HAL
open science

Ambiguïté temporaire des obstruantes voisées en parole chuchotée

Yohann Meynadier, Dufour Sophie

To cite this version:

Yohann Meynadier, Dufour Sophie. Ambiguïté temporaire des obstruantes voisées en parole chuchotée. XXXIIè Journées d'Études sur la Parole, Jun 2018, Aix-en-Provence, France. pp.125-133. hal-01980724

HAL Id: hal-01980724

https://hal.science/hal-01980724

Submitted on 14 Jan 2019

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.

Ambiguïté temporaire des obstruantes voisées en parole chuchotée

Yohann Meynadier, Sophie Dufour
Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France
yohann.meynadier@univ-amu.fr, sophie.dufour@lpl-aix.fr

RESUME

Meynadier & Dufour (2016) ont étudié la résolution perceptive d'ambiguïtés de voisement en parole chuchotée en français. Via un paradigme d'amorçage sémantique, des mots cibles visuels ont été présentés immédiatement après la fin de mots amorces audio sémantiquement reliées aux cibles. En parole chuchotée, un effet d'amorçage a été observé seulement pour les amorces contenant une obstruante sourde (/pe*t*al/ amorce FLEUR), mais pas pour les amorces à obstruante voisée : /pe*d*al/ ne facilitait ni le traitement de son associé sémantique VÉLO, ni celui de FLEUR, relié à l'autre membre de la paire minimale à obstruante sourde (/pe*t*al/). Dans cette étude, les cibles visuelles sont présentées 50 ms après les amorces audio. Contrairement à ce qui a été observé avec un délai de 0 ms, seules les mots chuchotés à obstruante voisée amorcent leur associé sémantique (/pe*d*al/–VÉLO), ce qui n'est plus le cas des amorces chuchotées à obstruante sourde (/pe*t*al/–FLEUR). Ces études suggèrent que des indices résiduels du voisement sont exploités par les auditeurs pour retrouver le mot chuchoté. Néanmoins, la reconnaissance du trait [+voisé] en parole chuchotée ne serait pas immédiate et nécessiterait un certain temps supplémentaire de traitement.

ABSTRACT

Temporary ambiguity of voiced obstruantes in whispered speech

Meynadier & Dufour (2016) have studied how French listeners resolve voicing ambiguities in whispered speech. Visual targets were presented at the offset of auditory semantically-associated primes. A similar priming effect to that observed with modal primes was found only with whispered primes included a voiceless consonant (/pe*t*al/ *"pe*t*al"* primed FLEUR *"flower"*). No priming effect was found with whispered primes included a voiced consonant (/pe*d*al/ *"pe*d*al"*), either on the target VÉLO "bike" related to /pe*d*al/ or on the target FLEUR related to /pe*t*al/. Here, visual targets were presented 50 ms after the offset of primes. While priming effects were no longer observed with whispered voiceless primes (/pe*t*al/–FLEUR), a priming effect emerged with whispered voiced primes (/pe*d*al/–VÉLO). Together, our results suggest that residual correlates of voicing present in whispered speech are exploited by listeners to recover the intended words. Nonetheless, they also showed that the reconstruction of the voiced feature is not immediate during whispered word recognition and requires a certain amount of time to be done.

MOTS-CLES : voisement, parole chuchotée, phonologie, reconnaissance lexicale
KEYWORDS: voicing, whispered speech, phonology, lexical recognition

1 Introduction

Cette étude constitue un second volet à nos travaux sur la reconnaissance des mots chuchotés en français, particulièrement focalisé sur l'ambiguïté potentielle des obstruantes voisées du fait de

10.21437/JEP.2018-15

l'absence totale de vibrations laryngées dans ce mode de phonation. Cette absence de vibrations laryngées provoque d'importantes modifications acoustiques qui affectent davantage les segments voisés que les segments sourds (Ito et al. 2005, Jovičić & Šarić 2008). Logiquement, la modification principale concerne le trait [+voisé] phonétiquement réalisé par un assourdissement complet des segments phonologiquement voisés. Ainsi, en français un mot tel que *pédale* /pe*d*al/ devient potentiellement ambigu parce que sa prononciation en voix chuchotée se rapproche de manière critique de celle de *pétale* /pe*t*al/, son seul voisin phonologique en paire minimale de voisement. Néanmoins, les assimilations complètes de voisement ne neutraliseraient pas totalement les oppositions par le seul trait de voisement. Ainsi, par exemple, Hallé et al. (2012) observent que les consonnes assimilées en voisement maintiennent des durées phonétiques propres à leur identité sous-jacente [±voisé], à savoir plus longues pour les [–voisé] que pour [+voisé]. Kohlberger & Strycharczuk (2015) rapportent un résultat similaire pour la consonne /s/ en contexte d'assimilation de voisement en parole chuchotée. Par ailleurs, quelques études montrent que des corrélats physiologiques du voisement sont préservés en voix chuchotée, tels que des différences de pression intraorale (Netsell 1969, Slis 1970, Garnier et al. 2014, Meynadier 2015) et d'ouverture glottique (Malécot & Peebles 1965, Mills 2009, Meynadier 2015). Sur le plan acoustique, alors que les mesures spectrales (par ex., transition du F1, intensité et fréquence de bruit) semblent peu ou variablement corrélées au contraste voisement (Tartter 1989, Jovičić & Šarić 2008, van de Velde & van Heuven 2011, Gilichinskaya 2012, Kohlberger & Strycharczuk 2015), les durées des consonnes et des voyelles préconsonantiques se révèlent y être plus systématiquement associées (Sharf 1964, Schwartz 1972, Parnell et al. 1977, Tartter 1989, Jovičić & Šarić 2008, Vercherand 2010, van de Velde & van Heuven 2011, Gilichinskaya 2012, Meynadier & Gaydina 2013), comme en voix modale.

