



HAL
open science

Analyse acoustique des occlusives produites par des jeunes locuteurs en dialecte wu de Suzhou

Ning Wang

► **To cite this version:**

Ning Wang. Analyse acoustique des occlusives produites par des jeunes locuteurs en dialecte wu de Suzhou. XXXIIe Journées d'Études sur la Parole (JEP), Jun 2018, Aix en Provence, France. hal-02879612

HAL Id: hal-02879612

<https://hal.science/hal-02879612>

Submitted on 24 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ACTES DES
32E JOURNEES D'ETUDES SUR LA PAROLE



ORGANISEES PAR LE LABORATOIRE PAROLE ET LANGAGE
CNRS - AIX-MARSEILLE UNIVERSITE

SOUS L'EGIDE DE L'ASSOCIATION FRANCOPHONE
DE LA COMMUNICATION PARLEE

AIX-EN-PROVENCE 4-8 JUIN 2018

JEP 2018

ACTES DES
32E JOURNEES D'ETUDES SUR LA PAROLE

4-8 JUIN 2018
AIX-EN-PROVENCE

ORGANISEES PAR LE LABORATOIRE PAROLE ET LANGAGE
CNRS - AIX-MARSEILLE UNIVERSITE

SOUS L'EGIDE DE L'ASSOCIATION FRANCOPHONE
DE LA COMMUNICATION PARLEE

EDITEURS

MARTIN COOKE (ISCA)

BRIGITTE BIGI (LPL)

JOELLE LAVAUD (LPL)

DIFFUSÉ PAR

INTERNATIONAL SPEECH COMMUNICATION ASSOCIATION (ISCA)

[HTTPS://WWW.ISCA-SPEECH.ORG/ISCAWEB/INDEX.PHP/ONLINE-ARCHIVE](https://www.isca-speech.org/ISCAWEB/INDEX.PHP/ONLINE-ARCHIVE)

ASSOCIATION FRANCOPHONE POUR LA COMMUNICATION PARLÉE (AFCP)

[HTTP://WWW.AFCP-PAROLE.ORG/](http://www.afcp-parole.org/)

CREDITS IMAGES

CANOPUS I, 1959, VICTOR VASARELY - AVEC L'AUTORISATION DE LA FONDATION VASARELY

SCHEMA DE PRINCIPE DU POLYPHONOMETRE, 1969 - BERNARD TESTON

TRACES IMPRIMES ISSUS DU POLYPHONOMETRE, 1969 - BERNARD TESTON

CONCEPTION COUVERTURE : YOHANN MEYNADIER & CLAUDIA PICHON-STARHE



Analyse acoustique des occlusives produites par des jeunes locuteurs en dialecte wu de Suzhou

WANG Ning¹

(1) Laboratoire LACITO, CNRS, 94801 Villejuif, France
wangning0531@gmail.com

RESUME

Le système phonologique consonantique du wu ne comporte aucune opposition de voisement pour les occlusives. Phonétiquement, on constate que les occlusives sourdes /p, t, k/ en position initiale peuvent être perçues voisées et légèrement soufflées quand elles sont suivies d'une voyelle à ton bas. En position intervocalique d'une unité complexe dissyllabique, leur réalisation peut être voisée. Cet article examine la condition du soufflement et du voisement dans le dialecte wu de Suzhou où on observe qu'à l'initiale et en registre bas, les jeunes locuteurs maintiennent le soufflement qui se propage à partir du milieu de la réalisation de la voyelle jusqu'à la fin ; à l'intervocalique, le degré de voisement est effectivement plus élevé mais seulement dans un contexte lexical et tonal spécifique. Ce degré ne réalise pas le voisement à 100%. De plus, on note que la réalisation intervocalique peut varier entre [b, d, g] et [β, ɖ, ɣ].

