

UNIVERSITE LUMIERE LYON 2

Numéro officiel :

Thèse de Doctorat en Sciences Cognitives  
option linguistique

Présentée par

**Julien Meyer**

Le 16 décembre 2005

**Description typologique et intelligibilité des langues  
sifflées, approche linguistique et bioacoustique**

Membres du jury :

Jean-Marie Hombert	Directeur de thèse
Jean-Claude Risset	Rapporteur
Denny Moore	Rapporteur
Didier Demolin	Rapporteur
John Esling	Examineur
Thierry Aubin	Examineur

Thèse préparée au sein du Laboratoire de Dynamique Du Langage,  
UMR CNRS 5596  
A l'Institut des Sciences de l'Homme, Lyon, France



*Aux siffleurs poètes et aux maîtres tambours, dont les traditions orales  
portent des clefs de réconciliation de l'être humain avec son milieu naturel*

*A Laure Dentel*

*A René-Guy Busnel*

*« ...Venez et nous suivez, qui n'avons mots à dire : nous remontons ce pur délice sans graphie où court l'antique phrase humaine ; [...] jusqu'à ces locutions inouïes, où l'aspiration recule au-delà des voyelles et la modulation du souffle se propage, au gré de telles labiales mi-sonores, en quête de pures finales vocaliques.»*

*Saint-John Perse, Neiges IV*

## Remerciements

Je remercie Jean-Marie Hombert et François Pellegrino de m'avoir accueilli au Laboratoire DDL et de m'avoir fait confiance en dirigeant mes recherches.

Messieurs les membres du jury, vous me faites l'honneur d'évaluer mon travail,

Je remercie Jean-Claude Risset d'avoir accepté de se replonger dans les langues sifflées, son travail en acoustique musicale m'a permis de mieux comprendre mon sujet de recherche.

Je remercie Denny Moore qui vient du Brésil et m'a fait comprendre que la linguistique est une enquête.

Je remercie Didier Demolin de venir du Brésil pour partager son intérêt pour la phonétique des sifflements.

Je remercie John Esling pour l'intérêt qu'il manifeste pour mon travail de thèse en venant du Canada.

Je remercie Thierry Aubin de venir apporter son regard de bioacousticien sur les phénomènes étudiés.

Tout mon respect et ma reconnaissance vont aux piliers scientifiques de ce travail qu'ont été pour moi René-Guy Busnel et Colette Grinevald.

Merci à tous les membres du Laboratoire DDL.

Je remercie Bernard Gautheron, Jacques Vielliard, Dora Pellicer, Michel André, Marie-Claire Busnel, Loïc Lemarrec, Torbson Johansson, Chô Ly pour leurs conseils.

Je remercie le CNRS pour m'avoir soutenu avec la Bourse B.D.I., Fanny Meunier pour son aide lors de l'achat d'une partie du matériel d'enregistrement et pour toutes les discussions que nous avons eues. Je remercie ABM et la Mairie de Paris pour nous avoir décerné le prix scientifique « *Paris Jeunes Aventure* » et mon frère Alexandre d'avoir défendu mon dossier à l'oral à ma place. Ma sœur Laure de l'avoir aidé à se préparer.

Ce travail, n'est que la partie émergente de ma collaboration avec des dizaines de personnes dans le monde entier. Je remercie, Luis el Maestro de Silbo pour son enseignement et les visites de l'île, Lino Rodriguez, Eugenio Darias, Juan Casimiro, la famille Betanzos, Sacha, la famille Fernando Peres Alonso, Sergio, Miguel, Alessandro Martinez, Maître Agbeli, Aju Jupoh, Noriko Higashide, Chief John Leo Tamata, Chief Rubens, Lorenzo Cappelli, prof. Jurg Gashé, David Dick, Mrs. Kula, Mr. Panayotis et sa femme, Maria Panayotis, Fabio et Ana, Thierry Fournel, Jean-Michel Beaudet, Patrick Kersalé, Madame Gisèle, Elise Thierry, Doña Sissi, Doña Julia, Janet Tchillinguirian, Mariana De Lacerda, Gessiane Lobato Picanço, Ross Caughley, Tschewong, Remi Saudax, Sylvano et Jailda, la Famille Santillan, Ralph Regenvanu, Carmen Figueredo. Je remercie tous les sujets des expériences.

Je remercie mes parents, leur vie m'a conduit à m'intéresser aux langues.

# Table des matières

<b>Introduction .....</b>	<b>1</b>
<b>Chapitre 1. Cadre de l'Étude des langues sifflées .....</b>	<b>5</b>
1.1. Historique de l'étude des langues sifflées.....	6
1.1.1. Préambule.....	6
1.1.2. La vague d'intérêt des années 50-70.....	6
1.1.3. Etudes plus récentes .....	7
1.1.4. Interprétation descriptive et théorique .....	8
1.1.5. Le regain d'intérêt local des dernières années .....	9
1.1.6. Positionnement de notre étude .....	11
1.2. Langues sifflées : langues en danger ?.....	14
1.2.1. Causes .....	14
1.2.2. Facteurs de maintien .....	15
1.2.3. Aperçu de la vitalité autour du monde .....	16
1.3. Problématique de la définition des langues sifflées .....	21
1.3.1. Le point de vue des chercheurs .....	21
1.3.2. Le point de vue des siffleurs .....	22
1.3.3. Positionnement des langues sifflées.....	22
1.4. Conclusion : réponse à la question « Qu'est ce qu'une langue sifflée ? ».....	24
1.4.2. Quelques quiproquos récurrents.....	25
<b>Chapitre 2. Influence du milieu naturel.....</b>	<b>28</b>
2.1. Introduction.....	29
2.2. Description des environnements naturels .....	31
2.2.1. Milieux écologiques et usages .....	31
2.2.2. Conclusion .....	39
2.3. Étude bioacoustique .....	40
2.3.1. Problématique .....	40
2.3.2. Production du sifflement.....	40
2.3.3. Bandes de fréquences typiques des sifflements humains.....	46
2.3.4. Analyse du spectre de fréquences des sifflements en milieu ambiant .....	49
2.3.5. Propagation du signal en milieu acoustique naturel et portée géographique.....	52
2.3.6. Distances limites d'intelligibilité de la voix parlée, de la voix criée et de la parole sifflée .....	60
2.3.7. Dynamique de la phrase sifflée.....	68
2.3.8. Dynamique temporelle: allongement.....	74
2.4. Conclusion .....	79
<b>Chapitre 3. Analyse typologique.....</b>	<b>80</b>
3.1. Intérêt de la typologie pour les études comparatives .....	81
3.2. Bilan des langues sifflées connues.....	82

3.2.1. Langues analysées, partiellement étudiées.....	82
3.2.2. Autres langues sifflées signalées.....	85
3.2.3. Conclusion .....	87
3.3. Description détaillée de quelques langues .....	88
3.3.1. Organisation générale .....	88
3.3.2. Organisation de la description de chaque langue.....	88
3.3.3. Langues sifflées s'appuyant sur l'articulation des voyelles et des consonnes.....	90
3.3.4. Langues sifflées utilisant une stratégie de transposition intermédiaire.....	129
3.3.5. Langues dont la transposition sifflée cible la hauteur fondamentale de la voix .....	143
3.4. Conclusion de la typologie comparative.....	156
<b>Chapitre 4. Intelligibilité.....</b>	<b>158</b>
4.1. Introduction.....	159
4.2. Adaptation étonnante du sifflement à l'audition humaine.....	163
4.2.1. Système périphérique.....	163
4.2.2. Psychoacoustique: une première approche du système central .....	167
4.2.3. Structuration perceptive du flux de parole.....	182
4.3. Intelligibilité de la parole sifflée: analyse progressive.....	191
4.3.1. Expérience de perception des voyelles sifflées par des sujets ignorant tout des langues sifflées .....	191
4.3.2. Analyse de la perception de non-mots et de logatomes par des siffleurs .....	209
4.3.3. Intelligibilité des mots sifflés.....	212
4.3.4. Intelligibilité des phrases sifflées (courte et longue distance) .....	217
4.4. Conclusion pour l'intelligibilité.....	232
<b>Conclusion : Bilan et perspectives .....</b>	<b>234</b>
<b>Bibliographie.....</b>	<b>238</b>
<b>ANNEXES .....</b>	<b>Erreur ! Signet non défini.</b>
A. Méthodologie d'étude des langues sifflées.....	<b>Erreur ! Signet non défini.</b>
A.1. Introduction.....	<b>Erreur ! Signet non défini.</b>
A.2. Enquête préliminaire .....	<b>Erreur ! Signet non défini.</b>
A.3. Travail local .....	<b>Erreur ! Signet non défini.</b>
A.4. Méthodologie d'analyse en laboratoire.....	<b>Erreur ! Signet non défini.</b>
A.5. Méthodologie d'approche phonétique .....	<b>Erreur ! Signet non défini.</b>
B. positionnement des langues sifflées .....	<b>Erreur ! Signet non défini.</b>
B.1. Musique et langage.....	<b>Erreur ! Signet non défini.</b>
B.2. Explication du tableau récapitulatif du § 1.3.3.2.....	<b>Erreur ! Signet non défini.</b>
C. Typologie des siffleurs .....	<b>Erreur ! Signet non défini.</b>
C.1. Variables qui permettent de distinguer les siffleurs .....	<b>Erreur ! Signet non défini.</b>
D. Données complémentaires pour la partie Typologie.....	<b>Erreur ! Signet non défini.</b>
D.1. Tableaux de statistiques .....	<b>Erreur ! Signet non défini.</b>
D.2. Descriptions d'autres langues sifflées.....	<b>Erreur ! Signet non défini.</b>
D.3. Systèmes sifflés partageant leur stratégie avec les techniques de langages tambourinés .....	<b>Erreur ! Signet non défini.</b>
E. Résultats de l'expérience de perception des voyelles sifflées .....	<b>Erreur ! Signet non défini.</b>
E.1. Résultats par sujets .....	<b>Erreur ! Signet non défini.</b>

E.2. Courbes d'analyses complémentaires.....	<b>Erreur ! Signet non défini.</b>
F. Deux expériences sur l'intelligibilité.....	<b>Erreur ! Signet non défini.</b>
F.1. Analyse pilote en pays Mazatèque.....	<b>Erreur ! Signet non défini.</b>
F.2. Les expériences révélatrices des Warren .....	<b>Erreur ! Signet non défini.</b>
G. Matériel de diffusion sonore pour les 3 expérimentations à distance.....	<b>Erreur ! Signet non défini.</b>

# INTRODUCTION

Pour communiquer à distance ou dans le bruit, de nombreuses populations vivant principalement dans des zones montagneuses et dans des forêts denses ont développé une version sifflée de leur langue locale. Les biotopes dont elles font partie favorisent l'éloignement en terme de temps de trajet ou de contact visuel. Le sifflement apporte une réponse adaptée à de telles situations en augmentant les distances limites de communication. Les scientifiques ont appelé ce phénomène « *parole* », « *langage* » ou « *langue sifflée* » (Busnel & Siegfried, 1990) car l'information que les siffleurs échangent dans ces conditions est suffisamment riche, complexe et variée pour nécessiter l'usage du vocabulaire, de la grammaire et de la syntaxe de la langue parlée locale. Le sifflement est utilisé comme source sonore à la place des vibrations des cordes vocales car il est bien adapté à la propagation des sons dans des conditions extrêmes. D'après les locuteurs, la parole sifflée est une des multiples manifestations complémentaires de la langue, au même titre que le chuchotement, la parole criée ou chantée. Elle permet un dialogue entre individus relativement isolés, comme un téléphone portable naturel et local. Elle est utilisée pour rendre service dans le cadre d'activités sociales vitales de la vie traditionnelle.

Aujourd'hui, certaines zones encore difficiles d'accès abritent des villages poursuivant la pratique quotidienne de ces langages ou à défaut, dont les plus anciennes générations entretiennent ce savoir oral dans leur mémoire. Notre travail s'appuie sur une enquête de terrain de plus d'un an menée en collaboration avec les responsables culturels d'une dizaine de communautés à travers le monde. Il tire aussi des informations de collaborations engagées avec les principaux acousticiens, linguistes, anthropologues ou musicologues étudiant ou ayant étudié la langue et le mode de vie de villages dont la population utilise ou a utilisé le sifflement langagier. Cette recherche a nécessité la mise en place d'une méthodologie d'investigation adaptée à la fois à l'acoustique des langages sifflés, à leur pratique en voie de disparition et à la volonté de nombreux locuteurs de participer à la sauvegarde vivante de cet aspect de leur tradition orale. En effet, à l'instar de la plupart des langues minoritaires, ces formes de communication sont aujourd'hui toutes menacées d'extinction ainsi que le patrimoine oral qu'elles véhiculent.

Dans chaque pays visité, nous nous sommes rendus dans des lieux où une forme sifflée d'une langue avait déjà été signalée et nous avons cherché de nouvelles sources linguistiques. Nous avons constitué plusieurs corpus plus ou moins spontanés en fonction des opportunités, par exemple en nous promenant avec une personne qui déclenchait des dialogues avec les individus croisés sur le chemin. Mais avant ce type de travail, le corpus que nous demandions de siffler dans un premier temps aux informateurs rencontrés était le suivant:

« Dans le monde entier, des hommes parlent en sifflant »

« Certains vivent dans les montagnes, d'autres dans les forêts »

« Les sifflements s'utilisent pour parler à distance »

« C'est un moyen de communication très riche en harmonie avec la nature et le mode de vie »

« On peut parler d'amour et de toutes choses de la vie »

« La mélodie ressemble aux chants des oiseaux »

« Ces langues ne doivent pas mourir car elles racontent l'histoire des peuples de ces terres »

« Ainsi que celle du langage humain »

Aucun des bons siffleurs n'a eu de mal à transposer en sifflements les paroles de ce texte, une fois les difficultés de la traduction résolues. Ceci montre qu'ils utilisent aisément des phrases non stéréotypées. Un interlocuteur placé à distance et aguerri en langue sifflée pouvait comprendre leur signification avec plus ou moins de facilité suivant sa pratique quotidienne et suivant la complexité sémantique ou métaphorique de la phrase. Les locuteurs Espagnols ou Grecs ayant une langue dont la structure phonologique et les réalisations phonétique se rapprochent le plus du français parmi toutes les langues analysées, n'ont pas eu de difficulté à siffler des phrases en langue française quand nous leur avons suggéré. Nous avons donc constaté que le sifflement de syllabes n'ayant pas de sens préalablement connu est possible, même si cela paraît incongru. Le français sifflé obtenu est marqué par « l'accent » local car chaque langue possède sa propre répartition d'unités sonores régulières.

Par contre, lorsque des locuteurs de langue tonale mazatèque du Mexique ont essayé de reproduire ce texte en espagnol sifflé, nous avons remarqué que le résultat n'avait rien de commun avec le *Silbo* espagnol. Nous avons ainsi vérifié que les siffleurs optent pour des stratégies différentes de conversion de la voix classique en parole sifflée en fonction de certaines différences phonologiques clefs de leur langue. En effet, les « *Maestros de Silbo* »<sup>1</sup> Espagnols de l'île de la Gomera aux Canaries reproduisent les voyelles et les consonnes à travers ce que nous appellerons leur « *prosodie interne* », qui est organisée rythmiquement dans la phrase. Les Mazatèques, de leur côté, ont sifflé l'intonation et le rythme de l'espagnol ce qui d'après des études sur la prosodie de la parole ne contribue qu'à hauteur de 10 à 15% de l'intelligibilité des langues sans ton de ce type (Sorin, 1989). Pourtant, dans leur langue, les Mazatèques atteignent plus de 80 à 90 % d'intelligibilité en transposant en sifflements la même partie de la voix. Ceci montre que, suivant la structure de la langue, les différentes bandes de fréquences de la voix classique ne portent pas la même contribution fonctionnelle pour l'intelligibilité de la parole.

Ces expériences interlinguistiques n'étaient pas complètement artificielles car un processus de transfert a effectivement eu lieu sur l'île de la Gomera aux Canaries depuis une langue berbère après la conquête Espagnole. De plus, un tel processus semble être actuellement encore en cours dans les montagnes Mazatèques. Il a commencé par les prénoms hispanisés il y a plus de cinquante ans (Cowan, 1948) et nous avons pu observer qu'il était encore hésitant aujourd'hui parmi les plus jeunes générations. Une des raisons principales tient à l'existence de tons phonologiquement distinctifs en mazatèque qui n'existent pas en espagnol. Dès lors, dans le cas d'un transfert de la pratique sifflée d'une langue à une autre, la rapidité

---

<sup>1</sup> Maîtres de Silbo

dépend de la structure phonologique de chacun des systèmes linguistiques qui se rencontrent. D'une manière générale, le développement d'une nouvelle stratégie de sifflement se fait suivant un processus évolutif lent qui sélectionne petit à petit les constituants phonétiques les plus saillants de la langue en termes d'intelligibilité. Ce phénomène repose sur plusieurs générations de pratique.

Lors de notre travail de thèse dont nous présentons ici un bilan, nous avons cherché à mettre évidence les caractères généraux partagés par les langues sifflées et à souligner les différences phonétiques dues aux propriétés phonologiques de chaque langue. Pour cela nous avons emprunté les outils scientifiques de différents domaines de recherche. Les particularités acoustiques des langues sifflées s'y prêtent bien. La présentation de cette recherche s'articule autour de quatre chapitres que nous décrivons dans les paragraphes suivants.

Dans un premier temps nous avons posé le cadre de notre étude. Nous avons d'abord fait l'historique de la recherche dans le domaine. A cette occasion, nous avons rappelé les principales connaissances sur ce sujet. En particulier la position de ce phénomène de communication par rapport à d'autres aspects plus connus de la communication humaine. Sur ce point, les trois principaux chercheurs ayant effectué des travaux dans le passé sur les langues sifflées (Cowan, 1948 ; Busnel & Classe, 1976) sont unanimes pour reconnaître que les langues sifflées représentent un style de parole complémentaire utilisé à distance (courte, moyenne ou grande). Nous avons ensuite mené une analyse sociolinguistique sur les facteurs influençant la vitalité des langues sifflées. Cet aspect de la réflexion s'est révélé crucial au quotidien lors de la phase de documentation sur le terrain. En particulier car il existe dans chaque lieu des profils types de siffleurs ayant des compétences différentes. De plus, l'histoire humaine entraîne aujourd'hui le phénomène de la disparition des cultures locales qui a fortement joué sur nos possibilités de documentation.

Dans un deuxième temps nous présentons les biotopes rencontrés, dans leur diversité. Nous verrons que la pratique sifflée est en partie conditionnée par les milieux écologiques. Nous introduisons les forêts denses qui viennent s'ajouter aux topographies montagneuses bien décrites dans les travaux des chercheurs précédents (Classe, 1956 ; Busnel, 1970 ; Busnel, 1974a). Puis nous présentons les techniques de production du sifflement ce qui permet d'initier une réflexion sur la bioacoustique du signal. D'autre part, les propriétés des langues sifflées qui définissent de véritables « *systèmes de télécommunication* » (Busnel et Classe, 1976, p.108) nous ont poussés à étudier quelques effets du filtre naturel du milieu ambiant sur le signal sifflé. Nous avons aussi cherché, grâce à la mise au point d'expériences pilotes, les liens qui peuvent exister entre les stratégies bioacoustiques utilisées par la voix parlée, la voix criée et la parole sifflée.

Dans le troisième chapitre que nous dédions à l'analyse typologique nous avons fait le bilan linguistique de notre enquête. Cette étape de notre recherche nous a permis de recenser les langues sifflées connues jusqu'à aujourd'hui et d'aller vérifier directement leur pratique actuelle pour certaines d'entre elles. Au final, la diversité des langues sifflées connues est grande et quelques nouvelles langues sifflées ont pu être repérées. Nous confirmons donc que le phénomène est international. Afin d'approfondir notre compréhension des modalités de transposition de la voix parlée en parole sifflée, nous avons sélectionné sept langues dont nous avons réalisé une analyse typologique détaillée. La comparaison systématique de la forme sifflée et de la

forme parlée nous a permis de mettre en évidence trois types de stratégies de transposition de la voix. Nous ne remettons pas en cause la différence entre langues tonales (transposition sifflée des tons) et langues non tonales (transposition sifflée de l'articulation des voyelles et des consonnes) mais nous montrons qu'il existe un groupe intermédiaire de langues qui choisissent une voie médiane entre ces deux options décrites précédemment. De plus, nous montrons que le sifflement ne définit pas de système indépendant de la forme parlée même s'il existe une réduction phonétique que nous analysons. Enfin, nous constatons qu'il y a une grande diversité de stratégies dans chaque groupe. Il s'ensuit que, des langues non tonales aux langues tonales, les différences se situent sur une échelle faite de degrés qui ont pour origine des propriétés de la structure phonologique de la forme parlée.

L'intelligibilité étonnante obtenue lorsque des phrases sont échangées en langue sifflée est l'objet de notre étude dans la dernière partie de cette thèse. Cette faculté s'appuie sur le rôle fonctionnel que jouent les caractères segmentaux et suprasegmentaux de chaque système linguistique. Elle repose aussi sur tout un ensemble de facteurs perceptifs qui ont fait l'objet d'analyses précises. C'est pourquoi nous présentons les résultats de psychoacoustique qui permettent de comprendre une partie du cadre de la perception des sifflements. Ensuite, nous regroupons les enseignements des domaines de « *l'analyse de la scène auditive* » (Bregman 1990) et de l'analyse de la prosodie du langage. Nous précisons ainsi quels types de relations entre les attributs de la perception marquent le flux de parole écouté par un siffleur. A partir de là, nous avons été en mesure de présenter une analyse progressive de l'intelligibilité des langues sifflées non tonales. Une expérience de psycholinguistique est alors présentée. Elle a consisté à tester l'identification des voyelles sifflées espagnoles par des sujets français. Cette étape de la réflexion nous a permis de revisiter la perception du timbre des voyelles humaines en suivant les indications implicites des siffleurs. Nous avons ensuite expliqué deux analyses de l'intelligibilité des logatomes et des mots (Moles, 1970 ; Busnel, 1970). Ces deux approches sont complétées par une étude de l'intelligibilité des phrases sifflées grecques. Enfin, comme ce type de communication sert avant tout à distance et dans le bruit ambiant, nous avons mis au point une expérience grandeur nature qui nous a permis de rediffuser certains de nos enregistrements turcs, espagnols et grecs. De cette manière nous avons pu observer, de manière contrôlée, les conséquences phonétiques de la dégradation du signal sifflé. Les langues sifflées offrent de ce point de vue un cas d'étude inédit du phénomène de *reconstruction cognitive de la parole*. En effet, les indices acoustiques qui arrivent à l'oreille du récepteur sont souvent suffisants pour permettre un taux d'intelligibilité élevé.

La somme des informations apportées par les langues sifflées est vaste, c'est pourquoi nous considérons le travail que nous présentons ici comme un début. Nous l'avons appuyé sur les bases scientifiques solides posées par les chercheurs ayant étudié les langues sifflées et par les linguistes ayant décrit les formes parlées des langues que nous analysons. La contribution des locuteurs, des siffleurs et des responsables culturels locaux a été un facteur essentiel de notre recherche car eux seuls vivent en profondeur la richesse de leur langue. Nous ne pouvons concevoir de travail de recherche sans eux. Leur collaboration est primordiale si l'on veut « *épouser du monde leur présence* » comme dit si bien Saint-John Perse.

# **CHAPITRE 1. CADRE DE L'ETUDE DES LANGUES SIFFLEES**

## 1.1. Historique de l'étude des langues sifflées

### 1.1.1. Préambule

S'il existe quelques textes très anciens qui relatent la présence de traditions sifflées utilisées pour communiquer à distance comme *le traité du Xiao*, écrit par Sun Gang en 765 (Picard, 1991), le premier témoignage historique attestant de l'existence d'une *langue sifflée* est celui de l'équipage du mercenaire Français Jean de Béthancourt parti à la conquête des îles Canaries en 1402. Dans un ouvrage intitulé *Le Canarien*, Bontier et Le Verrier (1602), évoquent le « *plus étrange langaige de tous les autres Pais de par deçà* »; pour eux les habitants de l'île « *parlent de beaulièvres<sup>2</sup> ainsi que fussent sans langue* » (cité par Busnel et Classe, 1976, p.6). Cet événement préfigure les premiers travaux scientifiques traitant essentiellement des aspects anthropologiques du phénomène, souvent suivant un style colonial (Quedenfeldt, 1887; Lajard, 1891; Labouret, 1923, Stern, 1955).

C'est seulement au début de la deuxième moitié du XX<sup>ème</sup> siècle que des descriptions précises ont été effectuées dans le domaine de la linguistique, d'abord pour la langue mazatèque (Cowan, 1948) puis pour le *Silbo*<sup>3</sup> espagnol (Classe, 1956, 1957); et dans le domaine de l'acoustique pour la langue béarnaise (Busnel et al., 1962a) puis pour la langue turque (Busnel et al., 1970) ou la langue mazatèque (Busnel, 1974a).

### 1.1.2. La vague d'intérêt des années 50-70

Quelques rares chercheurs firent du sujet des langues sifflées un des moteurs de leur réflexion sur les langues humaines et sur les communications en général. Les trois principaux furent Classe, Cowan et Busnel. Sous leur impulsion, l'intérêt pour les langues sifflées connut son apogée dans les années 70, en particulier lors de la publication des résultats de deux recherches qui font aujourd'hui office de référence : l'une issue du travail d'une expédition pluridisciplinaire organisée auprès des siffleurs Turcs de la région de Kusköy (Busnel et al., 1970), l'autre étant un bilan linguistique et acoustique des connaissances dans le domaine (Busnel et Classe, 1976).

Cowan permit de faire connaître plusieurs langues sifflées et initia la description de deux d'entre elles, dont la première version tonale connue, la langue sifflée mazatèque. Il bénéficia dans ces recherches du réseau d'information des linguistes missionnaires de l'organisation évangélique dont il est un membre : le

---

<sup>2</sup> De beaulièvres : avec deux lèvres, de bel en ancien français signifiant deux du latin bis. Ce témoignage semble avoir inspiré un livre de Cyrano de Bergerac ou un langage sifflé est explicitement expliqué pour la deuxième fois dans la littérature.

<sup>3</sup> L'appellation Silbo est le nom donné par les habitants de l'île de la Gomera à la langue sifflée espagnole pratiquée aux Canaries. Nous utilisons fréquemment ce terme par la suite pour désigner cette langue sifflée qui est la plus connue des scientifiques.

« Summer Institute of Linguistics »<sup>4</sup> qui est encore aujourd'hui très présente dans de nombreuses populations autochtones parlant des langues minoritaires.

De son côté, Classe eut une démarche unique puisqu'il devint lui-même un siffleur de Silbo et put ainsi décrire cette langue de « l'intérieur » de manière inégalée.

Busnel trouva une langue sifflée en France, ce qui paraissait impensable à l'époque. D'autre part, il aborda le problème de manière très différente de ses prédécesseurs. Inspiré par l'école allemande d'éthologie humaine (Lorentz et Eibl-Eibesfeldt), il s'intéressa aux comportements humains dans leur milieu écologique et utilisa pour les décrire les instruments technologiques les plus modernes. Ainsi un des premiers sonographes d'Europe fut importé pour étudier la langue sifflée des Pyrénées ; d'autre part, des analyses aux rayons X du *tractus* vocal des siffleurs furent effectuées et des documentaires audiovisuels détaillés furent édités à des fins didactiques. Les équipes pluridisciplinaires qu'il constitua, en intéressant à ce sujet des spécialistes de psychologie, d'acoustique, de traitement du signal, de biologie, d'éthologie et de linguistique permirent d'approfondir considérablement la compréhension des langages sifflés béarnais, turc et espagnol.

### 1.1.3. Etudes plus récentes

A partir des années 70, parallèlement au travail de ces trois auteurs, quelques scientifiques produisirent de manière ponctuelle des analyses de langues sifflées inconnues. Ainsi, de nouvelles langues furent partiellement analysées au Mexique comme le *kickapoo* (Voorhis, 1971) ou le *tepehua* (Cowan, 1976). Au Népal, le *chepang* fit l'objet d'une description précise (Caughley, 1976) : un nouveau type d'analyse, basé sur le poids des syllabes fut développé en raison de la particularité de la structure de la langue concernée. En 1976 un double volume, regroupant la plupart des travaux sur les langues sifflées et les langages tambourinés, fut édité avec certains documents inédits (Sebeok et Umiker-Sebeok, 1976). Plus récemment, le grec sifflé (Charalambakis, 1994 ; Xirometis et Spyridis 1994), le *wam* en Papouasie Nouvelle Guinée (Nekitel, 1992) et le *hmong* sifflé avec une feuille par des populations ayant quitté le Vietnam (Busnel et al, 1989) furent analysés partiellement. Enfin, de manière très succincte, le *benčnon* d'Ethiopie (Wedekind, 1981). Dans le domaine de l'audiovisuel l'ensemble des documents scientifiques furent publiés par le Service du Film de Recherche Scientifique (SFRS) en France à partir de films réalisés sur la langue béarnaise (Busnel, 1964), sur la langue turque (Busnel, 1967) ou sur le phénomène des langues sifflées dans son ensemble (Busnel & Siegfried, 1990).

---

<sup>4</sup> La SIL est une organisation qui utilise la linguistique pour traduire une version de Bible évangélique américaine dans des langues minoritaires. Son action conduit à des travaux de linguistique pour concevoir l'orthographe écrite de langues en majorité orales. Elle s'accompagne la plupart du temps d'une dénégation des valeurs culturelles traditionnelles qui lient la langue locale à ses racines, en particulier de la cosmogonie locale et toute la tradition orale qui la véhicule.

## **1.1.4. Interprétation descriptive et théorique**

### **1.1.4.1. Recherches complémentaires**

En plus des recherches sur des langues inconnues signalées dans les paragraphes précédents, une dizaine d'analyses complémentaires furent produites, la plupart sur le Silbo des îles Canaries en raison de sa renommée et de son accessibilité. Elles consistent en de nouvelles interprétations théoriques du phénomène des langues sifflées cherchant à définir le lien entre la parole sifflée et la voix parlée et parfois, paradoxalement, à définir un système plus ou moins indépendant.

La plus radicale fut celle de Trujillo (1978) qui expliqua que l'inventaire complet du système linguistique espagnol était réduit en Silbo à deux voyelles et quatre consonnes contrastives. Brusis (1973) remarqua qu'une ressemblance existait entre les modulations du deuxième formant de la voix et les modulations sifflées des consonnes du Silbo. Leroy (1970) de l'équipe pluridisciplinaire qui décrit la langue sifflée turque systématisa la description initiée par Classe (1956) sur les lieux d'articulation des consonnes et fit le lien avec les transitions typiques des formants de la voix. Rialland (2003) développa ce point de vue sur le Silbo pour le deuxième formant. Enfin une synthèse comparative de documents mérite d'être signalée, il s'agit du travail de comparaison philologique du DEA de Thierry (2002).

### **1.1.4.2. Description linguistique la plus connue**

L'ensemble des analyses a participé à la description du phénomène des langues sifflées tel qu'il est connu aujourd'hui par exemple dans l'encyclopédie Elsevier du langage et de la linguistique (Meyer et Gautheron, 2005) et qui pourrait être résumé ainsi :

Les langues sifflées sont pratiquées dans les montagnes pour réduire l'effort de communication. Elles s'appuient sur un système local parlé dont elles transposent certains éléments en sifflements.

Les linguistes et les acousticiens ont remarqué que la stratégie d'encodage des langues sifflées était corrélée avec certaines caractéristiques de la langue locale. Deux principales manières de siffler un langage ont été décrites:

- (1) Dans le cas des langues tonales comme le mazatèque du Mexique, la hauteur perçue de la fréquence principale du sifflement caractérise le phonème: dès lors, les sifflements sont concentrés sur des traits supra-segmentaux et reproduisent principalement la fréquence fondamentale de la langue parlée
- (2) Parallèlement, dans les langues sifflées issues des langues non tonales, l'étroite bande de fréquence des sifflements reproduit principalement les traits segmentaux du langage parlé: les voyelles sont émises à différentes hauteurs fréquentielles (la série [i e a o u] est ainsi descendante en Silbo) alors que les consonnes se prononcent en modulant les hauteurs vocaliques. Les transitions des consonnes sifflées sont par conséquent influencées par la hauteur de la voyelle voisine. D'autre part, quand elles sont représentées sur un spectrogramme, ces transitions consonantiques ont une ressemblance avec la modulation du second formant de la langue parlée. Les lieux d'articulation sont alors reportés à des niveaux de fréquences adaptés au sifflement.

De plus, dans de nombreux documents de recherche les langues sifflées sont qualifiées de substitut du langage, « *language surrogates* » en anglais. Ce terme, largement accepté dans la communauté scientifique pose problème car il réduit les langues sifflées à des systèmes fonctionnant en partie sur l'utilisation de phrases stéréotypées, à la manière de certaines pratiques de langages tambourinés, ce qui n'est pourtant pas le cas de 90% des systèmes décrits jusqu'à aujourd'hui. Cowan, Classe et Busnel ont tous les trois récusé cette définition réductrice introduite par Stern dans un article faisant une typologie des langages tambourinés et sifflés connus à son époque et elle fut popularisée par Sebeok et Umiker-Sebeok (1976). Comme nous le verrons dans les paragraphes suivants, pour eux les appellations « *style de parole* » ou « *langage complémentaire* » utilisé pour les communications à distance sont plus adaptées.

### **1.1.5. Le regain d'intérêt local des dernières années**

Ces dernières années ont été marquées par la reprise en main des langues sifflées par certaines communautés pour lesquelles elle est devenue un symbole culturel et où elle fait l'objet d'une fierté locale affirmée. Ce phénomène a des conséquences directes et indirectes sur les possibilités de recherche dans la mesure où un débat a lieu au niveau culturel et politique sur le sujet.

#### **1.1.5.1. L'île de la Gomera**

L'exemple le plus flagrant dans ce domaine est l'île de la Gomera où la pratique du Silbo a aujourd'hui sauté une génération puisque les anciens de l'île, confrontés au désintérêt pour le sifflement de la majorité de la génération de leurs enfants, ont réussi à imposer au Gouvernement des Canaries l'enseignement obligatoire du Silbo en classe de primaire en 1999. Le débat local a été long et a conduit certains siffleurs à enseigner bénévolement pendant plusieurs années. Mais le pouvoir local a finalement compris l'intérêt du Silbo, et, tout en officialisant l'enseignement à l'école fit une campagne de promotion qui donna lieu à l'organisation d'un premier congrès des langues sifflées (en avril 2003<sup>5</sup>), au financement de la première étude de neurosciences réalisée sur une langue sifflée par une équipe de l'Université de La Laguna sur l'île voisine de Tenerife (Carreiras et al, 2005). L'ensemble de ces démarches permit à la Gomera d'avoir une visibilité médiatique importante auprès de l'Union Européenne et de l'Unesco tout en engageant un processus culturel de revitalisation par l'éducation scolaire, ce qui reste encore unique dans le monde des populations siffleuses.

#### **1.1.5.2. Le Village de Kusköy en Turquie**

Le mode d'expression de l'intérêt actuel de la population turque pour le sifflement articulé s'est cristallisé autour de la tradition locale qui consiste à organiser un festival de début d'été dans chaque village de la région pour le départ des bergers vers les hauts plateaux proches de Trabzon et qui surplombent la mer Noire. Dans le village de Kusköy, reconnu comme celui où les siffleurs sont les plus habiles, le festival local est devenu ces dernières années le festival de « *Kusdili* », littéralement «de la langue des oiseaux ». D'autre

---

<sup>5</sup> Auquel nous avons pu participer à l'invitation des siffleurs puisque nous avons appris sa tenue lors de notre travail sur place en mars 2003.

part, en dehors de cette période, le responsable du festival est l'interlocuteur privilégié qui contrôle la venue des journalistes et des très rares scientifiques intéressés par le sujet<sup>6</sup>. Cette organisation politico-culturelle, qui témoigne de la place symbolique forte tenue par le sifflement à Kusköy, n'existe pas dans les villages voisins ni chez les bergers des hauts plateaux. En juin 2005, le Festival turc a eu l'honneur de révéler l'existence d'une nouvelle langue sifflée à la communauté scientifique internationale : la langue *Yupik* d'Alaska, puisque des membres de cette population étaient présents en tant qu'invités. De plus, lors de ce festival un concours est organisé auprès des élèves de l'école primaire : c'est un évènement qui préfigure peut être un enseignement scolaire local de la langue sifflée.

### 1.1.5.3. La Sierra Mazatèque

Lors de notre visite dans la Sierra Mazatèque, nous avons constaté une grande disparité d'intérêt entre les vallées pour le phénomène de la langue sifflée. Dans la ville de Huautla de Jimenez, certains acteurs de la vie intellectuelle Mazatèque ont essayé de développer des projets locaux sur ce thème mais n'ont pas obtenu de financement ou se sont heurtés aux réticences des gouvernements locaux d'autres villages.

Dans le village d'Eloxochitlan de Flores Magon, très attaché pour des raisons historiques à l'ensemble des aspects de la vie traditionnelle<sup>7</sup>, le conseil indigène local contrôle l'ensemble des actions faites sur le plan politique et culturel. Celui ci invite tous les dimanches l'ensemble des habitants qui le souhaitent à venir s'exprimer sur les sujets à l'ordre du jour, puis une décision doit être prise à l'unanimité. Ceci signifie que toute initiative dans le domaine culturel est âprement discutée par ce conseil et lorsque la décision est importante, une assemblée invite tous les villages du district à participer à la rédaction d'une déclaration d'action commune. En général, le conseil décide que l'action doit être réalisée par un groupe d'habitants d'un village ou du district avec, éventuellement, l'aide d'acteurs extérieurs comme un chercheur dans notre cas. Depuis 1994, le Mexique reconnaît officiellement la possibilité aux villages de s'organiser suivant leurs « us et coutumes »<sup>8</sup>. Notre participation à deux conseils indigène et une assemblée indigène trois dimanches de suite témoigne de l'intérêt local pour le sifflement et de l'hésitation de la communauté à faire participer un étranger à leurs activités. Cependant, en fin de compte, cette organisation permis un débat fertile et inscrivit l'usage de la *lengua silbada* parmi les priorités culturelles de la communauté qui garde le contrôle des initiatives à mener dans le domaine. En raison de la connaissance précise du terrain et de la participation au conseil des meilleurs siffleurs qui vivent dans des zones très reculées de la Sierra, on peut s'attendre à de

---

<sup>6</sup> Après l'expédition organisée par Busnel dans la fin des années 60, Laure Dentel et moi étions en 2004 le deuxième groupe de scientifiques à nous présenter sur place.

<sup>7</sup> Le village d'Eloxochitlan de Flores Magon, lieu de naissance du révolutionnaire Mexicain Flores Magon, suit l'organisation communautaire préconisée par ce fils du village : c'est une adaptation de l'organisation traditionnelle mazatèque à un Etat de type démocratique.

<sup>8</sup> Dans l'état d'Oaxaca, les us et coutumes sont reconnus dans la loi sur les droits des peuples et des communautés indigènes (1997) ; ils englobent un système de charges locales et de travail communautaire inspiré à la fois des habitudes précolombiennes et, suivant les lieux, d'un héritage colonial. A Eloxochitlan ils concernent aussi la reconnaissance de la médecine locale, du droit à la terre et de la défense de la langue sous tous ses aspects.

nouvelles révélations sur la pratique de la langue sifflée par la communauté elle-même. Par exemple, nous avons constaté que dans cette partie de la Sierra Mazatèque, contrairement à ce qu’avaient pu observer Cowan et Busnel dans les autres zones, la langue sifflée est utilisée à très grande distance (La Photo 16 du Chapitre 2 témoigne du fait que la technique de sifflement avec les doigts y est courante).