Les quelques études ayant examiné la perception du voisement des obstruantes dans des mots, non-mots ou syllabes montrent qu'elle est assez bien maintenue (au-dessus du seuil du hasard) en voix chuchotée, bien que moins performante qu'en voix modale (Dannenbring 1980, Tartter 1989, Munro 1990, Mills 2009, Vercherand 2010, Fux 2012, Gilinchskaya 2012). Cependant, si l'on s'y attarde, pour la plupart, cette bonne performance repose en grande partie sur un biais perceptif en faveur des obstruantes [–voisé] montrant un taux de reconnaissance souvent bien supérieur à celui des obstruantes [+voisé]. Par exemple, dans les études sur l'anglais le pourcentage d'identification correcte se situe entre 81 et 87 % pour les [–voisé] et entre 68 et 80 % pour les [+voisé] (Tartter 1989, Gilinchskaya 2012). En français, Fux (2012) observe un taux de 80 à 87 % pour les [–voisé] et de 16 à 47 % pour les [+voisés]. Ce biais perceptif en faveur des [–voisé] atténuerait ainsi l'hypothèse de corrélats résiduels du trait de voisement parfaitement exploités par les auditeurs pour reconnaître les mots chuchotés présentant un contraste de voisement. Plus généralement, cela suggère que les auditeurs, et plus particulièrement les auditeurs français, échoueraient à reconnaitre des mots chuchotés à obstruante voisée, comme *pédale* /pe*d*al/, et pourraient les confondre avec leur voisin phonologique à obstruante sourde, tel que *pétale* /pe*t*al/. Une limite partagée par toutes ces études est l'utilisation de tâches off-line, sans contrainte de temps de réponse, le plus souvent à choix forcé, favorisant l'implication de processus métalinguistiques ne permettant pas de refléter précisément les processus en temps réel de reconnaissance des mots parlés.

Pour cette raison, nous avons jugé utile de réexaminer la perception du trait de voisement à l'aide d'un paradigme d'amorçage sémantique, connu pour refléter des processus en temps réel de la reconnaissance des mots parlés, et en particulier l'activation des représentations lexicales (Tabossi 1996). Relié à notre question, ce paradigme consiste à présenter auditivement un mot amorce chuchoté potentiellement ambigu suivi d'un mot cible présenté visuellement, sémantiquement relié ou non à l'amorce, sur lequel les participants ont à réaliser une tâche de décision lexicale. Dans une première étude (Meynadier & Dufour 2016), nous avions observé que seules des amorces chuchotées à obstruante [–voisé] (/pe*t*al/) facilitaient le traitement ultérieur d'un mot cible relié

sémantiquement (FLEUR), alors qu'aucun effet d'amorçage de la cible reliée sémantiquement (VÉLO) n'apparaissait pour les amorces chuchotées à obstruante [+voisé] (/pé*d*ale/). De plus, ni les amorces à obstruante [–voisé], ni celle à obstruante [+voisé] précédant la cible reliée à l'autre membre de la paire minimale ne montrait d'effet d'amorçage (/pe*t*al/-VÉLO ou /pe*d*al/-FLEUR). Un point clé dans cette étude était que les cibles visuelles étaient présentées immédiatement après la fin acoustique des amorces auditives, c'est-à-dire avec un délai de 0 ms. Ces résultats laissaient ainsi suggérer qu'au moment où les mots cibles étaient présentés, si les obstruantes sourdes n'étaient plus ambiguës quant à leur trait sous-jacent, l'ambigüité liée au trait [+voisé] n'était quant à elle pas encore résolue. En effet, /pe*d*al/ n'amorçait pas sa cible VÉLO, mais n'était pas non plus confondu avec /pe*t*al/, puisque /pe*d*al/ chuchoté n'amorçait pas non plus FLEUR (cible reliée sémantiquement à /pe*t*al/). Dès lors, il s'avérait qu'un certain temps supplémentaire de traitement pouvait être nécessaire à la reconstruction du trait de voisement en parole chuchotée.

Afin de tester cette hypothèse, nous avons reconduit ici exactement la même expérience que dans notre étude de 2016, excepté que le délai entre la fin de la présentation des amorces audio et des cibles visuelles était allongé à 50 ms. Nous nous attendons donc qu'avec ce délai de traitement rallongé les mots chuchotés à obstruante [+voisé] amorcent uniquement leur cible sémantiquement reliée. Dès lors, le trait [+voisé] pourrait bien être identifié, probablement sur la base des corrélats phonétiques du voisement préservés en phonation chuchotée, à savoir plus précisément des durées consonantiques et/ou de la voyelle précédente. Afin d'évaluer l'implication des corrélats spectraux et temporels dans la reconnaissance du trait de voisement des obstruantes chuchotées, une analyse acoustique des mots amorces est proposée, avant de conclure sur une discussion générale de ces résultats de perception et de production.