ABSTRACT

Acoustical analysis of stops produced by young speakers of Suzhou dialect

The consonantal phonological system of wu has no opposition of voicing for stops. Phonetically, we find that the stops /p, t, k/ in initial position can be perceived voiced and slightly breathy when followed by a vowel with low tone. In intervocalic position of a complex disyllabic unit, the realization of /p, t, k/ can be voiced. This article examines the condition of the breathiness and the voicing in the dialect spoken in Suzhou and we observe that in initial and low register, the young speakers maintain the breathiness that start from the middle of the realization of the vowel until the end; in intervocalic position, the degree of voicing is actually higher but only in a specific lexical and tonal context. This degree doesn't arrive at 100%. Moreover, we note that the intervocalic realization of three stops can vary between [b, d, g] and [β, ɖ, ɣ].

MOTS-CLES: langue wu, dialecte de Suzhou, tonal haut et bas, voix soufflée acoustique et physiologique, voisement.

KEYWORDS: wu language, Suzhou dialect, low vs. high tone, acoustic/physiological breathiness, voicing.

1 Introduction

Si l'opposition de voisement au niveau des occlusives est pertinente en français entre /p/-/b/, /t/-/d/ et /k/-/g/, le mandarin quant à lui ne connaît qu'une opposition d'aspiration et oppose : /p/-/p^h/, /t/-

/t^h/ et /k/-/k^h/. La langue wu parlée à l'est de la Chine appartient à la même famille et au même groupe que le mandarin. Il n'est donc pas étonnant de retrouver la même corrélation d'aspiration au niveau des phonèmes occlusifs qu'en mandarin. Toutefois, ces phonèmes rencontrent dans certains contextes des particularités phonétiques et acoustiques qui n'existent absolument pas en mandarin. Le wu connaît des variétés régionales dont les plus importantes sont celle de Shanghai et celle de Suzhou. Cette étude s'intéresse au dialecte de Suzhou qui possède sept tons : *yin ping* « 44 » ; *yang ping* « 223 » ; *yin shang* « 51 » ; *yin qu* « 523 » ; *yang qu* « 231 » ; *yin ru* « 43 » ; *yang ru* « 23 » (Xin, 2011). En fonction de leur registre de départ initial : le registre haut *yin* regroupe les tons dont la hauteur initiale est égale ou supérieure à 3 sur l'échelle de hauteur tonale, le registre bas *yang* regroupe ceux dont la valeur tonale initiale est inférieure à 3. D'où quatre tons hauts et trois tons bas (Chao, 1930). Dans la littérature, la consonne initiale suivie d'une voyelle dont le ton est à un registre bas est considérée comme voisée. Or la perception voisée n'est pas manifeste puisque le spectrogramme prouve que, physiologiquement, elle est réalisée sans voisement. La perception voisée serait due à une voix légèrement soufflée ou aspirée (Chao, 1930 ; Cao et al., 1992 ; Shi, 2009 ; Gao et al., 2012) mais cette hypothèse reste controversée. A l'intervocalique d'une unité composée, nous constatons une neutralisation tonale conformément à la règle de sandhi qui veut que dans ce contexte, les syllabes, autres que la première syllabe d'une unité polysyllabique, héritent du ton de la première syllabe. Ce ton se propage vers la droite et les syllabes qui subissent le sandhi perdent leurs tons originaux. Par conséquent, une consonne intervocalique suivie d'une voyelle dont le ton, normalement, est au registre bas subit un changement tonal ce qui cette fois lui permet de se réaliser physiologiquement comme voisée (Shi, 2009). Par exemple : 番 [fe⁴⁴] + 茄 [ka²²³] = 番茄 [fe⁴⁴ga³¹] « tomate ».

Notre étude étudie ce phénomène en testant la production de jeunes locuteurs de wu vivant à Suzhou sur leur réalisation des phonèmes /p, t, k/ à l'initiale des morphèmes et à l'intervocalique d'unités complexes. L'objectif est d'en mesurer le *Voice Onset Time* (VOT) initial et le v-ratio intervocalique et d'appliquer la méthode acoustique (H1-H2 et HNR) pour confirmer ou infirmer nos hypothèses.

