

# Caractérisation acoustique des obstruantes phonologiquement voisées du dialecte de Shanghai

Jiayin Gao<sup>1</sup> Pierre Hallé<sup>1,2</sup>

(1) Laboratoire de Phonétique et de Phonologie (UMR 7018), CNRS / Paris 3, 75005 Paris

(2) Laboratoire Mémoire et Cognition (Inserm S894), Paris 5, 92100 Boulogne-Billancourt

jiayin.gao@gmail.com, pierre.halle@univ-paris3.fr

## RESUME

---

Dans le dialecte de Shanghai, les obstruantes sonores en position initiale/accentuée de mot sont phonétiquement non-voisées et se distinguent des autres (i.e., sourdes non-aspirées et aspirées) principalement par leur registre tonal bas. Selon la littérature, elles sont aussi associées à une qualité de voix légèrement soufflée dans cette position. Dans cette étude, nous utilisons des mots monosyllabiques et dissyllabiques pour réexaminer la question de la voix soufflée comme caractérisant les obstruantes sonores du dialecte de Shanghai non seulement en position accentuée mais aussi dans les autres positions, où le sandhi tonal peut modifier le registre tonal. Nos résultats confirment la présence de voix soufflée pour les obstruantes sonores dont le registre tonal reste bas. Nous trouvons aussi que des patterns de durée systématiques caractérisent les obstruantes sonores de façon robuste dans toutes les positions.

## ABSTRACT

---

### Acoustic characterization of phonologically voiced obstruents in Shanghai dialect

In Shanghaiese, phonologically voiced obstruents in word-initial, accented position are phonetically voiceless and are distinguished from the others (i.e., voiceless and/or aspirated) mainly by a low tone register. Slightly breathy voice is also reported in the relevant literature as an additional characteristic of these obstruents in accented position. In this study, we used both monosyllabic and disyllabic words to revisit the issue of breathy phonation as characterizing phonologically voiced Shanghaiese obstruents not only in accented position but also in non-accented position, where sandhi may affect tone register. We found breathy phonation most clearly when tone register is not affected. We also found that systematic duration patterns robustly characterize phonologically voiced obstruents, whether or not in accented position.

MOTS-CLES : dialectes Wu, shanghaien, voix soufflée, voix slack, durée consonantique

KEYWORDS : Wu dialects, Shanghaiese, breathy/slack voice, consonant duration

---

## 1 Introduction

### 1.1 Généralités

Les tons des dialectes chinois sont nés de distinctions segmentales en chinois archaïque puis ancien. Pour les consonnes en position de coda, cela a produit les quatre contours *ping*, *shang*, *qu* et *ru*, et pour les consonnes initiales, les deux registres *yin* et *yang* : le registre *yin* (haut) pour les sourdes et le registre *yang* (bas) pour les sonores (Sagart, 1999). Ce dernier processus est appelé “bipartition tonale” (“tone split”) (Haudricourt,

1961) ; il est assez répandu dans les langues du sud-est asiatique. La distinction de registre a ainsi remplacé la distinction de voisement dans la plupart des dialectes chinois.

Les dialectes Wu possèdent pourtant distinction de voisement et de registre tonal (*yin* vs. *yang*) de façon complémentaire. La triple distinction des obstruantes (sourdes aspirées, non-aspirées et sonores) est une des caractéristiques majeures des dialectes Wu (Chao, 1928). En position initiale ou accentuée de mot, c'est le registre tonal qui distingue les sourdes (registre *yin*) des sonores (registre *yang*), toutes réalisées sans voisement ; en position non-initiale et non-accentuée, c'est le voisement qui les distingue : la distinction tonale est neutralisée par sandhi tonal (Cao, 1987).

De plus, il semble que les sonores en position initiale sont produites avec une qualité de voix "soufflée". Chao (1928) et Liu (1925) ont proposé que les obstruantes sonores à l'initiale sont réalisées comme des obstruantes non-voisées suivies par une aspiration voisée ('qing yin zhuo liu' : "son sourd, écoulement sonore"). Une qualité de voix soufflée est en effet suggérée par des mesures acoustiques (H1-H2 ou H1-F1 : Cao et Maddieson, 1992) ou articulatoires (ouverture glottale : Ren, 1988 ; mais voir l'étude articulatoire de Gao et al., 2011). Il s'agit d'une qualité de voix légèrement soufflée 'slack' plutôt que 'breathy', avec une ouverture glottale bien moindre que celle, par exemple, des occlusives voisées aspirées en hindi.