2 Perception des obstruantes chuchotées

Les expériences de cette étude reprennent exactement le même matériel linguistique, le même design expérimental et les mêmes stimuli enregistrés pour l'étude de Meynadier & Dufour (2016). Nous y renvoyons le lecteur pour plus de détails. Ici, nous rappellerons l'essentiel de ce protocole et des différences avec l'étude antérieure, avant de présenter et de discuter les nouveaux résultats.

2.1 Matériel linguistique et design expérimental

Toutes les oppositions distinctives de voisement du français ont été testées dans ces expériences : /p t k f s ʃ/ *vs* /b d g v z ʒ/. Ces consonnes en opposition apparaissaient en position intervocalique de 2 x 20 mots en paire minimale (par ex., pé*t*ale–pé*d*ale, de*ss*ert–dé*s*ert, am*ph*i–en*v*ie). Ces 40 mots ont été soumis à un pré-test d'association sémantique libre passé par 8 hommes et 30 femmes (français, 20-47 ans), afin d'établir les couples amorce–cible reliée les plus fréquents (à savoir > 20% des réponses), par ex. pé*t*ale–FLEUR, pé*d*ale–VÉLO, de*ss*ert–CHOCOLAT, dé*s*ert–SABLE, am*ph*i–COURS, en*v*ie–DÉSIR. Les caractéristiques de ces amorces et de ces cibles, telles que le taux d'association sémantique, le nombre de phonèmes, de lettres et de syllabes, ainsi que leur point d'unicité et leur fréquence cumulée, ont été contrôlées (cf. Table 1, Meynadier & Dufour 2016). Pour mesurer l'effet d'amorçage attendu, chaque cible d'une amorce en paire minimale a été associée à un mot non relié sémantiquement (par ex., quittance–FLEUR/VÉLO, jumelle–CHOCOLAT/SABLE, héros–COURS/DÉSIR), constituant ainsi les conditions de contrôle comparées aux conditions de test où l'amorce est issue d'une paire minimale de voisement et reliée sémantiquement à la cible. L'éventualité d'un amorçage croisé étant aussi testé, les amorces en paire minimale ont également été associées à la cible reliée à l'autre membre de la paire minimale en opposition de voisement, par ex. pé*t*ale–VÉLO, pé*d*ale–FLEUR, de*ss*ert–SABLE, dé*s*ert–CHOCOLAT, am*ph*i–DÉSIR, en*v*ie–COURS. Les listes correspondant à ces huit conditions expérimentales sont rapportées dans la Table 1.

Amorces à obstruante [+voisé]					
Appariement de voisement congruent			Appariement de voisement incongruent		
contrôle	en paire minimale	CIBLE	*contrôle*	en paire minimale	CIBLE
grossi	briguer	mandat	*grossi*	briguer	nettoyer
jumelle	désert	sable	*jumelle*	désert	chocolat
bandits	revue	magazine	*bandits*	revue	non
obus	abat	viande	*obus*	abat	poisson
pincer	gaver	oie	*pincer*	gaver	erreur
lingot	cajou	noix	*lingot*	cajou	bonbon
entraver	aggraver	pire	*entraver*	aggraver	feuilles
douleur	modèle	mannequin	*douleur*	modèle	hôtel
outils	égaux	pareils	*outils*	égaux	montagne
héros	envie	désir	*héros*	envie	cours
guidon	badaud	passant	*guidon*	badaud	mer
talent	radeau	méduse	*talent*	radeau	pelle
fumer	caser	ranger	*fumer*	caser	briser
éduquer	embauché	travail	*éduquer*	embauché	argent
abouti	avaler	déglutir	*abouti*	avaler	canapé
tampon	combat	boxe	*tampon*	combat	math
quittance	pédale	vélo	*quittance*	pédale	fleur
impliquer	embraser	feu	*impliquer*	embraser	bisou
pinceaux	cabot	chien	*pinceaux*	cabot	voiture
verrue	condé	flic	*verrue*	condé	fromage

Amorces à obstruante [–voisé]					
Appariement de voisement congruent			Appariement de voisement incongruent		
contrôle	en paire minimale	CIBLE	*contrôle*	en paire minimale	CIBLE
grossi	briquer	nettoyer	*grossi*	briquer	mandat
jumelle	dessert	chocolat	*jumelle*	dessert	sable
bandits	refus	non	*bandits*	refus	magazine
obus	appat	poisson	*obus*	appat	viande
pincer	gaffer	erreur	*pincer*	gaffer	oie
lingot	cachou	bonbon	*lingot*	cachou	noix
entraver	agrafer	feuilles	*entraver*	agrafer	pire
douleur	motel	hôtel	*douleur*	motel	mannequin
outils	echo	montagne	*outils*	echo	pareils
héros	amphi	cours	*héros*	amphi	désir
guidon	bateau	mer	*guidon*	bateau	passant
talent	rateau	pelle	*talent*	rateau	méduse
fumer	casser	briser	*fumer*	casser	ranger
éduquer	empocher	argent	*éduquer*	empocher	travail
abouti	affalé	canapé	*abouti*	affalé	déglutir
tampon	compas	math	*tampon*	compas	boxe
quittance	pétale	fleur	*quittance*	pétale	vélo
impliquer	embrasser	bisou	*impliquer*	embrasser	feu
pinceaux	capot	voiture	*pinceaux*	capot	chien
verrue	comté	fromage	*verrue*	comté	flic