2 Analyse de la production des phonèmes occlusifs

2.1 Méthode

2.1.1 Participants

Quatre hommes et trois femmes de 20 à 29 ans (soit une moyenne de 23 ans) ont participé activement à cette étude. Tous sont locuteurs natifs de wu et ont une connaissance et une pratique élevée de leur langue maternelle.

2.1.2 Stimuli

Ils ont été amenés à oraliser vingt et une unités complexes (TABLE 1). Elles sont toutes à l'écrit des mots dissyllabiques de structure : C₁V₁C₂V₂ où C peut être /p/, /t/ ou /k/ et V /a/ ou /æ/. 1 indique la position initiale et 2 indique la position intervocalique. La présence de /æ/ s'explique par le faible nombre de morphèmes ayant en position intervocalique après /t/ la voyelle /a/. Le phonème /a/ est proposé car son F1 est relativement haut et de ce fait, il influence moins le premier harmonique H1 (le fondamental) et le deuxième harmonique H2 qui sont les deux indices importants à mesurer.

Initiale	Intervocalique
牌子 [pa ²² tsy ³³] « marque »	失败 [sef ⁴³ ba ³¹] « échouer »
败家 [pa ²² ka ³³] « faraud »	雪白 [se ⁴³ ba ^{f43}] « blanc »
白色 [pa ^{f23} sə ⁴³] « blancheur »	单摆 [te ⁴⁴ pa ³¹] « pendule »
摆设 [pa ⁵¹ sef ³] « décor »	一百 [iə ^{f43} pa ^{f43}] « cent »
百万 [pa ^{f43} mɛ ³¹ /vɛ ³¹] « million »	大伯 [dəu ²² pa ^{f3}] « oncle »
爸爸 [pa ⁴⁴ pa ³¹] « père »	葡萄 [be ^{f23} dæ ³¹] « raisin »
大学 [ta ²² g ^h o ^f h ³] « université »	海带 [he ⁵¹ ta ²³] « laminaire »
带鱼 [ta ⁴⁴ ng ³¹] « trichiure »	绊倒 [poe ⁴⁴ tæ ³¹] « buter »
茄子 [ka ²² tsy ³³] « aubergine »	番茄 [fe ⁴⁴ ga ³¹] « tomate »
加油 [ka ⁴⁴ yeu ³³] « encourager »	添加 [t ^h ie ⁴⁴ ka ³¹] « ajouter »
格子 [ka ^{f43} tsɿ ⁵¹] « carreaux »	

TABLE 1 - Les 21 unités composées produites par les 7 locuteurs wu de Suzhou.

2.1.3 Enregistrement

L'enregistrement a été fait avec une carte son (*Cakewalk USB AudioCapture*) et le *MicroMic AKG C520I* dans un studio professionnel à Suzhou en 2017. Les fichiers audios ont été enregistrés via le logiciel *Adobe Audition CC 2017*. Les participants ont fait les enregistrements assis pour leur confort mais surtout pour éliminer les balancements éventuels. Les locuteurs devaient prononcer avec un débit normal chaque mot deux fois, soit un total de 294 (21 termes x 2 fois x 7 locuteurs) stimuli. Tous les locuteurs n'ont pas produit les 21 mots, d'où seulement 228 stimuli d'enregistrés. Une session d'entraînement s'est déroulée avant l'enregistrement.

2.1.4 Paramètres acoustiques mesurés

En premier lieu, nous avons mesuré le *Voice Onset Time* (VOT) (Cho, 1999 et Lisker et *al.*, 1964) qui correspond à la durée entre le relâchement consonantique et le début des pulsations périodiques glottiques. Il existe trois valeurs possibles : 1) le VOT est égal à zéro pour les occlusives non-voisées et non-aspirées ; 2) le VOT est négatif pour les occlusives voisées ; 3) le VOT est positif pour les occlusives aspirées.