#### **1.1.5.4. Les communautés montagnardes en Asie du Sud-Est**

Depuis quelques années, certaines communautés montagnardes d’Asie ont initié un travail ayant de nombreux aspects scientifiques afin de promouvoir la documentation de leur tradition orale menacée. Comme pour les Mazatèques il s’agit d’une lutte identitaire où la documentation et la défense du patrimoine a pour enjeux la survie de tout un mode de vie. Comme pour La Gomera, cette lutte a été initiée par un petit groupe d’anciens. La langue sifflée est intégrée à cette démarche avec divers *instruments parleurs* car ils jouent un rôle important de vecteur de la mémoire orale dans ces cultures où des réciteurs traditionnels font office de professeurs de la tradition qui sert d’inspiration au quotidien. Par exemple, les musiciens interprètent en paroles jouées leur enseignement.

Lors de notre collaboration avec des organisations montagnardes Hmong et Akha en Thaïlande, nous avons constaté que ces communautés réalisaient des travaux de recherche très approfondis dont certaines techniques ont été inspirées de leur collaboration avec des anthropologues, des musiciens, des *reporter* ou des linguistes. Ces communautés, si elles ont quelques difficultés à financer leur action ont cependant publié des ouvrages extrêmement précis et inédits sur leur tradition orale qui sont organisés à la manière des ouvrages de recherche. Elles prennent donc le relais des chercheurs occidentaux qui les ont inspirés tout en présentant le point de vue interne de la communauté.

Le cas des Akhas illustre bien cette situation. L’ouvrage *Akha Oerzar*, Diara (2002) est par exemple issu d’un travail de plusieurs années de collaboration entre l’association Akha MPCD-SEAMP et le conseil des *Pirma* (professeurs réciteurs Akha), l’anthropologue hollandais Leo von Gesau et la linguiste suédoise Ingall Hanson. Cet ouvrage de 268 pages représente le premier volume de la tradition orale énoncée par les réciteurs traditionnels Akha. Il est devenu la référence des études sur cette culture. Il contient à la fois une transcription écrite de la langue ancienne récitée, de la langue akha courante et une traduction en anglais. Ce texte est la base de certains discours de la vie quotidienne qui sont utilisés par les siffleurs, par exemple lorsque la connaissance de la poésie locale les inspire.

### **1.1.6. Positionnement de notre étude**

#### **1.1.6.1. Enjeux scientifiques**

Les enjeux scientifiques soulevés par un tel sujet d’étude sont extrêmement larges en raison du relatif silence de la tradition scientifique sur le sujet depuis les années 70 dans tous les domaines (linguistique, anthropologie, linguistique, musicologie, acoustique) et de la rareté des corpus enregistrés, disponibles ou accessibles. Compte tenu de cette situation, nous avons décidé de poursuivre la recherche de nouvelles

langues sifflées, de visiter des lieux où des langues sifflées étaient connues, ou du moins signalées ; de constituer une nouvelle base de donnée sur le sujet tout en faisant un bilan comparatif au moyen des outils de la typologie linguistique, de la bioacoustique et de la sociolinguistique.

Comme nous en fournirons la preuve par la suite, il ressort de notre enquête que tout type de langue peut être sifflé. Dès lors, une grande variété de structures de langues est à analyser et surtout à comparer. De plus, presque toutes les langues sifflées sont aujourd'hui menacées de disparition<sup>9</sup>, ce phénomène apporte la possibilité d'analyser leur dégradation tout en renforçant l'exigence éthique locale des rares siffleurs authentiques actuels. Dans ces conditions, la logique de certaines communautés de prendre en main l'analyse de leur propre culture est un phénomène que nous avons décidé d'accompagner quand il existait et de suggérer quand il était absent afin de constituer, avec leur aide, un réseau international de recherche et de documentation ancré localement. Nous avons observé que c'est seulement sur cette base que de nouvelles perspectives peuvent être ouvertes afin d'approfondir la connaissance des langues sifflées avant que leur diversité ne disparaisse de la surface de la planète. Une attention particulière a été apportée à la constitution d'une méthodologie de travail que nous décrivons en Annexe **Erreur ! Source du renvoi introuvable.**

### **1.1.6.2. Enquête de terrain**

La réalisation de notre projet de recherche a nécessité l'organisation d'une expédition scientifique d'enquête de terrain et de recueil de données adaptées. Cette phase a duré 14 mois, avec une équipe de deux personnes<sup>10</sup>. Comme nous venons de le dire, une méthodologie d'enquête simple a été mise en place permettant de garder des liens locaux à long terme pour organiser l'archivage, l'utilisation scientifique et, éventuellement, la revitalisation culturelle des langues concernées. Le soutien moral obtenu auprès du département du « Patrimoine Immatériel de l'Humanité » de l'Unesco s'est concrétisé par deux réunions de travail avant le départ et au retour. Le budget réduit de l'enquête a été un atout pour rencontrer les acteurs culturels motivés par le projet.

Au final, une collaboration a été initiée avec 9 communautés différentes réparties dans le monde entier (Gomeros : espagnols aux Canaries, Mazatèques au Mexique, Boras en Amazonie Péruvienne, 2 cultures des îles d'Ambrym et Pentecôte à Vanuatu, Akhas et Hmong en Thaïlande et au Laos, Turcs à Kusköy « village des oiseaux », Grecs d'Antia), avec plusieurs linguistes de terrain (en particulier l'équipe du Museu Goldi de Bélem). Un espace de discussion a été créé avec 2 anthropologues : Jürg Gashé (IIAP, CNRS) : Amazonie péruvienne ; Mariano Pavanello (La Sapienza, Rome) : Ghana) et des institutions nationales de recherche ou d'éducation. Les cultures évoquées ici n'ont pas toutes une langue sifflée, certaines sont également concernées par les langues tambourinées sur lesquels nous avons fait une enquête préliminaire par la même occasion en raison des liens étroits, dans certains villages, avec la technique de sifflement.

---

<sup>9</sup> À part peut être à la Gomera où un programme de revitalisation scolaire a été initié.

<sup>10</sup> La présence de deux personnes est nécessaire pour les enregistrements des langues sifflées en raison de la nécessité de documenter le phénomène à longue distance dans des conditions réelles de pratique. L'autre personne de l'équipe est l'ingénieur Laure Dentel que je remercie ici pour m'avoir autorisé à utiliser ici l'ensemble des sons et des photos qu'elle a contribué à collecter.

### **1.1.6.3. Influence de la technologie de collecte et d'analyse des données**

D'un point de vue technique, nous avons bénéficié lors de notre recherche de nouveaux outils technologiques par rapport aux études des années 70 à la fois en matière de recherche documentaire, de technologie disponible pour la collecte de données ou même pour le traitement et l'analyse des données.

Par exemple, il est indéniable que l'augmentation des capacités de mémoire des ordinateurs et l'invention du réseau Internet ont accéléré la recherche documentaire et la recherche de lieux susceptible d'abriter des langues sifflées inconnues. L'outil internet a été également utile pour l'organisation du voyage d'enquête et pour la prise de contact avec différentes personnes relais dans certains pays. Cependant, il est important de remarquer que l'essentiel des contacts de qualité ont été découverts une fois sur place, par bouche-à-oreille. En ce qui concerne le matériel de collecte de données, celui-ci est aujourd'hui plus souple, plus discret qu'autrefois et il est également plus robuste dans des conditions climatiques difficiles. Enfin, lors de l'analyse des données, les technologies de traitement mises en place à travers la programmation ou les outils et logiciels de traitement du signal ont aussi permis de renouveler complètement la perspective de l'analyse et d'accélérer le traitement des sons (voir en Annexe **Erreur ! Source du renvoi introuvable.** pour plus de précision sur tous ces points).

## **1.2. Langues sifflées : langues en danger ?**

Les langues sifflées, retranchées dans des zones rurales à biodiversité riche, ont toutes perdu beaucoup de locuteurs au cours des trois dernières générations. Elles sont entretenues vivantes par des cultures orales ou des modes de vie minorisés (même lorsque des transferts de la pratique du sifflement ont eu lieu dans des langues dominantes). Nombre d'entre elles, encore en usage il y a trente ans comme le béarnais dans les Pyrénées françaises ou le tepehua de la Sierra Madre Oriental au Mexique, étaient éteintes lorsque nous nous sommes rendus sur place en 2003 et, à notre connaissance, toutes les autres sont malheureusement menacées d'extinction d'ici une à deux générations. Les locuteurs siffleurs les utilisant quotidiennement sont donc rares même lorsqu'elles ont été brandies comme un étendard identitaire comme sur l'île de la Gomera aux Canaries.

### **1.2.1. Causes**

#### **1.2.1.1. Situation générale**

Les causes premières qui entraînent une dégradation des langues sifflées sont celles qui brisent un équilibre social et écologique d'une société villageoise. Ce qui est particulier à l'usage du sifflement est qu'il est adapté à une vie en petites communautés dispersées ; il ne survit donc pas aux regroupements de populations qui résultent de l'exode rural et son intérêt est diminué par l'amélioration des systèmes de communication modernes d'une région. Ces deux facteurs viennent s'ajouter aux différentes raisons qui affectent la vitalité des langues minoritaires et des modes de vie ruraux. D'une manière générale, l'émergence de nouveaux modes de vie affecte durablement la langue s'ils touchent à ses racines en convainquant les locuteurs de les remettre en cause. Dès lors, l'impact d'un contact culturel avec l'extérieur va dépendre de la mutation, violente ou non, de ces trois domaines et donc de l'histoire géopolitique d'un lieu.

Notre enquête de terrain a souligné que les langues sifflées sont étroitement liées aux racines culturelles d'un mode de vie local à travers les métiers et les activités traditionnelles liées à un environnement naturel. De plus, elles constituent un patrimoine oral dont la réalité sonore ne peut pas être retranscrite facilement par écrit. Lorsqu'elles disparaissent, elles ne laissent pas la trace, même précaire ou approximative, d'un système d'écriture. Tout comme les langues tambourinées, au cours des derniers siècles et jusque dans les années 1940, elles ont été la cible d'un dénigrement tant culturel que scientifique<sup>11</sup> à cause de leur double faculté d'organiser une population en cas de danger et de permettre de maintenir la cosmogonie locale vivante grâce

---

<sup>11</sup> La tradition scientifique, portée principalement par la culture dominante, n'est parvenue à s'ouvrir au point de vue des autres cultures que suivant un processus très lent qui est toujours en cours. Il semble s'accélérer avec la mondialisation et il dépend toujours des individus chercheurs. A l'époque coloniale, les tambours parlant ou les sifflements ont été jugés comme les témoins de l'archaïsme des cultures qui les portaient alors que c'est précisément ce jugement qui est archaïque.

à leur implication dans les fêtes, la poésie ou les rituels. En effet, ces pratiques de communication ancrent la langue locale dans une représentation<sup>12</sup> traditionnelle d'un milieu.

### **1.2.1.2. Dégradation du tissu social traditionnel**

Les changements sociaux de mode de vie à l'origine de la dégradation des langues sifflées se manifestent de diverses manières:

- Usage de modes de communication modernes efficaces demandant un moindre investissement en effort physique ou en apprentissage: dans ce cas, le principe du moindre effort qui est à l'origine de la naissance des langues sifflées, est aussi une des raisons de leur disparition.
- Exode rural (dont les facteurs déclencheurs principaux sont la dépréciation du mode de vie traditionnel au profit d'un modèle idéalisé d'un mode de vie moderne et les difficultés financières).
- A long terme, le vieillissement de la population des villages.
- L'éducation scolaire se fait principalement dans la langue dominante.
- Dépréciation de la langue locale (externe d'abord puis intériorisée par les locuteurs).
- La tradition orale n'est plus transmise que de manière tronquée (rassemblements, fêtes, contes et poésie porteurs d'une transmission des us et coutumes), folklorisation des traditions.
- Changement (ou syncrétisme) de religions qui s'accompagne souvent de la condamnation d'une partie de la tradition orale.

Ces différents aspects ne sont pas au même stade d'avancement dans tous les lieux. Il n'en demeure pas moins que, pour l'observateur étranger, le sifflement est aujourd'hui relativement inaudible presque partout, sauf s'il accompagne les villageois lors de leurs activités quotidiennes.

### **1.2.2. Facteurs de maintien**

Le maintien d'une langue sifflée dans un village et dans ses alentours tient en premier lieu à l'isolement géographique de celui-ci. Nous avons vu qu'une topographie très accidentée et une végétation dense empêchant les moyens modernes d'agriculture et d'industrialisation de s'installer permettent à la parole sifflée de se poursuivre. De surcroît, le facteur écologique qui semble être à l'origine du phénomène est si fort qu'il a engendré un transfert du sifflement d'une langue mourante à une langue dominante si les habitants poursuivent leurs usages traditionnels, quand le choc des cultures a trouvé un nouvel équilibre comme nous l'avons signalé en introduction. Lorsque le processus de dégradation culturelle est entamé, c'est la survivance d'un mode de vie traditionnel qui explique l'usage des langues sifflées tant qu'aucun autre moyen de communication ne vient s'y substituer. Ainsi en Turquie aujourd'hui le sifflement n'est presque plus pratiqué dans le village de Kusköy mais il continue d'être pratiqué par certains bergers partant tout l'été dans les hauts plateaux alentours où il n'existe pas de route et où les lignes téléphoniques ne sont pas installées. C'est

---

<sup>12</sup> Le mot représentation prend une valeur différente suivant les implications épistémologiques de la cosmogonie des cultures (Descola 2004).

pourquoi de nombreux villageois en ont encore une pratique parfaite. D'autre part, lorsque la vitalité de la langue sifflée est entamée au point de n'être plus pratiquée par la quasi-totalité d'une génération, la volonté d'un petit groupe de personne pour trouver un moyen de la maintenir a souvent permis d'engager des processus de revitalisation dont les conséquences, encore variées aujourd'hui dépendent des stratégies choisies (éducatives, touristiques...). Ces expériences actuelles de revitalisation ont été évoquées dans l'historique. Dans la grande majorité des cas, elles sont engagées lorsque la situation devient critique. La prise de conscience de cette situation vient principalement de la confrontation avec d'autres modèles de vie. Le mode de confrontation le plus ambivalent est le tourisme. Ainsi, si le maintien de la langue sifflée est axée autour d'une activité importée comme le tourisme organisé, la pratique de la langue sifflée va être folklorisée et dénaturée en produisant une caricature de pratique. Dans certains lieux pourtant, le tourisme est officiellement présenté comme une solution alternative permettant aux traditions de se maintenir. C'est sans compter sur le fait que cette activité représente à la fois une manne financière source de discord et bien souvent un moyen politique d'intégration nationale d'une communauté isolée. Les dérives des activités touristiques vont dépendre des fins et des moyens des entrepreneurs engagés dans ce commerce. Dans le cas de notre travail de terrain, nous avons constaté que les locuteurs et les informateurs les plus authentiques n'étaient pas engagés dans des activités d'exploitation touristique de la culture. Bien souvent ils étaient extrêmement opposés à ces activités même s'ils étaient impliqués dans un processus de revitalisation alliant modernité et tradition. Lorsque le tourisme était installé dans leur région, leur souci était de juguler son flux et d'en contrôler les côtés négatifs.

### **1.2.3. Aperçu de la vitalité autour du monde**

#### **1.2.3.1. Carte de répartition des langues sifflées**

Le sifflement est aujourd'hui rarement observé comme une pratique étendue à plus de quelques vallées ou quelques villages dans une zone géographique donnée. Notre enquête nous a permis de dresser une liste des langues étudiées et des langues encore vivantes (Tableau 9 du Chapitre 3). Compte tenu de la durée de chacun de nos séjours dans chaque zone visitée (1 à 2 mois), de l'isolement qui caractérise les bons locuteurs siffleurs et de la variabilité de la qualité des sources d'information dont nous avons bénéficié, les résultats de cette enquête ne sont en aucun cas exhaustifs. Il n'en demeure pas moins qu'ils renouvellent les connaissances en la matière. Il s'avère que le phénomène est plus largement répandu que ce que n'en laisse paraître la documentation sur le sujet. Les langues sifflées ont un caractère quasi universel, même si elles n'ont pas été développées dans tous les endroits propices à leur usage. Elles ne sont pas le fait de trois ou quatre villages ni même d'une dizaine de communautés. Au contraire, elles se sont développées pour des usages, soit multiples, soit restreints à une activité, dans des centaines d'endroits et de communautés différentes. Si l'on ne considère que le Sud-Est asiatique et le Mexique, plus d'une trentaine de langues sont concernées. Quelques rares chercheurs l'avaient déjà suggéré et l'ensemble des témoignages que nous avons réunis l'atteste. Cette enquête a aussi confirmé que la vitalité du sifflement langagier est aujourd'hui en fort déclin. Ceci laisse penser que ce phénomène était encore plus largement répandu il y a quelques décennies et a fortiori il y a des centaines d'années. Les quelques travaux scientifiques ou historiques les plus anciens qui

se sont intéressés au sujet ont une grande valeur car leurs témoignages nous fournissent le moyen d'avoir une idée plus précise sur la répartition et sur l'évolution des langues sifflées.

**Figure 1 : Carte représentative de la diversité des langues sifflées signalées.**

Aujourd'hui le béarnais et le tepehua sont supposées « langues sifflées éteintes », le kickapoo et le ndiki sont très probablement dans le même cas. Le vide observé en Amérique rend compte de la forte dégradation des cultures traditionnelles. Le vide observé au Moyen-Orient ou dans les pays situés entre la Turquie et la Chine témoigne d'un manque d'investigation. Le vide observé en Australie cumule ces deux dernières explications.

### 1.2.3.2. Evaluation et analyse de la vitalité

Le phénomène de la lente mort des langues est de mieux en mieux connu et documenté en raison du nombre croissant de cas où les linguistes se trouvent confronté à cette situation. En effet, un groupe d'experts, réunit par l'Unesco en 2003 a estimé qu'aujourd'hui, parmi les 6000 langues que la terre abrite, une majorité d'entre elles est menacée de disparition. A l'heure actuelle la moitié perd des locuteurs. D'autre part, environ 97% de la population mondiale parlent 4% des langues du monde et, à l'inverse, 96% de langues ne sont parlées que par 3% des êtres humains.

D'autre part, dans tous les lieux que nous avons visités, nous avons pu observer les mêmes catégories de siffleurs. Celles ci sont extrêmement proches des catégories de locuteurs définis par les linguistes à propos des langues subissant une perte de vitalité et sont caractérisées par la disparité des compétences des locuteurs.

#### 1.2.3.2.1. Critères généraux

Pour analyser la vitalité de chaque pratique locale du sifflement articulé, nous nous sommes appuyés sur les critères dégagés par la réunion en 2003 d'un comité de l'Unesco sur les langues en danger. Il en est résulté plusieurs critères précis présentés dans le Tableau 1.

**Tableau 1 : Grille d'évaluation de la vitalité d'une langue d'après les critères établis par l'Unesco (2003)**

Facteurs	Description : Chaque critère est évalué suivant une graduation allant de 0 à 5 (0 : situation favorable à l'extinction, 1 : danger critique, 2 : danger sévère, 3 : danger reconnu, 4 : dégradation sans danger, 5 : situation sûre)
<b>A) Analyse de la vitalité et de l'état de danger de la langue</b>	
Facteur 1	Transmission intergénérationnelle du langage
Facteur 2	Nombre absolu de locuteurs
Facteur 3	Proportion de locuteurs dans la population totale
Facteur 4	Perte de domaine d'application
Facteur 5	Réceptivité du langage à de nouveaux domaines et présence dans les médias
Facteur 6	Matériel disponible pour l'enseignement du langage et dans la littérature.
<b>B) Attitude et politique linguistique</b>	
Facteur 7	Attitude et politique du gouvernement et des institutions nationales (statut officiel et utilisation inclus)

Facteurs	Description : Chaque critère est évalué suivant une graduation allant de 0 à 5 (0 : situation favorable à l'extinction, 1 : danger critique, 2 : danger sévère, 3 : danger reconnu, 4 : dégradation sans danger, 5 : situation sûre)
<b>Facteur 8</b>	Attitude des membres de la communauté linguistique vis à vis de leur propre langue
<b>C) Urgence de la documentation</b>	
Facteur 9	Etendue et qualité de la documentation

### 1.2.3.2.2. Diagnostic inspiré des critères de l'UNESCO et typologie des profils de siffleurs

Nous avons anticipé les difficultés de collectes de données en adoptant une grande partie de la méthodologie de travail mise au point par l'Unesco, dont l'association des personnes locales à notre enquête. Les premiers résultats de ce travail commun sont présentés pour quelques cultures.

#### **Bilan de vitalité : exemples de la Gomera, de la Grèce et du pays Mazatèque.**

Pour être adaptée à une langue sifflée, la grille d'évaluation doit être modifiée sur certains critères en raison de son usage uniquement à l'oral (absence de système écrit dédié spécifiquement aux sifflements). C'est pourquoi pour les facteurs 6 et 9, nous avons considéré l'existence d'une bibliographie sur le sujet, de la documentation sonore et vidéo et de leur disponibilité dans les écoles et les centres culturels locaux

Quand la version parlée de la langue concernée est en situation de dégénérescence, il est nécessaire d'interpréter notre tableau avec l'aide du tableau équivalent de la langue mère parlée.

**Tableau 2 : Vitalité de la langue sifflée de l'île de la Gomera ((bonne vitalité de la langue parlée espagnole)**

**Collaborateurs d'enquête : Luis Morales Mendez, Eugenio Darias, Lino Rodrigues, Isidro Ortiz.**

<b>Facteurs</b>	<b>Description :</b> Chaque critère est évalué suivant une graduation allant de 0 à 5 (0 : situation favorable à l'extinction, 1 : danger critique, 2 : danger sévère, 3 : danger reconnu, 4 : dégradation sans danger, 5 : situation sûre)
<b>A) Analyse de la vitalité et de l'état de danger de la langue</b>	
Facteur 1	4 (depuis 1999 : enseignement scolaire obligatoire entraînant un saut de génération) ; 2 auparavant (la majorité des siffleurs étaient issus de la génération des grands-parents)
Facteur 2	30 siffleurs ayant acquis la langue de manière précoce (estimation haute), 1000 enfants (de niveaux très inégaux)
Facteur 3	2 (une minorité la parle sur l'île, 10% avec les enfants)
Facteur 4	2 (usage quotidien rare pour la majorité des siffleurs)
Facteur 5	3 (le langage sifflé est adapté à de nouveaux domaines : chantiers de construction, enseignement informatique, personnes laryngectomisée, tourisme, démarchage dans d'autres îles.)
Facteur 6	4 (plusieurs types d'enseignement ont été adoptés, traditionnels (Lino Rodrigues) et théoriques (Isidro Ortiz)

<b>Facteurs</b>	<b>Description :</b>
	Chaque critère est évalué suivant une graduation allant de 0 à 5 (0 : situation favorable à l'extinction, 1 : danger critique, 2 : danger sévère, 3 : danger reconnu, 4 : dégradation sans danger, 5 : situation sûre)
<b>B) Attitude et politique linguistique</b>	
Facteur 7	4 depuis 1999 : enseignement scolaire et valeur culturelle reconnue <sup>13</sup> ; 3 auparavant : assimilation passive.
Facteur 8	4
<b>C) Urgence de la documentation</b>	
Facteur 9	4 (bonne documentation : films, reportages, études scientifiques accessibles à la population.)

**Tableau 3 : Vitalité de la langue sifflée grecque du village d'Antia (bonne vitalité de la langue parlée)**

**Collaborateurs d'enquête : l'école locale, Mr. Panayotis, Maria Panayotis et Mrs Kula.**

<b>Facteurs</b>	<b>Description :</b>
<b>A) Analyse de la vitalité et de l'état de danger de la langue</b>	
Facteur 1	1 (Les enfants que nous avons rencontré ne l'ont pas apprise. Dans la génération des parents seuls certains l'utilisent. Une grande partie de la génération intermédiaire a quitté le village pour Athènes)
Facteur 2	Une dizaine de siffleurs dont certains refusent de le montrer et d'autres disent qu'ils ne peuvent plus le pratiquer à cause de leurs dents
Facteur 3	10 %
Facteur 4	1 (les métiers traditionnels ne sont plus pratiqués comme avant sauf par les anciens)
Facteur 5	2 (adaptation du vocabulaire moderne mais pas d'emploi dans de nouvelles situations)
Facteur 6	0 (aucune tentative d'enseignement; l'idée même paraît saugrenue)
<b>B) Attitude et politique linguistique</b>	
Facteur 7	2 (peu d'intérêt des autorités, pas d'effet touristique organisé, mépris de certains scientifiques nationaux pour les populations rurales)
Facteur 8	2 (les habitants eux-mêmes ont intégré le fait que cette pratique fait partie du passé révolu)
<b>C) Urgence de la documentation</b>	
Facteur 9	2 (aucune documentation disponible sauf, à Athènes, un film documentaire réalisé en 1982 et les articles de recherches publiés. Le matériel sonore enregistré dans le passé n'est pas accessible.

**Tableau 4 : Vitalité de la langue sifflée Mazatèque**

**(version parlée pratiquée au quotidien mais subissant la domination administrative et scolaire de l'espagnol).**

**Collaborateurs d'enquête : Juan Casimiro, Inti Betanzos, Jaime Betanzos et le conseil Indigène d'Eloxochitlan**

<b>Facteurs</b>	<b>Description :</b>
<b>A) Analyse de la vitalité et de l'état de danger de la langue</b>	
Facteur 1	3 (plusieurs générations l'utilisent encore ; les jeunes de manière moins extensive.)

<sup>13</sup> La graduation atteint 5 à l'approche des élections : lancement du Journal « el silbo », lancement d'un site Internet abandonné depuis, projet de statue de siffleur et organisation d'un congrès des langues sifflées (Avril 2003).

Facteurs	Description :
Facteur 2	Encore en cours d'évaluation, voir ligne suivante pour un ordre de grandeur.
Facteur 3	2 (environ 10% le pratiquent mais de manière limitée en ville, une plus grande proportion à la campagne pour des conversations plus étendues : 20 % environ)
Facteur 4	3 (l'usage tend à être remplacé par des communications plus rapides par téléphone ou en voiture)
Facteur 5	2 (adaptation difficile du système à la nouvelle langue dominante, intégration difficile des prénoms espagnols (voir partie intelligibilité).
Facteur 6	1 (aucun matériel spécifique à la langue sifflée mais le système est expliqué par certains écrivains et responsables locaux)
<b>B) Attitude et politique linguistique</b>	
Facteur 7	3 (reconnaissance officielle des us et coutumes mais usage parfois détourné de la tradition)
Facteur 8	3 (quelques membres supportent la maintenance et l'intérêt pour la langue sifflée, certains responsables culturels ont décidé de le documenter (à Huautla de Jimenez et Eloxochitlan de Flores Magon)
<b>C) Urgence de la documentation</b>	
Facteur 9	2 (aucune documentation à part quelques enregistrements sonores. Manque de matériel et de moyens de documentation. Projet en cours d'initiation à Eloxochitlan et à Huautla de Jimenez)

## Profils types de siffleurs

Il est possible de distinguer plusieurs catégories principales de siffleurs. Ce type de distinction apparaît à la fois dans les lieux où la langue sifflée est encore largement répandue<sup>14</sup> et -de manière très prononcée- dans les lieux où la langue sifflée est presque morte. En effet, la raréfaction de l'usage du sifflement renforce la différence entre les locuteurs qui poursuivent des activités traditionnelles de plein air et ceux adoptant un mode de vie plus modernisé. A ce premier phénomène s'ajoutent ceux qui caractérisent la dégradation d'un système linguistique en situation de dégénérescence<sup>15</sup> (Dorian 1989, Grinevald 1997). C'est pourquoi nous nous sommes inspirés de l'étude des profils de locuteurs qui a été développée dans le cadre de typologies des locuteurs des langues en danger (Grinevald 2001, Bert 2001, Grinevald 2003) pour élaborer une typologie des siffleurs (disponible en Annexe **Erreur ! Source du renvoi introuvable.**).

<sup>14</sup> Si l'on se fie aux témoignages les plus anciens il semble que cela a toujours été le cas

<sup>15</sup> Voir l'analyse du béarnais en Annexe **Erreur ! Source du renvoi introuvable.**

## 1.3. Problématique de la définition des langues sifflées

### 1.3.1. Le point de vue des chercheurs

Comme nous l'avons vu, tous les auteurs ayant approfondi l'analyse du sujet reconnaissent que les langues sifflées servent les mêmes objectifs que la voix parlée, à partir des structures de langues dont ils ne sont qu'une transposition physique destinée à atteindre une plus grande distance que la parole classique parlée ou criée. Les taux d'intelligibilité élevés obtenus ne sont pas dus à un changement de la syntaxe ou du vocabulaire. Le degré de variabilité des énoncés témoigne du fait que le sifflement est un moyen d'expression de la langue comme un autre. On est bien loin de l'idée d'un système développant un langage de remplacement (*language surrogate*) tel qu'il est décrit par Sebeok et Umiker-Sebeok (1976) en introduction de l'ouvrage *Speech surrogates : drum and whistled systems*. Pour cette raison, Busnel et Classe initièrent la conclusion générale de leur travail commun de synthèse publié également en 1976, par les propos suivants:

*"It will no doubt have been noticed that not once throughout this monograph have whistled languages been referred to as speech surrogates. In the opinion of the present writers, there is nothing to recommend the use of this term. Whereas the sign language of deaf-mutes, for instances is truly a surrogate since it is a substitute for normal speech, whistled languages do not replace but rather complement it in certain specific circumstances. In other words, rather than surrogates they are adjuncts. Looking at the question from a slightly different point of view, when a Gomerio or a Turk whistles, he is in fact still speaking, but he modifies one aspect of his linguistic activity in such a way that major acoustic modifications are imposed upon the medium. Nevertheless the fact remains that he is still using the same medium, although in a vastly different shape. The procedure would be identical if one were to speak into a machine designed to convert the spoken words into whistled signals. [...] At the receiving end, the acoustic signal is mentally converted back into the original verbal image that initiated the chain of events"*(Busnel et Classe, 1976, p.107).

De son côté, Cowan affirmait à propos du Mazatèque *"Whistled conversations correspond very closely to spoken conversations. When asked what a person 'said' when he whistled, the natives will respond with a very specific and literal rendering in the spoken language [...]. Culturally whistling is treated as a natural and integral part of conversation [...]. The whistle is obviously based on the spoken language"* (Cowan 1948 in Sebeok et Umiker-Sebeok, p.1386-1390). Même s'il s'agit ici d'une langue tonale, la complexité des phrases échangées et le peu d'erreurs des siffleurs dans leur interprétation semblent suggérer que les remarques de Busnel et Classe s'appliquent aussi au mazatèque. Plus tard Cowan affirmera dans un article sur la langue tepehua: *« The question might be well asked, if whistled Tepehua should not be considered a style of speech (as whisper is for example), rather than a substitute for language, since it is used by the same person and involves the same physiological mechanism and linguistic system to achieve the same cultural purpose »* (Cowan 1976, in Sebeok et Umiker-Sebeok, p. 1407).

Le fait indéniable qu'un certain nombre de réductions phonétiques sont effectuées par le signal sifflé ne remet pas en cause l'opinion de Busnel, Classe et Cowan.

L'idée de langage de remplacement (*language surrogate*) n'est éventuellement recevable que dans le cas où le sifflement linguistique est basé sur un système discursif commun avec celui des langues tambourinées car les siffleurs utilisent alors un type de syntaxe particulier.

### **1.3.2. Le point de vue des siffleurs**

Pour tous les siffleurs que nous avons interrogés, il est évident que la langue sifflée n'est qu'un aspect de la langue locale à laquelle elle est intégrée : alors que le chuchotement est utilisé à courte distance, similairement, le sifflement est utilisé à longue distance. Pour justifier cette opinion, ils soulignent qu'ils disent en sifflement ce qu'ils pensent dans la langue et que le processus équivalent est en jeu lorsqu'ils reçoivent un message. Nous pouvons témoigner du fait que toutes les phrases de notre corpus ont pu être traduites littéralement sans ajout de périphrase ou de quelque technique de vocabulaire propre aux langues sifflées. Par contre, nous avons pu observer, en faisant écouter des sons de nombreuses cultures à nos interlocuteurs, que les siffleurs eux-mêmes ont tendance à ne reconnaître comme langue sifflée que les systèmes de transposition qui s'appuient sur la même stratégie que leur langue<sup>16</sup>.

### **1.3.3. Positionnement des langues sifflées**

#### **1.3.3.1. Propriétés clefs**

Un certain nombre de caractéristiques dont l'importance n'est pas évidente à première vue, jouent un rôle fondamental dans le phénomène des langues sifflées. Nous avons identifié les suivantes: une langue sifflée est un *phénomène naturel* issu de l'*évolution en réponse à une contrainte* particulière, elle transpose les caractéristiques de la voix parlée d'*une langue locale*, la *stratégie de transposition* repose sur le *système oral* d'une langue en particulier de la *voix parlée* ; l'*étendue du répertoire de signes distinctifs est finie* alors que leurs arrangements variés permettent de développer un *vocabulaire potentiellement infini*. La langue sifflée n'est *pas directement intelligible* à une personne connaissant la langue parlée mais elle requiert une phase d'apprentissage. La langue sifflée est presque toujours destinée à un *dialogue entre individus*.

#### **1.3.3.2. Comparaisons avec d'autres systèmes de communication**

La manière dont la parole sifflée partage ces propriétés clefs avec un certain nombre d'autres modes de communication élaborés permet de mieux comprendre son positionnement par rapport à la voix utilisée le plus communément. L'explication détaillée du tableau récapitulatif présenté ci-dessous est fournie en Annexe **Erreur ! Source du renvoi introuvable.**

---

<sup>16</sup> Résultat obtenu à partir d'écoutes inter-langues effectuées dans 4 communautés linguistiques.

**Tableau 5 : Positionnement de différents systèmes de communication par rapport aux propriétés clefs des langues sifflées.**

Mode de communication	Naturel	Artificiel	Basé sur la langue orale locale	Basé sur l'écriture	Non-intelligibilité systématique un locuteur de la langue	Style de parole de la langue	Usage à grande distance
Langue sifflée	X		X		X	X	X
Parole chuchotée	X		X			X	
Parole criée	X		X			X	X
Morse		X		X	X		X
Langues des signes	X				X		
Langue signée (français signé par exemple)		X		X	X	X	
Espéranto		X					(X)
Téléphone		X	X			(X)	X
Langue tambourinée Type I (système transposé de la parole)	X		X		X		X
Langue tambourinée Type II (système parallèle)	X				X		X

## **1.4. Conclusion : réponse à la question « Qu'est ce qu'une langue sifflée ? »**

### **1.4.1.1. Une forme de la langue locale**

Une langue sifflée est la version sifflée d'une langue locale avec laquelle elle partage le vocabulaire et la grammaire. Elle s'apparente dans la plupart des cas à un style de parole adapté à la distance. Elle partage alors également la syntaxe de la langue quotidienne. Le terme de « parole sifflée » convient donc aux différentes productions linguistiques des siffleurs. Il arrive plus rarement qu'elle repose sur un type de discours faisant appel à une syntaxe particulière intercalant des formes stéréotypées et des formes libres, comme la plupart des langues tambourinées.

### **1.4.1.2. Production**

Le sifflement fait office de source sonore à la place des cordes vocales de la voix parlée. La cavité de la bouche est presque toujours l'unique résonateur mais il arrive que certaines techniques lui joignent une cavité formée entre les deux mains jointes (voir Chapitre 2).

### **1.4.1.3. Origine évolutive naturelle**

Une langue sifflée est le résultat de l'adaptation de la perception et de la production humaine à la nécessité de communiquer dans des conditions d'isolation. C'est pourquoi, comme nous le verrons dans les chapitres suivants, les langues sifflées se sont développées principalement chez les communautés humaines vivant dans des forêts denses et des montagnes, qui favorisent l'éloignement en termes de temps de trajet ou de contact visuel.

### **1.4.1.4. Système de communication à distance optimisé au niveau bioacoustique**

Le sifflement réalise une optimisation bioacoustique. Il est non seulement adapté à la propagation à distance mais aussi cible une bande de fréquence pour laquelle, d'une part la réverbération et le bruit ont un effet minimisé (voir Chapitre 2) et d'autre part, l'audition est la plus sensible et sélective (voir Chapitre 4).

### **1.4.1.5. Lien avec des activités sociales**

Les conditions de communication à distance doivent être suffisamment fréquentes pour nécessiter l'emploi d'un système linguistique. La parole sifflée a trouvé sa place dans des communautés relativement peu peuplées en facilitant l'organisation de la vie sociale et parfois directement la transmission de la tradition orale. Le système linguistique développé permet principalement aux individus d'un groupe (famille, village) de coordonner des activités vitales liées:

- A l'approvisionnement en nourriture (chasse, pêche, cultures en terrasses et dans les champs des vallées);
- A la garde du bétail (bergers principalement);
- A la manifestation d'un danger ou d'un cas d'urgence;
- Au rythme de la journée (lever, nombreuses occasions de parler de l'heure, appel pour les repas, communications nocturnes);
- Aux moments forts de la vie (naissance, séduction, mariage, décès, fêtes, jeux et rituels);
- Pour parler en secret vis à vis des étrangers<sup>17</sup>.

Aujourd'hui, rares sont les cultures qui utilisent les sifflements pour parler dans tous ces contextes. La cause peut être une dégradation de la pratique mais aussi un choix culturel, comme réserver le sifflement à un domaine particulier de la vie sociale. Quoiqu'il en soit, le maintien d'un seul des contextes d'usage est l'occasion de poursuivre la pratique.

Les langues sifflées ont l'avantage de souligner de manière directe l'aspect crucial de l'usage contextuel social pour l'évolution d'une langue. Mais ce facteur n'est pas propre aux langues sifflées, c'est l'un des paramètres fonctionnels principaux du phénomène du langage en général. Un contexte donné va réduire les ambiguïtés dans une communication en favorisant l'usage d'un vocabulaire dédié (qui n'est pas forcément limité: potentiellement infini) et en donnant un cadre à la communication.

D'autre part, l'usage des langues sifflées, s'il est considéré à travers l'ensemble de ses manifestations dans toutes les cultures où il peut être observé, couvre toutes les possibilités de discours. De la même manière qu'il est rare d'observer dans un même lieu l'ensemble des formes humaines de discours parlé, il est très rare d'observer une culture ayant développé tous les aspects sifflés possibles.

## **1.4.2. Quelques quiproquos récurrents**

### **1.4.2.1. Discours et stratégie de transposition**

Dans certaines cultures, seul un mode bien précis de discours emploie la technique des langues sifflées. Ce phénomène a lieu principalement dans les populations ayant une tradition orale très riche. Il est parfois tentant de penser qu'un type de discours est une stratégie de transposition. En effet, dans certaines cultures, la syntaxe particulière du discours sifflé permet de réduire les confusions au point que certaines distinctions phonétiques deviennent inutiles. Cependant le fait que d'autres cultures développent à la fois le discours rituel et la parole la plus quotidienne en sifflement (par exemple les populations Ewe du Ghana (voir en

---

<sup>17</sup> L'usage d'une langue sifflée comme système de communication secret a été signalé dans de nombreux lieux. Il est dû à l'inintelligibilité du signal aux personnes étrangères. De nombreux témoignages des Pyrénées signalent que le sifflement a été utilisé localement pour éviter l'impôt de guerre en période d'occupation (Arripe 1990). A la Gomera les berbères l'utilisaient pour tromper l'ennemi Français puis Espagnol. En Papouasie Nouvelle Guinée, des siffleurs wam ont été recrutés par les Australiens lors de la seconde guerre mondiale pour tromper les écoutes Japonaises (Nekitel, 1992).

Annexe **Erreur ! Source du renvoi introuvable.**) et Akha d'Asie), semble indiquer que le type de discours n'est en général qu'une expression de la stratégie de transposition. Le choix de réduire les éléments utiles dans le sifflement est donc d'abord lié à des options culturelles faites par une population pour des raisons sociales. On peut légitimement penser que dans d'autres conditions son système sifflé aurait développé, pour la même langue, des formes sifflées phonétiquement plus riches. En fait, cette problématique permet de réfléchir de manière ciblée à l'évolution des langues.