TABLE 1 : Listes des stimuli amorce–cible et contrôle–cible utilisés dans les tests

30 couples amorce–cible constituées de mots non reliés sémantiquement (par ex. nuage–PION) ont été ajoutés aux stimuli test. La tâche des sujets étant de décider si la séquence de lettre (la cible) affichée à l'écran (et non le mot amorce entendu) était un mot du lexique français ou non, ont été créés 50 couples de mot amorce–pseudo-mot cible, dont 10 dérivés d'un mot réel relié sémantiquement à l'amorce (par ex. tomate–ROUZE) et 40 sans lien sémantique potentiel (par ex. visage–PLAME). Ces pseudo-mots ont été créés sur la base d'un mot réel présentant une substitution ou une permutation de lettres. A terme, les couples de mots composés d'une amorce en paire minimale de voisement associée à une cible reliée sémantiquement représentaient 20 % des stimuli de chaque liste proposée à chaque sujet. Afin que les sujets ne soient jamais soumis deux fois à la même amorce ou à la même cible, les stimuli ont été répartis en 8 listes affectées chacune à 8 groupes de 17 sujets différents (N = 136). L'expérience durait environ 30 mn.

L'expérience en parole chuchotée de la première étude (Meynadier & Dufour 2016) est reconduite ici.[1] Comme elle, les nouveaux tests portent sur les 8 conditions présentées dans la Table 1 : (i) deux où l'amorce en paire minimale est reliée sémantiquement à la cible congruente en appariement de voisement (pé*d*ale–VÉLO, pé*t*ale–FLEUR) ; (ii) deux où l'amorce en paire minimale est sans lien sémantique avec la cible qui correspond à celle reliée sémantiquement de l'autre membre de la paire minimale de voisement (appariement de voisement incongruent : pé*d*ale–FLEUR, pé*t*ale–VÉLO) ; (iii) quatre amorces contrôles pour chaque condition (i) et (ii) où aucun lien ne relie l'amorce et la cible (quittance–VÉLO, quittance–FLEUR). Seule la durée du délai entre l'amorce audio et la cible visuelle allongée à 50 ms (au lieu de 0 ms) diffère de la première étude. La taille de ce délai inter-stimuli est la plus fréquemment employée dans les études utilisant le paradigme de l'amorçage phonologique (par ex., où une amorce partage des phonèmes avec sa cible, cf. Dufour 2008 pour une revue).

2.2 Participants et passation des tests

Les participants, différents mais comparables à ceux de la première étude, ont été rémunérés. La passation des tests a suivi exactement les mêmes conditions que celle de l'étude de 2016, excepté

[1] L'expérience en parole modale (4 listes pour 4 groupes x 17 sujets, N = 68), où seules les deux conditions avec appariement de voisement congruent et les deux contrôles ont été testées, ne sont pas détaillées ici. Les résultats de ces tests sont rapportés dans la discussion générale.

que le délai entre l'amorce audio et la cible visuelle a été fixé à 50 ms.

2.3 Mesures et analyses statistiques

Les résultats ont été analysés au moyen d'un modèle linéaire généralisé à effets mixtes (Baayen 2008). Les analyses ont porté sur les temps de réaction (RT en ms) incluant uniquement les réponses correctes. Par ailleurs, les RT supérieurs à 1300 ms ont été exclus. Afin que le modèle respecte les critères de normalité de la distribution des résidus ainsi que l'homogénéité de la variance, une transformation logarithmique a été appliquée sur les RT (Baayen & Milin 2010), puis pour chaque participant les RT supérieurs et inférieurs à 2,5 écart-type de leur moyenne ont été également écartés. 4,52% des données ont été ainsi rejetées. Les valeurs rapportées pour F et *p* ont été obtenues à l'aide du package *lmertest* du logiciel R (Kuznetsova et al. 2016).

Le modèle incluait le *Voisement* ([+voisé], [–voisé]), l'*Appariement de voisement* (congruent, incongruent) et le *Type d'amorce* (en paire minimale, contrôle) comme facteurs à effets fixes (Table 1). La structure aléatoire du modèle incluait un intercept différent pour les participants et les items, ainsi que des pentes différentes à la fois par participants et par items pour le facteur *Type d'amorce*. Les comparaisons deux à deux, permettant d'évaluer les effets d'amorçage, ont été réalisées à l'aide de la fonction *glht* du package *multcomp* (Bretz et al. 2011) avec une correction de Bonferroni. Les RT moyens ainsi que les pourcentages de réponses correctes obtenus dans chaque condition sont présentés dans la Figure 1.