À l'intervocalique, nous avons ensuite mesuré le v-ratio qui est le pourcentage de voisement pendant la durée de la consonne (Hallé et *al.*, 2011) et cela pour deux raisons : 1) Sachant que la mesure du VOT ne se fait qu'en position initiale, le v-ratio propose une quantification du degré de voisement dès lors que la durée de la consonne peut être déterminée, ce qui est le cas à intervocalique ; 2) Le *Voice Termination Time* (V.T.T) (Agnello, 1975) marquant le résidu de la voyelle précédente pendant l'occlusion intervocalique (FIGURE 3, intervalle en vert clair) peut être pris en compte pour compléter le calcul à partir du passage transitoire. Cette durée est également appelée *edge vibrations* par Lisker & Abramson (1964) ou *bleed* par Davidson (2016).

La mesure de H1-H2 se base sur le paramètre acoustique *spectral tilt* (Fant, 1982 ; Bickley, 1982 ; Gordon et *al.*, 2001) : au sein d'un segment voisé, pour une voix soufflée, le signal de la source voisée, correspondant à une onde quasi sinusoïdale, renforce l'amplitude de H1. Par conséquent, la différence d'amplitude entre le premier et le second harmonique est plus large pour un segment aspiré.

L'*Harmonics to Noise Ratio* (HNR) estime la proportion de l'intensité du signal harmonique en fonction du bruit (Krom, 1993, Murphy, 1999). Une voix soufflée étant accompagnée de davantage de bruit turbulent, le HNR est nécessairement plus faible que pour une voix non-soufflée.

Pour faire ces analyses, nous avons été amenés à modifier le script de Landron (2015) pour le logiciel *Praat* (Boersma, 2011) servant à mesurer le VOT et le v-ratio. Le programme *VoiceSauce* de Shue (2017) a été utilisé pour mesurer le HNR et le H1-H2. La segmentation et l'alignement ont été faits et vérifiés manuellement.

La segmentation a été faite par le biais des conventions suivantes (voir FIGURE 1) :

- a) Pour C_1 , le début est arbitraire puisqu'aucune information sur le début de l'occlusion n'est donnée sur le spectrogramme et l'oscillogramme. Le seul but de fixer une frontière initiale permet de faire fonctionner le script.
- b) Pour C_2 , le début a été déterminée après la fin de F_2 de la voyelle précédente, l'endroit où la dernière courbe harmonique traverse la ligne zéro en descendant. Pour la fin de la réalisation de C_1 et C_2 , nous avons choisi le moment où le signal traverse, en montant, la ligne zéro.

- c) Pour la mesure du v-ratio, notre script calcule des séquences (*frame*). Il calcule le pourcentage de séquences voisées par rapport à toutes les séquences délimitées (plus de 27 ms). Le v-ratio est calculé sur toute la durée de la consonne. Par contre, la mesure de v-ratio n'indique pas où se situe le voisement. Pour tous types de C₂, nous nous attendons à un v-ratio positif en tenant compte de la *Voice Termination Time* (V.T.T).

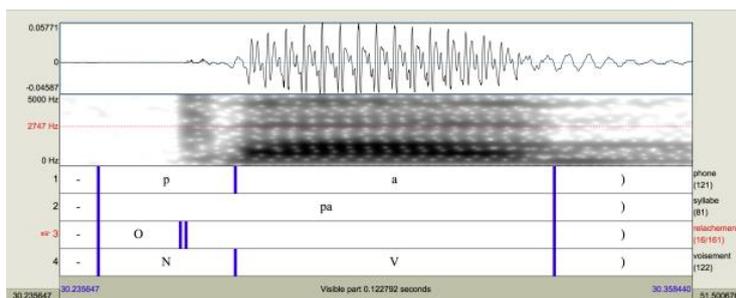


FIGURE 1 : Les tires 1,3 et 4 sont utilisées pour le calcul du VOT et du v-ratio. Sur la troisième tire, O signifie « Occlusion », R qui n'est pas visible ici signifie « Relâchement » et se situe dans l'espace très étroit après O. Sur la quatrième tire, N signifie « phase non-voisée » et V « phase voisée ». Le tiret - a été ajouté avant C₁ pour symboliser le silence tandis que la parenthèse fermante) remplace C₂. La délimitation arbitraire du début de l'occlusion pour C₁ est requise pour le calcul du VOT.