### **1.4.2.2. Langues sifflées et prosodie**

Les langues sifflées sont souvent associées à la prosodie de la langue. En linguistique cette association prête à confusion car la plupart des analyses de la prosodie s'appuient principalement sur l'observation de l'évolution de la fréquence fondamentale de la parole. En effet, les langues sifflées, similairement, ne s'appuient que sur une seule bande de fréquence mais elles ne privilégient pas forcément la fréquence fondamentale : tout dépend de la structure de la langue locale.

### **1.4.2.3. Langue musicale ?**

De nombreux auteurs ont qualifié les langues sifflées de « *langues musicales* » (Verneau 1923, cité par Busnel et Classe 1976 ; Eboué, 1935 ;...). Cette formule, s'appuyant sur le parallèle implicite qui existe entre prosodie et musique, n'évoque pas uniquement l'existence d'éléments musicaux dans la structure la langue comme il est fait couramment pour la voix, mais elle soutient l'idée d'une véritable « fusion » entre la musique et le langage dans le sifflement. De telles affirmations sous-entendent souvent que la langue sifflée ne s'appuie pas sur des aspects phonologiques de la langue mère mais sur l'équivalent d'une intonation qui serait figée pour faciliter la compréhension. La reconnaissance des phrases ne reposerait donc pas sur la structure des mots mais en priorité sur un effort de mémoire équivalent à celui qui est en jeu lorsqu'on se remémore les paroles d'une chanson à l'écoute de sa mélodie musicale connue.

Cette interprétation est très loin de la réalité même si, comme dans la parole classique, certaines expressions très utilisées ont parfois une prosodie ou une rythmique figée par l'usage (comme des locutions canoniques réduite en « b'jour » pour dire « bonjour »). Pourtant nous avons rencontré ce point de vue de manière récurrente chez des personnes qui ne sont pas familiarisées avec ces langues. Il semble que le sifflement leur donne l'impression de se rapprocher de la voix chantée alors qu'il s'appuie sur la voix parlée. L'usage de l'expression « *langue musicale* » a donc l'avantage de nous permettre de dissiper un malentendu dont l'étude des langues sifflées en linguistique a énormément souffert<sup>18</sup>.

D'autres confusions entre langues sifflées et musique viennent de l'impression de pouvoir rendre compte de la complexité du signal par la notation musicale. Pourtant de nombreuses démonstrations ont été faites à ce jour pour montrer que l'approximation réalisée par la notation musicale classique occidentale ne permettait pas de rendre compte de la complexité du signal sifflé, surtout dans le cas des langues sans tons et des

---

<sup>18</sup> Nous fournissons d'autre part en fin d'Annexe A une analyse des différences de degré de rythmicité et de mélodicité qui existent entre musique et langage.

langues tonales fortement modulées (voir Annexe **Erreur ! Source du renvoi introuvable.**). Les imprécisions de notation musicales ne sont pas dues uniquement au problème du passage à l'écrit : en effet, d'une manière générale il est difficile, pour un auditeur non habitué à entendre les sons d'une langue étrangère, de juger de sa musicalité sans y projeter les régularités mélodiques des langues et des musiques qu'il maîtrise, en particulier de sa culture maternelle<sup>19</sup>. Or, pour chaque langue les éléments de musicalité les plus pertinents en terme d'intelligibilité s'appuient sur des arrangements entre attributs perceptifs qui ont une part non négligeable de spécificité linguistique.

---

<sup>19</sup> Ainsi, dans le cas des langages tambourinés de Vanuatu, dont il semble qu'il créent un langage parallèle d'après une première analyse des enregistrements que nous avons fait sur l'île d'Ambrym, Deacon qui était un fin musicien notait : *"Quite often, what appears to me to be the same rhythm is said not to be, and on the other hand, I detect variations in rhythms which are said to be the same. It appears to be a question of the direction in which European and Melanesian musical sensibility has developed"* (Deacon 1934, p. 1210).

## **CHAPITRE 2. INFLUENCE DU MILIEU NATUREL**

## 2.1. Introduction

Les populations ayant développé des systèmes linguistiques ayant une forme sifflée vivent presque tous dans des zones montagneuses et des forêts denses. Pour les très rares peuples vivant dans des plaines non boisées et qui ont tout de même une adaptation sifflée de leur langue, cette dernière est peu pratiquée pour la conversation quotidienne<sup>20</sup>.

Dans les forêts, l'isolement en termes de contact visuel doit être vaincu pour permettre à des groupes de se déplacer sans se perdre. Dans ce cadre le cri aigu pourrait très bien faire l'affaire et il est parfois pratiqué comme nous avons pu l'observer chez les femmes Boras d'Amazonie. Mais il est mal adapté pour des chasseurs de gibier ou pour des pêcheurs, car il ne permet pas de se fondre dans les bruits de la faune locale. Le sifflement passe plus inaperçu en raison de la grande diversité des chants sifflés d'oiseaux de ces milieux naturels dont certains sont très proches du sifflement humain. D'autre part, comme nous le verrons ici, le sifflement est bien adapté à la propagation dans les forêts. L'effort des individus est donc optimisé face à la haute probabilité du signal de se réfléchir à la rencontre des obstacles physiques et donc de subir une déperdition d'énergie sonore (voir Figure 2) et un phénomène de réverbération<sup>21</sup>.

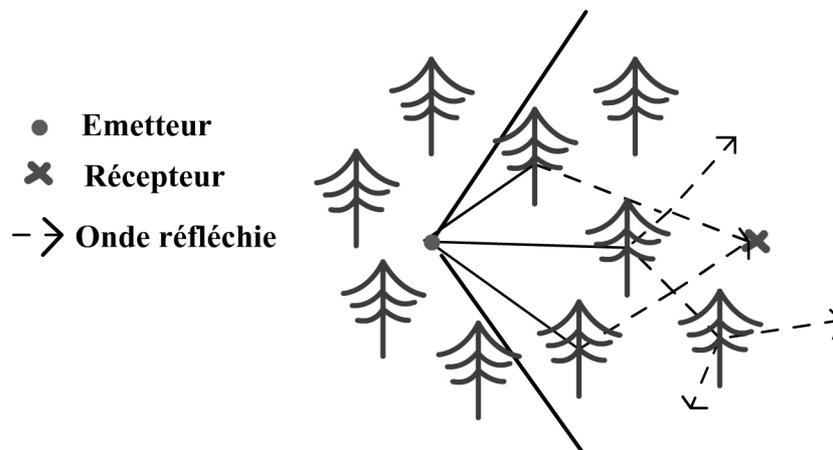
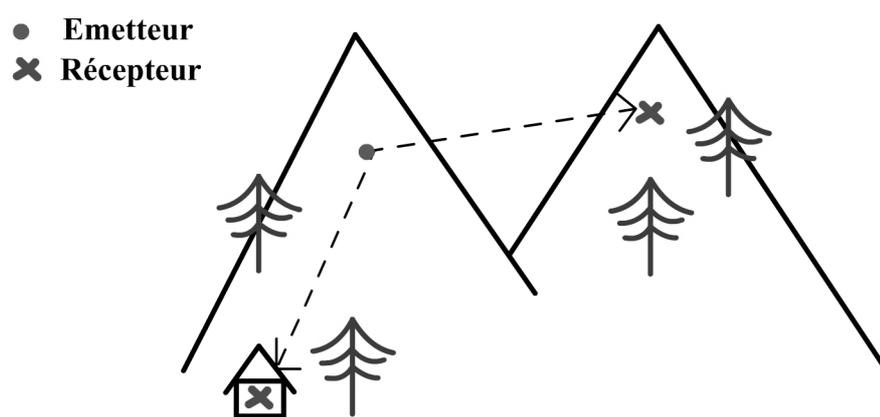


Figure 2 : Propagation du signal dans les forêts

Dans les montagnes, c'est le temps de trajet entre deux points qui est fortement contraignant. Pourtant, la portée du contact visuel est bien améliorée par rapport aux plaines et donc les occasions de communiquer à distance sont rendues plus fréquentes. Le sifflement est alors une solution adéquate permettant de réduire l'effort à fournir pour entrer en contact avec un autre individu que l'on aperçoit ou que l'on situe approximativement.

<sup>20</sup> Elle est généralement réservée au domaine du discours rituel.

<sup>21</sup> La réverbération est connue comme l'ensemble des réflexions des ondes acoustiques d'un son sur les objets du milieu qui l'entoure. La conséquence directe pour un auditeur est de recevoir, non seulement une onde directe mais un ensemble d'ondes réfléchies qui participent à la perception du signal, parfois en la dégradant, et parfois en l'améliorant..



**Figure 3 : Configurations types de communications sifflées dans les montagnes**

## **2.2. Description des environnements naturels**

Donc, la caractéristique commune principale des nombreuses populations ayant adopté les langues sifflées est de vivre dans des zones où la topographie est accidentée et/ou la végétation est dense. Dans certaines, une seule de ces caractéristiques prédomine : par exemple aux îles Canaries, dans les Pyrénées ou dans l'Est de la Turquie, ce sont les vallées encaissées des montagnes qui expliquent l'usage du sifflement. Par contre, en Amazonie ou dans certaines zones de la Papouasie Nouvelle Guinée, c'est la densité et l'omniprésence de la forêt qui le rend utile. Dans d'autres régions tropicales et équatoriales où les langues sifflées sont encore aujourd'hui répandues, ces deux facteurs écologiques se combinent suivant les altitudes (sommet de l'île de la Gomera aux Canaries, montagnes de la zone du Triangle d'Or en Asie du Sud-Est, Sierra Mazatèque au Mexique, montagnes en Papouasie Nouvelle Guinée et Irian Jaya). Dans les paragraphes suivant nous décrirons les comportements humains en lien avec quelques milieux que nous avons visités.

### **2.2.1. Milieux écologiques et usages**

#### **2.2.1.1. L'Amazonie, forêt dense**

La forêt amazonienne est indiscutablement un des écosystèmes les plus denses de la planète, tant en terme de biodiversité que pour la répartition de la végétation. Différentes populations y vivent encore relativement isolées malgré la pression exercée par tous types de trafics conduisant à une déforestation dramatique et à des déplacements de villages. L'Amazonie a la particularité d'être un bassin ayant un dénivelé très lent. Les fleuves qui y coulent sont donc très sinueux ce qui multiplie les occasions lors desquelles les individus sont séparés par un bras d'eau. Sur des territoires proches et non balisés, plusieurs populations parlant des langues différentes se côtoient. Il est donc important de pouvoir identifier les siens rapidement et à distance. D'autre part, dans ce cadre, la chasse et la pêche font partie des sources principales d'approvisionnement en nourriture. Ce sont donc des activités quotidiennes de toutes les populations. Dans ces conditions, le langage sifflé a de nombreuses raisons de se développer. Par exemple il est utile lors des communications d'un côté à l'autre d'un cours d'eau (Photo 1). La communication sur un cours d'eau est bien plus efficace que dans la forêt car le miroir de l'eau favorise la bonne propagation du son. La distance usuelle du sifflement en Amazonie est donc variable en fonction du lieu d'usage et de l'activité pratiquée, mais elle n'est pas aussi considérable que dans les montagnes.

Ask for photo

**Photo 1 : Amazonie péruvienne**

#### **2.2.1.2. Sierra Mazatèque : la forêt dans les montagnes**

La langue mazatèque n'a pas de mot distinct pour « montagne » et « forêt ». Une phrase de notre corpus : « *Certains vivent dans les montagnes, d'autres dans les forêts* », nous a permis de nous en rendre compte

rapidement. Le mot mazatèque qui synthétise ces deux concepts se traduit en espagnol par « *naturaleza* » et témoigne de l'influence de l'écologie d'un lieu sur la langue. En raison de sa latitude, cette chaîne de montagnes située dans le Nord de l'état d'Oaxaca au Mexique est couverte de forêt parfois jusqu'à ses sommets (Photo 2). Ni l'état, ni la population n'ont exercé une déforestation à outrance dans cette zone, contrairement à ce qui s'est passé dans les plus basses terres et a entraîné un changement de mode de vie et, très souvent, une forte migration vers la capitale ou vers les Etats-Unis (en particulier pour la population Mixtèque qui était répartie traditionnellement dans les basses terres (voir carte Figure 4)). La langue mazatèque est parlée dans une vaste région où certains villages sont encore très isolés, sans route pour y accéder. Cette situation explique le fait que plus de 14 variantes de la langue se soient développées. Le centre principal de la sierra Mazatèque est la ville de Huautla de Jimenez. Celle-ci est sur un flanc de montagne et ses habitants parviennent toujours à communiquer en sifflements entre les maisons, par quartiers (Photo 3). La portée usuelle du sifflement dépasse rarement 700 mètres. Il est plus courant d'utiliser la langue sifflée à courte et moyenne distance. Dans certaines vallées du district d'Eloxochitlan pourtant, nous avons rencontré des siffleurs utilisant couramment leurs doigts pour atteindre de plus grandes distances en sifflant (Photo 16). Nous avons observé que les femmes ne sifflent pas aussi volontiers que les hommes mais beaucoup d'entre elles maîtrisent la technique. Il semble également que dans la société Mazatèque, le sifflement soit peu pratiqué par les individus les plus âgés. Ces derniers expliquent qu'à leur âge on ne siffle plus. Pourtant nous avons souvent constaté qu'ils en maîtrisaient toujours la technique.

Une exclusivité de cette région est l'usage courant de la version sifflée en ville sur le marché (Photo 4), à la gare routière ou juste après la tombée de la nuit. C'est le seul cas de langue sifflée citadine que nous avons rencontré. Cela est dû à notre avis au fait que la ville de Huautla de Jimenez est en grande majorité Mazatèque et n'a pas perdu cette identité forte.



Figure 4 : Carte des groupes culturels de la région d'Oaxaca. (source : exposition Paris la Villette 2002 *Indiens Chiapas=>Mexico=>Californie*)

Ask for photo

**Photo 2 : Vue générale de la sierra Mazatèque d'un sommet.**

Ask for photo

**Photo 3 : Un quartier de la ville de Huautla de Jimenez**

Ask for photo

**Photo 4 : Marché de la ville de Huautla de Jimenez : lieu de communication sifflée dans le bruit**

### **2.2.1.3. Les contreforts de l'Himalaya : du Yunnan au Nord de la Thaïlande.**

Plusieurs dizaines de minorités différentes ayant chacune une langue propre vivent dans les régions montagneuses du Yunnan Chinois, du nord de la Thaïlande, du Laos, du Vietnam et de la Birmanie. D'après les informations que nous avons pu réunir au Laos, en Thaïlande et au Yunnan en 2004, plusieurs d'entre elles utilisent les langues sifflées. Certaines de manière extensive et d'autres de manière limitée. Cette vaste région est considérée comme un des berceaux de nombreuses populations asiatiques qui ont ensuite émigré vers le sud (comme les Thaïs par exemple). Dans les plus basses terres, les villages sont au milieu des forêts de montagne.

Nous avons visité plusieurs de ces communautés linguistiques et avons pu initier et réaliser un travail avec trois d'entre elles. Parmi celles ci, deux ont une langue sifflée utilisant la feuille d'un arbre: les Akhas (Photo 21) et les Hmongs. Dans cette zone la version avec la feuille est particulièrement prisée pour la séduction (échanges amoureux) car elle est réputée comme plus mélodique. Les sons passent facilement à travers les murs des maisons locales (Photo 6), il est possible de venir courtiser discrètement la nuit sans être repéré par les parents de la jeune fille ou alors d'exercer sa poésie lors d'une déclaration amicale alors que l'on passe près de chez quelqu'un (com. pers. Jupoh 2004). Autour des villages où nous avons séjourné, la montagne est presque toujours couverte de forêt. Le sifflement avec feuille qui porte aisément jusqu'à 500m est donc largement suffisant.

Ask for photo

**Photo 5 : Répartition actuelle de la population Akha et photo d'une femme Akha sifflant avec une feuille**

Ask for photo

**Photo 6 : Maisons traditionnelles, sur le flanc d'une montagne couverte de forêt**

En montant vers l'Himalaya, la végétation est plus rare et les montagnes deviennent plus escarpées. Les formes sifflées y sont donc utilisées à plus grande distance. Par exemple, dans un village au Laos, à la frontière chinoise, nous avons rencontré un siffleur utilisant ses doigts pour produire un sifflement plus puissant. En général les communications sifflées se font lors des travaux aux champs (riz, canne à sucre), pour la chasse ou dans les villages.

#### **2.2.1.4. Turquie : les vallées descendant des hauts plateaux vers la Mer Noire**

La langue sifflée turque continue d'être pratiquée dans la région de la petite ville de Görele, non loin de la ville de Trabzon à 35 km à l'intérieur des terres par rapport à la Mer Noire. Les vallées qui descendent des hauts plateaux abritent de nombreuses habitations encore isolées. En 2003, les accès aux villages où nous avons rencontré quelques siffleurs n'étaient pas encore goudronnés, mais des travaux étaient en cours, témoins des changements progressifs dans la région. Les habitations sont réparties sur les flancs des montagnes et sont relativement éloignées les unes des autres. Les vallées sont étroites et un torrent coule au fond de chacune faisant un bruit de fond non négligeable pour les habitations environnantes (voir expérience du § 2.3.5.2.1). Les habitants utilisent donc le turc sifflé à la fois pour se parler entre les maisons (Photo 7), dans les champs et dans les alpages où les bergers montent à partir de juin. La distance de communication est comprise entre 100 et 2000 m en dehors du village. Mais il peut atteindre parfois des distances bien plus grandes. Cependant, ces occasions se présentent surtout sur les hauts plateaux quand le paysage des vallées encaissées se dégage (Photo 8). Toutes les techniques connues sont employées en Turquie.



**Photo 7 : Village de Kusköy vu du fond de la vallée, au-dessus du torrent. Le relief est très accidenté**



**Photo 8 : Vue sur les hauts plateaux depuis les montagnes en bordure de la mer Noire**

### **2.2.1.5. Le village béarnais d'Aas dans la vallée d'Ossau : la biodiversité maintenue**

Le village d'Aas a longtemps maintenu le sifflement béarnais qui a aujourd'hui disparu. La topographie et l'isolement des vallées d'Ossau et d'Aspe en sont les raisons principales. Dans ce contexte, les vallées font office de guide d'ondes pour le signal sifflé (Photo 9). Les bergers montaient l'été dans les alpages et restaient en contact entre eux grâce au *sifflet*. Ils pouvaient également ainsi garder des nouvelles du village. Il est intéressant de remarquer que, conformément aux observations de Busnel (comm pers 2005) et de Arripe aucun autre village de la région n'utilisait le sifflement. Il semble que de nombreux habitants des villages voisins en ignoraient l'existence dans les années 60.



**Photo 9 : Vue des alpages au dessus des vallées d'Aspe et d'Ossau**

## **2.2.1.6. Les îles montagneuses: une isolation salvatrice**

### **2.2.1.6.1. La Gomera : de multiples contextes**

L'île de la Gomera est la deuxième île la plus petite des Canaries. On dit que l'intérieur des terres n'a jamais vraiment été conquis par les Espagnols. Elle est également restée isolée du tourisme de masse jusqu'à aujourd'hui. C'est une île qui a une topographie extrêmement accidentée jusqu'à son sommet où l'une des deux seules forêts de l'ère tertiaire de la planète subsiste encore grâce à un microclimat de brumes accrochées fréquemment par les pourtours montagneux. Contrairement à ce qui s'est passé dans les autres îles, le sifflement n'a pas disparu de la Gomera, principalement à cause des services qu'il a toujours rendu aux habitants. Suivant les régions de l'île, des « *dialectes sifflés* » se sont même développés. Il s'agit d'un phénomène unique que nous pensons dû en partie aux différences de topographie, de végétation et de climat. Ainsi, dans la région du village de Chipude, situé en altitude, le climat est souvent brumeux et les reliefs sont moins accentués que dans les régions des villages d'Agulo où de Vallehermoso. Il se trouve qu'à Chipude la pratique du sifflement est traditionnellement plus grave que dans les deux autres régions. Dans le monde entier, c'est à la Gomera que les plus grandes portées de discussions sifflées ont été observées. Il est fréquent de dépasser un kilomètre et il arrive de parler à plus de 8 km, quand le climat et la vallée sont favorables.



**Photo 10 : Region de Chipude**

**Zone de paturages pour les moutons à proximité de la forêt du Tertiaire, maintenue au sommet de l'île par un microclimat de brumes passagères**



**Photo 11 : Région de Vallehermoso les zones de récolte sont réparties sur les flancs les moins abrupts des vallées encaissées**

#### **2.2.1.6.2. Le village d'Antia sur l'île grecque d'Eubée : une petite vallée fertile dans un environnement sec**

Le village d'Antia, dans le sud de l'île d'Eubée est encore très isolé. Il est le seul dans la région à posséder une langue sifflée. Il est situé dans une combe en pente douce où coule une rivière permettant à la végétation de se développer alors que les alentours sont arides. Il est dominé par le sommet de l'île. La langue sifflée grecque est à la fois pratiquée dans le village et dans les environs où les bergers se déplacent souvent à pied ou à dos de mule. La portée peut atteindre 2 km. Certains habitants pratiquent la langue sifflée avec les doigts mais cette technique puissante est plus rare que la technique labiodentale qui suffit pour atteindre les alentours du village et pour communiquer entre les maisons.



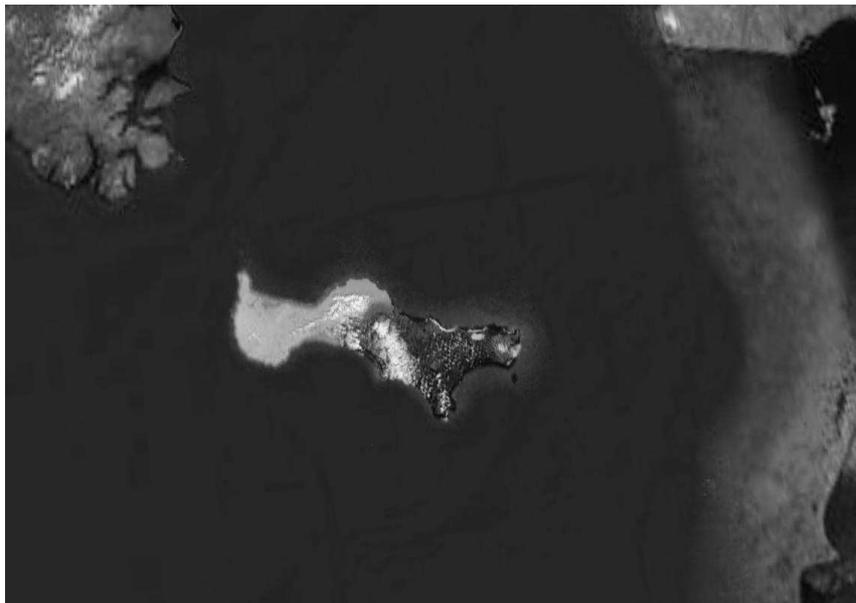
**Photo 12 : vue générale de la vallée et du village d'Antia sur le flanc gauche**



**Photo 13 : Les maisons du village permettent une communication à moyenne distance avec la technique labiodentale**

### 2.2.1.6.3. St Lawrence : une population Yupik isolée entre deux continents

Nous ne sommes pas encore allés sur l'île St Lawrence dans le détroit de Béring, mais sa situation symbolique et extrêmement isolée permet de comprendre pourquoi elle a permis à ses habitants de conserver leur usage de la langue sifflée et d'être depuis juin 2005 la seule langue sifflée en Amérique du Nord et en Sibérie connue du public scientifique.



**Photo 14 : Image satellite de l'île St Lawrence dans le détroit de Béring**

**L'Alaska est à droite(Est) et la Sibérie est à gauche (Ouest). Cette île, de la taille de la Sardaigne possède plusieurs massifs montagneux (Source : GoogleEarth)**

## **2.2.2. Conclusion**

L'étirement spatial des besoins de communication donne un rôle important au milieu acoustique qui sert de canal de transmission. Le sifflement est clairement justifié, car, pour se parler dans de tels milieux les interlocuteurs ont besoin de signaux clairs, portant loin et qui ne soient pas perturbés par le bruit ambiant de la nature. Il y a un lien direct entre l'écologie du lieu, sa topographie et la distance habituelle de pratique qui influence en partie la bande de fréquence du signal sifflé utilisé. Mais la variété des usages, malgré les similarités de milieux montre aussi l'influence de la culture locale qui réserve un rôle au sifflement qu'elle détermine en partie. Dès lors, les pratiques linguistiques sifflées sont spécifiées par une histoire de *couplage* entre la nature et une société locale. C'est un équilibre qui bien souvent se brise quand le monde rural qui l'a fait naître disparaît, soit à cause du départ des populations soit à cause du changement de culture, soit à cause de la destruction du milieu.

## **2.3. Etude bioacoustique**

### **2.3.1. Problématique**

Les différents milieux écologiques que nous avons décrits exercent un filtre naturel sur le signal acoustique, auquel s'ajoute le bruit de fond ambiant. Dans ces conditions, les formes sifflées des langues sont très efficaces pour étendre la portée de la parole locale. Cet aspect est un de leurs principaux intérêts scientifiques<sup>22</sup>. Il convient donc d'en analyser les caractéristiques bioacoustique afin de comprendre ce qui les rend particulièrement adaptées à des usages préfigurant les télécommunications modernes<sup>23</sup>.

### **2.3.2. Production du sifflement**

#### **2.3.2.1. Techniques de production**

Les techniques de sifflement varient en fonction de la distance de communication et des usages locaux: un ou plusieurs doigts dans la bouche (longues distances: de 500 m à plusieurs kilomètres<sup>24</sup>), rétroflexion de la langue (moyenne distance : de 200 à 500m) ou sifflement bilabiale (courtes distances jusqu'à 200 m).

##### **2.3.2.1.1. Techniques de longue distance**

Tout un ensemble de techniques est employé pour siffler à une grande distance. Le choix de l'une ou l'autre dépend des pratiques courantes locales, du goût et des capacités de chacun.

#### **Techniques avec un doigt**

Le doigt vient se poser sur la langue ou sous la langue, au choix. N'importe quel doigt peut être utilisé et il peut être droit ou courbé (Photo 15). La technique du doigt courbé est la technique la plus répandue sur l'île de la Gomera.

**Photo 15 : Siffleur Gomero utilisant la technique la plus commune de l'île  
Le doigt courbé et la main en porte-voix**

---

<sup>22</sup> Deux autres aspects fondamentaux qui apparaîtront lors de l'analyse bioacoustique concernent la capacité des humains à encoder plusieurs types de structure de langues ainsi qu'à détecter et décoder le message. Chacun de ces aspects fait l'objet d'une analyse détaillée dans les parties suivantes de cette thèse.

<sup>23</sup> Cet aspect a des conséquences importantes sur l'organisation de la vie en société, à échelle locale pour les villages utilisant les langues sifflées ou les langages tambourinés. Nous fournissons en Annexe **Erreur ! Source du renvoi introuvable.** une analyse des liens historiques et physiques de ces systèmes avec les télécommunications modernes.

<sup>24</sup> Valeurs de distances valables pour des milieux avec un bruit ambiant de 40 dB.

## **Techniques avec deux doigts**

Toutes les combinaisons de doigts sont possibles. Les plus fréquemment utilisées sont les techniques utilisant le pouce et l'index d'une même main ou celle utilisant les deux petits doigts (auriculaires). A chaque fois, les deux doigts forment un V appuyant sur la langue. Le sifflement produit est très puissant. Cette technique est utilisée par certains siffleurs Grecs, Gomers, les Turcs et Mazatèques (Photo 16), quand ils sont isolés.

**Photo 16 : Technique de sifflement Mazatèque avec deux doigts se rejoignant en un V**

## **Lèvre intérieure tirée**

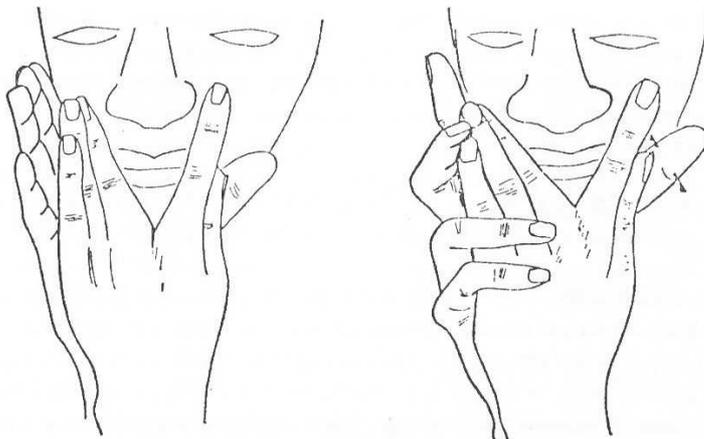
Cette technique qui consiste à siffler en tirant sur la lèvre inférieure et en aspirant simultanément (Photo 17) n'avait jamais été observée auparavant. Elle est en usage dans toute l'Amérique latine chez les plus jeunes générations et pour tous types de sifflements, linguistiques ou non.

**Photo 17 : Technique de sifflement par pression sur la lèvre inférieure tirée**

**L'air est inspiré alors que toutes les autres techniques utilisent l'expiration de l'air**

## **Sifflement par vibration de la membrane entre le majeur et l'index.**

Nous ne l'avons jamais observée personnellement mais elle serait utilisée encore actuellement chez les Ari en Ethiopie (com. pers. Fournel 2004) ou chez les Gaviaõ en Amazonie (com. pers. Moore 2003). Elle a également été décrite chez les Banen du Cameroun par Dugast (1976) (Figure 5).



**Figure 5 : Position des doigts pour le sifflement chez les Banens (ou Ndiki) du Cameroun par Dugast (1976, p713)**

**Dans cette population, la cavité créée entre les deux mains permet en partie de régler la fréquence du sifflement**

## **Variantes**

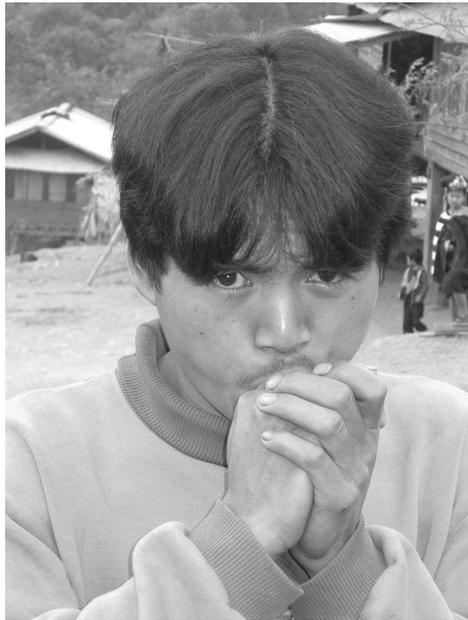
De nombreuses autres méthodes ont été observées dans le passé, la plupart sont des variantes de celles que nous avons observées. Nous en citerons trois en exemple :

## Bout de la langue retourné contre les dents (rétroflexée) et un doigt posé sur la langue

C'est une combinaison avec une technique de moyenne distance qui produit un sifflement très puissant car il combine l'effet d'accentuation de la vibration créé par la cavité entre la langue et les dents de la mâchoire inférieure et par le doigt. Nous ne l'avons observé qu'en Turquie (en 2004) où le sifflement moyen est le plus aigu de tous ceux que nous avons rencontré.

## Sifflement dans le creux des mains

Le sifflement dans le creux des mains peut permettre de parler à condition de maîtriser une partie de l'articulation en changeant la taille de la cavité et de l'ouverture de l'air qui s'échappe. Il est plus couramment utilisé pour imiter certains types d'oiseaux que pour parler. Nous l'avons rencontré à un seul endroit : au Laos dans un village Akha (Photo 18), mais cette technique n'est pas très courante dans cette population. Le locuteur nous a expliqué que son père qui lui a enseigné cette technique pouvait parler plusieurs langues sifflées car il avait vécu dans plusieurs villages lorsqu'il avait quitté le Vietnam. Une technique similaire a été observée en Ethiopie chez les Benčnon (Wedekind 1981).



**Photo 18 : Méthode créant une cavité en dehors de la bouche**

**Elle est plutôt rare à observer pour la parole sifflée car le contrôle de l'articulation demande une très grande dextérité**

## Sifflement dans le creux du coude

Toutes les parties du corps accessible facilement avec la bouche peuvent être mises à profit. Dugast (1955) cite l'exemple d'un locuteur Banen lui ayant montré qu'il pouvait exploiter la membrane de peau dans le creux du coude de la même manière que celle entre l'index et le majeur de la main pour produire un son plus grave.

### 2.3.2.1.2. Techniques de moyenne et courte distance

#### **Bout de la langue retourné contre les dents (langue retroflexée)**

Cette technique est très répandue parmi les langues tonales comme parmi les langues non tonales. Les mains restent libres. Elle permet un sifflement très puissant. De plus la maîtrise de l'articulation est plus proche de celle utilisée pour la voix parlée (Photo 19 et Photo 20).

**Photo 19 : Siffleur Grec utilisant la technique de la langue courbée**

**La langue formant une cavité en venant buter contre les dents e la mâchoire inférieure.**

**Photo 20 : Même technique que précédemment aussi appelée langue rétroflexée**

#### **Sifflement avec feuille**

La technique avec la feuille (Photo 21) permet d'obtenir un sifflement de nature différente car la vibration de cet *outil* provoque dans le signal d'un ensemble d'harmoniques fréquentielles. Nous ne l'avons rencontré que dans le cas de langues sifflées asiatiques ayant de nombreux registres de tons et des contours de tons. Le son porte moins loin que les sifflements produits avec les doigts pour parler à longue distance, mais il est suffisamment efficace pour atteindre 300 m dans une forêt et plus de 500 m dans les milieux montagneux. Nous l'avons observé pour deux langues asiatiques pour lesquelles il ne dépasse jamais une fréquence maximum légèrement inférieure à 1500 Hz.

**Photo 21 : Technique de la feuille sifflée exécutée par une femme Akha**

#### **Sifflement bilabial**

Répandue parmi les chasseurs communiquant à courte distance, cette technique permet de réaliser parfois des occlusions avec les lèvres. Le sifflement bilabial est très rarement utilisé dans les pays utilisant des langues non tonales, tout d'abord car il ne permet pas d'atteindre une puissance suffisante mais aussi car la bande de fréquence qu'il permet de couvrir n'est pas aussi large que dans les autres techniques. Or les sifflements des langues non tonales couvrent souvent un intervalle de fréquence plus large que ceux des langues transposant l'intonation ou les tons (voir 2.3.4).

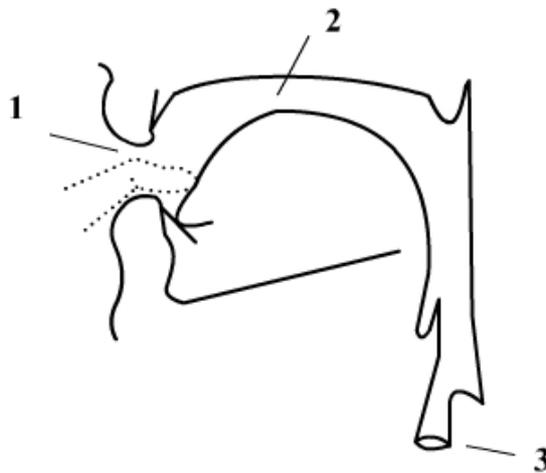


**Photo 22 : technique bien connue du sifflement bilabiale**

## **2.3.2.2. Particularités de la source**

### **2.3.2.2.1. Mode d'émission**

Le sifflement est produit par un effet choc du souffle d'air comprimé à l'intérieur de la cavité de la bouche ou juste à sa sortie. Il ne fait pas intervenir la vibration des cordes vocales (point 3, Figure 6), le sifflement est d'ailleurs incompatible avec ce phénomène. La vibration sonore obtenue par le passage de flux d'air concentré est un signal quasi périodique de type sinusoïdal. Plus le courant d'air expulsé (ou inspiré) est rapide, plus le son obtenu est aigu et puissant. L'effort exercé par la cage thoracique n'est pas le seul élément intervenant pour augmenter la puissance du son. Le réglage de la taille des trous minimaux par lesquels l'air passe lorsqu'il est expulsé intervient également (points 1 et 2, Figure 6). Quand les mâchoires sont fixées par les lèvres serrées et éventuellement le doigt (point 1, Figure 6), le trou de la bouche est stable. La mâchoire inférieure ne fait plus qu'esquisser l'articulation des mots. Les mouvements de la langue et de la glotte sont les autres principaux éléments servant à régler le son. Dans certaines limites, cette articulation peut être comparée à celle utilisée pour la forme parlée équivalente, elle permet de faire varier le volume de la cavité de résonance et de la pression de l'air expulsé par les poumons afin de moduler la fréquence du phénomène bioacoustique obtenu. Tout un ensemble de processus physiologiques et acoustiques s'équilibre dans une cohérence informationnelle commune avec la parole classique. L'objectif avoué est d'obtenir une intelligibilité optimale de la part de l'interlocuteur.



**Figure 6: Articulation des sifflements**

### **2.3.2.2.2. Effets sur les modulations du signal**

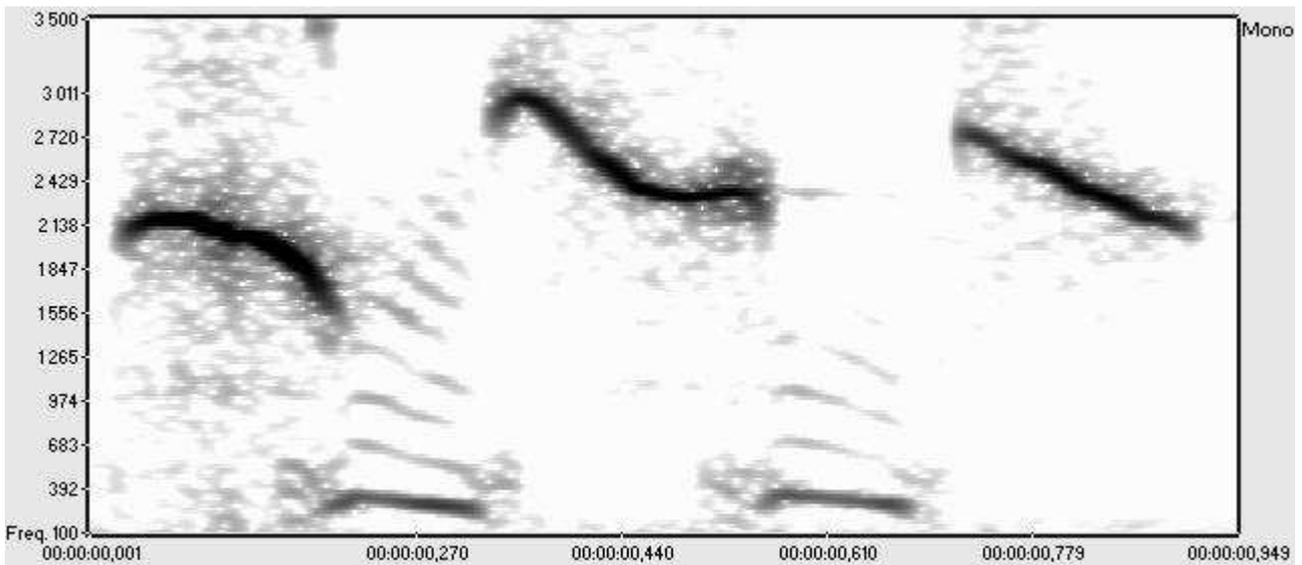
Cette description nous a permis de conclure que les facteurs principaux intervenant lors de la production du sifflement ainsi que leurs effets sont les suivants:

**Tableau 6 : Facteurs principaux intervenants dans la production du sifflement**

Facteur de contrôle	Effet(s)
Mouvements de la mâchoire inférieure	Articulation souvent similaire à la parole, en plus limité
Rétroflexion de la langue, intromission d'un doigt	Point fixe (ou relativement fixe) pour contrôler les mouvements de la langue, augmentation de la puissance du flux d'air résonnant
Mouvement de la langue	Contrôle du flux d'air et des types de résonance suivant une logique proche de celle de la parole. Contrôle du flux d'air allant jusqu'à l'occlusion
Coups de glotte, à la manière des ventriloques	Contrôle du flux d'air: pour certaines occlusions labiales qui ne peuvent pas être réalisées en sifflement et pour contrôler la fréquence du sifflement
Puissance du souffle	Pour s'adapter à la distance de communication et rendre l'amplitude du signal plus intense

Les coups de glotte et de la langue sont utilisés de manière complémentaire par de nombreux siffleurs. Ils permettent d'obtenir des modifications similaires du signal sifflé en terme de fréquence. Il semblerait que la stratégie d'utiliser plus fréquemment la glotte soit particulièrement employée par les locuteurs ayant une dent manquante car le contrôle de la fréquence est alors plus aisé loin de cette perturbation du flux d'air. On peut comparer ces observations aux stratégies employées chez des flûtistes professionnels. La technique la plus académiquement reconnue comme correcte est celle qui consiste à moduler un souffle régulier issu des poumons avec uniquement les mouvements de la langue pour contrôler les notes. Les coups de glotte étant bannis en théorie, ils sont cependant pratiqués très régulièrement.