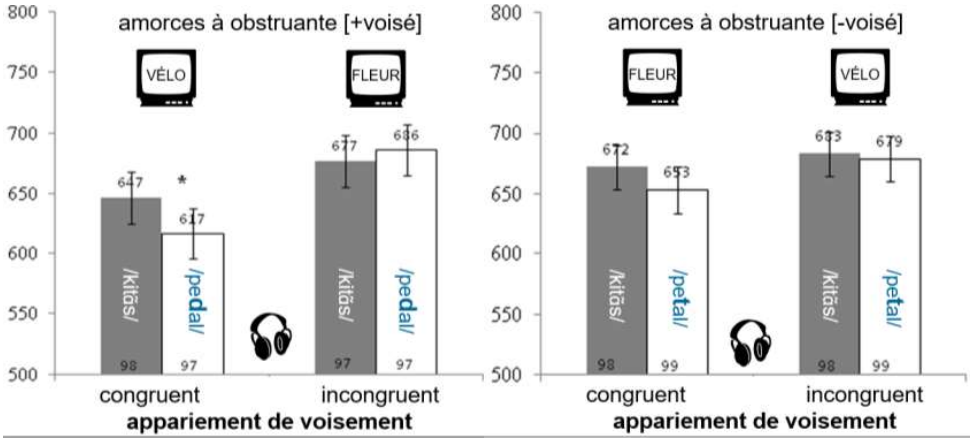

Figure 1 : RT moyen (ms) et erreur-type des réponses pour les amorces chuchotées contrôles (barres grises) vs en paire minimale à obstruantes [+voisé] et [–voisé] (barres blanches) selon les conditions d'appariement de voisement (congruent vs incongruent). Les pourcentages de réponses correctes sont donnés au bas des barres. Un astérisque (*) indique un effet d'amorçage significatif.

2.4 Résultats

Le modèle révèle un effet simple du *Type d'amorce* (paire minimale, contrôle) [F=5.22, p<.05] avec des RT plus rapides sur les mots cibles précédés d'une amorce en paire minimale que d'une amorce contrôle. Les effets simples de *Voisement* ([+voisé], [–voisé]) et d'*Appariement de voisement* (congruent, incongruent) n'atteignent pas la significativité. Seule l'interaction *Type d'amorce* Appariement de voisement* est significative [F=6.64, p<.05]. De façon globale, cette interaction est due à un effet d'amorçage plus important dans le cas d'un appariement congruent que dans le cas d'un appariement incongruent. La double interaction *Voisement*Type d'amorce* Appariement de voisement* n'est pas significative [F=1.86, p=.18].

Cependant, de façon cruciale, les sous-analyses dans chaque condition de *Voisement* ([+voisé] et

[–voisé]) révèlent une interaction significative *Type d'amorce*Appariement de voisement* seulement pour les amorces à obstruante [+voisé] (pé*d*ale) [F=6.56, p<.05] et non pour celles à obstruante [–voisé] (pé*t*ale) [F=0.44, p>.20]. Pour les amorces à obstruante [+voisé], les comparaisons deux à deux avec une correction de Bonferroni montrent un effet d'amorçage significatif seulement pour les couples amorce–cible appariés en voisement (congruent, pé*d*ale–VÉLO) [z=3.03, p<.05]. Dans ce cas, le RT (617 ms) est en moyenne 30 ms plus rapide que dans la condition contrôle (647 ms, quittance–VÉLO). En revanche, aucun effet d'amorçage n'a été observé pour les couples amorce–cible non appariés en voisement (incongruent, pé*d*ale–FLEUR comparé au contrôle quittance–FLEUR) [z=-1.01, p>.20]. Pour les amorces à obstruante [–voisé], aucun effet d'amorçage n'est observé, tant pour les couples amorce–cible appariés en voisement (congruent, pé*t*ale–FLEUR comparé au contrôle quittance–FLEUR) [z=2.06, p>.20] que pour ceux non appariés (incongruent, pé*t*ale–VÉLO comparé au contrôle quittance–VÉLO) [z=0.92, p>.20].

Pour les pourcentages de réponses correctes, un modèle *logit* à effets mixtes n'a montré aucun effet.

En résumé, contrairement aux résultats de notre première étude où amorce et cible se suivaient sans délai, un délai de 50 ms permet l'émergence d'un effet d'amorçage avec des amorces chuchotées à obstruante [+voisé]. Cependant, l'effet d'amorçage précédemment observé (Meynadier & Dufour 2016) pour les mots amorce à obstruante [–voisé] ne l'est plus avec ce délai de traitement plus long.

3 Analyse acoustique des obstruantes voisées vs sourdes

L'analyse porte sur les 40 (2x20) mots en paire minimale de voisement utilisés comme amorce audio dans les expériences de perception et produits en phonation modale et chuchotée par l'un des auteurs de l'étude (Figure 2). Ce sont les mêmes stimuli que ceux de la première étude. Nous renvoyons à Meynadier & Dufour (2016) pour les détails relatifs à leur enregistrement et leur segmentation manuelle. Cette liste de 40 mots comprend 4 paires avec /p–b/, 5 avec /t–d/, 2 avec /k–g/, 5 avec /f–v/, 3 avec /s–z/ et 1 avec /ʃ–ʒ/.