2.2 Résultats

2.2.1 HNR et H1-H2

Pour simplifier la présentation des syllabes avec différents tons, quatre étiquettes sont proposées : C₁V₁^B pour un ton bas ; C₁V₁^H pour un ton haut ; C₂V₂^(B) pour un ton inhérent bas mais modifié par sandhi ; C₂V₂^(H) pour un ton inhérent haut modifié par sandhi.

Les deux indices ont été mesurés au début (1/3), au milieu (2/3) et à la fin (3/3) de la réalisation des voyelles initiale et intervocalique. Pour H1-H2, la moyenne obtenue pour les trois intervalles a montré que V₁^B était manifestement plus soufflée que lors de la production de V₁^H (TABLE 2). Par contre, la caractéristique soufflée ne joue pas de rôle distinctif à l'intervocalique.

À propos du HNR, la mesure a été faite entre 0 à 3500 Hz. Le résultat montre que seul le HNR05 (entre 0-500 Hz) à l'initiale est significatif au milieu et à la fin de la réalisation (TABLE 2). C'est-à-dire que le HNR de V₁^B est moins élevé que celui de V₁^H et la différence se trouve au-dessus du F₁ pour V₁ qui est soit /a/ soit /æ/ (F₁ se situe entre 168Hz et 242Hz). Paradoxalement, il n'y a pas de différence au début alors qu'elle devrait subir une influence transitoire plus forte.

Mesure (dB) (n=58,54)	1/3	2/3	3/3
H ₁ -H ₂	1.45*	1.78*	2.32**
HNR05	-0.88	-3.97*	-6.08**

TABLE 2 : La différence de $V_1^B-V_1^H$ en fonction des mesures H1-H2 et HNR05 en trois points de la voyelle.
Signification : * pour $p < .05$, ** pour $p < .01$.

2.2.2 VOT et v-ratio

Le VOT de $C_1V_1^B$ est plus court que celui de $C_1V_1^H$ mais la différence n'est pas significative. Certes, nous avons remarqué quelques réalisations pré-voisées chez un locuteur, mais la majorité des locuteurs ne l'a pas produit à l'initiale. La moyenne du VOT reste positive pour $C_1V_1^B$ et $C_1V_1^H$ (TABLE 3). Ainsi, l'utilisation du terme « sonore » ne correspond pas ici à la notion traditionnelle de voisement.

Le v-ratio de $V_2^{(B)}$ s'avère plus élevé que celui de $V_2^{(H)}$. (TABLE 3). Cela suggère une réelle opposition au niveau du degré de voisement. Mais ce qui est remarquable c'est le taux de v-ratio assez élevé qui se manifeste lors de la production de $C_2V_2^{(H)}$. Il est probable que les C_2 étaient largement pré-voisées en structure $C_2V_2^{(H)}$ même si ce type de C_2 a été considéré comme sourd dans les études précédentes (Chao 1935, Cao 1987, Cao et al., 1992, Hu, 2001, Shi, 2009). De plus, cette proportion considérable pourrait s'interpréter par la prise en compte du VTT (cf. 2.1.4) dans un contexte intervocalique facilitant la poursuite des vibrations des plis vocaux.

Par ailleurs, en contexte $C_2V_2^{(B)}$, certaines réalisations phonétiques des consonnes sont implosives ainsi /p/ est réalisé [ɓ] avec des amplitudes de voisement très importantes (FIGURE 2) ce qui n'est absolument pas attesté pour la réalisation de /p/ en contexte $C_2V_2^{(H)}$ (FIGURE 3).