Cet aspect de la prononciation a été souligné par Classe (1956) et nous pouvons le confirmer par le fait que sur de nombreux enregistrements réalisés en champ proche, les coups de glotte sont présents sur le sonagramme Figure 7.



**Figure 7 : Début de phrase turque : les trois syllabes « Kom -yun-köp » se succèdent  
On observe que la labiale [m] comme le [k] donnent lieu à un coup de glotte visible en champ proche.**

Manifestement, la glotte contrôle de la puissance du souffle en permettant de maîtriser la compression d'air et son relâchement. La différence de forme acoustique entre le coup de glotte du [m] et du [k] est un bon indice permettant de l'observer. Une manière de le vérifier serait d'utiliser la méthode laryngoscopique comme dans des études sur la voix (Esling et al, 1998).

## **Contrôle des voyelles ou des tons**

Les voyelles sont produites en reproduisant approximativement la forme de la bouche de la prononciation parlée. Le contrôle des tons est obtenu en contrôlant les zones de résonance de la bouche. Un son plus aigu est réalisé vers le devant du palais et un son plus grave plus en profondeur vers l'arrière.

### **2.3.3. Bandes de fréquences typiques des sifflements humains**

Si l'on considère toutes les langues et toutes les méthodes de sifflements, on obtient une bande de fréquence générale de la fréquence fondamentale de sifflement allant de 800 Hz à 4000 Hz, ce qui reste étroit comparé à la bande de fréquence utilisée pour la parole (100 à 16000 Hz) où même à la bande passante du téléphone (400 à 4000Hz). Mais pour une phrase donnée couverte par la fréquence fondamentale à l'émission est environ d'une largeur de 3000 Hz dans les langues non tonales et moins de 1000 Hz avec des variations significatives suivant la structure de la langue et la technique de sifflement utilisée (labial, labiodental, avec doigt(s), avec feuille). La technique de la feuille est un peu particulière car elle produit un sifflement avec de nombreuses harmoniques ayant une amplitude non négligeable, par conséquent elle est parfois qualifiée de voix sifflée<sup>25</sup> (comm. pers. Tran Quan Hai 2005).

#### **2.3.3.1. Facteurs principaux de variabilité de la bande de fréquence**

##### **2.3.3.1.1. Structure de la langue**

###### **Généralités**

Alors que le registre fréquentiel des valeurs extrêmes de la fréquence fondamentale des formes sifflées des langues non tonales peut s'étendre sur plus de 2 octaves (0.9 à 3.8 kHz à la Gomera en espagnol, 1.2 à 3.5 kHz à Aas en béarnais (Busnel & Classe, 1976), 1.4 à 4 kHz à Kusköy en turc, 1.2 à 3.4 kHz à Antia en grec); celui des langues à tons est plus étroit (1 à 3 kHz en langue mazatèque en tenant compte de plusieurs distances de communication (4 tons, Mexique), 1 à 2 KHz en surui (2 tons, Amazonie), 0.8 à 1.5 kHz en akha (3 niveau de tons, technique exclusivement avec une feuille), 0.9 à 1.5 kHz en hmong (8 tons, technique

---

<sup>25</sup> Ce type de voix sifflée est différent de celui obtenu par la technique vocale parfois appelée « sifflet laryngé » pour laquelle l'onde sonore est produite par une fermeture de la glotte imparfaite sans vibration des cordes vocales « laissant une fente étroite au travers de laquelle passe l'air en sifflant comme dans un biseau » (Zemp 1996).

de la feuille exclusivement)). Dans une même phrase, la bande de fréquence dépasse rarement un octave. Nous n'avons jamais observé de dépassement de cette limite en langue sifflée tonale ou en chepang ou même en turc. C'est très rare en grec et, par contre, certains siffleurs de silbo espagnol le font couramment. Nous verrons dans la partie « Intelligibilité » (Chapitre 4) que des considérations sur la perception de la Hauteur Fondamentale (HF) permettent de donner une explication à la limite d'un octave.

## **Implications pour la compréhension de la transposition sifflée**

### **Différentes parties de la voix ?**

Nous avons vu dans l'Historique qu'il était reconnu que les parties les moins modulées des langues sifflées issues d'une langue sans tons transposent le timbre des voyelles en des hauteurs différentes alors que pour les langues tonales elles transposent plutôt les tons porté par les voyelles. Les observations que nous venons de faire sur la largeur de bande de fréquence corroborent cette affirmation car le spectre d'une voyelle est constitué d'harmoniques plus aiguës que la fréquence fondamentale. Etant donnée la structure de la voix humaine, les fréquences harmoniques sont en effet des multiples de la fréquence fondamentale. Lorsque cette dernière varie légèrement, plus les harmoniques sont élevées plus l'étendue de leur variation fréquentielle correspondante est grande.

### **Enseignement pour les langues non tonales**

La grande variabilité des voyelles chez un même siffleur est expliquée par son articulation qui couvre une zone de résonance dans la bouche. Cette zone ciblée n'est pas toujours atteinte de la même manière. D'autre part, l'ampleur différente de la variation fréquentielle entre les distinctions vocaliques et les distinctions d'intonation ou d'accent (liées au Fo comme les tons) suggère qu'il est possible, dans certaines limites de garder la répartition étagée des voyelles sifflées tout en marquant l'accentuation. Nous verrons par la suite que c'est effectivement le cas en grec, en silbo et en turc.

### **Enseignement pour les langues tonales**

Dans beaucoup de phrases de langues tonales, l'étendue de la bande des fréquences sifflées n'excède pas 800 Hz, c'est à dire qu'elle est en dessous d'un octave et duplique la bande de fréquence des tons de la parole en l'élargissant un peu. D'autre part, celle des sifflements réalisés avec une feuille est plus stable que celle des sifflements sans outils autres que les doigts. En effet, les feuilles semi-rigides utilisées (citronnier ou café par exemple) disposent d'un mode de vibration plus fixe que le sifflement contrôlé par la bouche. Ainsi, la langue akha à trois niveaux de tons et la langue hmong à 5 niveaux de tons ont la même largeur de bande de fréquence fondamentale. Ces résultats sont des données nouvelles qui peuvent alimenter le débat sur la nature fixe ou non de l'étendue des hauteurs tonales des langues à tons. Par exemple, Pike (1948) a observé - et de nombreux auteurs ont approuvé son point de vue - que toutes les langues à tons utilisent le même intervalle de hauteurs tonales qu'elles subdivisent en autant de registres nécessaires pour leur système tonal. Maddieson (1972) avait initialement proposé une autre approche qui considérait que les tons d'une langue à

trois tons se répartiraient comme les trois tons centraux d'une langue à cinq tons. Il fait ainsi de la valeur centrale de l'étendue tonale un bassin attracteur des tons extrêmes. Anderson (1978) a remarqué que les systèmes tonals asiatiques semblent peu "marqués" par rapport aux niveaux tonals centraux. Si l'on considère nos données concernant les deux langues asiatiques akha et hmong, les observations de Pike et Anderson sont validées. Par extension, il semble bien que les tons des langues ne soient pas espacés de la même manière pour chaque langue.

Maddieson a soulevé de nombreuses autres questions sur la répartition des tons à l'intérieur de l'intervalle qu'ils couvrent. En particulier sur les limites fréquentielles dans lesquelles chaque ton peut osciller.

Une pratique esthétique d'*ornementation* du ton en langue sifflée akha peut être intéressante de ce point de vue. Le sonagramme de la Figure 8 montre le résultat de cette technique. En prenant les mesures les plus larges possibles on observe que les modulations du ton bas ou du ton haut ont une oscillation maximum de 240Hz. L'intensité de la fréquence n'est élevée que sur un intervalle de 200 Hz. Or, dans cette phrase d'une langue à 3 niveaux de tons modulés, les tons sont justement écartés de 200Hz. Le fait que les oscillations ne dépassent pas la fréquence médiane qui sépare l'intervalle entre deux tons est un phénomène stable sur tous nos enregistrements qui a forcément une raison perceptive forte. Il faut ajouter que l'oscillation rapide autour d'une même fréquence participe sûrement à la perception privilégiée de cette fréquence.

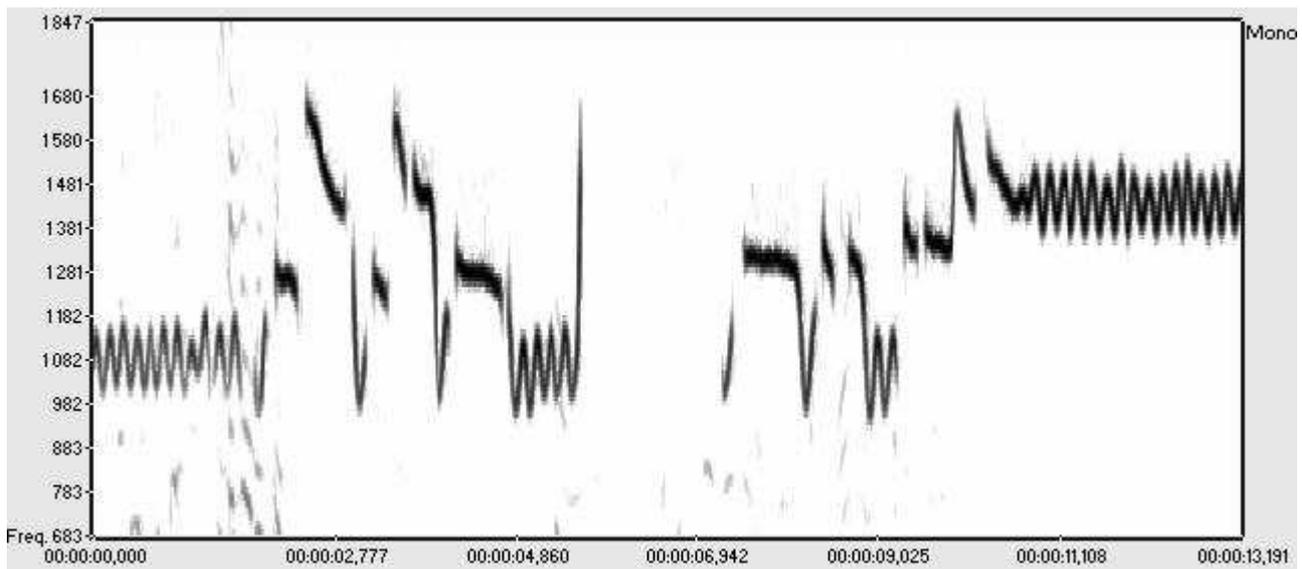


Figure 8 : Fréquence fondamentale d'une phrase akha

Elle est ornementée sur les tons bas et, en fin d'extrait, sur un ton haut (siffleuse de la Photo 21)

### 2.3.3.1.2. Variabilité en fonction de la technique

#### Influence de la technique

Le facteur le plus évident qui influence la hauteur de la fréquence fondamentale est la technique de sifflement choisie par le siffleur. Nous avons vu que ce choix est fortement corrélé avec la distance à laquelle se trouve son interlocuteur. Plus l'interlocuteur est éloigné plus le siffleur aura tendance à siffler fort pour atteindre un niveau d'amplitude qui permet aux phrases de rester clairement intelligibles. Or, par un phénomène de bioacoustique non linéaire et non négligeable, plus le souffle est fort, plus la fréquence issue

de l'organe siffleur est élevée. Nous avons mesuré<sup>26</sup> une augmentation de la fréquence de 100 Hz à 250 Hz<sup>27</sup> entre un sifflement réalisé pour atteindre 100 m et un sifflement effectué pour atteindre une distance de 500 m. Pour augmenter la pression de l'air et donc la puissance du sifflement, l'effort à fournir est réduit par l'utilisation d'un doigt ou par la création d'une cavité entre la langue et les dents.

### **Conséquences de l'aspect relatif des hauteurs fréquentielles**

L'encodage de l'information sifflée est en grande partie réalisé par des valeurs relatives. Ce phénomène souligne que de nombreux aspects sonores de la parole humaine sont en fait relatifs. Cette propriété explique la souplesse du système de communication humaine. C'est la raison principale pour laquelle il est possible de convertir la hauteur parlée en hauteur sifflée. C'est aussi en vertu de cette caractéristique que les voix des enfants, des femmes ou des hommes peuvent porter les mêmes informations tout en ayant des spectres de paroles différents (c'est un des facteurs qui pose des problèmes à la théorisation phonétique).

Pour le mazatèque, Cowan (1948) a expliqué que les siffleurs se donnaient une clef de hauteur, l'équivalent du « la » des musiciens qui permet d'accorder les instruments avant de jouer une partition. Nous pensons que, même pour des langues tonales, cette fixation de la hauteur de communication est plutôt donnée par la distance de communication. Par contre, une fois la phrase débutée le siffleur restera dans les limites que nous avons définies ci dessus (un octave).

Dans le cas des langues non tonales, l'appellatif « aah » ou « ooh » qui précède certaines phrases de béarnais sifflé ou du « Silbo Gomero » de la région de Chipude est essentiellement réalisé pour attirer l'attention. De manière similaire à ce qui se passe en mazatèque, certains éléments de l'information acoustique contenue dans la première phrase d'un dialogue permettront à l'auditeur d'évaluer la distance de communication et donc d'adapter la puissance et la fréquence de sa réponse. Dans certains cas, les limites de 3 octaves qui sont respectées sont liées à des considérations de production (Silbo espagnol)

### **2.3.4. Analyse du spectre de fréquences des sifflements en milieu ambiant**

Le spectre de fréquences du sifflement est caractéristique d'un signal sinusoïdal presque pur. A courte distance on peut observer sur sonagramme (Figure 9) plusieurs harmoniques de la fréquence fondamentale<sup>28</sup>. Dans un sifflement, la fréquence fondamentale est élevée, c'est pourquoi les harmoniques qui en sont des multiples directs sont largement plus espacées que celles de la voix parlée (Figure 9). On observe que, dans toutes les techniques sauf celle de la feuille, l'intensité présente dans les harmoniques est rapidement atténuée. Alors qu'à courte distance plusieurs harmoniques sont audibles, dès la moyenne distance (150 m) seule la fondamentale émerge du bruit de fond (Figure 10).

---

<sup>26</sup> Mesure réalisée sur la langue mazatèque dans un milieu montagneux peu boisé.

<sup>27</sup> Sachant qu'il y a des différences entre les phrases et les locuteurs.

<sup>28</sup> Pour un enregistrement à quelques mètres du siffleur, on observe rarement plus de 3 harmoniques clairement visibles sur spectrogramme. Seul le sifflement avec feuille fait exception.

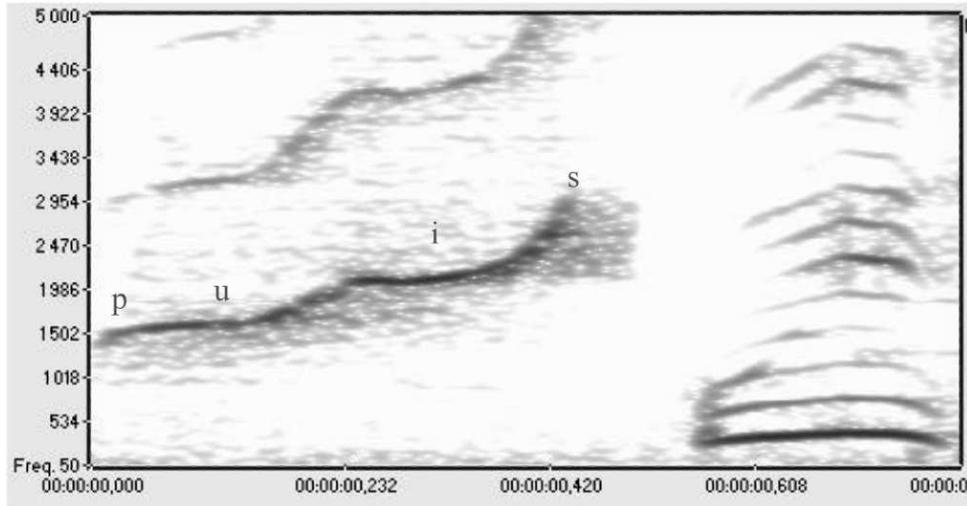


Figure 9 : Syllabe /puis/ en grec sifflé puis parlé. Les éléments de la voix qui sont transposés sont indiqués (l'explication complète est fournie au §4.3.1)

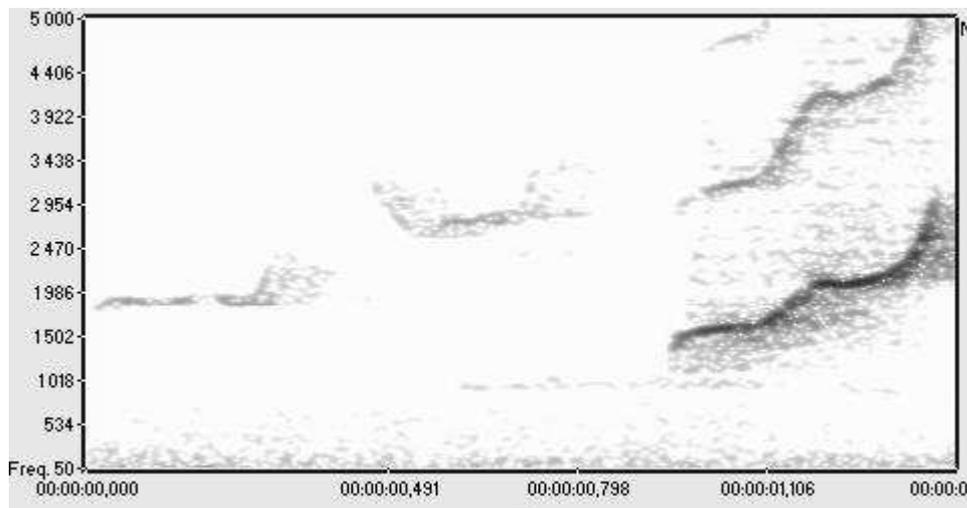


Figure 10 : Conversation grecque sifflée

Le premier interlocuteur est situé à 150 m. Le second (avec deux harmoniques pour la même syllabe /puis /) est situé à 10 m de la prise de sons

### 2.3.4.1.1. Particularités du timbre de la feuille sifflée

#### Une stratégie adaptée aux langues à contours de tons

Le spectre du sifflement obtenu avec une feuille qui vibre entre les lèvres est plus complexe que pour les autres types de sifflements. Même à distance, les harmoniques sont très peu atténuées par rapport à la fréquence fondamentale (Figure 11). Nous pensons que le choix de la feuille est bien adapté à la structure des langues tonales. En effet, dans ce type de langues, la transposition de la fréquence fondamentale ne permet pas une très grande amplitude de variation, du moins pas aussi grande que dans les langues non tonales. Le type de spectre de fréquences produit par la feuille sifflée développe une réponse judicieuse à cette situation en créant une redondance d'information à la fois en fréquence et en temps ce qui permet une plus grande acuité perceptive sur les modulations des tons et sur les intervalles de silence. Dès lors il est logique de

retrouver cette technique dans les langues tonales utilisant des modulations fréquentielles distinctives à l'intérieur même d'une syllabe, typiquement des langues à *contours de tons* comme le akha ou le hmong<sup>29</sup>.

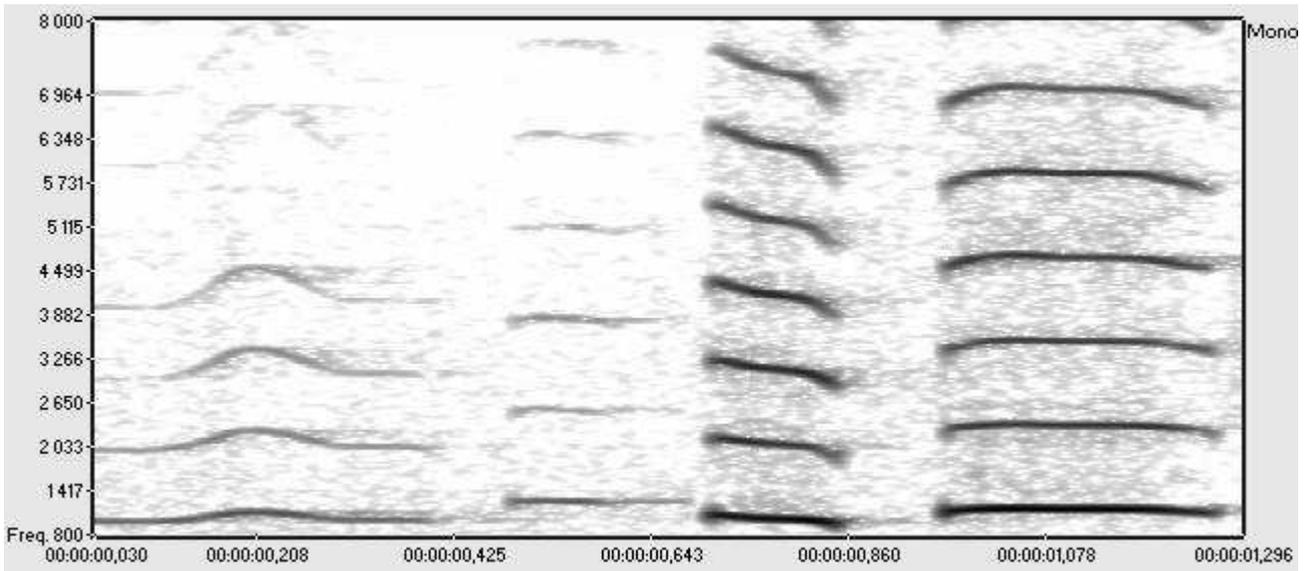
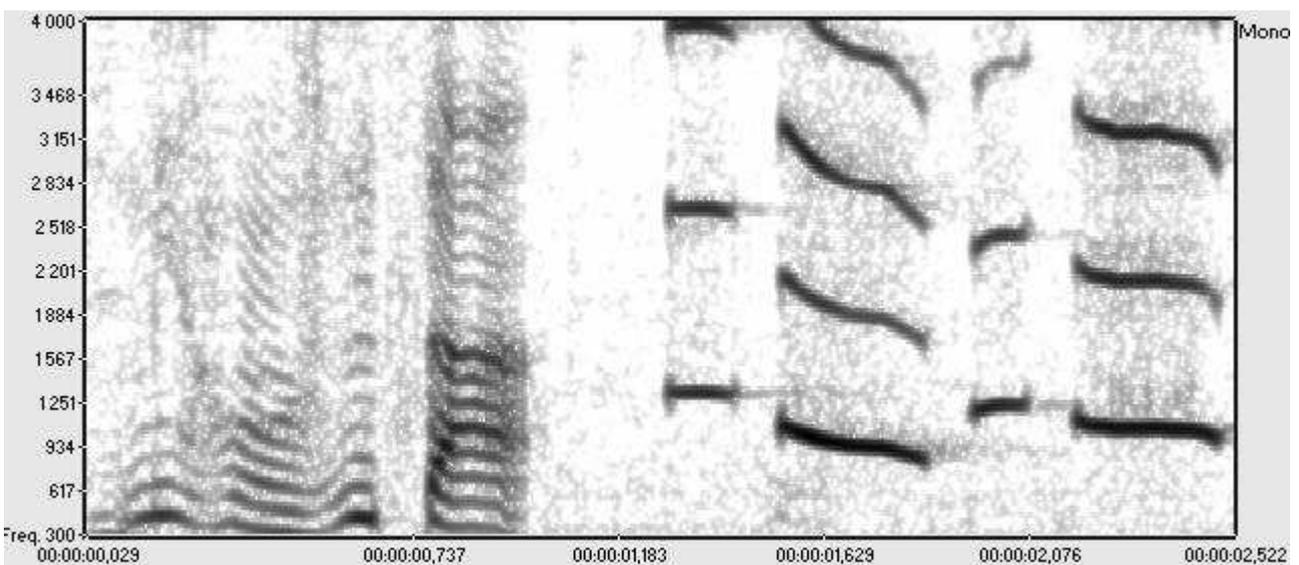


Figure 11 : Syllabes akha en feuille sifflée, enregistrées à 100 m (2 premières syllabes) et 10 m (2 suivantes)

### Comparaison du spectre de la feuille sifflée et de la voix parlée

La dynamique de la fréquence fondamentale du sifflement avec feuille est une transposition de celle de la voix parlée. Les écarts entre harmoniques de la feuille sifflée recréent la dynamique complexe du signal de la voix. Ainsi, même à une distance de 300 m le sifflement avec une feuille produit une copie de la voix parlée assez saisissante, même si celle-ci n'a pas les regroupements caractéristiques des formants que l'on peut souvent observer dans la voix parlée. L'appellation *voix sifflée* est donc tout à fait adaptée au spectre de sifflement avec feuille.



<sup>29</sup> Voir chapitre 3 pour une description du système phonétique de la langue hmong et en Annexe **Erreur ! Source du renvoi introuvable.** pour une description du système phonétique de la langue akha.

Figure 12 : Voix parlée et feuille sifflée du même mot akha.

« Mir gangq goer tar » (notation RPA voir explications dans l'annexe Erreur ! Source du renvoi introuvable.)

## 2.3.5. Propagation du signal en milieu acoustique naturel et portée géographique

### 2.3.5.1. Atténuation en amplitude avec la distance

#### 2.3.5.1.1. Théorie physique et réalité sifflée

Les lois physiques découvertes sur la propagation acoustique dans l'air établissent qu'en théorie une atténuation du signal en intensité de 6 dB se produit à chaque fois que la distance est doublée à partir de la source. De plus, l'absorption est proportionnelle au quart de la fréquence. Cette atténuation est aussi souvent supposée relativement indépendante de l'humidité et dépend principalement de la température qui détermine la rapidité de propagation du son dans l'air (Figure 13).

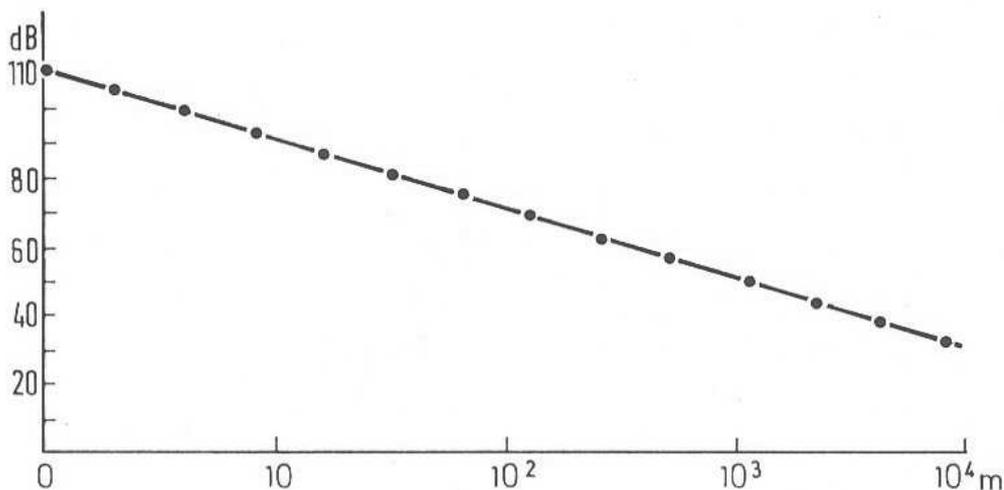
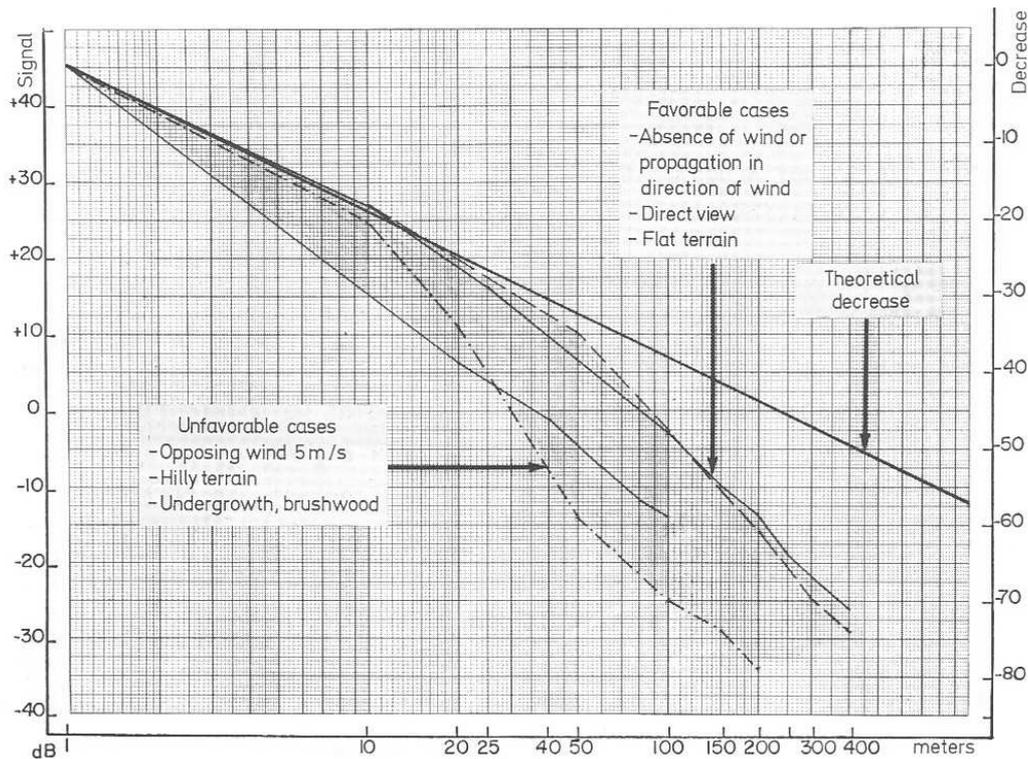


Figure 13 : Atténuation théorique d'un son de 110 dB

(amplitude moyenne d'un sifflement puissant émis pour atteindre une grande distance) en fonction de la distance de propagation (in Busnel & Classe 1976 p.39)

En réalité, tous les auteurs ayant eu affaire au phénomène de propagation ont appliqué des facteurs correctifs à la théorie. En effet, Busnel et Classe (1976) rappellent qu'en pratique, les gradients atmosphériques de températures dépendent du vent et de l'humidité à tel point que l'influence de ces facteurs est difficile à quantifier. C'est seulement dans certaines conditions que les gradients de température sont égalisés, comme la nuit ou sous la pluie. La transmission est alors optimisée puisqu'il n'existe plus de couches d'air qui réverbèrent le signal et entre lesquelles il perd une partie de son intensité.



**Figure 14 : Mesures expérimentales de la propagation du son dans la nature, dans différentes circonstances**

Busnel a montré par une série de mesures qu'entre la droite théorique et la réalité des exemples pratiques, il y a une différence qui révèle que la valeur de -12 dB pour chaque doublement de distance par rapport à la source serait plus réaliste que celle de -6 dB (Busnel & Classe 1976). C'est pourquoi un signal sifflé à 110 dB qui devrait être en théorie de 34 dB à 10000 m l'est en fait entre 1000 et 2000 m suivant les milieux.

### 2.3.5.1.2. Pratique des siffleurs

Les habitants de chaque emplacement que nous avons visité ont une connaissance approfondie des conséquences des conditions météorologiques et topographiques locales. Ceci leur permet d'évaluer rapidement si les conditions de communication à distance sont optimales ou minimales et donc de choisir en conséquence la puissance avec laquelle ils doivent siffler afin de se faire comprendre.

### 2.3.5.2. Filtre de la Nature et bruit ambiant

Les différentes observations que nous venons de faire montrent que la réduction du spectre fréquentiel avec la distance est essentiellement due à la perte d'intensité du signal : c'est la raison pour laquelle l'ensemble des éléments encodant l'information linguistique dans une langue sifflée sont compris dans la fréquence fondamentale. La manière dont celle-ci va être affectée par le filtre formé par le milieu naturel et la capacité à émerger du bruit de fond seront des facteurs clefs pour que la communication puisse avoir lieu.

#### 2.3.5.2.1. Adaptation aux conditions extrêmes

##### Le bruit

Dans les conditions habituelles d'usage du sifflement, le bruit de fond sera extrêmement variable en fonction de la situation géographique, de la topographie, du type de terrain, de la végétation, des circonstances

météorologiques, de la période de la journée (effet réverbérant des couches d'airs créées par la chaleur du soleil dans la journée) et des bruit biologiques environnants (animaux, vent, rivières et torrents). Tous ces paramètres créent un éventail large de situations, allant des conditions optimales à des situations très bruitées. Les conditions optimales sont réunies pendant les nuits sans vent. Les niveaux de bruit alors observés sont de l'ordre de 30 à 35dB et l'analyse spectrale de différents niveaux de bruits montre que ces niveaux ne sont atteints que pour des fréquences inférieures à 150 Hz. En effet, le bruit de fond n'a pas une répartition uniforme dans le domaine fréquentiel<sup>30</sup>. Si le vent souffle, il y a un bruit supplémentaire, dont le niveau peut s'élever aux alentours de 50 dB. On a alors un niveau de bruit d'environ 30 dB à 1kHz et de 20 dB à 4kHz. Dans les montagnes ou les forêts où nous avons enregistré les langues sifflées en 2003 et 2004, la caractéristique principale du bruit ambiant est de n'être pas (ou très peu) constitué de ces bruits modernes, qui sont la conséquence de notre civilisation industrielle. S'il existe des routes ou des voies non goudronnées empruntées par les voitures, celles-ci sont peu nombreuses et suffisamment éloignées les unes des autres. Ceci signifie que le bruit de fond est extrêmement réduit en comparaison des milieux urbains. Les bruits de fond les plus intenses que nous ayons rencontré sont ceux de la forêt amazonienne, des grillons en montagne au début de l'été, ou encore des torrents en Turquie. Or, comme nous l'avons montré, les sifflements humains peuvent atteindre 110 à 120 dB d'intensité proche de la source dans la bande de fréquence 1-4kHz. Par conséquent, le signal a la possibilité d'émerger du bruit ambiant relativement facilement, jusqu'à des distances significatives. Il arrive assez fréquemment que des sifflements d'oiseaux viennent interférer avec le sifflement humain. Sauf à de rares occasions, les chants d'oiseaux se situent dans des plages fréquentielles légèrement différente, d'autre part, ils ont une dynamique plus régulière car ils reposent souvent sur la répétition de chants contenant des motifs stéréotypés. Quoiqu'il en soit, s'il y a interférence, l'auditeur bénéficie alors de l'effet "*cocktail party*" (Cherry 1953) qui lui permet de focaliser son attention sur la dynamique, la direction, l'éloignement et la signature qui caractérisent le sifflement de son interlocuteur.

### Sifflement des merles à la Gomera

A la Gomera, les merles possèdent la faculté de copier le sifflement humain. De tels phénomènes ont rarement été observés dans la nature en d'autres lieux, mais sont relatés dans de nombreux témoignages sur cette île des Canaries et ont pu être reproduits en Angleterre par le linguiste André Classe (Busnel et Classe, 1976). De nos jours, ils sont de plus en plus rares en raison de la perte de vitalité de la langue sifflée dans la génération des 20 à 50 ans, et donc de la diminution du nombre d'occasions pour les merles d'entendre le sifflement humain.

### Expérimentation dans le contexte d'un bruit de torrent en Turquie

En Turquie, de nombreux villages sont traversés par des torrents dévalant les différentes vallées depuis les hauts plateaux où les bergers partent l'été avec leur troupeau de moutons. Ces torrents produisent un bruit qui

---

<sup>30</sup> Le fait que les sons de fréquence basse se propagent plus loin que les sons de fréquence haute pour des raisons liées à l'ampleur de la longueur d'onde par rapport à la taille des objets rencontrés en est une des raisons. L'autre raison est liée à la nature de certains sons très fréquents de la nature, comme le bruissement des végétaux sous l'effet du vent.

empêche toute communication parlée et même criée à leur proximité ou de part et d'autre du cours d'eau, particulièrement au printemps, époque à laquelle nous nous sommes rendus sur place. Pourtant, certains métiers nécessitant une proximité avec le torrent (comme l'élevage de truites en pisciculture) sont courants dans la région et font que les habitants possèdent leur maison en bordure du torrent. A l'occasion de notre rencontre avec le propriétaire d'une petite pisciculture, dans une vallée voisine de celle de Kusköy, nous avons réalisé des enregistrements pour évaluer l'efficacité du sifflement dans ce contexte extrême.

Les Photo 23 et Photo 24 montrent la pisciculture et le torrent photographiés 4 mètres en avant du point d'enregistrement situé sur la terrasse d'une maison. Le propriétaire et un autre visiteur ont sifflé trois phrases différentes depuis une position 3 mètres en retrait du début du pont en bois, situé à proximité des bassins et surplombant le torrent. La distance de communication a été réglée à 30 m.



**Photo 23 : Zone d'expérimentation dans le bruit du torrent en contrebas, pisciculture en Turquie**

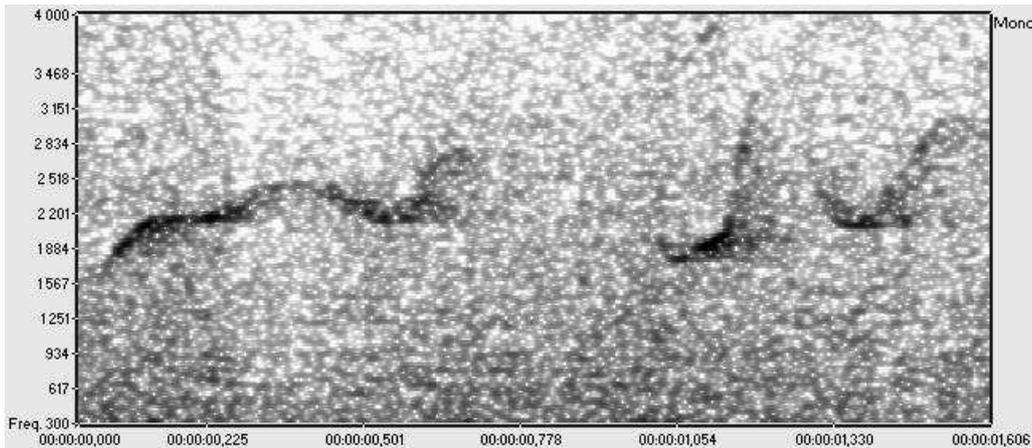


**Photo 24 : Torrent passant en contrebas de la pisciculture**

**Prise de vue depuis le point d'enregistrement. Malgré les apparences, ce torrent produit un son assourdissant qui empêche toute communication claire, même criée depuis la maison (d'où la photo est prise) jusqu'à l'autre côté**

Le torrent fait un bruit que nous avons évalué a posteriori à plus de 70 dB en le comparant au bruit des torrents des Alpes de force égale enregistrés à la même distance. En raison de la fonte des neiges d'avril, le torrent turc était particulièrement fort et présentait un rapide en aval de la photo dont on perçoit clairement la puissance sur l'enregistrement sonore.