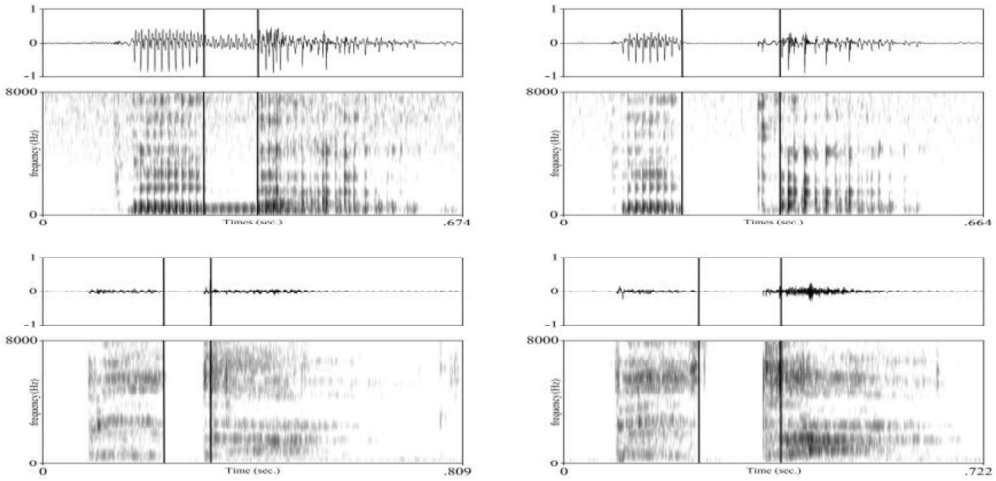

Figure 2 : Onde acoustique et spectrogramme des amorces /pe*d*al/ (à gauche) et /pe*t*al/ (à droite) en parole modale (en haut) et chuchotée (en bas). Les obstruantes /d/ et /t/ sont segmentées par les lignes verticales sombres.

3.1 Mesures acoustiques

Les mesures acoustiques effectuées sont les suivantes : (i) Durée de la consonne intervocalique (par ex. /pe*d*al/) ; (ii) Durée de la voyelle préconsonantique (par ex. /p*e*dal/) ; (iii) Intensité RMS de l'explosion ou du bruit de friction de l'obstruante intervocalique ; (iv) Centre de gravité du spectre

de bruit de l'obstruante intervocalique ; (v) Dispersion fréquentielle du spectre de bruit de l'obstruante intervocalique. Ces mesures ont été calculées et extraites sous Praat. La durée des voyelles préconsonantiques a été mesurée du début à la fin du F2, celle des obstruantes de la fin du F2 de la voyelle précédente à la fin de l'explosion ou du bruit de friction. Les mesures spectrales (iii, iv, et v) sont des indices de distribution de l'énergie dans le spectre de bruit de la consonne. Avant extraction, le spectre de bruit a été filtré passe-haut à 1 kHz afin d'exclure les basses fréquences intenses dues au voisement, ne conservant que les composantes du bruit entre 1 et 12,5 kHz. Pour réduire l'influence des voyelles, ces indices spectraux ont été calculés sur les 40 ms au centre des fricatives et sur les 10 ms initiales de l'explosion des plosives. L'intensité RMS (dB) du bruit a été calculée comme l'intensité moyenne sur cet intervalle. Le Centre de gravité (Hz) est le premier moment central (moyenne) du spectre de bruit, c'est-à-dire la fréquence moyenne pondérée en puissance du spectre. La valeur du Centre de gravité augmente avec la quantité d'énergie concentrée dans les hautes fréquences. La Dispersion fréquentielle du spectre de bruit est calculée par la racine carrée du deuxième moment central (variance) du spectre, c'est-à-dire l'écart type de l'énergie répartie dans les fréquences du spectre. Une valeur de dispersion élevée indique une distribution de l'énergie sur une large gamme de fréquences, tandis qu'une valeur faible signale une énergie concentrée autour du Centre de gravité.

3.2 Analyse statistique des corrélats acoustiques du voisement

Les 20 amorces à obstruante [+voisé] ont été comparées aux 20 amorces à obstruante [–voisé] en paire minimale par un test de Student sur les cinq paramètres acoustiques présentées ci-avant : durée de l'obstruante, durée de la voyelle préconsonantique, intensité RMS, centre de gravité et dispersion spectrale de l'explosion des plosives ou du bruit de friction des fricatives. Ces analyses portaient sur les productions en parole chuchotée et modale.

Parole modale – Les obstruantes voisées sont plus courtes et ont une intensité plus faible (89 ms, 49 dB) que le sourdes (151 ms, 57 dB) [respectivement, $t(38)=7.63$, $p<.0001$; $t(38)=2.77$, $p<.01$]. Les voyelles sont plus longues avant une obstruante voisée (154 ms) qu'avant une sourde (118 ms) [$t(38)=3.18$, $p<.01$]. Aucune différence significative ne concerne le Centre de gravité des obstruantes ([+voisé] : 3714 Hz ; [–voisé] : 3960 Hz) [$t(38)=0.59$, $p>.20$], ni la Dispersion fréquentielle ([+voisé] : 1871 Hz ; [–voisé] : 1862 Hz) [$t(38)= 0.07$, $p>.20$].