VOT (ms)		v-ratio %	
$C_1V_1^B$ (n=58)	$C_1V_1^H$ (n=54)	$C_2V_2^{(B)}$ (n=56)	$C_2V_2^{(H)}$ (n=60)
10.97	15.94	88.25**	73.92**

TABLE 3 : À gauche, la moyenne du VOT de $C_1V_1^B$ et $C_1V_1^H$. À droite, la moyenne du v-ratio entre $C_2V_2^{(B)}$ et $C_2V_2^{(H)}$. Signification : * pour $p < .05$, ** pour $p < .01$.

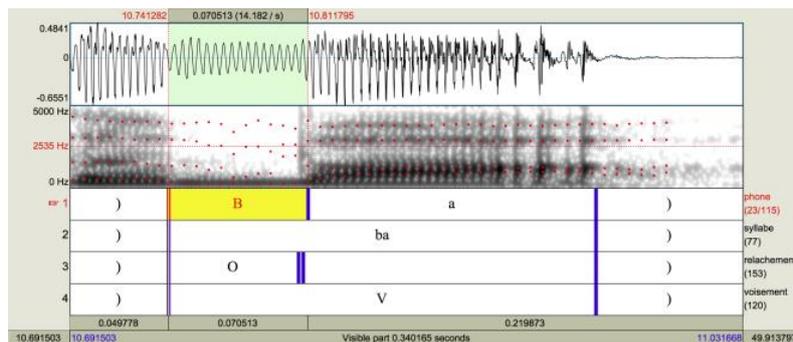


FIGURE 2 : Réalisation de [ɓ] par une locutrice wu de Suzhou. B noté dans la tire « phone » sert à indiquer une consonne intervocalique

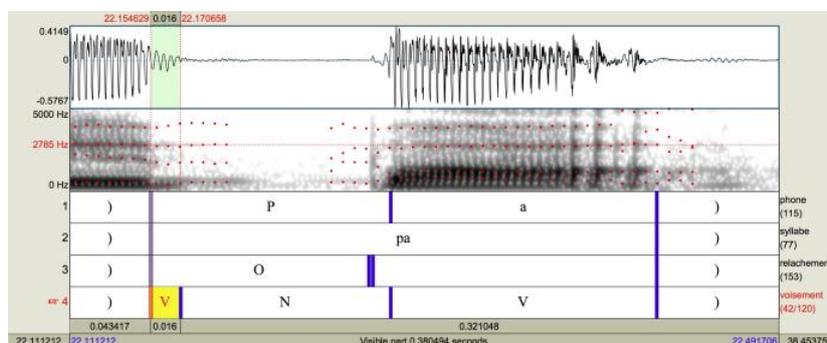


FIGURE 3: Le phonème /p/ est réalisé [p] par une locutrice wu de Suzhou. P note dans la tire « phone » la consonne intervocalique.

3 Discussion et conclusion

Dans ce travail, nous avons exploré les caractéristiques acoustiques des phonèmes occlusifs du dialecte wu de Suzhou. Nos résultats montrent que :

1) À l'initiale d'un morphème, le VOT ne sert pas à distinguer une série sonore dû à ses valeurs positives comme l'avaient déjà confirmé les scientifiques dans la littérature (Cao et *al.*, 1992 ; Shi, 2009 ; Gao et *al.*, 2012). En revanche, la différence des deux indices H1-H2 et HNR pourrait suggérer une plus grande proportion d'air turbulent lorsque V_1 porte un ton bas. D'où le terme de léger soufflement ou d'aspiration légère largement répandu.