Notre étude est une étude acoustique des obstruantes du dialecte de Shanghai en attaque de mot monosyllabique, et de première ou deuxième syllabe de mot dissyllabique. Dans ce dernier cas, les sonores sont réalisées voisées mais le registre tonal réalisé sur la syllabe cible est déterminé par le ton de la syllabe précédente (sandhi : voir 1.2) : la qualité de voix reste-t-elle soufflée ? En plus des obstruantes (occlusives et fricatives), nous examinons les nasales /m/ et /n/ et l'attaque "vide" /Ø/.

## 1.2 Les tons du dialecte de Shanghai

Dans cette section, nous présentons brièvement les tons du shanghaien en isolation (i.e., en "citation") et en contexte polysyllabique après application des règles de sandhi. Nous nous limitons au sandhi à "dominante gauche", où le ton d'une syllabe forte se propage à droite sur les syllabes suivantes. Selon les descriptions traditionnelles, le shanghaien moderne possède 5 tons en citation : trois *yin* et deux *yang*. Deux tons dits "rentrants" ('ru sheng') sont portés par des syllabes courtes terminées par un coup de glotte. Dans cette étude, nous nous intéresserons uniquement aux tons non-rentrants : deux tons *yin*, notés ci-après T1 et T2 et un ton *yang* noté T3. Dans la notation à cinq niveaux de Chao (1930), ces trois tons ont les contours suivants en citation (Xu et Tang, 1988) : 53 (ou 52), 34 et 23 (pour T1, T2 et T3, respectivement), globalement en accord avec les données acoustiques de Rose (1993). Pour ce qui est du sandhi, nous nous limitons ici aux mots dissyllabiques avec première syllabe dominante. La propagation vers la droite du ton de la première syllabe modifie le contour tonal des syllabes suivantes, indépendamment de leur ton d'origine, ainsi que celui de la première syllabe, comme indiqué en (1) :

(1) 53 + S → 55-21 ; 34 + S → 33-44 ; 23 + S → 22-44 (quel que soit le ton de S)

Après application de (1), la première syllabe ne change pas de registre (si l'on considère que le registre *yin* correspond à une valeur moyenne supérieure ou égale à 3 et le *yang* à une valeur inférieure à 3) ; la seconde syllabe S est *yang* après T1, *yin* après T2 ou T3.

## 2 Étude de production

### 2.1 Méthode

#### 2.1.1 Participants

Nous avons enregistré à Paris 10 locuteurs natifs du shanghaien (5 hommes et 5 femmes) âgés de 22 à 30 ans (moyenne 26 ans) ayant vécu en France depuis quatre ans en moyenne. Aucun de ces locuteurs n'avait souffert de troubles du langage.

#### 2.1.2 Matériel

Le matériel utilisé consiste en 60 mots cibles monosyllabiques ou dissyllabiques insérés au début de la phrase porteuse *\_ gə ə zi ŋo nin tə ə* (“\_ ce caractère/mot, je le connais”). La syllabe cible apparaît dans trois conditions : comme mot monosyllabique (**mono**) et comme première (**S1**) ou deuxième (**S2**) syllabe de mot dissyllabique. Elle est composée d'une attaque (occlusive : /p<sup>h</sup>, p, b, t<sup>h</sup>, t, d/, fricative : /f, v, s, z/, nasale : /m, n/, ou attaque vide Ø) et de la rime /ɛ/.

Pour la condition S1, la syllabe cible (S1) peut porter l'un des tons d'origine T1, T2, ou T3. Le sandhi (1) ne change pas son registre tonal. Pour la condition S2, le ton d'origine de la première syllabe est soit T1 soit T2. La syllabe cible (S2) prend le contour 21 après T1 et 44 après T2. Ainsi, les syllabes *yang* comme /be/ restent *yang* après T1 mais deviennent *yin* après T2 tandis que les syllabes *yin* comme /pe/ deviennent *yang* après T1 mais restent *yin* après T2. La Table 1 montre les syllabes cibles pour les 3 conditions, et leurs contours tonals en citation et réalisé (après sandhi pour S1 et S2). Notons qu'ici, “syllabe *yang*” ou “*yin*” renvoie aux syllabes dont le ton en citation est *yang* ou *yin* ; “initiale *yang*” ou “*yin*” renvoie aux sonores ou sourdes : /be/ est *yang* ; /pe/ est *yin*.