Parole chuchotée – Les obstruantes voisées sont plus courtes (98 ms) que les sourdes (158 ms) [$t(38)=7.06$, $p<.0001$]. Les voyelles sont plus longues avant obstruante voisée (172 ms) qu'avant sourde (134 ms) [$t(38)=4.20$, $p<.001$]. Aucune différence significative n'est observée pour l'intensité RMS ([+voisé] : 52 dB ; [–voisé] : 55 dB) [$t(38)=1.22$, $p>.20$], le Centre de gravité ([+voisé] : 3370 Hz ; [–voisé] : 3896 Hz) [$t(38)=1.16$, $p>.20$] ou la Dispersion fréquentielle du spectre de bruit ([+voisé] : 1416 Hz ; [–voisé] : 1672 Hz) [$t(38)=1.34$, $p=.19$].

En accord avec les études antérieures (cf. §1), pour les stimuli amorces de nos expériences de perception, les corrélats de durée consonantique et vocalique se montrent plus robustes que les corrélats spectraux pour différencier les obstruantes [+voisé] des obstruantes [–voisé] en parole modale comme chuchotée.

4 Discussion et conclusion

Par rapport à notre première étude de 2016, l'introduction d'un délai de traitement plus long (50 ms au lieu de 0 ms) entre amorce chuchotée et cible visuelle provoque des effets d'amorçage diffèrents à la fois concernant les mots à obstruante [+voisé] que ceux à obstruante [–voisé].
Concernant les mots chuchotés à obstruante [+voisé], l'étude de 2016 avait mis en évidence que le traitement de la cible sémantiquement reliée présentée immédiatement après la fin acoustique de

l'amorce (pé*d*ale–VÉLO) n'était pas facilité, contrairement à ce qui était observé en parole modale. Cela suggérait que l'auditeur n'avait pu résoudre l'ambiguïté phonétique de l'obstruante [+voisé] produite sans vibration glottique en parole chuchotée. Dans cette nouvelle étude, un effet d'amorçage significatif émerge : la cible reliée sémantiquement à l'amorce est reconnue environ 30 ms plus vite qu'en condition contrôle (quittance–VÉLO). Cela confirme notre hypothèse qu'un temps plus long de traitement est nécessaire à l'auditeur pour identifier l'obstruante [+voisé] chuchotée.

Concernant les mots à obstruante [–voisé], l'étude de 2016 avait mis en évidence que le traitement de la cible sémantiquement reliée présentée immédiatement après la fin acoustique de l'amorce (pé*t*ale–FLEUR) était facilité. Cet effet d'amorçage était d'amplitude similaire (env. 26 ms) à celui observé en parole modale (comparaisons post-hoc par paire avec une correction de Bonferroni ; amorces modales à obstruante [+voisé] : $z=3.49$, $p<.01$; à obstruante [–voisé] : $z=3.02$, $p<.05$). Cela suggérait que la propriété sous-jacente [–voisé] était immédiatement disponible pour l'auditeur et que les obstruantes [–voisé] chuchotées n'étaient pas ambiguës. Dans cette nouvelle étude, cet effet d'amorçage n'est plus observé. L'absence d'effet est comparable aux résultats obtenus en parole modale où avec un délai de 50 ms entre amorce et cible les effets d'amorçage disparaissent aussi bien pour les mots à obstruante [–voisé] [$z=2.03$, $p>.20$] que pour ceux à obstruante [+voisé] [$z=1.98$, $p>.05$]. Un tel résultat est très probablement dû au fait que l'effet d'amorçage sémantique est connu pour être éphémère (Marslen-Wilson et al. 1996, Andruski et al. 1994). Ainsi, en augmentant le délai entre amorce et cible, nous prenions le risque de voir cet effet disparaître quand les amorces ne sont pas ambiguës, à savoir pour les amorces modales et pour les amorces chuchotées à obstruantes [–voisé]. Inversement, ce délai de 50 ms entre la fin de l'amorce et le début de la présentation de la cible permet de manière cruciale aux amorces chuchotées 'ambigües', à savoir à obstruante [+voisé], d'activer pleinement leurs représentations lexicales, faisant émerger l'effet d'amorçage.

La résolution tardive de l'ambiguïté des obstruantes chuchotées [+voisé] repose probablement sur un processus d'extraction des indices phonétiques associés au trait de voisement et toujours présents dans le signal chuchoté caractérisé par l'absence de vibration glottique, propriété principale associée au trait [+voisé] en français. L'analyse de ces traces acoustiques (cf. §3), confirmant les études antérieures en parole chuchotée, laisse supposer que les auditeurs peuvent s'appuyer sur des indices résiduels de durée de l'obstruante et/ou de la voyelle qui la précède. Cette étude montre donc que les auditeurs sont pleinement capables d'exploiter les indices résiduels du trait de voisement en parole chuchotée de façon à reconnaitre le mot attendu. De façon cruciale, il est apparu que l'extraction du trait de voisement en parole chuchotée requière un certain temps, probablement nécessaire à la réparation phonologique du trait ou du phonème sous-jacent lors d'une étape pré-lexicale de traitement (Hallé et al. 1998).