Mais les mesures de ces deux facteurs n'ont pas donné des résultats homogènes notamment lorsque l'on cherche à préciser à partir de quel moment le soufflement commence à intervenir. Le résultat de HNR, qui a montré une différence significative entre V_1^B et V_1^H à partir des 2/3 de la durée vocalique, est à l'antipode de celui de H1-H2, qui a pu relever une différence significative dès le début de la voyelle. En fait, ces deux mesures s'intéressent à différents composants acoustiques : 1) La mesure de H1-H2 propose une voix plus soufflée quand le signal de la source voisée correspondant à une onde quasi sinusoïdale renforce l'amplitude de H1, d'où H1-H2 plus saillant. 2) La mesure de HNR dans notre étude, estime la proportion du bruit turbulent de la partie vocalique, d'où le HNR plus bas pour une voyelle plus soufflée. Selon une étude précédente (voir Gao et *al.*, 2012) sur le dialecte de Shanghai produit par des jeunes locuteurs (moyenne 26 ans), le H1-H2 permet de localiser le soufflement à l'*onset* de la voyelle et l'effet se propage jusqu'au milieu. Par contre, elle n'a pas fourni les données du HNR. Nos résultats hétérogènes sur le début de l'apparition du soufflement nous invitent à réfléchir au niveau physiologique : a) Sachant qu'au niveau physiologique, une grande proportion de turbulence d'air sortant pendant l'ouverture des plis vocaux demande une abduction écartée des plis vocaux suite à l'action crico-aryténoïdienne. « Aspiration is presumably a result of a somewhat open glottis » (Abramson et *al.*, 2017 : 76), est-ce qu'au début de la voyelle, la manœuvre laryngale compliquée est capable d'effectuer une abduction assez grande afin de créer une voyelle plus soufflée ? b) Sachant que ce soufflement se produit dans le contexte où le registre tonal est relativement bas, serait-il possible que la caractéristique soufflée

se stabilise et se renforce après le passage transitoire entre consonne et voyelle ? Dès que les plis vocaux courts et épais sont configurés en position stable et relâchée par le muscle crico-thyroïdien donnant ainsi un accès facile à la partie basse et médium de la voix, le soufflement se manifesterait.

2) À l'intervocalique d'une unité complexe dissyllabique, nous n'avons pas remarqué de différence significative pour H1-H2 et HNR entre $C_2V_2^{(B)}$ et $C_2V_2^{(H)}$. Par contre, une réalisation phonétique voisée a été constatée pour C_2 dans les deux contextes et le degré de voisement a une corrélation forte avec le ton intrinsèquement bas et modifié par sandhi de la V_2 . Par contre, nous n'avons pas obtenu la moyenne concernant le v-ratio pour $C_2V_2^{(B)}$ la hauteur attendue de 100% montrant que C_2 est toujours voisée à l'intervocalique (Cao et al., 1992 ; Shi, 2009 ; Gao et al., 2012). La FIGURE 4 montre un /p/ partiellement voisé repéré dans plusieurs réalisations de trois locuteurs. De plus, la partie encadrée ne devrait pas être une simple pause. Il est fort probable que cette partie puisse se traduire par une série d'ouverture et de fermeture glottique irrégulière entre la diminution considérable des amplitudes suivie de la voyelle précédente ([e] dans cette séquence) et le pré-voisement de la consonne intervocalique.

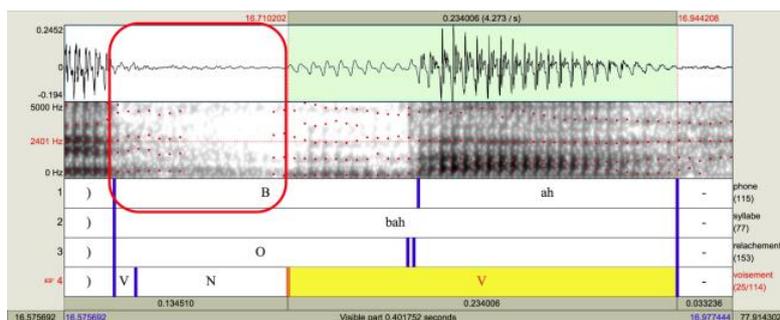


FIGURE 4: Réalisation partielle d'un [b] par une locutrice wu de Suzhou. B noté dans la tire « phone » indique la consonne intervocalique. La partie encadrée pourrait correspondre à une série d'ouverture et de fermeture glottique avant que le voisement ne se stabilise.