Condition	Ton en citation	Ton réalisé	Syllabes	N
mono	yin (53 ou 34)	yin (53 ou 34)	/ɛ, p <sup>h</sup> ɛ, pɛ, t <sup>h</sup> ɛ, tɛ, fɛ, sɛ, mɛ, nɛ/	9
	yang (23)	yang (23)	/ɛ, bɛ, dɛ, vɛ, zɛ, mɛ, nɛ/	7
S1	yin (53 ou 34)	yin (55 ou 33)	/ɛ, p <sup>h</sup> ɛ, pɛ, t <sup>h</sup> ɛ, tɛ, fɛ, sɛ, mɛ, nɛ/	9
	yang (23)	yang (22)	/ɛ, bɛ, dɛ, vɛ, zɛ, mɛ, nɛ/	7
S2	yin (53 ou 34)	yin (44) après T2	/ɛ, p <sup>h</sup> ɛ, pɛ, t <sup>h</sup> ɛ, tɛ, fɛ, sɛ, mɛ/	8
		yang (21) après T1	/ɛ, pɛ, tɛ, fɛ, sɛ, mɛ/	6
	yang (23)	yin (44) après T2	/ɛ, bɛ, dɛ, vɛ, zɛ, mɛ, nɛ/	7
		yang (21) après T1	/ɛ, bɛ, dɛ, vɛ, zɛ, mɛ, nɛ/	7

TABLE 1 – Matériel utilisé divisé selon les 3 conditions ; en bleuté : changement de registre tonal.

### 2.1.3 Enregistrement et analyses des données

Les locuteurs ont été enregistrés directement sur ordinateur avec le logiciel Sound Studio (16 bits, 44.1 kHz) à partir d'un micro-casque AKG C520L relié à une carte son externe EDIROL. Les locuteurs ont prononcé chaque phrase deux fois, d'où un total de 1200 phrases (60 cibles x 2 répétitions x 10 sujets). Les syllabes cibles ont été analysées comme suit. Pour les consonnes initiales, nous avons mesuré la *durée* et le *v-ratio*, ou proportion de signal voisé (0 : non-voisé; 1 : entièrement voisé). Pour la voyelle /ε/, nous avons mesuré la *durée*, un *contour tonal* schématisé en 5 points par pas de 20% de la durée vocalique (Kratochvil, 1985) et la différence H1-H2 (dB) en début, milieu et fin de la voyelle. H1-H2 est d'autant plus élevé que la voix est plus soufflée.

## 2.2 Résultats

### 2.2.1 Contours tonals

Les contours moyens obtenus pour les syllabes cibles sont montrés dans la Figure 1. Ces contours sont normalisés pour la durée.

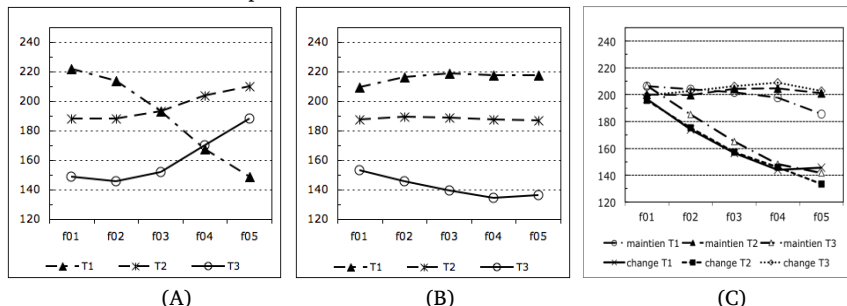


FIGURE 1 – Contours tonals moyens (Hz) sur la syllabe cible pour (A) la condition mono, (B) la condition S1 et (C) la condition S2 avec T1 ou T2 en première syllabe.

Les données correspondent bien aux descriptions de la littérature. En citation, T1 est haut tombant, pratiquement du haut jusqu'au bas de la plage de variation de F0, soit 51 ou 52 selon l'échelle à 5 niveaux ; T2 et T3 sont tous deux montants et se distinguent par le registre, plus haut pour T2 que T3 : 34 pour T2 et 23 pour T3 (Xu et Tang, 1988) semblent des notations raisonnables. En première syllabe de mot dissyllabique (condition S1), les contours réalisés sont pratiquement plats et pourraient être notés 55 (T1), 33 (T2), et 22 ou 21 (T3), en bon accord, avec les valeurs proposées par Xu et Tang (1988). En deuxième syllabe de mot dissyllabique (condition S2), les trois tons perdent bien leur contour d'origine ; après T2, ils prennent un contour haut plat que l'on pourrait noter 44 ou 33 ; après T1, ils prennent un contour bas tombant que l'on pourrait noter 32 ou 31 (plutôt que 21 selon Xu et Tang, 1988).