Remerciements

Ce projet de recherche *Whispeech* a été soutenu par les financements ANR-16-CONV-0002 (ILCB), ANR-11-LABX-0036 (BLRI) et Initiative d'Excellence d'Aix-Marseille Université (A*MIDEX).

Références

ANDRUSKI J.E., BLUMSTEIN S.E., BURTON M. (1994). The effects of subphonetic differences on lexical access. *Cognition* 52, 163-187.

BAAYEN R.H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge: Cambridge University Press.

BAAYEN R.H., MILIN P. (2010). Analyzing reaction times. *International Journal of Psychological Research* 3(2), 12-28.

BRETZ F., HOTHORN T., WESTFALL P.H. (2011). *Multiple comparisons using R*. Boca Raton: CRC Press.
DANNENBRING G.L. (1980). Perceptual discrimination of whispered phoneme pairs. *Perceptual and Motor Skills* 51(3), 979-985.
DUFOUR S. (2008). Phonological priming in auditory word recognition: When both controlled and automatic processes are responsible for the effects. *Canadian Journal of Experimental Psychology* 62, 33-41.
FUX T. (2012). *Vers un système indiquant la distance d'un locuteur par transformation de sa voix*. Thèse de doctorat, Université de Grenoble, Grenoble.
GARNIER M., BOUHAKE S., JEANNIN C. (2014). Efforts and coordination in the production of bilabial consonants. *Proceedings of the 10th Int. Seminar on Speech Production*, 138-141. Cologne.
GILICHINSKAYA Y.D. (2012). *Perception of final consonant "voicing" in phonated and whispered speech*. UMI Dissertation Publishing, ProQuest LLC, Ann Arbor, USA.
HALLE P., ANDROJNA K., SEGUI J. (2012). L'assimilation de voisement en français : elle vaut pour les non-mots autant que les mots. Actes des XXIXè JEP, 441-448.
HALLE P., SEGUI J., FRAUENFELDER U., MEUNIER C. (1998). The processing of illegal consonant clusters: a case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance* 24, 592-608.
ITO T., TAKEDA K., ITAKURA F. (2005). Analysis and recognition of whispered speech. *Speech Communication* 45(2), 139-152.
JOVIČIĆ S.T., ŠARIĆ Z. (2008). Acoustic analysis of consonants in whispered speech. *Journal of Voice* 22*,* 263-274.
KOHLBERGER M., STRYCHARCZUK P. (2015). Voicing assimilation in whispered speech. *Proceedings of the 18th International Conference on Phonetic Sciences.* Glasgow.
KUZNETSOVA A., BRUUN BROCKHOFF P., HAUBO BOJESEN CHRISTENSEN R. (2016). Tests in linear mixed effect models. Consulté le 20/11/2017, *https://cran.r-project.org/package=lmerTest.*
MALÉCOT A., PEEBLES K. (1965). An optical device for recording glottal adduction-abduction during normal speech. *ZPSK* 18, 545-550.
MARSLEN-WILSON W.D., MOSS H.E., VAN HALEN S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1376-1392.
MEYNADIER Y. (2015). Aerodynamic tool for phonology of voicing. *Proceedings of the 18th International Conference on Phonetic Sciences*, paper#0497. Glasgow.
MEYNADIER Y., DUFOUR S. (2016). Accès lexical et reconnaissance du voisement en voix chuchotée. XXXIè JEP, 19-27. Paris.
MEYNADIER Y., GAYDINA Y. (2013). Aerodynamic and durational cues of phonological voicing in whisper. *Proceedings of the 14th Interspeech*, 335-339. Lyon.
MILLS T.L.P. (2009). *Speech motor control variables in the production of voicing contrasts and emphatic accent*. Ph Dissertation, University of Edinburg, Ednburgh.
MUNRO M.J. (1990). Perception of "Voicing" in Whispered Stops. *Phonetica* 47(3-4), 173-181.
NETSELL R. (1969). Subglottal and intraoral air pressures during the intervocalic contrast of /t/ and /d/. *Phonetica* 20(24), 68-73.
SLIS L.H. (1970). Articulatory measurements on voiced, voiceless and nasal consonants. *Phonetica* 21(4), 193-210.
TABOSSI P. (1996). Cross-modal semantic priming. *Language and Cognitive Processes* 11, 569-576.
TARTTER V.C. (1989). What's in a whisper? *The Journal of the Acoustical Society of America*, *86*, 1678-1683.
VAN DE VELDE D.J., VAN HEUVEN V. (2011). Compensatory strategies for voicing of initial and medial plosives and fricatives in whispered speech in Dutch. *Proceedings of the 17th ICPhS*, 2058–2061. Hong-Kong.
VERCHERAND G. (2010). *Production et perception de la parole chuchotée en français : analyse segmentale et prosodique*. Thèse doctorale, Université Paris VI, Paris.