Nous avons pu observer que le sifflement articulé est bien adapté à la communication dans ce type de bruit. Ainsi sur la Figure 15, la forme des modulations sifflées du pisciculteur turc émerge clairement.



**Figure 15 : Parole sifflée turque émergeant du bruit d'un torrent (communication à 30m)**

## **Effet de la réverbération dans les montagnes**

Dans certaines vallées mises à profit par les siffleurs pour communiquer, un phénomène de prolongation du sifflement et d'écho peut être entendu : c'est une manifestation extrême de la réverbération due aux parois montagneuses. Ce facteur peut perturber l'intelligibilité car il modifie le signal d'origine. Deux expériences nous ont permis de l'observer de manière claire.

### **Expérience 1 : Echo en conditions d'usage naturel à La Gomera**

Une série d'enregistrements a été réalisée à la demande d'un des siffleurs dans un de ses lieux favoris pour la propagation des sons et de l'écho. L'enregistrement typique d'une phrase dans ces conditions est présenté Figure 16.

#### **Photo 25 : Luis sifflant avec la plus grande puissance possible**

**Il siffle depuis un surplomb au-dessus d'une vallée pour provoquer une réverbération maximale. L'habitude de mettre sa main en porte voix vise à concentrer le champ de parole. Nous avons observé cet usage qu'à la Gomera.**



Photo 26 : Type de terrain ou l'expérience à été menée à la Gomera

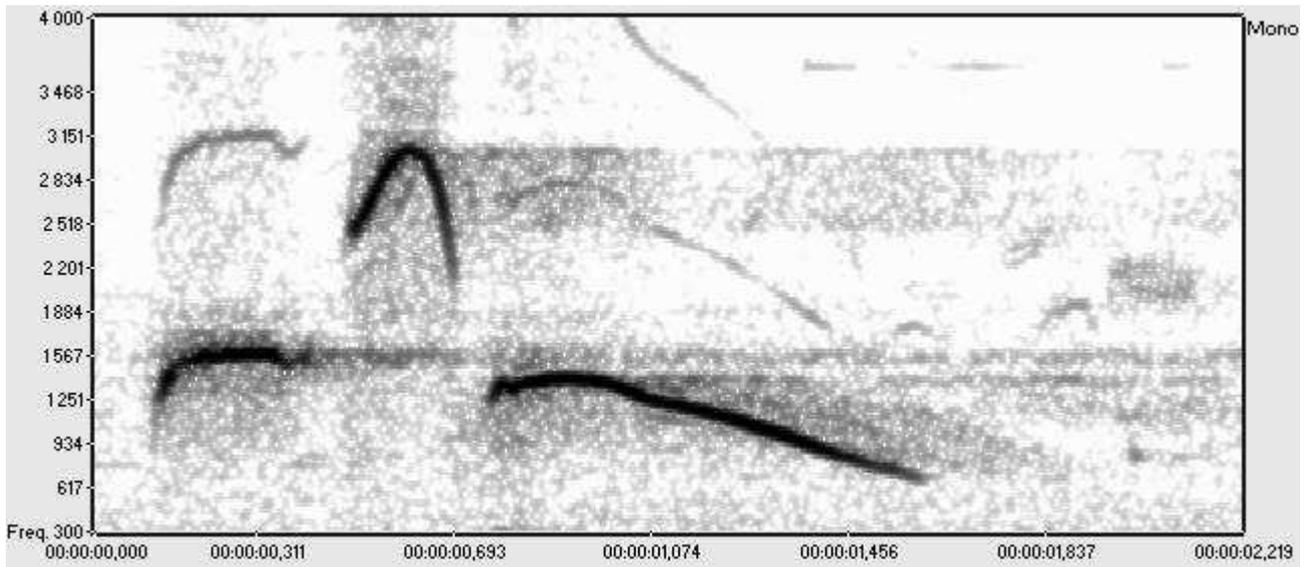


Figure 16 : Mot espagnol « a-mi-go » en enregistré à 10m

On remarque que chaque bande de fréquence provoque une traînée caractéristique d'une très forte réverbération du milieu

## Expérience 2 : Diffusion contrôlée de phrases sifflées dans une vallée du Vercors

### a) Principe

Avec le dispositif de diffusion de sons décrit en détails en Annexe G et §4.3.4.3.2, nous avons diffusé des sifflements de plusieurs langues (grec, silbo et turc) toujours environ au même niveau (100 dB), et nous les avons enregistré à des distances différentes (100, 150, 300, 550m et un enregistrement contrôle à 4 m<sup>31</sup>).

<sup>31</sup> L'enregistrement contrôle permet de vérifier l'effet du haut-parleur sur le signal grâce à une comparaison avec l'enregistrement d'origine. La distance d'enregistrement à 4m de la source est justifiée par le souci de se positionner en champ lointain, c'est à dire dans une zone où l'influence des bruits de la source d'émission est négligeable.

b) Analyse de l'évolution de la réverbération

Un phénomène robuste apparaît : la réverbération diminue avec la distance par effet d'absorption de la nature. Ceci a pour effet de clarifier les sons à 300m par rapport à ceux de 100m. On observe aussi une dégradation du son dont nous initierons l'analyse dans la dernière partie de cette thèse (en particulier en ce qui concerne les conséquences phonétiques de cette dégradation).

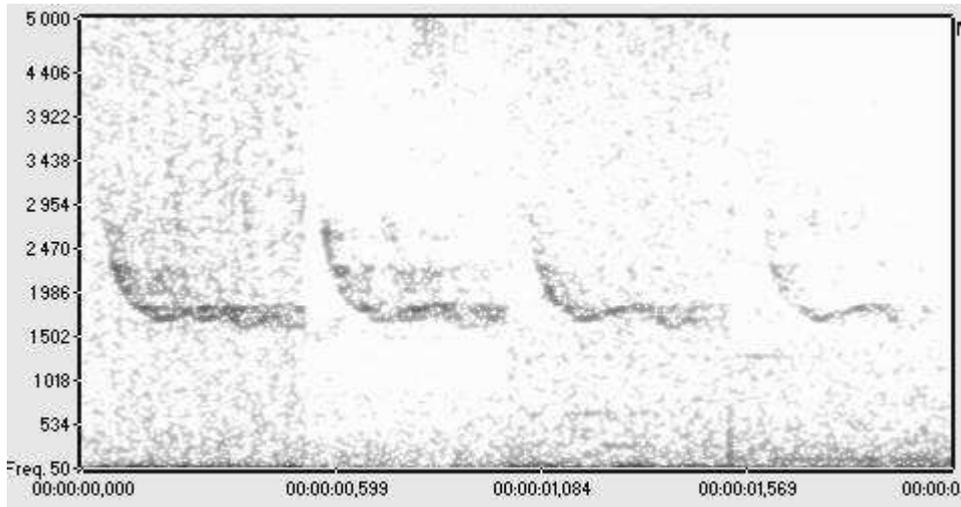


Figure 17 : Mot grec /thelo/ enregistré à 100, 150, 300 et 550 m

Expérience complémentaire :

A la distance de 100 m, un enregistrement complémentaire a été effectué en changeant l'orientation du haut-parleur de 45 degrés par rapport à l'orientation initiale comme indiqué sur le schéma en §4.3.4.3.2. Ceci a été effectué à l'origine pour observer les modifications d'intensité engendrées par un tel changement (effet directionnel). Mais nous avons aussi constaté des changements significatifs en ce qui concerne la réverbération (Figure 18).

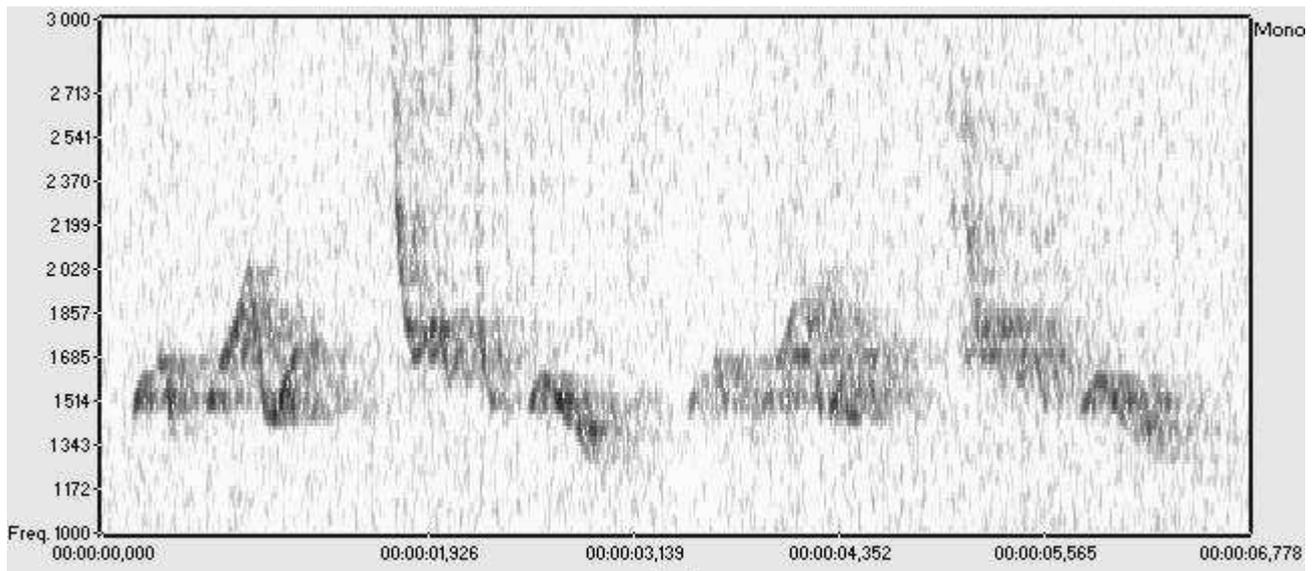


Figure 18 : Phrase grecque /parakalo thelo na fao/ « s'il te plaît, j'ai faim ». Enregistrée en face puis avec un angle de 45° par rapport à la source

La réverbération est bien plus élevée pour le second enregistrement (45°). Le récepteur ne reçoit plus l'onde directe la plus intense car il n'est plus dans le champ privilégié par la source. Le signal est beaucoup perturbé même s'il reste intelligible très facilement. Cette expérience explique pourquoi les siffleurs cherchent toujours à se faire face ainsi que l'habitude courante à la Gomera de faire un porte-voix avec ses mains (cela pour restreindre l'effet de réverbération du signal en le concentrant vers l'auditeur).

### 2.3.5.2.2. Analyse de la réponse du filtre de la nature dans les montagnes et les forêts

Une corrélation est souvent observée entre les contraintes environnementales et les comportements acoustiques de communication (Aubin, 2004 ; Mathevon et al 2004), la parole sifflée en est un exemple humain.

#### **Dans les forêts**

Le domaine de la bioacoustique a fourni depuis quelques années une analyse des sifflements de nombreuses espèces vivant dans les forêts, en particulier des oiseaux. Marler (1955) observa que ce mode de production sonore était toujours utilisé pour les communications à distance. Comme nous l'avons vu, une partie de son efficacité est due à une bande de fréquence étroite située à une fréquence supérieure à celle de la zone la plus intense de bruit de fond du milieu ambiant. Des analyses en forêt de la réverbération naturelle ont été effectuées plus récemment. Une étude particulièrement pertinente pour notre compréhension du développement des sifflements par les chasseurs en forêt a démontré qu'il existe un domaine central de fréquences entre 1kHz et 3 kHz (bande de fréquence des sifflements humains) pour lequel la réverbération varie relativement peu avec la distance (Padgham 2004).

#### **Dans les montagnes**

Il existe étonnamment peu de données sur la propagation des sons en milieu montagneux. D'une manière générale on peut affirmer que l'acoustique des montagnes est très dépendante de la configuration locale des surfaces des vallées et des massifs. Cependant le comportement similaire des siffleurs dans différentes régions montagneuses démontre que les vallées sont toujours utilisées comme des guides d'ondes efficaces car elles favorisent la propagation des signaux sifflés. De plus, comme dans le cas des forêts, les propriétés de réverbération dans les montagnes pourraient avoir de nombreuses propriétés communes en ce qui concerne le filtre fréquentiel naturel qu'elles constituent. Pourtant, ce filtre n'a jamais été étudié en détail dans une recherche du domaine public (com. pers. Padgham 2005). Nous avons tenté lors d'une expérience pilote de l'évaluer en enregistrant un bruit blanc à plusieurs distances et des signaux carrés d'une fréquence de 700 Hz et 1600 Hz. Notre dispositif (voir Annexe **Erreur ! Source du renvoi introuvable.** et §4.3.4.3.2), bien qu'adapté à la diffusion de sifflements ou de la parole s'est révélé ne pas être assez précis pour évaluer précisément le filtre de la nature.

## **2.3.6. Distances limites d'intelligibilité de la voix parlée, de la voix criée et de la parole sifflée**

### **2.3.6.1. Introduction**

Le fait d'augmenter la distance entre deux individus en situation de communication linguistique entraîne des variations spectrales inégales, une diminution du rapport signal sur bruit due à l'atténuation de l'intensité des sons à l'oreille du récepteur et une augmentation du temps de réverbération. Ces trois phénomènes participent à l'évaluation de la distance de l'interlocuteur (Canevet 1989). Or, pour augmenter la portée de sa voix, l'individu en augmente le niveau de manière presque inconsciente et, à mesure que la distance augmente, change progressivement de registre pour émettre une voix criée. Ce phénomène a été décrit principalement dans les années 50 sous le nom d'effet "Lombard": *"some years ago it was noticed that a speaker with normal hearing would unconsciously raise the level of his voice to compensate for the level of noise in which he was speaking [...] the Lombard reaction, while a general one varies in magnitude, as do other voice measures, among individuals"* (Dreher & O'Neill, 1957 p.1320). Handley et Steer (1949) ont remarqué qu'en présence de bruit, les locuteurs avaient tendance également à réduire leur débit de parole et à prolonger les syllabes. Le passage de la voix parlée à la voix criée est une intensification substantielle de tous ces phénomènes. Mais il existe une distance à laquelle le cri va épuiser les cordes vocales qui auront tendance à érailler la voix, d'abord de manière douce puis de manière accentuée au point qu'elle peut ensuite dérailler si l'effort est trop grand. La fatigue biologique provoquée dans ce cas limite est rapide. Les conversations criées à distances sont donc forcément de courte durée. Chez les populations connaissant les langues sifflées, le cri est très rarement pratiqué, on lui préfère le sifflement, car l'effort dans ce cas n'a pas de conséquences fâcheuses sur la santé des cordes vocales. Afin de confirmer ces observations sur les limites de la voix parlée, de donner un ordre de grandeur des distances en jeu et d'évaluer l'avantage des langues sifflées sur le cri en montagne, nous avons réalisé trois séries de mesures lors de trois expériences en plein air menées, dans des milieux montagneux.

### **2.3.6.2. Expérimentations : portées limites la voix parlée et de la voix criée**

Peu d'études existent, à notre connaissance, sur la parole à distance. La plus connue est celle de Gardner (1969) sur l'évaluation de la distance. Notre approche est différente, puisqu'il s'agit pour nous de trouver, autant que possible, les ordres de grandeurs des limites d'intelligibilité de la voix parlée et de la voix criée dans des conditions de plein air, du type de celles rencontrées dans les milieux ruraux. Les résultats présentés ici sont issus d'*expériences pilotes* permettant d'expliquer l'intérêt du sifflement. Les résultats obtenus sont

en cours de validation avec plus de sujets, mais il nous a semblé intéressant d'avancer ici les données obtenues pour l'instant (avec un nombre très restreint de sujets) car elles expliquent en partie les choix acoustiques des siffleurs. La difficulté du contrôle des paramètres de la voix et des paramètres de bruit en milieu ouvert explique pourquoi de telles analyses n'ont jamais été menées en linguistique. Seul notre intérêt pour le sifflement et pour l'étude du langage dans des conditions naturelles de pratique et d'évolution peuvent justifier le fait de conserver autant de paramètres variables lors d'une étude sur le langage. L'élaboration du protocole a subi quelques améliorations depuis que nous les décrivons succinctement en conclusion de cette expérience.

### **2.3.6.2.1. Expérience 1 : ordre de grandeur de la limite de la voix parlée**

#### **Cadre de l'étude**

Un grand champ en terre dure et en forme de cirque a été choisit dans les Alpes du Nord, à 1350m d'altitude. Le choix d'un tel milieu se justifie par le souci d'être dans des conditions optimales de propagation des sons sans toutefois avoir les désavantages de réverbération d'un cirque fait de facades rocheuses.

#### **Contrôle des sujets**

Deux sujets ont prêté leur voix à l'expérience (un homme et une femme). Leur audition a été testée préalablement comme normale par audiogramme.

#### **Protocole**

L'un des sujets parlait de manière continue avec pour instruction de raconter sa journée depuis l'aube tout en gardant le regard fixé sur le sonomètre qu'il tenait dans les mains de manière à maintenir les maxima de sa parole dans une zone d'amplitude de 10dB de largeur (60-70 dB dans un premier temps en partant de l'auditeur). Il ne devait pas regarder son interlocuteur pour ne pas corriger inconsciemment le niveau d'émission. Le sujet émetteur reculait par à-coups de trois pas en arrière en s'éloignant de l'auditeur et continuait de parler. L'auditeur devait signaler quand il ne comprenait plus les phrases de l'émetteur. A ce moment là, le locuteur et l'auditeur cherchaient ensemble, par dichotomie, la distance qui correspondait au point où l'intelligibilité de la parole était perdue. Une fois ce point trouvé et la distance mesurée avec un mètre, le locuteur repartait du point trouvé en changeant de niveau de voix (70-80 dB). L'expérience a été arrêtée après 80 dB de voix car ensuite, il est admis que l'on bascule progressivement vers la voix criée. Pour chaque mesure le processus a été répété deux fois avec des résultats très stables en terme de distances.

#### **Caractéristiques de la voix des deux sujets**

VoixF= Voix de Femme (Fomédian = 200 Hz)

VoixM= Voix d'Homme (Fomédian= 120 Hz)

## Caractéristique du bruit ambiant

Le bruit ambiant a été mesuré à plusieurs reprises aux alentours de 35-40 dB. Tout élément perturbateur (moteurs d'avion, beuglement de vache) a occasionné l'interruption momentanée de l'expérience.

## Résultats

Plusieurs phénomènes intéressants ont été observés. Ils sont résumés dans le tableau ci dessous :

**Tableau 7 : Distances limites d'intelligibilité de la voix parlée contrôlée**

Sons		Distance limite d'intelligibilité (en m)	
Type de voix	Amax	Distance	Zone tampon (variabilité)
VoixF parlée	60-70 dB	25,50	N/A
VoixF parlée	70-80 dB	53,30	N/A
VoixM parlée	60-70 dB	22,90	1
VoixM parlée	70-80 dB	46,80	N/A

Tout d'abord, à Fo égal, une augmentation de niveau de 10 dB entraîne un allongement de la distance limite d'intelligibilité. Dans les conditions de l'expérience, les mesures ont révélé un peu plus qu'un doublement de la distance.

D'autre part, il semble que un Fo plus élevé (voix de femme) semble entraîner un allongement de la distance limite d'intelligibilité. Cet effet est à vérifier mais semble plausible car plus le Fo est élevé plus l'ensemble de la voix se place de manière évidente audessus du bruit de fond.

Une zone tampon a été trouvée à l'intérieur de laquelle la perte d'intelligibilité se fait pour un même sujet. Elle est relativement courte et d'environ 1m dans tous les cas mais nous ne l'avons mesuré de manière précise qu'à une seule distance. A l'intérieur de cette zone, la perte de l'intelligibilité de certains mots intervient dans un premier temps puis par effet de cascade et de manière très rapide la perte d'intelligibilité totale. Passé cette distance et même en accroissant son attention qui était déjà élevée, l'auditeur ne parvient plus à reconstruire le sens des phrases. Cette expérience confirme de nombreuses approches de l'intelligibilité à distance (en particulier en salles de concert) qui ont analysé que ce phénomène n'était pas linéaire.

Les tendances que nous recherchions sont les suivantes : la voix parlée dont les maxima d'amplitude restent entre 60 et 70 dB est intelligible jusqu'à environ 22-25 m dans un milieu naturel optimal, mais son écoute confortable n'est effective que jusqu'à environ 15-20 m ; la voix parlée jusqu'à 80 dB est intelligible jusqu'à environ 50 m dans un milieu naturel optimal mais son écoute n'est confortable que jusqu'à environ 30-35m.

## 2.3.6.2.2. Expérience 2 : ordre de grandeur de la limite de la voix criée

### **Cadre de l'étude**

Le milieu choisi était une vallée dans le Vercors (voir § 4.3.4.3).

### **Objectif**

L'expérience a été mise en place afin de comparer les distances limites d'intelligibilité de la voix criée aux distances limites d'intelligibilité de la parole sifflée dans le même milieu. Seul un ordre de grandeur devait être obtenu. Par contre une attention toute particulière a été portée à l'enregistrement et à la mesure des niveaux d'émission afin d'en extraire des tendances d'évolution de stratégie du cri.

### **Sujet**

Le sujet volontaire (un homme) pour crier possède une voix parlée ayant un Fo bas.

### **Protocole**

Le crieur reste fixe alors que l'expérimentateur recule de 50 m en 50 m<sup>32</sup>. Quand l'expérimentateur est prêt, le crieur prononce 4 longues phrases en Français à deux reprises afin que l'expérimentateur enregistre tout en jugeant en direct de l'intelligibilité dans le casque puis sans le casque<sup>33</sup>. L'expérience s'est arrêtée lorsque l'expérimentateur ne pouvait plus comprendre les mots des phrases.

### **Résultats**

Les données mesurées sont présentées dans le tableau suivant. L'intelligibilité a été évaluée dans le casque sur le terrain lors de la prise de son pour qu'il y ait correspondance avec les mesures faites sur les enregistrements. Le bruit de fond est de 40 à 45 dB sans vent (présence de grillons)

---

<sup>32</sup> Distances mesurées avec un mètre déroulant de 50 m

<sup>33</sup> A priori le micro utilisé n'étant pas directionnel, aucune différence ne devrait apparaître mais il se peut qu'un confort d'écoute lié à la localisation plus précise sans le casque améliore les performances. L'écoute dans le casque était nécessaire afin de pouvoir interpréter les sons enregistrés.

Tableau 8 : Données caractéristiques de l'émission et de la perception d'une voix criée à plusieurs distances

Distance (en m)	Emetteur			Récepteur	
	Niveau d'émission (en dB)		Allongement par rapport à 50 m (en %)	Elévation des voyelles par rapport à 50 m (en Hz)	Intelligibilité
	Moyen	Max			
	effets de voix				
50	85	95	0	0	Bonne, écoute très confortable
	Voix criée maîtrisée				
100	88	97	0,5	55	Bonne, écoute confortable
	Voix maîtrisée, fin de mots accentués				
150	93	107	5	105	Bonne, écoute attentive Perte quelques rares syllabes
	Voix parfois légèrement éraillée en fin de mots				
200	95	108	40	155	Ecoute attentive nécessaire Reconstruction de certaines syllabes
	voix flanche parfois en fin de mots				
250	98	108	45	200	Difficile mais possible avec une grande attention, perte de nombreuses syllabes
	saturation de la voix, fatigue biologique				
300	98	108	Non mesuré	Non mesuré	Perte totale d'intelligibilité

Le phénomène d'effet Lombard a été confirmé de même que la stratégie d'allongement des syllabes. Une progression a été observée comme suit : alors qu'à 50 m, le crieur ne force pas et adapte son augmentation de niveau, à 100 m il commence à accentuer certaines fin de mots. A 150 m il commence à allonger ses phrases et atteint des maxima de niveau qui frôlent ses limites. A 200 m, il atteint ses limites de niveau, des pics d'intensité limités à 110 dB peuvent rapidement lui dégrader la voix s'il insiste ; à ce stade, il reste intelligible mais au prix d'un allongement des mots de 40% en moyenne. A 250 m, le crieur s'abîme la voix, l'allongement supplémentaire est dû à des temps de récupération après les fins de mots, mais la communication reste possible dans des conditions toutefois très difficiles qui demandent une grande attention d'écoute. A 300 m l'intelligibilité est perdue. Les écoutes, avec ou sans casque, ont donné les mêmes résultats à toutes les distances.

Le sonagramme, présenté Figure 19, montre un exemple d'extrait d'une des phrases enregistrées de 50 à 250m. On remarque que le bruit fait par les grillons n'affecte pas la zone où se placent les principales harmoniques soulignées par le cri. On observe aussi qu'une réverbération importante affecte le signal dès 100m mais celle-ci n'a pas été mesurée précisément.

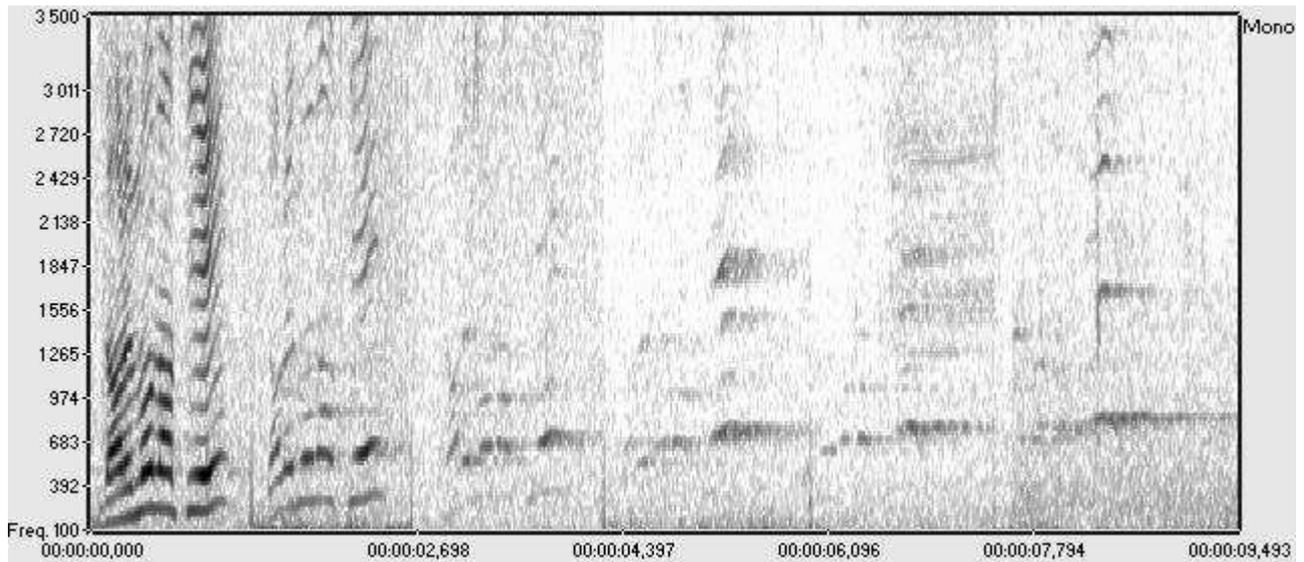


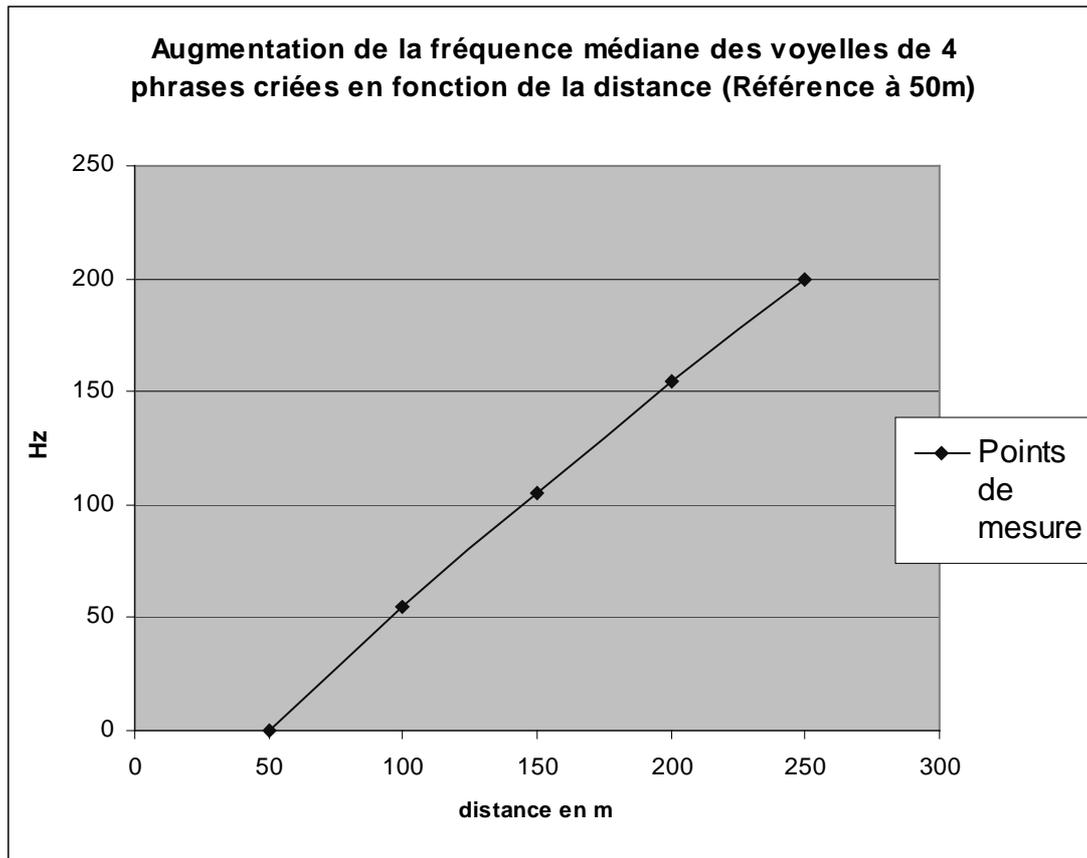
Figure 19 : extrait de la même phrase parlée à 10 m (référence) puis criée à 50, 100, 150, 200 et 250 m

On observe une forte dégradation des harmoniques de la voix avec le maintien de certaines d'entre elles sur lesquelles le locuteur insiste en situation de communication à distance.

## Conclusion

Trois phénomènes sont particulièrement remarquables dans l'ensemble de ces résultats :

- Le crieur allonge les phrases à partir du moment où il sature au niveau de l'amplitude.
- Le récepteur commence à perdre de l'intelligibilité à partir du moment où la voix du crieur sature, mais cette perte est comblée un moment par l'allongement
- Le crieur augmente la fréquence médiane des syllabes des mots qu'il prononce de manière quasi-linéaire par rapport à la distance (pente d'environ 1 Hz/m) comme nous pouvons le voir sur la Figure 20



**Figure 20 : Mesure de la fréquence de la voix criée à plusieurs distances**

Finalement, l'ordre de grandeur que nous cherchions est le suivant : en montagne dans une vallée formant un guide d'onde de qualité intermédiaire, la voix criée ne reste intelligible que jusqu'à environ 200 m.

### 2.3.6.2.3. Critique des protocoles et perspectives

Les deux dernières expériences ont permis de donner un ordre de grandeur des distances atteintes en milieu ouvert dans des environnements montagneux définis. Tout nous porte à croire que les distances atteintes dans ces milieux ont des caractéristiques très variables en fonction des voix des sujets et du terrain mais que les rapports de ces distances sont assez constants si l'on considère un lieu donné. Seul une étude étendue à un grand nombre de milieux et à un grand nombre de sujets permettrait de tirer des conclusions solides sur les paramètres qui font varier l'intelligibilité et la portée de la voix parlée et criée dans ces conditions. Le contrôle du niveau d'émission gagnerait à être contrôlé plus précisément, par exemple en utilisant des phrases préenregistrées et normalisées (toujours avec un enregistrement contrôle à 4 m)<sup>34</sup>.

### 2.3.6.2.4. Voix criée Aiguë : un usage traditionnel récurrent

Lors de notre enquête en Amazonie péruvienne, nous avons pu observer, dans les communautés *Boras* qui nous ont hébergés, l'usage épisodique de la parole criée aiguë pour parler à distance. L'élocution paraît

<sup>34</sup> A partir de 4 m on peut considérer que l'on n'est plus en champ proche.

différente de celle de la parole normale et il est difficile de la reproduire. Il semble que ce phénomène soit assez répandu dans d'autres groupes linguistiques notamment en milieux montagneux. Lors de notre enquête sur les langues sifflées en Asie, un de nos contacts nous a signalé l'existence de ce type de pratique à plusieurs endroits (au Zanskar par exemple): *"There is no kind of music or whistle in long distant communication, but they use to talk in long distant by shouting, like from one side of the valley to another side of the valley or from the roof top to their far away field. But for us it is not possible to understand or hear what they say, when they call to us from very far"* (com. pers. Dorjee 2004). Cette personne est originaire de la région et parle la langue locale mais il semble que, comme les langues sifflées, la parole à distance soit peu intelligible aux personnes qui n'en ont pas un usage courant.

### **2.3.6.3. Portées usuelles de la parole sifflée**

#### **2.3.6.3.1. Observations de terrain : portée et niveau d'émission**

Le niveau d'amplitude de production de sifflement est nécessairement associé à la distance de dialogue entre les interlocuteurs car c'est en réglant la puissance du flux d'air dans sa bouche que le siffleur adapte l'amplitude de son signal pour atteindre sa cible perceptive. Les sifflements à courte distance (jusqu'à 200m) sont produits de 60 à 80 dB en moyenne (ce qui est de l'ordre des amplitudes utilisées dans la parole humaine). Pour atteindre des distances moyennes (jusqu'à 500 m), les amplitudes utilisées sont de l'ordre de 80 à 100 dB, le sifflement avec la feuille se situe en général autour de 100 dB et permet d'atteindre des distances de 500m au maximum (sauf dans le cas où un guide d'onde très favorable se présente). Pour les grandes distances les siffleurs peuvent monter jusqu'à une intensité de 110 à 120 dB (mesure à 1 m de la source). Les distances maximales mesurées ont été obtenues à La Gomera. Busnel et Classe font état de communications à plus de 8 km (Busnel et Classe 1976). Ces valeurs maximales sont exceptionnelles. Les conditions d'usage sont plus souvent de l'ordre de 1 à 2 km.

#### **2.3.6.4. Conclusions**

Dans cette partie nous avons présenté des expérimentations en plein air permettant de comprendre dans un premier temps l'effet de l'environnement sur le sifflement et l'avantage de la bande de fréquence de ce dernier, car malgré de fortes perturbations l'intelligibilité reste élevée. Dans un deuxième temps nous avons analysé l'avantage du sifflement sur la voix parlée et criée tout en montrant la continuité qui existe entre les différentes stratégies de parole à mesure que la distance s'accroît. En effet, nous avons montré que le sifflement est l'extension naturelle des stratégies de la parole criée car il n'utilise pas les cordes vocales qui sont saturées par l'effort du cri dès 100 dB<sup>35</sup> et car il se place à une fréquence élevée permettant de minimiser le rapport signal sur bruit. Nous verrons par la suite que l'allongement de la phrase est aussi une caractéristique du sifflement. C'est pourquoi nos résultats mettent en évidence l'avantage du sifflement sur le cri tant en terme d'effort qu'en termes de confort et d'efficacité d'écoute.

---

<sup>35</sup> Permettant d'atteindre 200 à 250 m dans la vallée où nous avons réalisé nos mesures.

Les distances mesurées sont fonction du milieu et ne peuvent pas être généralisées mais leur ordre de grandeur permet de se rendre compte rapidement de l'avantage du sifflement sur les techniques utilisant la voix. La Figure 21 illustre cette conclusion.

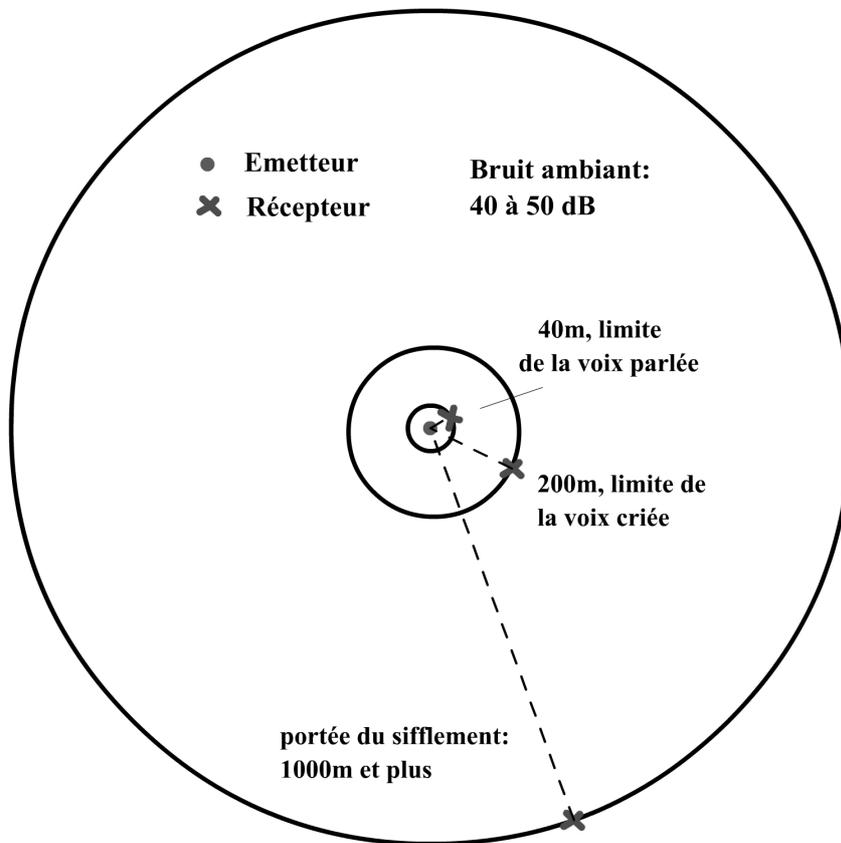


Figure 21 : Bilan des ordres de grandeur des limites d'intelligibilité de la voix parlée, de la voix criée et de la parole sifflée

## 2.3.7. Dynamique de la phrase sifflée

### 2.3.7.1. Dynamique de l'amplitude

#### 2.3.7.1.1. Analyse préliminaire

Comme dans la parole classique, et complémentairement aux modulations de fréquences, le signal de la parole sifflée contient des modulations d'amplitude. Nous avons observé deux catégories de dynamique de modulations d'amplitude. Celle des sifflements par rapport au bruit de fond, que nous appelons dynamique générale de l'amplitude et celle des sifflements les plus puissants par rapport aux sifflements les moins puissants, que nous appelons dynamique intra-sifflements.

## Dynamique générale de l'amplitude

La dynamique générale est relative au bruit de fond et à la puissance utilisée. Pour montrer la différence d'émergence du bruit de fond à 2 distances différentes d'enregistrement nous avons repris l'exemple du dialogue spontané grec (voir sonagramme Figure 10). On observe que la parole sifflée du premier locuteur (situé à 150m) émerge de plus de 20 dB du bruit de la bande de fréquence des sifflements (Figure 22). Par contre il n'émerge pas du bruit de fond général (Figure 23). Pourtant l'interlocuteur comprenait parfaitement tous les propos. Il semble donc que c'est le bruit local qui soit important au niveau de la perception. Nous en verrons les raisons dans la partie dédiée à l'intelligibilité. Les sifflements enregistrés à courte distance émergent de plus de 40 dB de la bande des sifflements et de plus de 20dB du bruit de fond général (Figure 24).