### 2.2.2 v-ratio

Dans les conditions mono et S1, les occlusives, qu'elles soient sourdes ou sonores ont une occlusion entièrement non-voisée (donc, v-ratio=0). Les /v/ ont un v-ratio plus

élevé que les /f/ ( $0.70 > 0.10$ ),  $p < .0001$  (pour mono et S1 regroupées,  $t$  de student bilatéral). La tendance est similaire pour les /z/ et les /s/ ( $0.18$  vs.  $0.09$ ) mais n'est pas significative,  $p = .10$ .

Dans la condition S2, les occlusives /b, d/ ont un v-ratio moyen de  $0.86$  contre  $0.31$  pour /p, t/ ; les fricatives /v, z/ ont un v-ratio de  $1$  (voisement complet) contre  $0.23$  pour /f, s/. Le v-ratio tend entre  $0.2$  et  $0.3$  pour les obstruantes sourdes correspondant sans doute à l'extinction du voisement de la voyelle précédente.

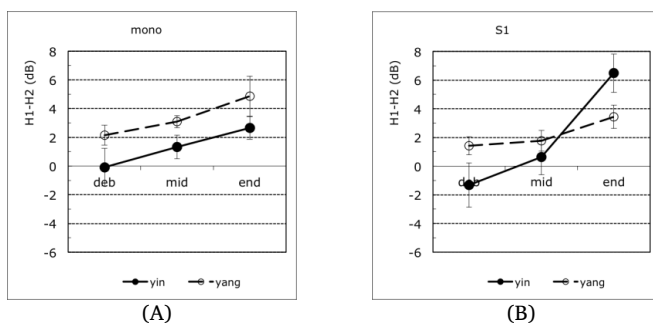
Les nasales ont un v-ratio proche de  $1$  (plus de  $0.9$ ) dans toutes les conditions.

### 2.2.3 H1-H2

Nous avons comparé les différences H1-H2 entre les syllabes *yin* et *yang* dans les trois conditions mono, S1 et S2. Pour la condition S2. Nous avons distingué les cas où le registre tonal de la syllabe S2 change (S2-change : T1 + *yin*, T2 + *yang*) de ceux où il ne change pas (S2-maintien : T1 + *yang*, T2 + *yin*). Dans les cas où le registre tonal est maintenu, on peut s'attendre à un avantage des syllabes *yang* sur les *yin* pour H1-H2 ; dans les cas où le registre change, il est possible que l'avantage s'inverse s'il est lié au registre tonal réalisé, ou qu'il soit maintenu s'il est lié au registre tonal d'origine.

Pour les initiales nasales, il n'y a pas d'avantage systématique des syllabes *yang* sur les *yin* en termes de H1-H2 pour les trois conditions : autrement dit, la qualité de voix n'est pas plus soufflée pour les syllabes *yang* que les *yin*, ce qui montre que la voix soufflée n'est pas liée uniquement au registre tonal.

Pour les obstruantes, il y a un avantage *yang* > *yin* d'environ  $2$  dB sur toute la voyelle (condition mono) ou au moins jusqu'en milieu de voyelle (condition S1 et condition S2-maintien). Le pattern tend à s'inverser (*yin* > *yang* d'environ  $1$  dB) en cas de changement de registre tonal (condition S2-change) : la valeur H1-H2 est plus élevée pour les syllabes d'origine *yin* que les syllabes d'origine *yang*, comme montré dans la Figure 2. Ceci suggère que la qualité de voix soufflée est liée au registre tonal réalisé de la syllabe et que cette qualité persiste au moins jusqu'en milieu de la voyelle. (Notons que nos résultats montrent une tendance inverse de celle rapportée par Iseli et al., 2006, qui trouvent une corrélation positive entre F0 et H1-H2 pour des voix masculines.)



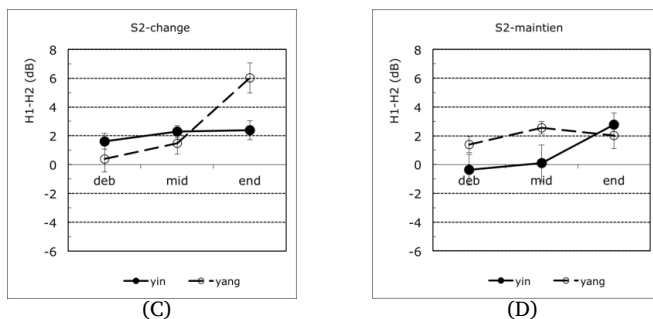


FIGURE 2–Valeurs de H1–H2 pour syllabes *yin* et *yang* attaque obstruante, en début, milieu et fin de voyelle, dans les conditions (A) mono (B) S1, (C) S2-change et (D) S2-maintien.