En résumé, nous pouvons donc déduire qu'à l'initiale, la caractéristique soufflée se manifeste d'une manière robuste dans le dialecte de Suzhou malgré le contact intense avec le mandarin. Cette particularité phonétique a été confirmée par les données acoustiques testées à Shanghai (Gao et al., 2012). Donc, le wu resterait en synchronie et d'un point de vue phonétique une langue sinitique différente du mandarin. De plus, à l'intervocalique, on a observé une augmentation du degré de voisement chez les jeunes locuteurs ce qui les conduit à produire des implosives [b, d, ɟ] en contexte $C_2V_2^{(B)}$. Par contre, il nous reste encore des questions à résoudre notamment à propos du mécanisme complexe permettant le pré-voisement partiel et le mécanisme amenant le soufflement initial. Pour y répondre, nous recourrons à EVA2TM (Computerised Vocal Assessment, SQLab) qui quantifie le débit d'air et la pression intra-orale dans nos prochaines études.

Références

- ABRAMSON A.S., Whalen D.H. (2017). Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics* 63, 75-86.
- AGNELLO J.G. (1975). Voice onset and voice termination features of stutterers. *Vocal tract dynamics and dysfluency: the proceedings of the first annual Hayes Martin Conference on Vocal Tract Dynamics*. New York: Webster, L.M. et Furst, L.C..
- BICKLEY C.A. (1982). Acoustic analysis and perception of breathy vowels. *MIT Speech communication group working papers* 1, 71-81.
- BOERSMA P. (2011). Praat, a system for doing phonetics by computer. *Glott International* 5(9/10), 341-345.
- CAO J., IAN M. (1992). An exploration of phonation types in Wu Dialects of Chinese. *Journal of Phonetics*. 20,79-92.
- CHAO Y. (1930). A System of Tone Letters. *Le Maître Phonétique* 45, 24-27.
- CHAO Y. (1935). *Xiandai wuyu de yanjiu [L'étude du dialecte Wu contemporain]*. Beijing. Shangwu Yinshuguan [La maison d'édition commerciale].
- CHO T. LADEFOGED P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27, 207-229.
- DAVIDSON, L. (2016). Variability in the implementation of voicing in American English obstruents., *Journal of Phonetics* 54, 35–50.
- FANT G. (1982). Preliminaries to analysis of the human voice source. *Quarterly Progress and Status Report, Speech Transmission Laboratory, KTH, Stockholm, Sweden*. 23(4), 1-27.
- GORDO M., LADEFOGED P. (2001). Phonation types: a cross-linguistic overview. *Journal of Phonetics*. 29(4), 383-406.
- GAO J., HALLÉ P. (2012). Caractérisation acoustique des obstruantes phonologiquement voisées du dialecte de Shanghai. Actes de *les Journées d'Études sur la Parole*, 57-64.
- HALLÉ P., ADDA-DECKER M. (2011). Voice assimilation in French obstruents: A gradient or a categorical process? *Tones and features: A festschrift for Nick Clements*. De Gruyter, 149-175.
- KROM G.D. A Cepstrum-Based Technique for Determining a Harmonics-to-Noise Ratio in Speech Signals. (1993). *Journal of Speech, Language, and Hearing Research*. 36(2), 254-266.
- LANDRON S. (2017). *L'opposition de voisement des occlusives orales du français par des locuteurs taiwanais*. Thèse en phonétique pour le grade de Docteur. Paris. Université Paris Nouvelle Sorbonne.
- LISKER L., ABRAMSON A.S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. 20 (3), 384-422.
- MURPHY P.J. (1999) Perturbation-free measurement of the harmonics-to-noise ratio in voice signals using pitch synchronous harmonic analysis. *The Journal of the Acoustical Society of America*. 105(5), 2866-2881.
- SHI F. (2009). *Shiyan yuyinxue tanjiu [L'exploration de la phonologie expérimentale]*. Beijing. Beijing Daxue Chubanshe [Peking University Press].
- SHUE Y. (2010). The voice source in speech production: Data, analysis and models. *UCLA dissertation*.