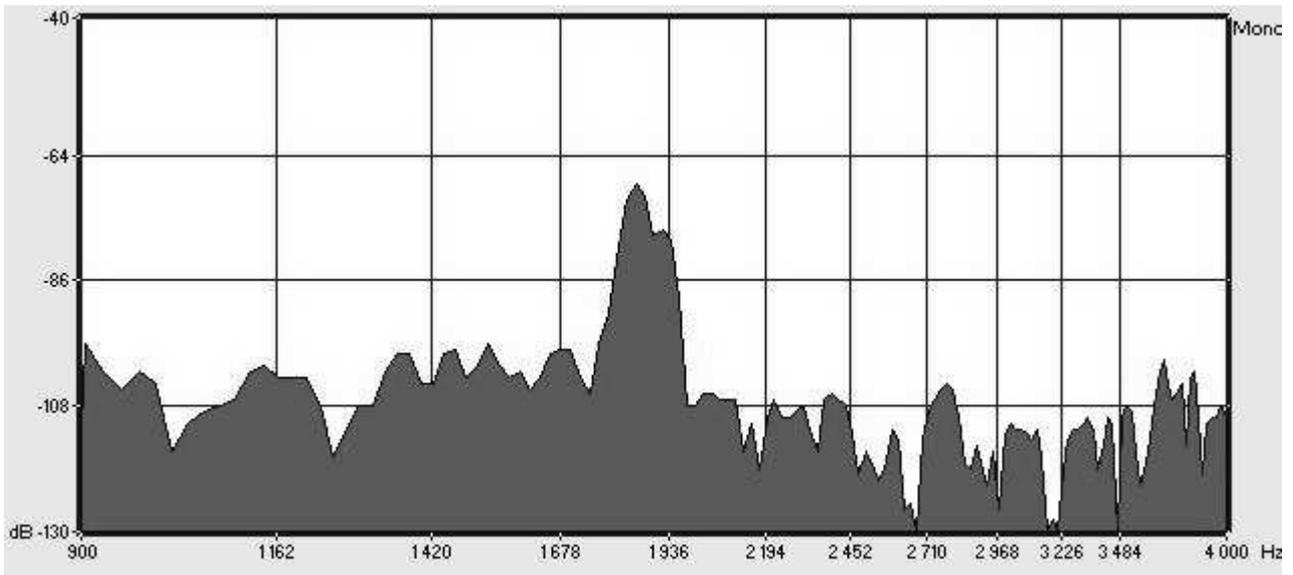


Figure 22 : Emergence de la parole sifflée du premier siffleur Grec (situé à 150m dans le dialogue Figure 10, exemple de voyelle à 1800 Hz) du bruit de fond de la bande des sifflements (900 à 4000 Hz)

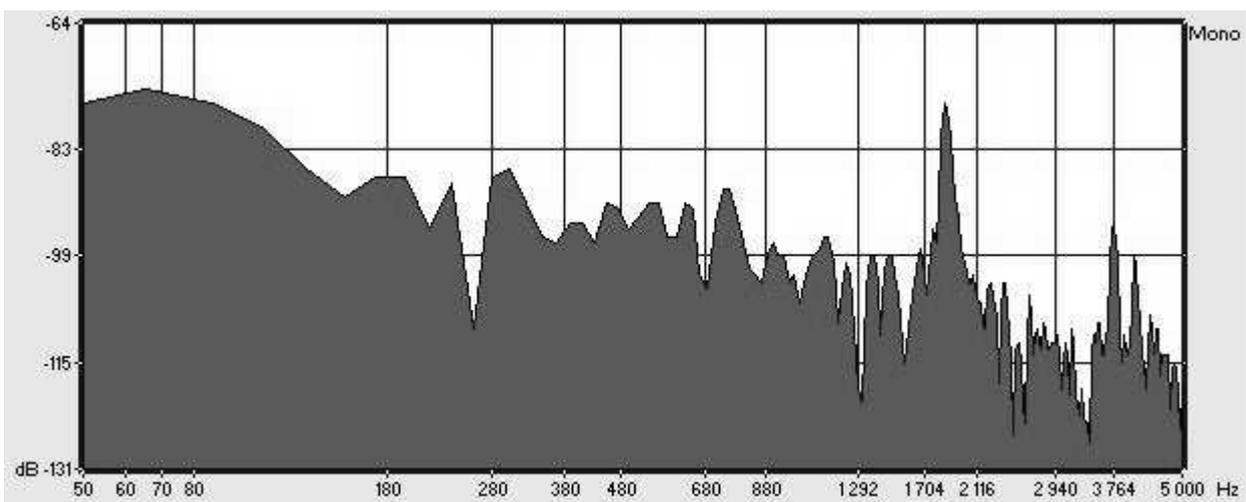


Figure 23 : Emergence du sifflement (voyelle à 1900 Hz) du premier siffleur Grec (à 150 m) de l'ensemble du bruit de fond

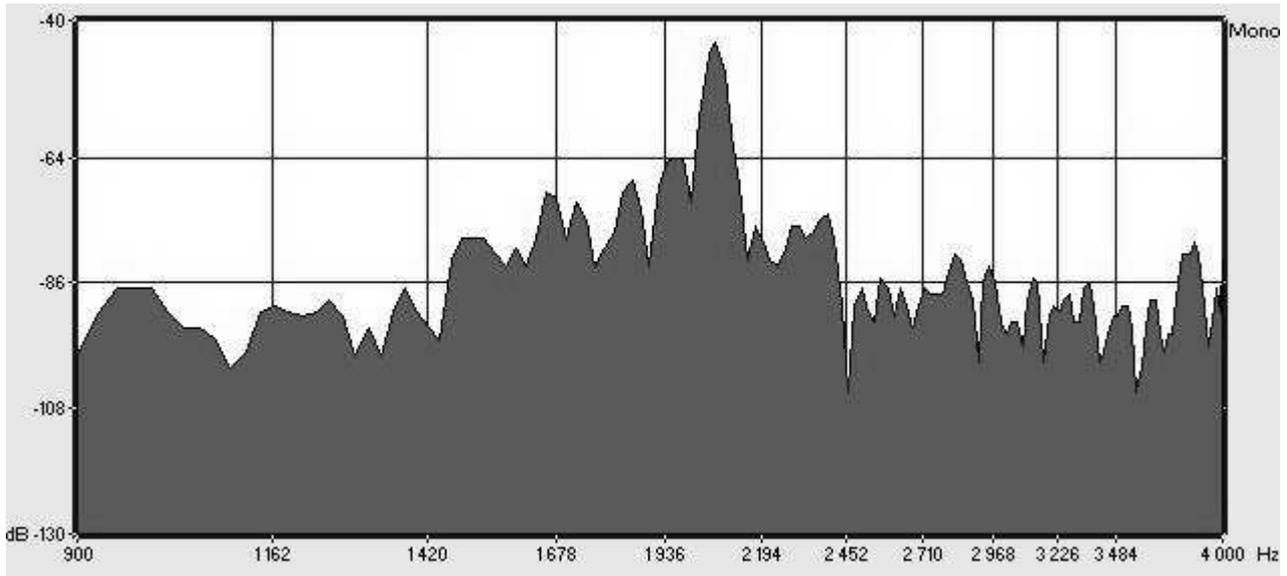


Figure 24 : Voyelle du dialogue grec, émergence du bruit de fond de la bande usuelle des sifflements (900 à 4000 Hz) de la parole du deuxième siffleur (proche)

## Dynamique intra sifflements

La dynamique intra sifflements couvre un intervalle de moins de 20 dB dans une même phrase (couramment moins de 12 dB) qu'elles que soient les langues. Une telle bande d'amplitude utile est justifiée par la nécessité d'émerger du bruit. L'enveloppe de l'amplitude d'une phrase sifflée définit une répartition de l'énergie sonore qui permet de caractériser le débit d'élocution à travers des ensembles énergétiques qui marquent le plus souvent les différentes syllabes (ou unités de paroles).

### 2.3.7.1.2. Comparaison avec la voix parlée

La dynamique de l'amplitude de la voix parlée est plus étendue car elle dépasse souvent 50 dB dans une même phrase. Malgré cela, une étude comparée des variations relatives d'amplitudes des mêmes énoncés avec une voix parlée et avec la parole sifflée fait apparaître un grand nombre de corrélations.

## Approches précédentes dans ce domaine

Très peu d'analyses comparées de l'amplitude du sifflement et de la voix ont été réalisées dans le passé. Cependant, certains scientifiques ont signalé l'intérêt du phénomène. Ainsi Busnel, qui a ouvert la voie dans ce domaine, explique que *"Complémentairement à la modulation de fréquence, il y a dans le sifflement une modulation d'amplitude, qui, avec les temps de silences constitue ce que j'appellerai peut être à tort, la ligne mélodique. Ce n'est pas mon propos d'en faire une étude exhaustive, mais il paraît intéressant d'en donner un exemple à l'intention des phonéticiens. [...] (On constate sur les tracés) que la séparation des sous ensembles reste identique et donc que la ligne mélodique bien qu'étalée dans le sifflement reste la même, ce qui apporte un élément de reconnaissance de plus au travers du conditionnement au vocabulaire courant"* (Busnel 1970 p. 56).

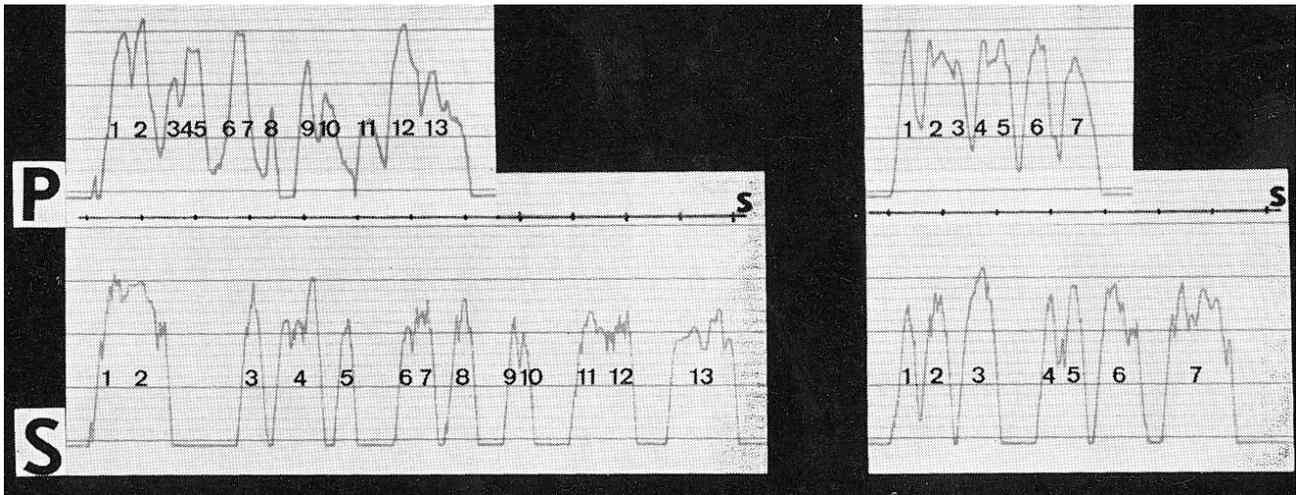


Figure 25 : Représentation des variations d'amplitude de deux phrases turques

P indique la version parlée et S la version sifflée de chacune des phrases (Busnel 1970, p.75). Le découpage par sous-ensembles<sup>36</sup> est indiqué par des chiffres

Deux autres types d'études ont apporté quelques éléments de réflexion sur le sujet :

-Tout d'abord Busnel et al (1989) ont remarqué dans une analyse préliminaire du hmong: "*sur l'oscillogramme d'une phrase on voit qu'entre les groupes d'énergie qui correspondent aux syllabes, des interruptions totales du signal marquent la place des consonnes sourdes. Mais on rencontre aussi très souvent sur ces frontières syllabiques des diminutions relatives d'énergie que l'on est en droit d'attribuer à des qualités consonantiques [...]. Une étude approfondie de ces divers indices reste à entreprendre et nécessiterait, pour être rigoureuse, d'être faite à partir d'enregistrements sur un langage sifflé encore bien vivace.*" (Busnel et al 1989, p.43).

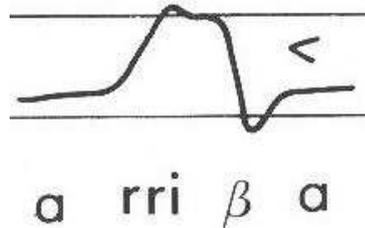
-Deuxièmement, en ce qui concerne le silbo, quelques éléments de détails peuvent être extraits des analyses publiées :

a) En ce qui concerne le voisement : Classe (1957) a fourni des éléments de détail sur le silbo, indiquant que les voisements sont parfois réalisés par une amplitude continue. Il a par exemple remarqué que c'était le cas de la consonne [β]. Plus récemment Rialland (2003) a confirmé l'approche de Classe grâce à une analyse sur oscillogramme : "*We observe that the Ho (fréquence fondamentale du sifflement) is not interrupted in the realisation of whistled [aβa] and [aγa]*".<sup>37</sup> Nous avons pu confirmer cette remarque pour le grec et le turc et le schéma de la Figure 26 l'illustre pour l'espagnol.

<sup>36</sup> Ces sous-ensembles que nous appelons unités de parole sont la plupart du temps syllabiques ou bisyllabiques.

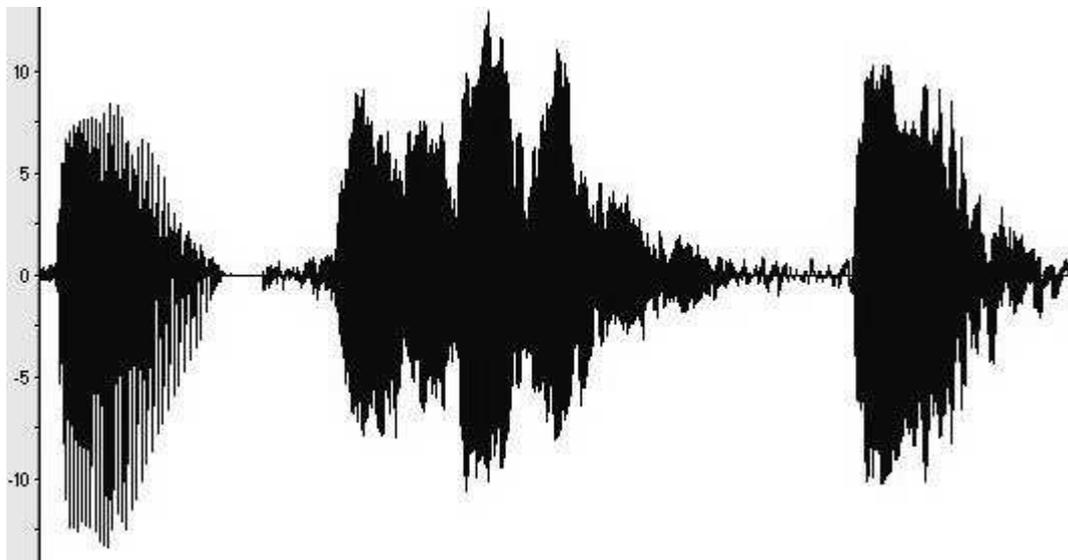
<sup>37</sup> Comme nous le verrons dans la partie typologie, d'après Classe (1956, 1957), [γ] n'est jamais réalisé en Gomero parlé, contrairement à l'espagnol Castillan, ce qui explique qu'elle soit réalisée comme un [β].

b) Classe (1955) a remarqué que le [r] et le [rr] étaient réalisés différemment (Busnel et Classe, 1976). Rialland a confirmé ces remarques: *"Enveloppe modulations do not only carry over the stop vs. continuant distinction in Silbo. They also transpose tapped r and trilled rr quite well, reproducing their typical alternation of peaks and valleys. [...] The envelope also contains cues to the distinction between nasal and voiced stop continuants but these cues are unstable and the existence of this contrast in Silbo is debatable"* (Rialland, 2003, p.2132).



**Figure 26 : Illustration de deux continuités d'amplitude (Busnel et Classe, 1976, p.66)**

c) En ce qui concerne les différences, Rialland (2003) a trouvé qu'il existe quelques différences notables entre le sifflement et la voix, en particulier dans le cas des consonnes qui utilisent des fréquences très hautes de la voix que le sifflement semble éliminer: *"the signal envelope is different in speech and Silbo as spoken fricatives and affricates involve friction noise, triggering a peak in the envelope"*. (Ibid. p.2133). Nous n'avons pas pu confirmer cette différence en grec ni même en silbo comme l'illustre la Figure 27.



**Figure 27 : Illustration de l'enveloppe de la fricative [f] en version parlée**

(premier groupe d'énergie : mot grec /fao/) et sifflée (mot grec /fao/ pour le 2<sup>ème</sup> groupe d'énergie puis syllabe /far/ du mot espagnol « farmacia » pour le 3<sup>ème</sup> groupe d'énergie). On constate une grande similarité dans la forme d'attaque de l'enveloppe

## Analyse complémentaire sur les langues tonales.

Pour les langues tonales nous avons observé que les unités de parole visibles sur oscillogramme sont toujours réparties de manière syllabique ou bisyllabique en respectant les unités de paroles de la voix parlée. Par contre la forme sifflée segmente plus clairement l'énonciation.

Nous avons initié une analyse comparative de l'amplitude des langues à tons sifflés et parlés. Comme le montrent les oscillogrammes Figure 28 et Figure 29, au premier abord aucune corrélation directe n'apparaît mais ceci n'exclut pas qu'une analyse plus fine ultérieure décèle des éléments de comparaison. En effet, même si les rapports relatifs des énergies sonores et de l'enveloppe de chaque syllabe sont souvent différents de ceux de la voix parlée, il se peut qu'une cohérence interne à chaque syllabe puisse émerger à partir de considérations perceptives, par exemple.

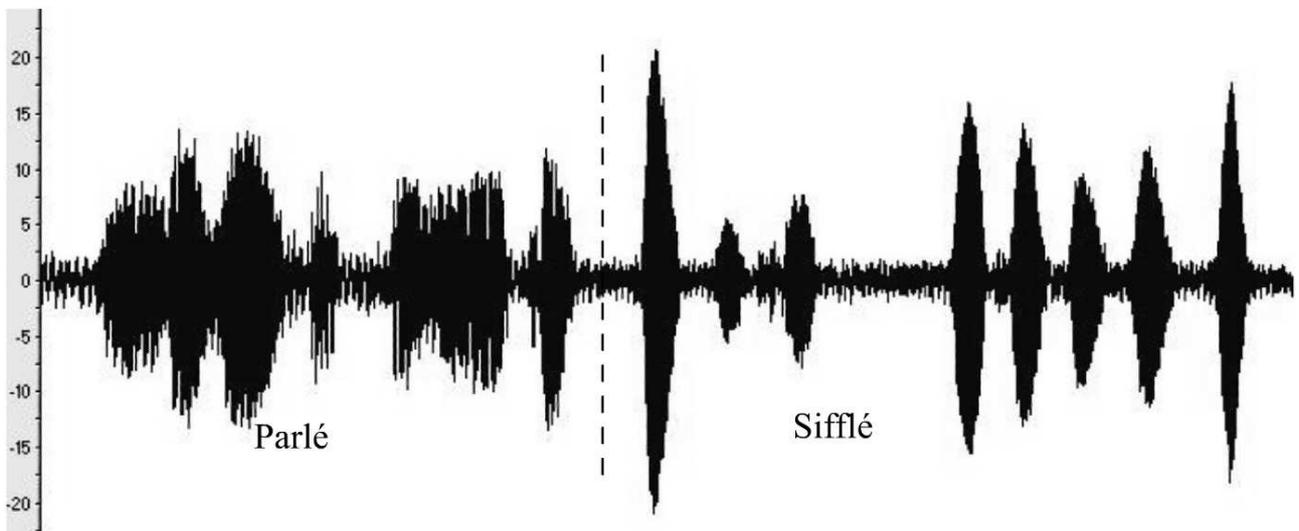


Figure 28 : Oscillogramme parlé puis sifflé d'une phrase Mazatèque

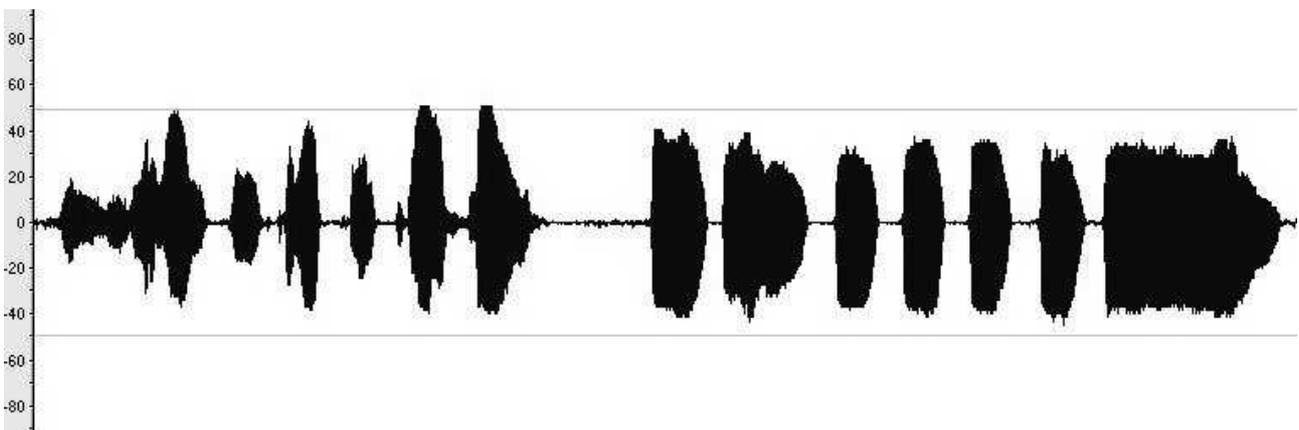


Figure 29 : Oscillogramme parlé puis en feuille sifflée d'une phrase hmong : « Lub sij hawm twg koj thiaj tshuab nplooj ?<sup>38</sup> » (Notation RPA)

<sup>38</sup> Traduction de Chô ly : « A quel moment siffles-tu avec une feuille ? »

### 2.3.7.1.3. Conclusion et perspectives

L'abilité du sifflement langagier à reproduire l'amplitude de la voix dans une amplitude sifflée de dynamique plus étroite souligne à nouveau un aspect relatif de la parole humaine. Un des avantages principaux de ce phénomène est le fait que la parole sifflée émerge plus facilement du bruit de fond dans sa totalité.

L'amplitude sifflée, comme celle de la voix parlée, présente des caractéristiques de variabilité inter-locuteurs ou entre plusieurs énonciations d'une même phrase par un même locuteur. Cependant très souvent, des éléments observés en détails dans l'amplitude d'une phrase énoncée en voix parlée peuvent être observés dans son équivalent sifflé. Une analyse exhaustive reste à faire dans ce domaine, nous avons par exemple mis en évidence que les correspondances sont plus flagrantes dans les langues non tonales qui sifflent les traits consonantiques spectraux des consonnes. Les éléments que nous avons réunis lors d'expériences de diffusion sonore peuvent permettre d'avancer dans cette recherche (enregistrements de mêmes phrases énoncées par plusieurs personnes, analyse des effets stables sur les enregistrements réalisés à distance). Nous pensons que l'analyse de l'amplitude doit être couplée avec des considérations temporelles et perceptives car de nombreuses études du langage soulignent son rôle en tant que marqueur de continuité rythmique.

## 2.3.8. Dynamique temporelle: allongement

### 2.3.8.1. Comparaison avec la langue parlée

Lorsque l'on compare la forme sifflée et la forme parlée d'une même phrase, on constate presque toujours que la phrase sifflée est plus étalée dans le temps, signe d'un débit d'information temporel moins rapide. L'effet perceptif le plus direct de ce changement de débit est une clarification de l'énoncé en terme de détachement des unités de parole. Par contre, le rythme général est maintenu de manière robuste. Nous avons abordé l'explication du rôle de cette stratégie à propos de la voix criée. Le sifflement segmente encore plus les éléments qui étaient distingués dans la parole. Il est important de remarquer que le phénomène n'est pas systématique pour des communications à courte distance (moins de 150 m). Il est par contre toujours utilisé à longue distance. Cowan avait constaté cet aspect sur le mazatèque en expliquant que le rythme de la langue sifflée était différent de celui de la langue parlée. Il paraît plus vraisemblable de dire que le rythme propre de la langue est maintenu mais allongé suivant un processus particulier qui n'affecte pas l'intelligibilité et qui sert à optimiser la communication à distance ou dans le bruit.

### 2.3.8.2. Calcul du taux d'allongement

#### 2.3.8.2.1. Analyses précédentes

##### Comparaison de plusieurs langues

Moles (1970, p80) a analysé l'allongement de phrases turques: *"Nous avons fait un certain nombre d'expériences pour fixer objectivement l'allongement moyen du message résultant du passage de la langue parlée à la langue sifflée"*.

A partir de diagrammes que nous reproduisons Figure 30 Moles déduit qu'en turc sifflé "l'allongement paraît de l'ordre de 22% à 25 %, en gros, un quart". Il remarque également que si pour toutes les langues un allongement peut être constaté, il n'est pas le même suivant les cultures: "par comparaison avec les autres langages sifflés que nous avons étudié (Gomera, Vallée d'Aas), [...] l'allongement moyen paraît nettement plus faible: pour la vallée d'Aas, il était de l'ordre du double, non compris le signal d'appel qui précède la phrase et que nous n'avons pas retrouvé à Kusköy" (ibid., p.81).

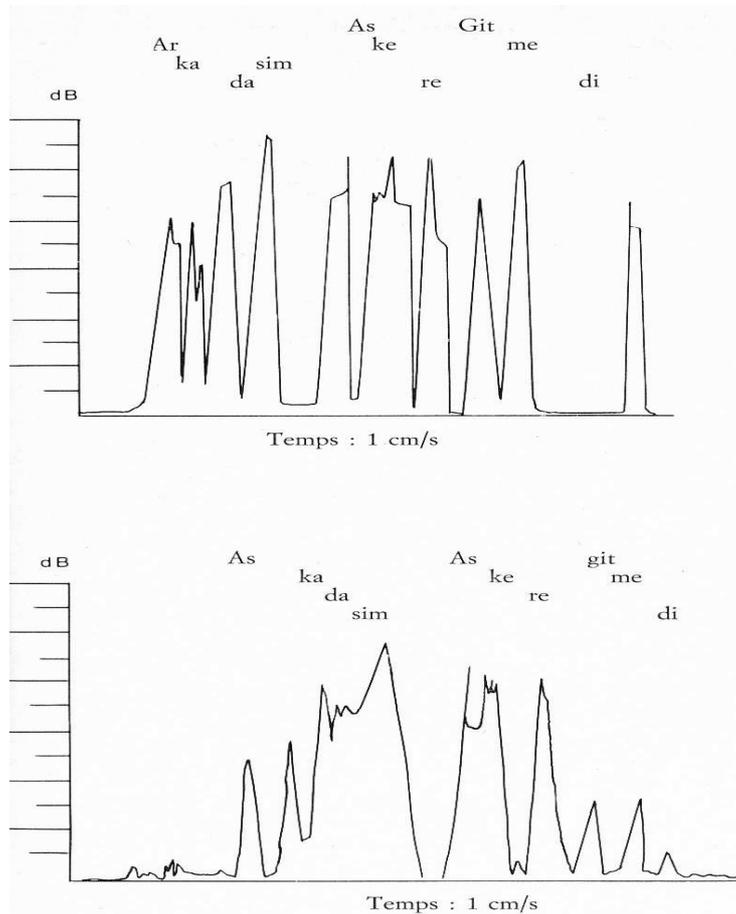


Figure 30 : Diagramme temporel obtenu sur sonomètre dans les années 60-70

### Maintient de caractéristiques d'élocution personnelles :

Par contre, d'un siffleur à l'autre, le taux d'allongement varie très peu: "Dans l'ensemble nous n'avons pas relevé de différences notables entre les différents siffleurs [...]; ce point est intéressant puisque le débit d'élocution de ceux-ci pour les mêmes phrases était notablement différent et qu'il paraît donc y avoir un lien entre le passage du mécanisme de parole au mécanisme de sifflement" (ibid., p.81) Le signal sifflé de ce type de langue s'appuyant essentiellement sur l'articulation, chaque personne garde ses propres habitudes de dynamique d'élocution quand il passe à la langue sifflée.

## 2.3.8.2.2. Analyses complémentaires à partir de notre corpus

### Des différences linguistiques corrélées aux distances de pratiques

Pour mesurer le taux d'allongement, nous avons observé le signal mot à mot puis l'ensemble de la phrase. Pour chaque mot nous avons mesuré la durée de la version parlée et la durée de la version sifflée. A partir de ces deux valeurs nous calculons l'allongement en pourcentage. Notre corpus n'est pas le plus propice à faire des calculs statistiques d'allongement car nous avons fait varier les distances de communication et les personnes traductrices. Or, au vu de nos résultats, l'allongement change avec la technique et la distance de communication. D'autre part, comme nous l'avons vu, chaque personne a un débit d'élocution propre.

Cependant comme Moles, nous avons mis à jour des différences de pratique suivant les langues. Elles sont en fait corrélées avec les usages les plus communs. Par exemple, les Mazatèques utilisent fréquemment le sifflement de près, ils allongent donc en général beaucoup moins le sifflement que les Gomero qui eux pratiquent en majorité le sifflement à grande distance.



**Figure 31 : Même phrase prononcée deux fois par le même locuteur mazatèque suivant deux techniques**  
**La technique très longue distance est celle de gauche et la technique courte distance est celle de droite**

### Intérêt de l'introduction de nouveaux types de langues

Certaines des langues tonales que nous avons introduites dans le corpus d'étude ont des distinctions phonologiques de quantités<sup>39</sup>. En effet la distinction entre voyelles longues et voyelles courtes existe en surui et en akha et est fidèlement respectée en sifflements. Le corpus de surui est trop restreint pour nous permettre de conclure car il n'est constitué que de mots indépendants. Par contre le corpus du akha montre que les voyelles longues sont régulièrement prolongées de manière proportionnellement plus importante que les voyelles courtes.

<sup>39</sup> Dans cet emploi linguistique du terme, quantité : durée

## Analyse détaillée de l'allongement

### Les modalités de l'allongement

#### a) *Allongement sur les voyelles*

L'observation plus détaillée de l'allongement de la parole sifflée permet de remarquer qu'il n'affecte pas tout le signal dans les mêmes proportions : ce sont principalement les voyelles qui sont concernées. Il semble que c'est cela que Moles veut dire lorsqu'il remarque: "*Le pourcentage d'allongement correspond d'une part à une diminution du débit d'information, d'autre part à une répétition de la forme sonore par rapport à un délai de perception immédiate*".(Moles 1970, p.80). D'après nos mesures, la seule forme sonore qui est maintenue -ou répétée- est la fréquence de la voyelle.

On observe plusieurs phénomènes récurrents :

- plus la communication est lointaine, plus le siffleur allonge les voyelles.
- les allongements les plus conséquents sont en général en fin de mots.
- à la Gomera, il n'est pas rare qu'un siffleur maintienne une voyelle sur deux secondes, marquant une pause sifflée dans le cours de la phrase puis il reprend une élocution plus conventionnelle. Nous avons observé ce type de comportement dans le cas de communications dans des endroits très réverbérants (guides d'ondes de vallées aux parois rocheuses dénudées). Il nous a été relaté aussi pour un usage différent : l'annonce d'un décès qui est un sifflement très lent et immédiatement caractérisé par les siffleurs.
- En akha la technique de la feuille sifflée utilise souvent des voyelles longuement prolongées mais seulement sur les voyelles longues. C'est sur ce type de voyelle que des ornements peuvent avoir lieu à des fins esthétiques.

L'allongement des voyelles a donc deux raisons principales : un effet d'amélioration de l'intelligibilité du signal à distance et parfois une recherche d'un effet esthétique voire émotionnel.

#### b) *Allongement des consonnes*

Les consonnes ne semblent pas affectées en première approximation par l'allongement. Pour mesurer un tel effet de manière précise, il est nécessaire d'adopter une technique de calcul commune pour la voix parlée et la parole sifflée. Ce type d'approche n'a pas pu être développé au cours de cette thèse mais est en cours d'élaboration, en particulier grâce à une analyse de la réduction d'information avec la distance dont nous présenterons les principes dans la dernière partie de cet exposé et une caractérisation paramétrique du signal grâce à un logiciel dédié au sifflement humain (introduit en fin d'Annexe **Erreur ! Source du renvoi introuvable.**). Quelques études sur la voix parlée portant sur des effets de ralenti/accélération temporel adaptatif ont montré la sensibilité des consonnes aux modifications temporelles. Par exemple Verfaille (2003) a souligné qu'on ne peut ralentir une voix parlée ou chantée que sur les voyelles et qu'il est nécessaire de conserver les consonnes telles quelles pour conserver la compréhension de l'énoncé. Si les consonnes sont ralenties il a observé des transferts de propriétés (par exemple les /k/ deviennent des /g/ et les /t/ des /ts/).

c) *Discussion*

L'allongement de la parole sifflée par rapport à la voix parlée donne, à notre avis plusieurs indications intéressantes :

-D'autre part, le phénomène d'allongement n'est pas propre à la langue sifflée mais il concerne le phénomène du langage en général. Nous l'avons également mesuré dans nos expériences sur la voix criée lorsque celle ci a pour but d'atteindre de très longues distances (Figure 31). Dans ce cas également, une observation plus détaillée des différences d'allongement indique que ce sont les voyelles seules qui portent ce phénomène.

-Le fait que seules les voyelles soient affectées de manière évidente indique que les consonnes et les voyelles exploitent le domaine spectro-temporel de manière différente. Cet aspect n'est pas nouveau. La conséquence directe de cette observation est que pour chacun de ces aspects de la parole il serait préférable d'utiliser des méthodologies d'observation adaptées. En effet, les modulations consonantiques semblent nécessiter une information temporelle précise tout en réalisant des modulations fréquentielles relatives alors que les voyelles semblent jouer un rôle essentiel au niveau fréquentiel tout en ayant des durées relatives.

## **2.4. Conclusion**

Lors de cette partie nous avons pu confirmer avec des exemples de terrain puis des arguments de physique acoustique que le sifflement était avant tout bâti pour la communication à grande distance.

Nous avons montré que le choix du sifflement s'inscrit dans la continuité des stratégies développées par la voix lorsque l'interlocuteur s'éloigne. Ce phénomène quasi-intuitif procède par l'effet de l'augmentation de l'intensité (effet Lombard), de la fréquence et de la durée d'élocution des voyelles. Au court de cette triple augmentation des paramètres variables fondamentaux de la parole, la voix atteint ses limites biologiques au niveau des cordes vocales entre 90 et 100 dB. Le sifflement est une réponse à cette situation qui permet logiquement de poursuivre la logique développée et donne un plus grand confort d'élocution dans les distances habituellement couverte par la voix criée, ce qui permet des dialogues plus longs. Cette stratégie de communication qui demande un léger entraînement est principalement pratiquée dans les milieux sociaux se trouvant souvent en situation de crier. Le sifflement, utilisé par de nombreux oiseaux dans les mêmes conditions, fournit alors un véritable système de télécommunication optimisé au niveau bioacoustique tout en maintenant la souplesse de la parole humaine. L'intensité, la fréquence et la durée sont des paramètres qui vont donc s'adapter chacun à la distance, au bruit et au terrain dont le siffleur a une connaissance implicite et explicite riche. Leurs aspects relatifs permettent une compression en amplitude et en fréquence. Ils vont aussi permettre à la structure de la langue de s'émanciper dans un nouveau domaine tout en gardant la trace de l'origine qui est la voix. Le sifflement donne en effet un cadre d'expression à chacun des systèmes linguistiques en fonction des particularités qui le caractérisent au niveau fonctionnel, c'est à dire, en premier lieu au niveau de l'intelligibilité.

Cette étude nous invite donc à réaliser deux analyses complémentaires, l'une sur la structure des langues sifflées et l'une sur la perception des sons et l'intelligibilité du langage.

Enfin nous avons les moyens de préciser la définition des sifflements de parole humaine : chez les êtres humains le sifflement est essentiellement caractérisé par la fréquence fondamentale qui consiste en une bande de fréquence étroite (variant typiquement entre 1000 et 4000 Hz) qui dure un certain temps et est modulée en amplitude et en fréquence. La parole sifflée résiste bien à la réverbération et lutte efficacement contre le bruit ambiant. Elle présente une dynamique en amplitude inférieure à celle de la voix classique et souvent une réduction du débit d'élocution

## **CHAPITRE 3. ANALYSE TYPOLOGIQUE**

### 3.1. Intérêt de la typologie pour les études comparatives

La description typologique est un moyen puissant pour faire émerger les comportements communs des systèmes linguistiques des langues. Notre approche se situe dans la ligne des analyses du langage qui considèrent que les lois d'organisation à l'intérieur des systèmes phonologiques peuvent en grande partie être mises en évidence avec les propriétés phonétiques (concept d'« *integrative phonology* » Ohala (1991, p.10-11)). Cette démarche est justifiée car toutes les langues du monde sont issues de contraintes productives et perceptives communes à tous les êtres humains. Or, on sait aujourd'hui que le nombre d'éléments utilisés dans les langues varie dans des limites relativement restreintes. Dans ce cadre, de nombreux travaux ont pu comparer des langues entre elles, constater leurs parentés et observer que des stratégies communes d'encodage sont parfois choisies par des langues éloignées : “*Some system types are apparently impossible, others common, and others rare; and almost all seem to be built along certain basic principles*” (Lass 1984, p.134).

Aucune forme sifflée de langue n'a été observée indépendamment d'une forme parlée car ce sont des transpositions de la voix classique. Il s'ensuit que leur description typologique est un des moyens les plus naturels d'obtenir un nouveau point de vue sur les principes de base qui guident les systèmes linguistiques humains. Ils ont d'autres avantages: leur réalité phonétique est suffisamment éloignée de la voix parlée pour permettre de dégager de nouvelles problématiques et leur cadre d'usage est assez proche de la langue quotidienne pour définir un système très varié. D'autre part, tous les aspects que nous avons mis en évidence dans les chapitres précédents semblent montrer que nous avons réuni assez d'éléments pour envisager de revisiter les implications de l'expression utilisée par Classe en titre d'un de ses articles: « *Langues sifflées, squelettes informatifs du langage* » (Classe, 1963, p.129), expression que nous considérons encore comme une question.

## 3.2. Bilan des langues sifflées connues

### 3.2.1. Langues analysées, partiellement étudiées

En réalité, peu de langues sifflées sont couramment connues du milieu scientifique. La liste que nous avons établie à partir d'une enquête bibliographique (Meyer, 2004) puis d'une enquête de terrain, témoigne pourtant du fait que ces langages sont largement plus répandus que l'on ne pourrait le croire. L'ensemble des formes sifflées utilisées spontanément couvre un large éventail de familles linguistiques. Le Tableau 9 contient l'ensemble des langues dont nous avons pu vérifier l'existence présente ou passée, soit directement en rencontrant les populations, soit grâce à des documents sonores, soit à travers des publications de recherche.