Pour les attaques vides, dans les deux conditions mono et S1, les données H1–H2 sont analogues à celles obtenues pour les obstruantes dans la condition S1 : davantage de voix soufflée pour les syllabes *yang* que les *yin* jusqu’en milieu de voyelle. Notons que cette différence *yang–yin* pourrait être due à un /f/ étymologique sous-jacent en attaque de syllabe *yang*. Dans la condition S2, l’avantage *yang > yin* est très net jusqu’en milieu de voyelle (de 9 à 5 dB) en cas de maintien de registre tonal, confirmant les données pour les conditions mono et S1 (avantage de 5 à 3 dB). L’avantage est moins net en cas de changement de registre tonal (~2 dB). Il semble donc que la qualité de voix soufflée des syllabes *yang* avec attaque vide, peut-être la trace d’un /f/ sous-jacent, tend à disparaître en cas de changement de registre tonal.

#### 2.2.4 durées

segment	mono		S1		S2-maintien		S2-change	
/f, v/	156*	87	111*	52	123*	50	122*	61
/s, z/	172*	148	145*	122	145*	48	120*	60
/p, b/	–	–	–	–	122*	51	96*	45
/t, d/	–	–	–	–	122*	68	105*	66
initiales	<u>164*</u>	<u>117</u>	<u>128*</u>	<u>87</u>	<u>128*</u>	<u>54</u>	<u>111*</u>	<u>58</u>
rime /ε/	204	242*	171	188*	159	164	151	174*

TABLE 2 – durées des consonnes *yin* (en blanc) et *yang* (en bleuté), et de la voyelle /ε/ (après fricatives ou occlusives) ; pour mono et S1, les durées des occlusions non-voisées ne peuvent être mesurées ; ‘\*’ : différences *yin–yang* significatives.

Nous trouvons des patterns de durées systématiques pour les consonnes initiales (sauf les nasales) ainsi que pour les voyelles. Ces patterns, récapitulés dans la Table 2, sont donc une caractéristique robuste distinguant les syllabes *yin* et *yang*.

Comme le montre la Table 2, les obstruantes *yin* ont une durée significativement plus longue que les obstruantes *yang* dans toutes les conditions, qu'il y ait changement de registre ou non. La voyelle / $\epsilon$ /, commune à tout le matériel, est plus longue après une initiale *yang* que *yin*. Nous avons conduit une analyse de variance sur les durées des consonnes (condition S2), avec les facteurs *Ton* (*yin* vs. *yang* : registre tonal en citation de la syllabe cible) et *Changement* (changement vs. maintien du registre tonal). Les facteurs *Ton* et *Changement*, ainsi que l'interaction *Ton* x *Changement* sont très significatifs,  $ps < .0005$ , montrant que les initiales *yin* sont plus longues que les *yang*, et que cet avantage des *yin* sur les *yang* est plus marqué lorsque le registre tonal est maintenu. (Il reste cependant très significatif en cas de changement de registre tonal.)

### 3 Discussion et conclusion

Dans ce travail, nous avons revisité les caractéristiques acoustiques des obstruantes du shanghaien. Nos résultats pour les v-ratios sont globalement conformes aux descriptions existantes de la littérature. Les occlusives *yang* et la fricative / $z$ / sont dévoisées en première syllabe de mot et voisées en seconde syllabe de mot (à l'intervocalique), mais la fricative / $v$ / tend à être voisée même en première syllabe de mot, contrairement à la description traditionnelle "son sourd, écoulement sonore".

Nos résultats pour H1–H2 confirment la présence de voix soufflée en début et milieu de rime, mais pas en fin de rime, en accord avec notre étude pilote (Gao et al., 2011). La voix soufflée en début de voyelle ne semble pas simplement résulter de la transition entre attaque et rime puisqu'elle est observée au moins jusqu'au milieu de la rime (voir aussi Cao et Maddieson, 1992). Elle n'est pas non plus une propriété de la syllabe entière, comme proposé par Chen (2010), puisqu'elle n'est pas observée en fin de rime. Les données obtenues pour la condition S2, où l'on peut comparer les situations où le registre tonal de la syllabe cible est maintenu avec celles où il est changé, suggèrent que la voix soufflée est plutôt associée au registre tonal réalisé qu'au registre d'origine. Autrement dit, la voix soufflée serait une caractéristique liée à un contour bas (à l'exception des syllabes avec initiale nasale), plutôt qu'à une initiale sonore (i.e., *yang*).

