de manière conjointe.

8. Conclusion

Nous avons comparé les propriétés acoustiques des segments de /a/ aux quatre tonèmes du mandarin dans le même environnement phonétique en voix normale et en voix chuchotée.

La durée du segment vocalique en voix chuchotée est supposée être plus perceptible qu'en voix normale. La stratégie utilisée est la suivante : pour garder la distinction entre les différents tons en l'absence de la sonorité, la durée du timbre vocalique en voix chuchotée reflète l'influence du tonème sur elle, tout en la mettant en évidence.

Dépourvu de la sonorité, les tonèmes ne peuvent pas garder leurs traits distinctifs qu'ils possèdent en voix normale ; pour pallier ce déficit, l'intensité et les formants des segments vocaliques chuchotés emploient une autre stratégie : ils respectent un pattern contenant une séparation des deux zones et une hiérarchie en

fonction des tonèmes parmi les trajectoires. Les trajectoires du F1 de /a/ chuchotées n’y correspondent pourtant pas.

Bien que les différences entre les trajectoires de l’intensité et des formants des quatre tons soient quantitativement faibles, elles révèlent, d’après nous, des gestes d’ajustement articulatoire à différents niveaux.

En général, en voix chuchotée, la capacité d’indiquer la tonalité de l’intensité est réduite. Mais la découverte de ce pattern nous révèle que la coordination entre l’intensité et les formants occupe une place importante dans la perception de la tonalité.

9. Perspectives

Nous exploiterons ultérieurement nos données acoustiques de la même expérience sur les voyelles /i/ et /u/, dans l’espoir d’observer un pattern similaire de modification des paramètres acoustiques ou d’en découvrir d’autres. De plus, nous avons émis dans cette recherche des hypothèses sur l’origine articulatoire du pattern acoustique mis au jour et nous sommes particulièrement intéressés par une expérience articulatoire dans l’objectif d’observer les ajustements aux niveaux glottique et supraglottique en voix chuchotée.

Références

- Blicher Deborah, Diehl Randy, Cohen Leslie, 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.
- Chao Yuanren, 1930: “A system of tone letters”. *Le Maître Phonétique*, **45**, 24—27.
- Chen Trevor, Massaro Dominic, 2008: “Seeing pitch: visual information for lexical tones of Mandarin-Chinese”. *Acoustical Society of America*, **123**(4), 2356—2366.
- Crevier-Buchman Lise, Vincent Coralie, Maeda Shinji *et al.*, 2008: « Comportements laryngés en voix chuchotée, étude en camera ultra rapide ». Congrès de la Société Française de Phoniatrie et des Pathologies de la Communication France. (En ligne : <http://halshs.archives-ouvertes.fr/halshs-00683808>, accessible : 13.04.2014).
- Fónagy Judith, 1969: « Accent et intonation dans la parole chuchotée ». *Phonetica*, **20**, 177—192.
- Higashikawa Masahiko, Minifie Fred, 1999: “Acoustical-perceptual correlates of whisper pitch in synthetically generated vowels”. *Journal of Speech Language and Hearing Research*, **42**, 583—591.

- Howie John Marshall, 1976: *Acoustical studies of mandarin vowels and tones*. Cambridge University Press.
- Lehiste Ilse, 1970: *Suprasegmentals*. Cambridge: Massachusetts Institute of Technology Press.
- Massaro Dominic, Cohen Michael, 1983: "Integration of visual and auditory information in speech perception". *Journal of Experimental Psychology: Human Perception and Performance*, **9**, 753—771.
- Meyer-Eppler Werner, 1957: "Realization of prosodic features in whispered speech". *Acoustical Society of America*, **29**, 104—106.
- Pike Kenneth, 1956: *Tone languages*. Ann Arbor: University of Michigan Press.
- Segerbäck Börje, 1965 : « La réalisation d'une opposition de tonèmes dans des dissyllabes chuchotés ». *Studica Linguistica*, **19** (1—2), décembre, 1—54.
- Tartter Vivian, 1991: "Identifiability of vowels and speakers from whispered syllables". *Perception & Psychophysics*, **49**, 365—372.
- Tseng Chiu-Yu, Massaro Dominic, Cohen Michael, 1986: "Lexical tone perception in Mandarin Chinese: evaluation and integration of acoustic features". *Linguistics, Psychology and the Chinese Language*, **67**, 91 [University of Hong Kong Press].
- Whalen Douglas, Xu Yi, 1992: "Information for Mandarin tones in the amplitude contour and in brief segments". *Phonetica*, **49**, 25—47.