**Tableau 9 : Bilan des langues sifflées rencontrées lors de notre enquête 2003-2004**

Langue	Famille	Tons, accent, voyelles	Stratégie Sifflée (sachant que toutes reproduisent le rythme de la parole)	Aspects de la forme sifflée étudiés dans le passé	Vitalité actuelle et origine	Travaux de référence
<b>a) Langues non tonales</b>						
1. Silbo Gomero: espagnol sifflé de l'île de la Gomera (Canaries, Espagne)	Romane, Indo européenne	Sans tons, Accent tonique	-Hauteur Brute <sup>40</sup> des voyelles et des consonnes -quelques éléments d'intonation (accent)	Linguistique Acoustique Neurologique	En perte de vitalité depuis 1970; en cours de revitalisation depuis 1994 (enseignement scolaire obligatoire en primaire) Le silbo espagnol est une pratique héritée de la langue sifflée des Guanches Berbères qui vivaient aux Canaries avant l'arrivée des Espagnols.	Classe (1955); Busnel et Classe (1976); Rialland (2003) Carreiras et al (2005)
2. Espagnol de Tlaxcala (Mexique)	Romane, Indo Européenne	Sans tons	-Hauteur Brute des voyelles et des consonnes -quelques éléments d'intonation (accent)	Ethnologique Linguistique	Nous n'avons pu rencontrer cette forme sur les lieux décrits dans le passé. Mais il se peut qu'elle survive dans un des villages de la région. La région a subi un fort exode à cause du déboisement et de la désertification. Cette forme sifflée de l'espagnol semble avoir été héritée de formes sifflées Otomis.	Wilken (1979)
3. Béarnais d'Aas (Pyrénées, France)	Romane, Indo européenne	Sans tons, Accent jouant un rôle important	-Hauteur Brute des voyelles et des consonnes -quelques éléments d'intonation (accent)	Acoustique Ethnologique	Langue morte en 1999, en perte de vitalité depuis la deuxième guerre mondiale.	Busnel et al (1962a), Arripe (1985)
4. Turc (Turquie, région de Trabzon au bord de la mer Noire).	Oural Altaïque, Turcic	-Sans tons, - Harmonie vocalique forte Accents	-Hauteur Brute des voyelles et des consonnes -quelques éléments d'intonation (accent)	Acoustique Linguistique Psycholinguistique (tests d'intelligibilité)	Langue en perte de vitalité récente. Quelques efforts de revitalisation sont engagés localement: par exemple un festival est organisé chaque année en l'honneur de « kUSDILI »: la parole sifflée.	Monographie pluri-disciplinaire: Busnel et al (1970)

<sup>40</sup> La notion de « Hauteur Brute » évoque l'attribut de la perception du timbre des voyelles, la notion de « Hauteur Fondamentale » évoque l'attribut de la perception de la fréquence fondamentale. Ces deux notions sont expliquées en détail au chapitre sur l'Intelligibilité (§4.2.2.2.1).

Langue	Famille	Tons, accent, voyelles	Stratégie Sifflée (sachant que toutes reproduisent le rythme de la parole)	Aspects de la forme sifflée étudiés dans le passé	Vitalité actuelle et origine	Travaux de référence
					A cette occasion un concours scolaire à lieu.	
5. Grec (Grèce, île de Eubée)	Indo européenne	Sans tons, Accent à degré de liberté intermédiaire	-Hauteur Brute des voyelles et des consonnes -quelques éléments d'intonation (accent)	Linguistique (uniquement vocalisme)	Langue en perte de vitalité récente mais avec écart générationnel très fort. L'école a fermé en 2004 à cause de l'exode rural et du vieillissement de la population du village.	Xirometis et Spyridis (1994) Charalambakis (1994)
6. Tepehua (Mexique)		Sans tons	-Hauteur Brute des voyelles et des consonnes -quelques éléments d'intonation (accent)	Linguistique	Langue presque morte en 2003 dans les zones où elle a été décrite dans le passé: une seule vallée sur les trois vallées de culture Tepehua n'a pas été explorée lors de notre passage. Il se peut que la langue sifflée survive dans cette vallée.	Cowan (1976)
7. Yupik (île St Lawrence, Sibérie, Alaska)	Eskimo-Aleut	Sans tons Trois types d'accent	-Hauteur Brute des voyelles, -Éléments d'intonation -Articulation des consonnes	Aucune, Langue sifflée révélée au public scientifique en 2005	A cause de l'isolement sibérien de la population, la zone où la langue sifflée Yupik est pratiquée est une des rares où la langue Yupik est encore la langue maternelle des enfants.	
8. Kickapoo (Mexique)		Classée sans ton système accentuel marquant des hauteurs tonales	-Intonation (poids) des voyelles -Hauteur Brute des consonnes	Linguistique (introduction) Ethnologique		Voorhis (1971)
9. Chepang (Népal)	Tibeto-Birmane	Classée sans ton système accentuel marquant des hauteurs tonales	-Intonation (poids) des voyelles -Hauteur Brute des consonnes	Linguistique	Très peu de siffleurs poursuivent la pratique qui est essentiellement dédiée à la chasse. Les informateurs siffleurs de Ross Caughley sont encore vivants. Il est actuellement très difficile de se rendre sur place à cause de troubles politiques dans cette zone du Népal	Caughley (1976)
10. Wam (Papouasie Nouvelle Guinée)		Sans tons	D'après le papier de Nekitel : intonation (poids des voyelles) Pour les consonnes nous ne pouvons conclure à l'heure actuelle	Ethnologique Linguistique	Le système sifflé a été adopté par des langues voisines comme le Abu?Arapesh. En 1988, à l'époque de la collecte des données de Nekitel, le système était encore largement répandu	Nekitel (1992)
11. Wayāpi (Amazonie: fleuves Jari et Araguari (Brésil), fleuve Oyapock (Guyane française))	Tupi	signalée comme non tonale	Non indiqué	Ethnologique	Encore utilisée en 1988 dans les villages éloignés des villes dans l'Oyapock. Utile lors de la pêche, de la chasse ou même dans le village. Dans la partie Brésilienne, le patrimoine oral Wayāpi donne lieu à des projets de documentation	Beaudet (1997)
<b>b) Langues tonales</b>						
12. Surui (Amazonie, Brésil)	Monde	2 tons	-Hauteur des tons -Hauteur brute des consonnes	Aucune, La version sifflée a participé à la description linguistique de la version	Utilisé encore aujourd'hui pour les communications à distance en forêt (chasse, pêche)	

Langue	Famille	Tons, accent, voyelles	Stratégie Sifflée (sachant que toutes reproduisent le rythme de la parole)	Aspects de la forme sifflée étudiés dans le passé	Vitalité actuelle et origine	Travaux de référence
				parlée (Com. pers. De Lacerda 2004)		
13. Gaviaõ (Amazonie, Brésil)	Mónde	2 tons	-Hauteur des tons -Quantité des voyelles (durée des tons) - Peut être certaines consonnes (à confirmer)	Aucune, La version sifflée a participé à la description linguistique de la version parlée (Moore 1998)	Utilisé encore aujourd'hui dans la forêt pour les communications à distance	
14. Pirahã (Amazonie, Brésil)	Sud Américaine, Paez	2 tons	-Non indiqué	Ethnologique, la version sifflée a aidé la description linguistique de la version parlée (Everett 1986)	Le sifflement est essentiellement utilisé pour la chasse	
15. Banen ou Ndiki (Cameroun)		4 tons	-Hauteur des tons		En 1955 Dugast signalait déjà que le sifflement n'était bien pratiqué que par les anciens et quelques rares jeunes	Dugast (1955)
16. Moba (Togo)		tonale	-Hauteur des tons	Linguistique	Employé par les enfants principalement dans les années 1980	Com. pers. Rialland (2003)
17. Ewe (Ghana, Côte d'Ivoire, Togo)	Niger-Kordofaine, Kwa	tonale	Hauteur des tons Système basé sur la technique de transposition des langues tambourinées	Aucune pour le sifflement, Aspect étudié en 2003. In existe de nombreuses études sur le système tambouriné		
18. Mixteque (Mexique)	Nord Américaine, Oto-mangué	3 tons	-Hauteur des tons Modulations des tons	Aucune, première observation scientifique en 2003	Le sifflement est encore bien pratiqué dans quelques communautés de montagne. Il est inexistant dans les populations Mixtèques des plaines. Dans certains endroits il ne sert plus qu'aux prénoms, ce qui semble être un vestige de la pratique du passé.	
19. Mazatèque (Mexique)	Nord Américaine, Oto-mangué	4 tons	-Hauteurs des tons	Linguistique Acoustique	Le sifflement est encore relativement répandu dans les montagnes de la Sierra Mazateca, mais, comme partout, il perd de la vitalité, surtout chez les jeunes. Il n'est pas pratiqué dans les populations mazatèques de plaines.	Cowan (1948) Busnel (1974a)
20. Ari (Ethiopie)		tonale	Hauteurs des tons	Ethno-musicologie	Le sifflement est encore couramment pratiqué en particulier par les enfants dans les montagnes. Il donne lieu à des jeux sonores aussi bien qu'à des communications linguistiques	Fournel (2002)
21. Bencon (Ethiopie)		5 niveaux de tons	Hauteur des tons	Linguistique	Non mentionné	Wedekind (1981)
22. Akha	Tibéto-birmane	6 tons 3 niveaux de tons ; Tons de	Hauteurs des tons, Modulations des contours, durée des tons	Aucune, première observation scientifique en	La tradition de la feuille sifflée est surtout connue pour la séduction et pour le travail aux champs. Nous	

Langue	Famille	Tons, accent, voyelles	Stratégie Sifflée (sachant que toutes reproduisent le rythme de la parole)	Aspects de la forme sifflée étudiés dans le passé	Vitalité actuelle et origine	Travaux de référence
		durées différentes		2003	avons pu observer qu'elle est toujours pratiquée en Thaïlande et au Laos, mais de moins en moins par les jeunes générations.	
23. Chin (Birmanie)		tonale	Hauteurs des tons, Modulations des contours	Ethnologique	Non mentionné	Stern (1957)
24. Bai (Chine, Yunnan)	Sino-tibétaine, Sinitique	8 tons ; contours tonals	Hauteurs des tons, Modulations des contours	Ethno-musicologie	En forte perte de vitalité depuis les années 90	Xian Ming (2002)
25. Hmong (Chine, Vietnam, Thaïlande)	Sino-tibétaine, Hmong-Mien	8 tons répartis sur 5 niveaux; 8 contours	Hauteurs des tons, Modulations des contours	Acoustique (Guyane, Paris)  Ethno-musicologie (Laos, Vietnam)	Feuille sifflée utilisée lors de la séduction ou de la chasse, mais elle sert dans d'autres contextes variés dans les villages les plus traditionnels. Nous avons pu observer en Thaïlande et au Laos qu'elle est encore connue par les générations ayant plus de 30 ans, mais de moins en moins par les jeunes générations. D'après les travaux des ethnomusicologues, la situation est similaire au Vietnam.	Acoustique: Busnel et al (1989)  Ethno-musicologie: Brunet (1972) Kersalé (1997)

### 3.2.2. Autres langues sifflées signalées

De nombreuses autres langues sont susceptibles d'avoir un système sifflé naturel. Quelques auteurs ont publié des listes relatant la présence de langues sifflées dans de nombreuses autres communautés linguistiques ou ont bien voulu nous informer des résultats de leur enquête personnelle sur le sujet.

#### 3.2.2.1.1. Amériques

Ainsi, Hasler (1960) signale la présence de langues sifflées au Mexique dans certains villages Totonèques, Otomis, Amuzgos, Chinantèques, Zapotèques et Chol. Lors de l'enquête que nous avons menée dans ce pays, nous avons rencontré des spécialistes des langues chol, chinantèque et amuzgo du CIESAS (*Centro de Investigaciones y Estudios Superiores en Antropología Social*). Cependant aucun n'a pu nous indiquer un village de siffleurs. Nous sommes également allés dans les régions Totonèques, Otomis et Zapotèques et avons rencontré des responsables culturels locaux. Mais notre enquête est restée là également infructueuse. Seuls quelques informateurs Otomis âgés ont pu nous montrer qu'ils avaient autrefois maîtrisé cette pratique mais leur âge et leur dentition lacunaire ne leur permettait pas de réaliser des enregistrements.

En Amazonie, de nombreuses langues semblent avoir un système sifflé, parfois seulement dans certains villages. Un grand nombre de ces langues nous ont été signalées par les linguistes du « Museu Goeldi » de Bélem.

### 3.2.2.1.2. Afrique

#### **Afrique de l'Ouest et Afrique centrale**

En Afrique de l'Ouest, de nombreuses langues ont été sifflées et le sont peut être encore. Une des particularités de cette zone géographique tient au fait que les rythmes sifflés et joués sur tambours sont souvent similaires car la stratégie de transposition est la même. Les cultures mentionnées par Labouret (1923)<sup>41</sup> sont les suivantes: Ewe, Tshi, Marka, Ule, Daguri, Birifor, Burunsi, Bobo, Bafia, Bape. Plus tard, Huttar (1996) citant Cowan parle des cultures Nchumburu, Yoruba, Gbaya, Doohwayoo, Mofu. D'autre part, toujours en Afrique, Moreau (1997) a réalisé une étude sociolinguistique d'un type de communication sifflée particulier car il n'est qu'épisodiquement lié à la phonétique. Il s'agit du Diola au Sénégal. Dans la même région, elle a signalé la présence de systèmes sifflés chez les Bainuk et les Manjak. Basden (1966) signale la présence de systèmes sifflés chez les Ibos du Nigéria. Parmi toutes ces langues africaines, nous avons pu vérifier l'existence du sifflement articulé chez les Ewe grâce à notre rencontre avec un maître tambour africain en voyage en Europe. Les cultures les plus à même d'avoir conservé ces pratiques sont celles qui ont une volonté culturelle forte de rester attachées à leurs racines. En effet, il existe parfois aujourd'hui un regain d'intérêt pour les racines culturelles spoliées durant la période coloniale, puis par le modernisme, en particulier pour les tambours qui véhiculent aussi une partie des sources musicales de l'Afrique. On peut donc penser que les pratiques sifflées seront ravivées par l'intermédiaire des pratiques tambourinées.

#### **Afrique de l'Est**

Des équipes d'ethnomusicologues travaillant avec l'Unesco ont également récemment signalé la présence de nombreuses langues sifflées dans les régions montagneuses du sud de L'Ethiopie. Elles sont en cours d'étude (com. pers. Fournel 2004).

#### **Asie**

Toute la zone du Yunnan, et du Nord Vietnam, Laos, Thaïlande, Birmanie, utilisait de manière assez répandue le sifflement jusqu'aux années 80 (Xian Ming 2002), aujourd'hui, cette pratique se dégrade rapidement et ne peut être rencontrée que dans des villages isolés. Plus de 26 minorités culturelles sont concernées.

#### **Océanie**

A l'Institut de Papouasie Nouvelle Guinée, Niles (com pers. 2003) a signalé l'existence de nombreux systèmes sifflés comme le gadsup, le bisumarien, le telefol et le folopa. Certains ont donné lieu à des enregistrements publiés.

---

<sup>41</sup> Dans une revue de synthèse marquée par le type de préjugés racistes qui ont longtemps entretenu l'obscurantisme à propos des langues sifflées et des langages tambourinés.

## **Usage secret**

Dans certaines cultures, le langage sifflé est réservé à des pratiques initiatiques comme dans l'état du Sarawak en Malaisie chez les Punans (Bausani 1974) ou chez les Toma et Kissi en Guinée (Germain 1984). Cependant, dans la plupart des cultures la langue sifflée est utilisée à la fois pour les communications quotidiennes et pour les invocations à caractère ludique ou religieux.

### **3.2.3. Conclusion**

Le contenu du Tableau 9 témoigne de la grande variabilité des systèmes linguistiques qui ont développé une forme sifflée. Notre présentation a été faite en respectant la distinction *langues tonale/langues non tonales*, mais la forme sifflée de chaque langue permet d'observer que cette distinction se fait en réalité de manière graduelle en fonction du poids des syllabes et du rôle des accents. Les langues caractérisées par une recherche d'équilibre entre les deux modes de transposition parlé-sifflé décrits jusqu'à aujourd'hui font émerger une troisième tendance.

Si nous faisons un bilan des langues du Tableau 9, les sept premières langues (1. à 7.) adoptent une stratégie de transposition essentiellement basée sur les caractères segmentaux des voyelles et des consonnes, certaines laissent une place remarquable à l'accentuation (comme dans le grec). Un second groupe de langues se distingue du reste, car leurs formes sifflées reposent sur l'articulation des consonnes (caractères segmentaux) conjointement à l'intonation des voyelles ou des tons portés par la syllabe (caractères suprasegmentaux). Les langues 8. 9. et 12. sont concernées par cette stratégie de transposition intermédiaire. Il se peut que 10, 13. et 14. soient aussi de ce type, mais une analyse plus approfondie à partir d'enregistrements serait nécessaire pour le confirmer. Ce groupe est constitué à la fois de langues classées comme tonales (surui) et non tonales (chepang et kickapoo, wam). Le troisième groupe est constitué uniquement de langues tonales (langues 15 à 25). Le sifflement s'appuie alors essentiellement sur des attributs dits suprasegmentaux. Ce groupe n'est pas non plus uniforme. Il regroupe des langues aux structures très différentes comme le mazatèque utilisant des registres de tons et le hmong à contours de tons. Un sous-groupe de cette stratégie de transposition comprend des langues sifflées influencées par la technique des langues tambourinées (langues 15. 17 et peut être 16.). Ce sous-groupe est essentiellement constitué de langues d'Afrique Centrale et d'Afrique de l'Ouest.

Enfin il faut bien distinguer les langues sifflées de la pratique très répandue chez les linguistes qui consiste à faire siffler les tons de la langue afin de les identifier plus facilement car, très souvent, la capacité des locuteurs à siffler les tons ne s'appuie pas sur l'usage d'une forme sifflée de la langue.

### **3.3. Description détaillée de quelques langues**

#### **3.3.1. Organisation générale**

Les regroupements déduits du Tableau 9 et tous les degrés d'adaptations de la pratique sifflée à la structure phonologique des langues dévoilent que les siffleurs exécutent une description phonétique naturelle et subtile de leur langue. Afin d'approfondir la compréhension de ce phénomène nous avons sélectionné 7 langues que nous estimons représentatives de l'ensemble des stratégies développées.

Groupe 1 : 3 langues sans tons dont la transposition sifflée est basée sur le spectre vocalique et les transitions spectrales consonantiques. Il s'agit du grec, du turc, du silbo espagnol.

Groupe 2 : 2 langues ayant un statut intermédiaire de transposition. Le chepang (non tonale) et le surui (tonale)

Groupe 3 : 2 langues tonales dont le système de transposition sifflé s'appuie essentiellement sur les hauteurs séquentielles tonales, sur leurs modulations et sur des formes de contours tonaux. Il s'agit du mazatèque, et du hmong.

Pour chaque groupe, nous avons ajouté en Annexe **Erreur ! Source du renvoi introuvable.** des éléments de description d'autres langues :

Le béarnais et le yupik (groupe 1), le kickapoo (groupe 2), le mixtèque, le bençon et le akha (groupe 3), puis les techniques de transposition sifflées basées sur les techniques des langages tambourinés. Enfin, un bilan comparatif est effectué pour chaque groupe puis une discussion générale compare toutes ces stratégies.

#### **3.3.2. Organisation de la description de chaque langue**

Pour chaque langue nous avons fait une description des éléments segmentaux essentiels (triangle vocalique et tableau de réalisation phonétique des consonnes) et suprasegmentaux (accentuation, intonation ou système tonal) de la version parlée. Puis une étude de la version sifflée est développée. En fonction des possibilités offertes par le corpus réunit certaines analyses des éléments transposés en sifflements ont pu être menées de manière approfondie ou plus succincte. La description des langues tonales du premier groupe est assez uniforme de ce point de vue car les corpus utilisés ont été en grande partie constitués par nos soins et ont donc pu être adaptés aux questions que nous nous posons. Pour les langues du groupe 2 nous nous sommes adaptés aux corpus fournis par des linguistes de terrain. Pour les langues du groupe 3, le système tonal sifflé est décrit en détail et nous ouvrons des pistes pour des études ultérieures.

##### **3.3.2.1.1. Triangles vocaliques et tableaux phonétiques**

Les tableaux phonétiques présentés dans cette partie servent à rendre compte de l'ensemble des représentations phonologiques de chacune des langues. Ils sont issus du travail de nombreux linguistes s'étant immergés dans une langue pour comprendre ses règles suivant un travail d'enquête et de classification

statistique systématique. Etant donné que certaines langues sont décrites depuis peu de temps (langue Surui d'Amazonie par exemple), leur description est susceptible d'évoluer avec l'approfondissement des connaissances des linguistes de terrain. D'autres langues ont une grande variabilité locale en raison du relatif isolement des populations qui les parlent, c'est le cas par exemple du mazatèque qui possède selon les sources de 10 à 14 variantes : dans ce cas nous avons indiqué pour quelle variante de la langue les tableaux ont été établis. Quand nous nous sommes rendus sur place, nous avons essayé dans tous les cas de donner une représentation phonétique fidèle aux parlers des villages que nous avons visités.

Les linguistes considèrent que les sons du langage qui apparaissent dans les manifestations phonétiques peuvent être regroupés en unités prototypiques qu'on appelle les phonèmes de la langue. Les phonèmes sont aussi considérés comme les plus petites entités segmentales distinctives (ainsi en français [l] et [ɫ] sont deux réalisations d'un même phonème /l/. Ces deux types de prononciations sont regroupés car ils ne permettront jamais de distinguer deux mots). Nous verrons au chapitre suivant que la réalité perceptive des phonèmes n'est pas aussi simple que la représentation écrite le suggère.

Pour certaines langues, il arrive qu'une consonne ait des variantes glottalisées ou aspirées distinctives. Dans ce cas, en fonction des descriptions faites par les linguistes, soit le phonème de la glottalisation et de l'aspiration apparaissent dans le tableau phonétique des consonnes, soit nous signalons l'existence de ces variations.

Les triangles vocaliques que nous présentons respectent les normes IPA tout en cherchant à être au plus près de la réalité de la répartition des voyelles les unes par rapport aux autres. Ils serviront donc de base pour expliquer les réductions vocaliques qui ont lieu en sifflement.

### 3.3.2.1.2. Description des formes sifflées

Etant donné que les sifflements s'expriment essentiellement par des modulations de fréquence et d'intensité, pour les présenter dans chaque langue nous avons choisi trois options qui se combinent entre elles :

- Pour les voyelles et les tons, nous avons utilisé des données statistiques de répartition fréquentielle.
- Pour les modulations caractéristiques des consonnes et de certains tons (glides et contours), nous avons parfois recours à une schématisation de la forme du signal.
- Des outils tels que les sonagrammes, des représentations paramétriques fréquentielles ou des oscillogrammes nous permettent de présenter des cas précis de sifflement.

D'une manière générale, nous n'avons pas cherché à extraire des règles phonologiques propres au système sifflé, celles-ci émergeront toutes seules si elles existent. Nous verrons que, comme une langue sifflée n'existe jamais indépendamment d'une version parlée (à notre connaissance), c'est toujours vers le système phonologique de la voix parlée que le siffleur tendra avec plus ou moins de succès suivant sa pratique<sup>42</sup> et

---

<sup>42</sup> La variabilité des performances en fonction de la pratique invite le linguiste à se baser sur les réalisations des meilleurs siffleurs. Or ceux-ci ont une précision extrême dans leurs meilleurs jours qui, compte tenu de l'adaptation de l'oreille à la perception des sifflements (voir §4.2) impose la mesure dans l'établissement d'un système réduit.

suivant le mode de transposition opté pour la langue concernée. C'est pourquoi nous pouvons déjà supposer que les réductions phonétiques qui ont lieu en sifflement sont ancrées dans la répartition phonétique de la version parlée d'origine et sont souvent également présentes en parole naturelle (par exemple en parole rapide).

### **3.3.3. Langues sifflées s'appuyant sur l'articulation des voyelles et des consonnes**

#### **3.3.3.1. Le Silbo: espagnol sifflé de l'île de La Gomera**

##### **3.3.3.1.1. Introduction**

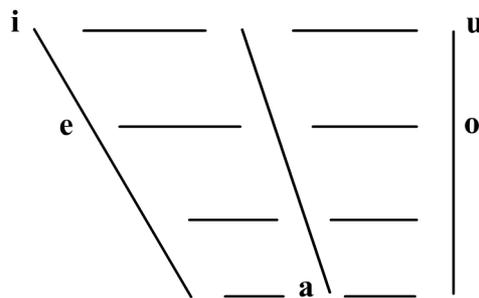
L'espagnol parlé dans les Canaries a de légères différences avec l'espagnol castillan. Il ressemble à un mélange entre l'espagnol andalou et certaines formes d'accents d'Amérique du Sud. C'est pourquoi nous avons choisi de présenter les données de l'analyse phonétique de la version parlée réalisée par Classe (1957, 1976) directement à partir de l'espagnol de l'île de la Gomera. En ce qui concerne les données sur le sifflement celles-ci sont issues du corpus que nous avons constitué en mars-avril 2003 avec l'aide des 2 siffleurs<sup>43</sup> les plus habiles et les plus authentiques ayant acquis le sifflement en même temps que la langue parlée et qui le pratiquent encore quotidiennement.

##### **3.3.3.1.2. Voyelles**

###### **Version parlée**

Voyelles de la Gomera : i, /e, ε/, /α, a/, /o, o, u/

Diphtongues: [ai], [ei], [au], [oi].



**Figure 32 : Triangle vocalique de l'espagnol parlé à la Gomera**

<sup>43</sup> Nous remercions Luis Morales Mendes et Lino Rodriguez.

## Remarques

- Quand [a] et [o] sont en contact il sont souvent assimilés: "Hablo a morales" est souvent prononcé [ablaa: morale],
- En position finale, [o] apparaît souvent comme un [ʔu], par exemple el trigo est prononcé [el trigu],
- En espagnol castillan, les voyelles constituent environ 50% du matériel phonétique et leur distribution n'est pas uniforme: [a] 32%, [o] 21%, [e] 20%, [i] 12%, [u] 7%.

## Forme sifflée

### Généralités sur la répartition

Les voyelles sont sifflées à différentes hauteurs fréquentielles : /i/ est sifflé à la plus haute fréquence, puis dans l'ordre décroissant des fréquences moyennes: /e/, /a/ et /o, u/. Cependant, chaque voyelle n'a pas une fréquence fixe. Celle ci dépend en premier lieu de la distance de communication et donc en partie de la technique de sifflement utilisée mais également des particularités anatomiques de chaque siffleur. De plus, pour une distance et pour un siffleur donné, chaque voyelle couvre un intervalle de valeurs fréquentielles qui permet au siffleur de rendre compte des subtilités accentuelles de la phrase ou d'adapter sa pratique afin de produire le sifflement le plus intelligible possible. Les statistiques de la répartition des fréquences vocaliques de deux siffleurs issus de deux régions différentes de l'île, l'un sifflant à 300 m, l'autre sifflant à plus d'un kilomètre sont présentées Figure 33 et Figure 34. Le **Erreur ! Source du renvoi introuvable.** en Annexe **Erreur ! Source du renvoi introuvable.** donne les valeurs d'origine de ces deux figures.

Voyelles sifflées en Silbo (espagnol): Lino R. siffle à 300 m.

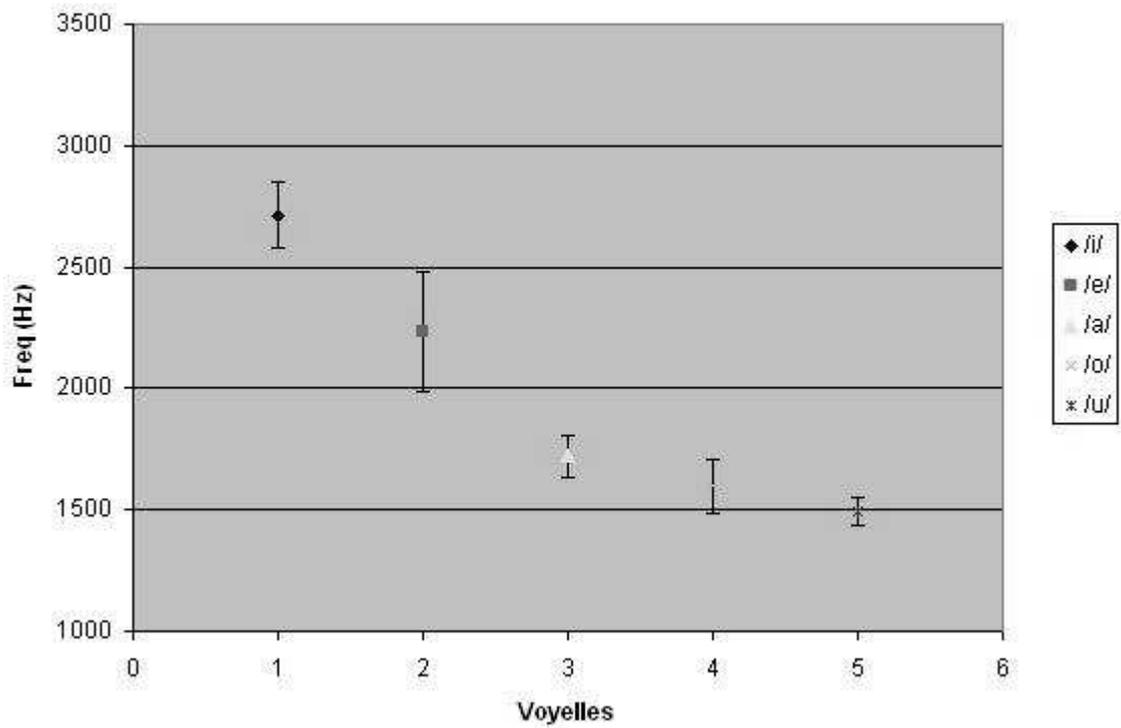
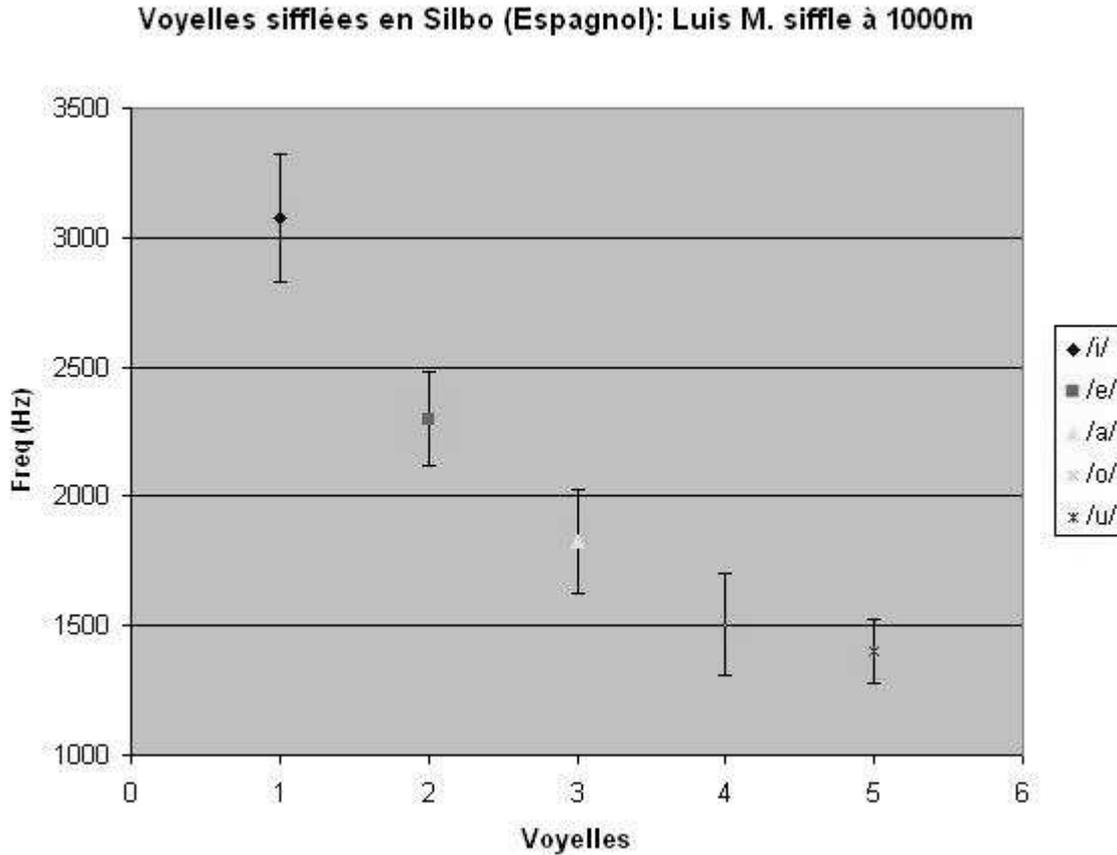


Figure 33 : Répartition des voyelles en Silbo, Siffleur 1



**Figure 34 : Répartition des voyelles en Silbo, Siffleur 2**

La bande de fréquences couverte par chaque voyelle est variable suivant les locuteurs, elle peut atteindre une largeur proche des 1000 Hz. Les écart-types représentés montrent que les voyelles /i/, /e/ et /a/ sont distinguées de manière statistiquement significatives. Le groupe /o, u/ est bien distingué du reste.

### Explication de certaines confusions et distinctions

Les bandes de fréquence de /o/ et de /u/ interfèrent largement au point qu'elles semblent ne former qu'une unité même si les /u/ sont en moyenne sifflés plus bas que les /o/. La rareté des [u] en espagnol peut favoriser cette confusion. La bande de fréquence /o,u/ ainsi constituée interfère avec la partie basse de la bande de fréquences des /a/. Les transpositions sifflées des voyelles [a], [o] et [u] se chevauchent souvent. Mais comme dans la langue parlée, elles sont clairement distinguées quand le contexte ne permet pas de lever les confusions possibles. Bien qu'en général [u] soit de fréquence inférieure à [a] et [o] et qu'il soit distingué par une plus forte concentration d'énergie dans la première partie de l'enveloppe d'amplitude, il arrive qu'il soit sifflé plus haut que [o], c'est par exemple le cas quand il est compris entre deux [o].

Certains chevauchements sont explicables par le fait que le siffleur s'applique particulièrement sur les zones les plus ambiguës en fonction du contexte phonétique. Dès lors, parfois, quand il n'y a pas de confusions possibles, il s'autorise des chevauchements. Par exemple, en position finale, il arrive souvent que /e/ atteigne la fréquence d'un /i/.

## Nasales ?

De la même manière que dans la parole, les voyelles du sifflement ne sont jamais nasalisées. Lorsqu'un /n/ suit une voyelle il s'exprime par une légère montée en fréquence, sauf peut être dans certains cas après un /i/ (voir consonnes)

## Quantité

La quantité des voyelles est assez uniforme sauf pour améliorer la communication et faire des effets esthétiques. En général, il n'y a pas de grande différence entre voyelles longues et voyelles courtes.

## Conclusion

Nos données confirment que les voyelles sont sifflées à des hauteurs relatives. Pour une distance, une technique et un siffleur donné, chaque voyelle couvre un intervalle de fréquences. Quatre bandes de fréquences se distinguent malgré des chevauchements dans les réalisations extrêmes de chacun des intervalles. Dans l'ordre des fréquences décroissantes, il s'agit de I: (i), E: (e), A: (a) et O: (o,u). Ces regroupements sont liés à l'articulation sifflée. Nous ne pouvons pas dire qu'il n'y a que 4 voyelles sifflées tout d'abord car certains siffleurs distinguent clairement le [o] et le [u] quand ils ont besoin de le faire et ensuite car le sifflement ne définit pas de nouvelles paires minimales puisqu'il garde toujours en référence le système d'origine.

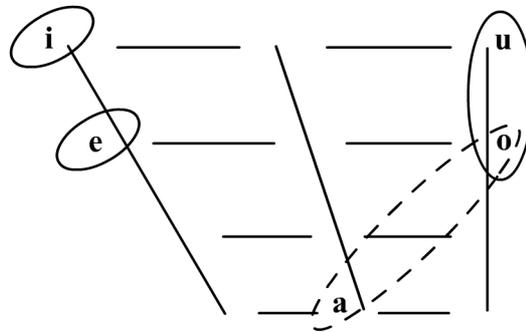


Figure 35 : Triangle vocalique sifflé du Silbo

### 3.3.3.1.3. Diphtongues

Les diphtongues sont traitées exactement comme des paires de voyelles, formant une modulation rapide allant de la fréquence de la première voyelle à la fréquence de la seconde voyelle. Les modulations plus lentes (*glides*) sont réservées aux consonnes.

### 3.3.3.1.4. Accentuation

Dans la pratique du Silbo l'accent est conservé surtout s'il facilite l'intelligibilité du signal. Il est marqué de deux manières différentes suivant les contextes, soit l'accentuation est réalisée par un allongement de la

durée de la voyelle, soit à travers une élévation de fréquence et d'amplitude. Les règles de l'accent tonique de l'espagnol sont respectées en Silbo, dans la mesure où ils le sont en Gomero: en principe, tout mot terminé par une voyelle est accentué sur l'avant dernière syllabe, les mots terminés par une consonne portent l'accent sur la dernière syllabe, sauf si la consonne finale est un /s/ ou un /n/ (marques du pluriel). Dans tous ces cas l'accent est marqué par une augmentation de la fréquence et de l'amplitude de la voyelle sifflée. Il existe quelques exceptions à cette dernière règle, ce sont les mots proparoxytons (Classe 1956). Leur accentuation est marquée par un allongement des durées des syllabes.

D'autre part, l'intonation d'une question modifie également la hauteur des voyelles de façon parfois très sensible sur la dernière syllabe de la phrase. Classe (1956) cite à ce propos l'exemple de la phrase « *Como te llamas tu?* » prononcée [komo te jama tu?] pour laquelle [u] peut alors être sifflé plus haut que qu'un [a], tout en restant dans les fréquences basses.

### 3.3.3.1.5. Consonnes

#### Version parlée

Semi-voyelles: /j, dʒ/, w.

Consonnes: p, /b, β/, t, /d, ð/, k, /g, ɣ/, f, s, x, m, /n, ñ/, l, r, r̄

#### Particularités de la Gomera

Comme pour les voyelles il y a des confusions propres à l'espagnol parlé à la Gomera. Classe (1957) cite les formes suivantes: « Pilar » prononcé [pilað], « el sur » prononcé [el sul], « comer » prononcé [komen], « el que » prononcé [er ke], « silbo » prononcé [sirbo] et beaucoup d'autres. Certains sons du castillan ne sont pas ou sont peu représentés en Gomero comme [ɣ] presque toujours prononcé [g]; [ʎ] et [θ] qui sont remplacés par [j] et [s]. La jota [x] qui est remplacée systématiquement par un [h]. Et le groupe [gw] toujours réduit à [w]. Ces particularités propres à la version parlée de la Gomera sont conservées en sifflement.

En espagnol castillan les consonnes ont une fréquence moyenne de 2,2% par rapport à l'ensemble des sons. Les six plosives sont proches de la moyenne mais les autres ont des répartitions très inégales. [n] se distingue avec 5,62%. Le tableau phonétique récapitulatif de l'ensemble des réalisations des consonnes de la Gomera est présenté en Tableau 10.

**Tableau 10 : phonétique des consonnes de l'espagnol Gomero**

	<i>Bilabiale</i>	<i>Labio dentale</i>	<i>Labio vélaire</i>	<i>dentale</i>	<i>alvéolaire</i>	<i>Post alvéolaire</i>	<i>Retroflexe</i>	<i>Palatale</i>	<i>Vélaire</i>	<i>Uvulaire</i>	<i>Glottale</i>
Occlusive	p b			t̥	t d				k g		
Implosive											
Click											
Trille					r						
Tap					r						
Flap											
Fricative		f			s				x		h
Affriquée						tʃ					
Nasale	m				n			ɲ	ŋ		
Latérales fricative											
Latérale approximante					l						
Approximante			w					j			

### Forme sifflée

Le système des consonnes sifflées est basé entièrement sur la phonétique et non sur les oppositions phonémiques. Des exemples de modulations sont présentés sur la Figure 36. Pour chaque consonne plusieurs types de réalisations sont présentées avec les figures des autres modulations dues à des consonnes dans d'autres langues.