Les durées consonantiques et vocaliques que nous avons dégagées caractérisent les deux séries d'obstruantes, *yin* et *yang*, de façon bien plus tranchée que la qualité de voix plus ou moins soufflée : les obstruantes sourdes (i.e., *yin*) sont plus longues que les obstruantes sonores (i.e., *yang*), en première ou en deuxième syllabe de mot. En deuxième syllabe de mot, un changement de registre tonal par sandhi atténue sans annuler cette différence. De façon complémentaire, les voyelles (ici / $\epsilon$ /) sont plus courtes après les obstruantes sonores qu'après les sourdes.

Les différences de durée entre obstruantes sourdes et sonores sont une tendance universelle (O'Shaughnessy, 1981). Dans le dialecte de Shanghai, le voisement est un trait neutralisé dans certains contextes (en position initiale de première syllabe de mot) pour ce qui est des indices primaires de voisement (e.g., v-ratio). Pourtant, les patterns de durée liés au voisement *sous-jacent* sont maintenus. Nous n'avons pu étendre ces patterns aux occlusives en position initiale absolue, puisque leur occlusion, non-voisée dans cette situation, n'est pas mesurable à partir du signal acoustique. Nous nous proposons d'effectuer à l'avenir des mesures physiologiques pour combler cette lacune.

## Références

- CAO, J-F. (1987). The ancient initial “voiced” consonants in modern Wu dialects. *In Proc. 11<sup>th</sup> ICPHS 1987*, Tallinn, pages 169–172.
- CAO, J-F. et MADDIESON, I. (1992). An exploration of phonation types in Wu dialects of Chinese. *Journal of Phonetics*, 20, pages 77–92.
- CHAO, Y-R. (1928). *Studies in the modern Wu dialects*. Beijing: Qinghua University Press.
- CHAO, Y-R. (1930). A system of tone-letters. *Le Maître Phonétique*, 45, pages 24–27.
- CHEN, Z-M. (2010). An acoustic study of voiceless onset followed by breathiness of Wu dialects : based on the Shanghai dialect. *Yuyan yanjiu*, 30(3), pages 20–34.
- GAO, J-Y., HALLÉ, P., HONDA, K., MAEDA, S. et TODA, M. (2011). Shanghai slack voice : acoustic and ePGG data. *In Proc. 17<sup>th</sup> ICPHS 2011*, Hong Kong, pages 719–722.
- HAUDRICOURT, A. G. (1961). Bipartition et tripartition des systèmes de tons dans quelques langues d’Extrême-Orient. *Bulletin de la Société de Linguistique de Paris*, pages 163–180.
- ISELI, M., SHUE, Y. et ALWAN, A. (2006). Age- and Gender-Dependent Analysis of Voice Source Characteristics. *In Proc. ICASSP mai 2006*, Toulouse, I-389.
- KRATOCHVIL, P. (1985). Variable norms in Beijing prosody. *Cahiers de Linguistique Asie Orientale*, 14, pages 135-174.
- LIU, F. (1961). *Étude expérimentale sur les tons du chinois* (Thèse). Société d’édition “Les Belles lettres”, Paris.
- O’SHAUGNESSY, D. (1961). A study of French vowel and consonant durations. *Journal of Phonetics*, 9, pages 385–406.
- REN, N-Q. (1988). A Fiberoptic and transillumination study of Shanghai stops. Paper presented at *International Conference on Wu dialects*, Hong Kong.
- ROSE, P. (1993). A linguistic-phonetic acoustic analysis of Shanghai tones. *Australian Journal of Linguistics*, 13, pages 185–220.
- SAGART L. (1999). The origin of Chinese tones. *In Proc. of the Symposium/Cross-Linguistic Studies of Tonal Phenomena/Tonogenesis, Typology and Related Topics*, pages 91–64. Tokyo University of Foreign Studies.
- XU B-H., et TANG Z-Z. (1988). *Shanghai shiqu fangyan yanjiu* [A description of urban Shanghai dialects]. Shanghai Jiaoyu Chubanshe.