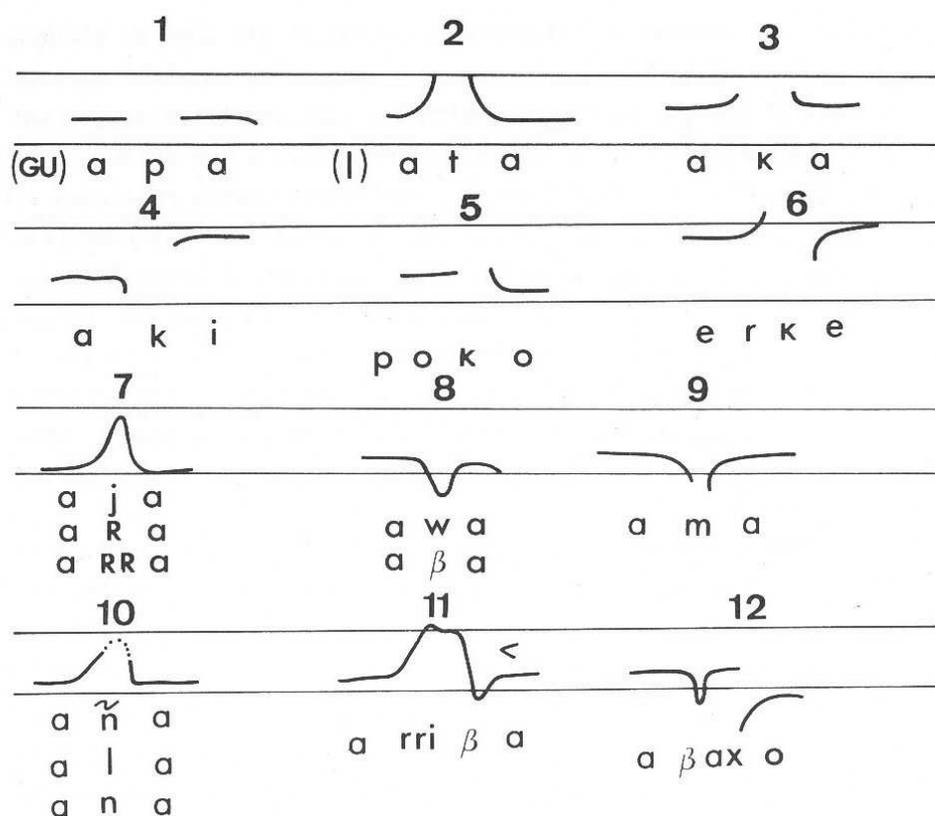


Figure 36 : Exemple de modulations schématisées à partir des spectrogrammes et de l'impression auditive (Busnel et Classe, 1976, p. 66)

**Les directions de modulations peuvent changer pour certaines consonnes (/k/ par exemple)**

-[p], [k] sont tous les deux réalisés par des interruptions soudaines du sifflement presque identiques. Elles sont cependant distinguables à courte distance. Ces interruptions sont le plus souvent réalisées par des occlusions glottales. Classe (1976) reconnaît ne pas pouvoir apporter la preuve de ce fait mais avoue lui-même siffler ainsi. Certains enregistrements que nous avons réalisés en champ proche l'ont déjà confirmé (Figure 7).

-[g] est sifflé comme [k] et [b] est sifflé comme [p], sauf qu'il n'y a pas d'occlusion glottale pour [b].

-[t] et [d] sont réalisés de la même manière, sous la forme d'une montée fréquentielle à partir de la hauteur vocalique qui précède (parfois une descente à partir du /i/), suivie d'un silence et d'une descente vers la hauteur vocalique suivante (sauf pour /i/ chez certains siffleurs).

-[s] ressemble à [t] sur un spectrogramme car l'articulation du [s] entraîne la montée de la langue puis sa redescende mais la pente est moins abrupte. La raison principale est que le lieu d'articulation du [s] est plus rapide à atteindre en position de sifflement et la pression de l'air est moins explosive au relâchement que pour un [t].

-[tʃ] est sifflé comme [t] mais le silence est plus long, ce qui suffit à les distinguer.

-[m] est caractérisé par une dépression de la fréquence, suivie parfois (mais pas systématiquement) d'un court silence puis d'une remontée vers la hauteur de la voyelle de la syllabe suivante. Bien que la prononciation bilabiale ne puisse pas être effectuée, les siffleurs réalisent ainsi une forme sifflée caractéristique. [f] a la

même forme mais comme il est soufflé, il laisse souvent un silence plus long et surtout sa modulation est bien plus lente à tel point qu'elle a une influence perceptible sur la voyelle suivante.

-Lorsque [f] est en début de mot, il y a seulement une longue montée fréquentielle modulant la voyelle. [β] a la même forme mais comme il est voisé dans la version parlée, il n'y a pas d'interruption dans la version sifflée. [x] et [w] sont également réalisées de manière continue avec une dépression intermédiaire, la pente de la dépression est bien marquée.

-[n, ɲ, l, r, r̄, j, ð] sont tous sifflés de manière similaire. La modulation de fréquence effectue une montée plus douce que pour un /t/ ou un /s/ puis une convexité qui entame une redescente qui peut être rapide ou lente. Si la voyelle de départ et la voyelle d'arrivée sont à la même hauteur, la bosse ainsi créée est bien marquée. Par contre s'il y a un écart fréquentiel important entre les deux voyelles, la bosse est atténuée au niveau de la voyelle la plus haute (*Federico* est sifflé [fedeiko]). Il n'est pas rare que le sommet de cette modulation soit marqué par un silence, en particulier pour [n] mais les pentes de l'attaque ou du relâchement sont plus légères que celles d'un [t] ou d'un [d] et sont perçues comme telles par les siffleurs. En fin de mot ces consonnes sont souvent éludées sauf dans le cas de [r, l, n] qui sont alors juste présentes sous la forme d'une montée fréquentielle.

### Particularité des nasales

Ainsi on peut dire que même si le sifflement des consonnes nasales ne se fait pas par l'ouverture de la cavité nasale (car cela provoquerait une perte trop grande de pression de l'air qui empêcherait la production d'un sifflement), [m] est tout à fait identifiable phonétiquement et [n] est souvent différenciée du groupe de consonnes dont elle fait partie car la forme sifflée de sa modulation de fréquence est en général associée à une modulation d'amplitude au sommet de la convexité qui provoque un léger silence. D'après le sifflement, c'est donc une forme intermédiaire entre [t] et [l].

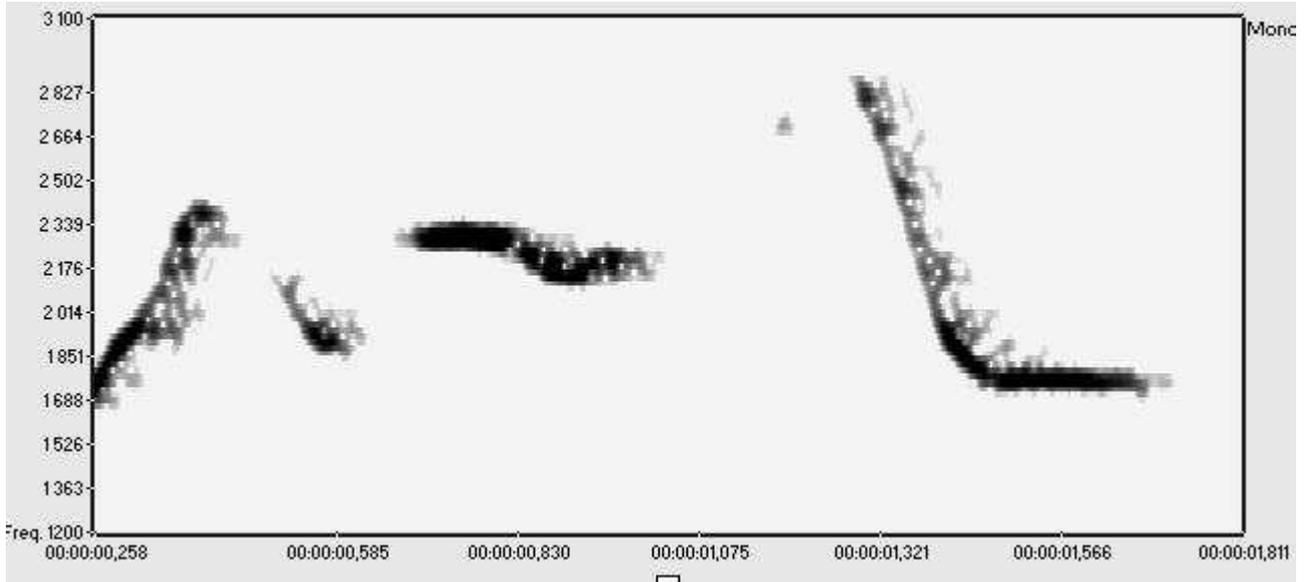
### Particularité des consonnes voisées

En ce qui concerne les consonnes voisées, l'absence de vibration des cordes vocales crée une perte de voisement. Cependant comme le décrit Classe: « *The loss of voicing is compensated by a rather gentle attack, which makes for an audible distinction. At close range, some silbadores may be heard producing a quite loud laryngeal buzz, the vocal chords are then quite close together and the air pressure in the mouth is reduced* » (Classe, 1957, p.978). Les consonnes voisées ont en effet une augmentation et une baisse d'amplitude plus légère que celles des non voisées, cela se traduit par une attaque moins franche que pour les consonnes non voisées, ce qui est souvent perceptible même à distance, malgré la dégradation du signal.

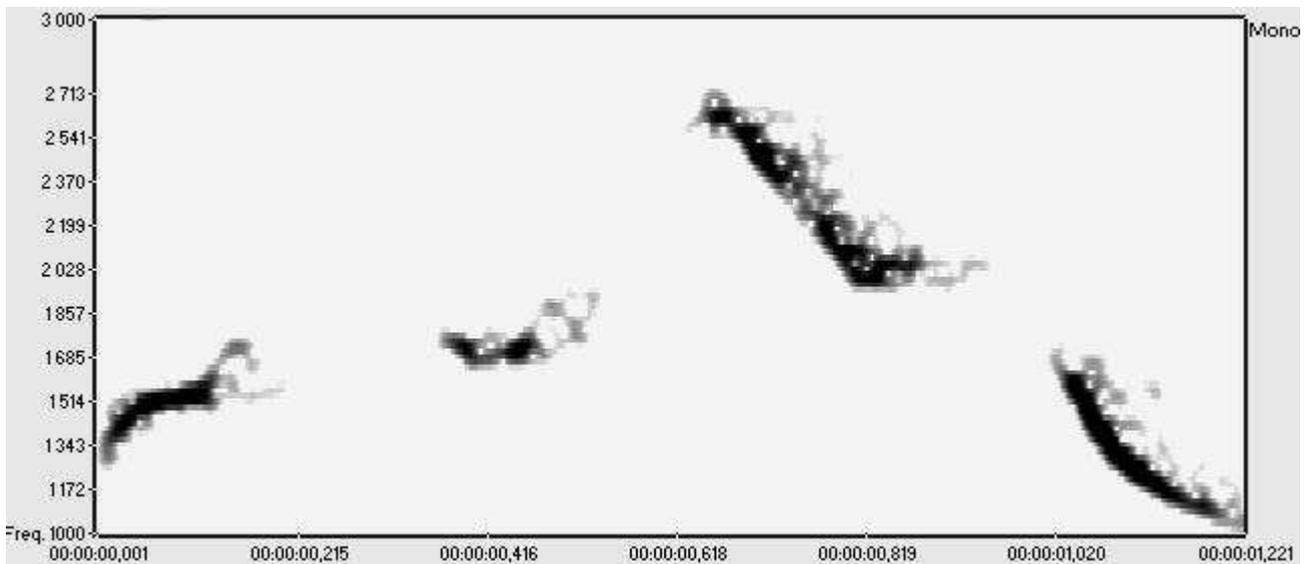
#### 3.3.3.1.6. Les clusters de consonnes

Le nombre de clusters possibles en espagnol n'est pas très élevé. En général, la modulation de chaque élément du cluster est maintenue, au moins dans une forme réduite. Cela se produit dans le mot /erke/

(Figure 36), on peut l'observer pour le mot « *farmacia* » (Figure 37) ou le mot « *montañeta* » (Figure 38). Il existe pourtant quelques cas où dans la pratique, il semble qu'il y a une consonne éludée. En particulier lorsque [r] suit une plosive, il est omis. Ce qui n'affecte pas l'intelligibilité générale. D'après Classe (1955) ceci permet de prévenir les confusions de [pr] et [kr] avec [pj] et [kj] par exemple. Par contre lorsque le [r] précède une autre consonne il est prononcé. D'autre part, si [s] fait partie d'un cluster il est éludé en fréquence mais de même que pour un [r] éludé, il marque la longueur du silence de manière perceptible.



**Figure 37 : Mot de silbo « *farmacia* » réparti suivant les 4 unités sonogramiques: « fa »-« r »-« ma »-« cia »**



**Figure 38 : Mot « *montañeta* » répartie suivant les 4 unités sonogramiques: « mon »-« ta » -« ñe »-« ta »**

### 3.3.3.1.7. Historique de la description du Silbo:

Le Silbo est la langue sifflée la plus connue et la plus largement étudiée. Elle a fait l'objet de plus d'une dizaine de publications scientifiques à elle seule. Les premiers témoins qui relatèrent par écrit son existence étaient les moines faisant partie de l'équipage du mercenaire Jean de Béthancourt. Pour eux, les insulaires. Il

fallut attendre les premiers travaux d'anthropologie de Von Frisch (1867), de Quedenfeldt (1887) et de Verneau (1891,1923) pour que des descriptions plus précises, bien que toujours erronée soient avancées. Ce fut Lajard (1891) qui le premier expliqua correctement la technique de sifflement, en remarquant qu'elle consistait en l'articulation des consonnes et des voyelles plutôt qu'en la reconnaissance de formes figées, voire idéographiques. Ce n'est que dans les années 50 que les premières véritables études linguistiques furent publiées par Classe (1955,1957), à la fois en phonétique et en phonologie. Elles restent la référence en la matière. Il remarqua le premier les regroupements consonantiques et vocaliques effectués par les siffleurs et indiqua qu'ils *étaient purement phonétiques et ne visaient pas à former des paires minimales phonologiques*. Il ne présenta pas de données statistiques c'est pourquoi nous l'avons fait.

Plus tard, grâce au développement des techniques de présentation des données, sous forme de sonagramme, il devint possible de comparer le spectre de fréquence de la voix à celui du sifflement. Alors que la recherche sur les formants de la langue parlée se développait, Brusis (1973) remarqua que les hauteurs fréquentielles des voyelles sifflées du Silbo suivaient le même type de répartition que le second formant des voyelles parlées espagnoles. Nous verrons par la suite que cette comparaison entre le deuxième formant de la parole et le sifflement est valable pour les voyelles du Silbo mais n'est pas aussi claire dans d'autres langues sifflées. Les recherches sur les lieux d'articulation en langue parlée et en turc sifflé (Leroy 1970) inspirèrent une partie des recherches sur le Silbo. Ainsi Rialland (2003) synthétisa ces observations sur le lieu d'articulation et celles de Brusis sur le deuxième formant, elle remarqua en effet que les modulations du deuxième formant de la parole sont affectées par le lieu d'articulation des consonnes de manière similaire à la modulation des consonnes sifflées. Nous verrons dans la partie comparative qui suit l'exposé descriptif des langues non tonales, qu'en ce qui concerne les consonnes, cette comparaison est surtout valable pour les consonnes occlusives prononcées lentement.

Une autre étude marqua l'histoire de l'analyse du Silbo, en particulier dans le milieu scientifique et éducatif des Canaries: Trujillo (1978), de l'Université de la Laguna à Tenerife, développa une théorie réductionniste affirmant que les siffleurs n'utiliseraient que deux voyelles (haute et basse) et quatre consonnes (« CHE », « Ye », « Ge », « Ke »). Cette théorie est radicalement contestée tant pour les voyelles que pour les consonnes, par la grande majorité des siffleurs (dont le professeur et « Maestro » de Silbo Lino Rodriguez), mais également par les données que nous avons recueillies (puisque nous avons vu qu'il y avait au minimum trois bandes sifflées clairement distinguées). Malgré cela elle est utilisée pour sa valeur didactique lors de l'enseignement scolaire en primaire, par l'un des professeurs de Silbo (Le « Maestro » de Silbo Isidro Ortiz). Elle a également été prise comme référence dans l'unique publication de neurosciences sur les langues sifflées, réalisée dans la même université que celle de Trujillo (Carreiras et al 2005).

### **3.3.3.2. Le grec du village d'Antia (île d'Eubée)**

Le grec parlé dans le village d'Antia de l'île d'Eubée (Evia) a un accent distinct de tous les autres parlers de l'île. Malgré cela ses réalisations phonétiques sont cohérentes avec celles du grec moderne. Les tableaux phonétiques que nous présentons pour décrire la langue parlée s'inspirent en grande partie de la synthèse de données réalisée pour la base de donnée Upsid disponible au laboratoire DDL. En ce qui concerne la langue sifflée, hormis des considérations ethno-linguistiques générales (Xirometis et Spyridis 1989, Charalambakis

1994), seule la répartition des voyelles (Xirometis et Spyridis 1994) a été étudiée dans le passé. La description que nous proposons ici a été réalisée à partir d'enregistrements réalisés en Avril-Mai 2004 dans le village d'Antia grâce à la collaboration d'un siffleur et d'une siffleuse ayant appris le sifflement en même temps que la version parlée<sup>44</sup>.

### 3.3.3.2.1. Voyelles

#### Forme parlée

Les voyelles prononcées dans le village d'Antia sont les suivantes /i/, /ε, e/, /α/, /ɔ, o/, /u/. Etant donné que [e] et [o] sont très rares à Antia, la description du triangle vocalique de la Figure 39 est tout à fait appropriée.

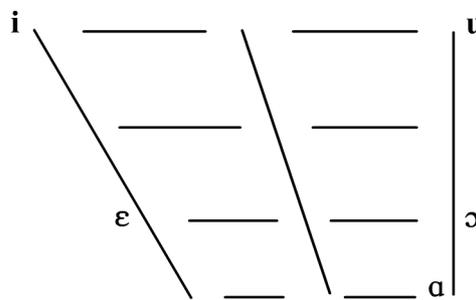


Figure 39 : Triangle vocalique du grec

#### Forme sifflée

Les cinq voyelles grecques phonologiques (i, ε, α, ɔ, u) sont sifflées en cinq intervalles de fréquences se chevauchant de manière inégale. Ils forment trois groupes distinctifs qui sont, dans l'ordre des fréquences décroissantes : (i), (u, ε) et (α, ɔ). Les résultats obtenus à partir des deux meilleurs siffleurs rencontrés dans le village sont présentés Figure 40 et Figure 41. Les données d'origines sont issues du **Erreur ! Source du renvoi introuvable.** situé en Annexe C. Il est important de remarquer que les données que nous présentons concernant Mr. Panayotis sont le résultat d'un dialogue *enregistré sur le vif lors d'une conversation spontanée* avec sa voisine Mrs. Kula. De tels enregistrements, suffisamment longs pour produire des résultats scientifiques sont très rares. Ils supposent en effet d'être présent avec son matériel sans perturber la discussion et d'avoir le réflexe d'enregistrer assez vite. Nous avons eu également l'opportunité d'avoir une traduction simultanée par la fille de Mr. Panayotis.

<sup>44</sup> Nous remercions Mr. Panayotis et Mrs Kula pour leur collaboration

Voyelles sifflées en Grec: Mr. Panayotis siffle à 150 m

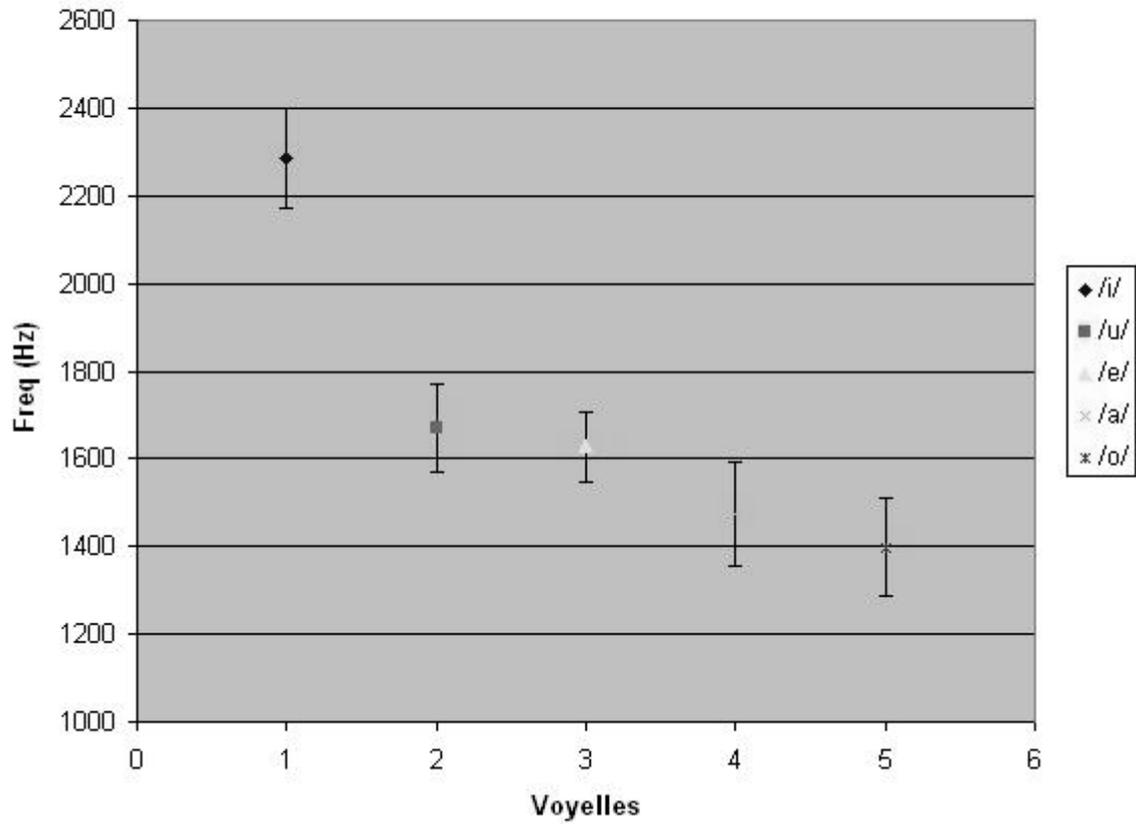
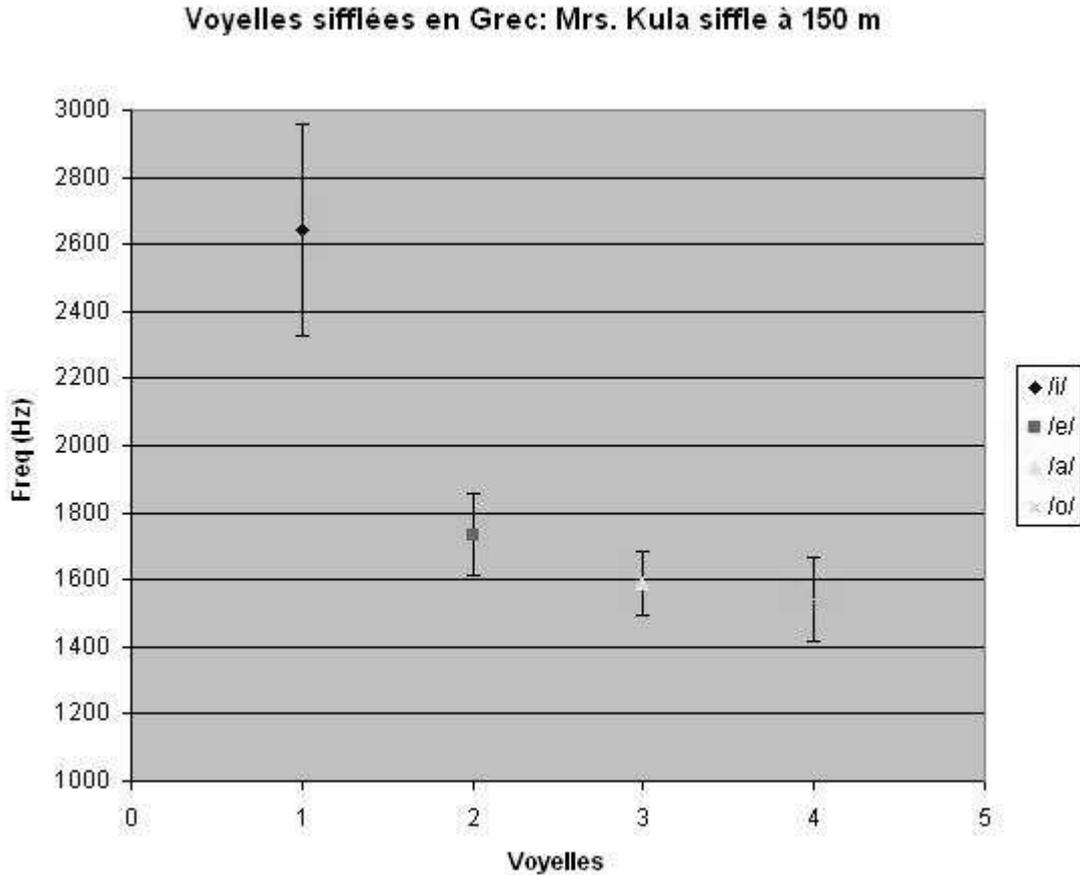


Figure 40 : Répartition des voyelles du grec. Siffleur 1



**Figure 41 : Répartition de 4 voyelles du grec (/u/ absent du corpus), Siffleuse 2**

Les transpositions sifflées de [i] n'interfèrent jamais avec les autres valeurs. Les deux siffleurs ont la même répartition générale des voyelles. Dans le cas des sifflements de Mr. Panayotis, /u/ et /e/ se chevauchent au point qu'il devient impossible de les distinguer autrement que par le contexte lexical. En raison du peu d'occurrence de /u/ dans ce corpus constitué de 70 mots, nous avons également fait siffler l'ensemble des lettres du grec. Mais à nouveau, le /u/ est peu présent. Nos résultats sont cependant confirmés par ceux de Xirometis et Spyridis (1994) dont le corpus constitué de 34 mots sifflés n'avait également que 5 occurrences de /u/ (voir leurs résultats

**Tableau 11 : Répartition fréquentielle des voyelles grecques sifflées (Xirometis et Spyridis 1994)**

**F1 correspond à la fréquence fondamentale du sifflement, F2 à la première harmonique et F3 à la troisième harmonique. Nous ne nous intéresserons qu'aux valeurs de F1.**

Vowel		Stressed			Unstressed		
		F1	F2	F3	F1	F2	F3
a	M	1650	3275	4800	1500	3030	4433
	F	1775	3462	5275	1600	3112	4625
ε	M	1725	3375	5050	1750	3500	—
	F	1850	3650	5537	1733	3300	5700
i	M	2600	5000	7500	2550	5057	7166
	F	2825	5525	6900	2642	5342	—
o	M	1500	2975	4400	1700	3300	—
	F	1766	3566	5366	1581	3081	4910
u	M	1816	3400	5000	1600	3250	5100
	F	1850	3650	5500	1725	3350	5000

D'autre part, les bandes de fréquences des /a/ et des /o/ sifflés se chevauchent également au point de ne pas être discernables phonétiquement. Ceci s'explique car le [α] grec est une voyelle arrière plus proche phonétiquement du [ɔ] que du [ɑ]. Son caractère arrondi disparaît avec la fixation des lèvres quand il s'agit de la siffler.

On remarque de plus que le sifflement féminin est plus aigu que le sifflement masculin, c'est une tendance qui est également confirmée par les résultats d'autres recherches sur le grec (Xirometis et Spyridis, 1994) et des recherches sur le turc, le béarnais ou le gomero (Busnel & Classe, 1976). Ceci est dû à des propriétés physiologiques de l'organe phonatoire, comme dans le cas de la voix parlée.

### Regroupements vocaliques

Les données que nous avons obtenues montrent que le sifflement grec effectue les regroupements vocaliques présentés sur la Figure 42. Il est tant de dire qu'il y a trois bandes de fréquences sifflées en grec d'Antia et donc seulement trois voyelles: I: (i), E: (u,ε), O : (α, ɔ) (voir triangle vocalique Figure 42). Mais nous n'irons pas jusque là pour les mêmes raisons qu'en espagnol sifflé. D'autre part, comme le montrent les bandes de fréquence de chacune des voyelles de ces regroupements, les siffleurs auront souvent besoin du contexte pour distinguer les voyelles /u/ et /ε/ d'une part, et les voyelles /α/ et /ɔ/ d'autre part.

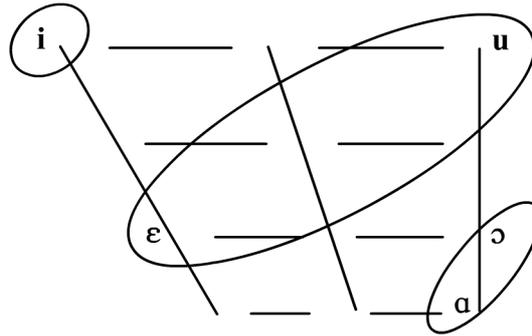


Figure 42 Triangle vocalique du grec sifflé

### 3.3.3.2.2. Diphtongues

Les diphtongues sont traitées exactement comme des paires de voyelles. La modulation de fréquence obtenue est parfois aussi lente que pour les consonnes (exemple Figure 43).

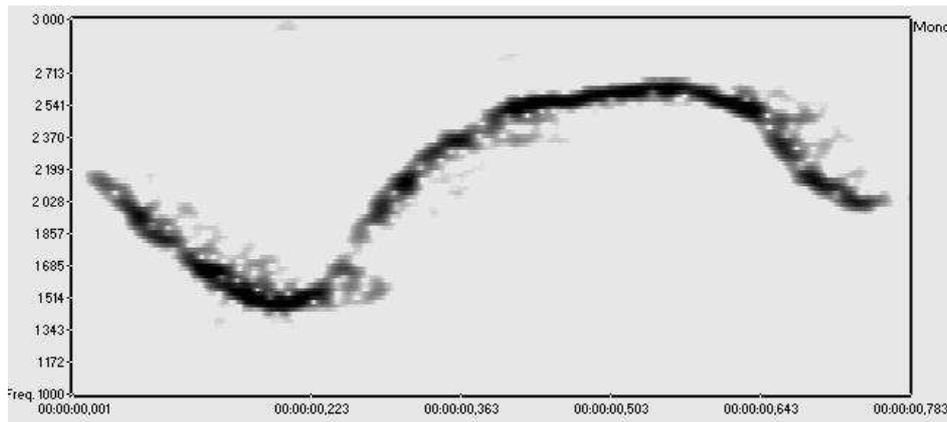


Figure 43 : extrait de grec sifflé : syllabe /proi/. On voit la longue modulation de passage de la fréquence basse du /o/ à la fréquence haute du /i/. La pente initiale est la consonne tronquée /r/, le p est un silence.

### 3.3.3.2.3. Accentuation

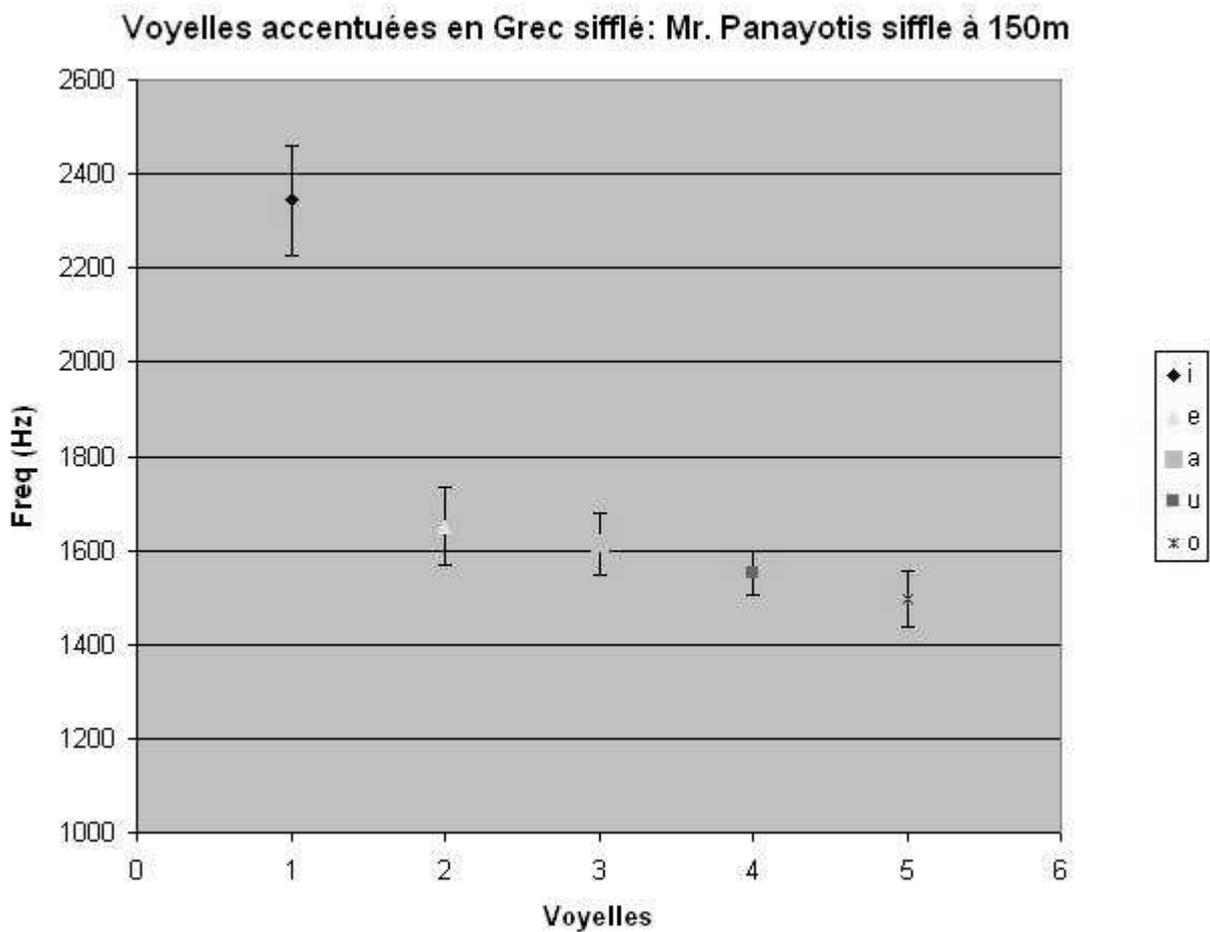
L'accentuation en grec moderne a un degré intermédiaire de liberté. Les distinctions du grec ancien (aigu, grave et circonflexe) liées à la hauteur et au timbre des voyelles ont été neutralisées en grec moderne en un seul accent. « *Il est généralement admis que dans un contexte intonatif neutre les voyelles accentuées sont plus longues, plus hautes et plus intenses que les inaccentuées [...]. L'accent de mot en grec moderne est libre, i.e. non prédictible, pour deux raisons. D'abord il existe dans cette langue des paires minimales qui ne se différencient que par la place de l'accent. La seconde raison est qu'il est fixe dans la mesure où les morphèmes n'assignent la place de l'accent qu'à l'intérieur d'une zone accentuable : indépendamment du nombre de mot* » (Dimou Athanasia et Dommergues 2004, p.177).

Les siffleurs marquent les accents dans 80% des cas par une augmentation d'intensité accompagnée d'une élévation sensible de la fréquence de la voyelle sifflée. Pourtant, sur la Figure 44, ce phénomène prosodique

entraîne un abaissement de la fréquence du /u/ sifflé alors que les autres voyelles accentuées sont situées dans la partie haute de l'intervalle de fréquence de la voyelle correspondante.

Il est important de remarquer que de nombreuses d'autres considérations intonatives ou articulatoires font que des voyelles non accentuées occupent également la zone des fréquences haute de chaque intervalle. La distinction d'accent, si elle est perceptible par le récepteur dans le contexte de la phrase est difficilement évaluée sur un corpus de phrases comme le nôtre. Les résultats présentés par Xirometis et Spyridis (1994) montrent que dans le cas d'une liste de mots sifflés, les voyelles accentuées ne sont pas mieux distinguées. De plus leurs données ne confirment pas nos observations sur le /u/.

Le peu de stabilité du /u/ peut être expliqué si l'on considère que c'est une voyelle arrondie par les lèvres dans la voix parlée, ce phénomène est perturbé dans la parole sifflée par la fixation des lèvres.



**Figure 44 : Répartition des voyelles accentuées grecques**

**La place du /u/ change et les fréquences sont légèrement plus élevées que dans le cas de la répartition des voyelles non accentuées.**

## 3.3.3.2.4. consonnes

**Forme parlée**

Les consonnes du grec parlé sont présentées dans le Tableau 12. Cette description est très proche de celle que l'on rencontre dans le village d'Antia. Les consonnes nouvelles par rapport au Gomero sont le flap [ɾ] (qui n'est jamais réalisée comme un [r] contrairement à d'autres endroits de la Grèce), la présence de la jota [x], de [θ], de [v] et de [ts tz].

**Tableau 12 : phonétique des consonnes grecques**

	<i>Bilabiale</i>	<i>Labio dentale</i>	<i>Labio vélaire</i>	<i>Dentale</i>	<i>Alvéolaire</i>	<i>Post alvéolaire</i>	<i>Retroflexe</i>	<i>Palatale</i>	<i>Vélaire</i>	<i>Uvulaire</i>	<i>Glottale</i>
occlusive	p b			t̪ d̪					k g		
implosif											
click											
trill											
Tap											
flap					ɾ						
Fricative		f v		θ ð	s z				x γ		
Affriquée					(ts tz)						
Nasale	m			ɲ							
Lateral fricative											
Lateral approximant				l							
approximant								j			

**Forme sifflée**

Comme le sifflement des consonnes est lié à l'articulation on retrouve le même type de répartition des consonnes que pour le Silbo. Cependant, alors que la technique de sifflement la plus répandue à La Gomera utilise l'intromission de doigts dans la bouche, la technique la plus utilisée à Antia est celle qui consiste à plier la langue contre les dents de la mâchoire inférieure. Nous n'avons rencontré qu'une seule personne utilisant parfois le sifflement avec les doigts. Ceci est essentiellement dû au fait que la topographie de la région d'Antia est moins accidentée que les vallées très encaissées de l'île de la Gomera. La puissance de sifflement nécessaire pour se parler à distance est donc moindre. Au niveau de la prononciation des

consonnes, la configuration de la bouche n'est donc pas perturbée par le doigt en particulier pour les mouvements de la mâchoire inférieure qui peuvent être exécutés plus facilement.

-[p] est réalisé par une interruption du sifflement, de même que [b] dont l'interruption est moins radicale en terme d'amplitude. [k] et [g] forment également une interruption clairement due à des coups de glotte.

-[ɣ] est sifflé comme un [g] mais a un silence moins marqué.

-[χ] n'était pas présent en Silbo car il était remplacé par un [h]. Il est ici sifflé parfois de manière continue avec une légère dépression convexe. Il semble stabiliser les voyelles qui l'entourent.

-[t] et [d] ont les mêmes modulations fréquentielles que celles qui les caractérisent en Silbo mais la modulation est très rarement vers le bas. Cette distinction est une transposition des différences de spectre des voyelles et consonnes /i/ et /t/ grecques par rapport aux /i/ et /t/ de l'espagnol gomero(nous reviendrons sur ce phénomène dans la comparaison des langues)

-[m] est parfois sifflé comme en Silbo, mais il est aussi souvent marqué par un simple silence.

-[l, n, j, ð et ɾ] sont sifflés en une forme continue concave. Le sommet de cette modulation fréquentielle a une modulation tendant à réduire l'amplitude surtout dans le cas de [n].

-[s, z] sont presque sifflés comme un [d] avec une modulation de fréquence dont la pente d'attaque et de relâchement sont moins longues

-[θ].est presque toujours sifflé comme un [s] mais ses interruptions sont plus douces en terme de modulations d'amplitude et de fréquence

-[ts tʰz] ont une forme similaire mais ils sont plus proche du [t] et marquent un silence plus long.

## **Clusters de consonnes**

Les modulations de fréquence et d'amplitude des consonnes sont cumulées et très souvent tronquées. Par exemple, le /r/ de /proi/ sur la Figure 43 en début de sonagramme forme une descente fréquentielle sans avoir de convexité au départ. Contrairement au Silbo, on constate sur cette même figure que le /r/, n'est pas éludé après une plosive.

### **3.3.3.2.5. Historique de l'étude de la langue sifflée grecque**

La langue sifflée d'Antia est encore la seule signalée à ce jour en Grèce, malgré un usage très répandu du sifflement non articulé dans de nombreuses zones pastorales du pays et parmi les peuples de bergers nomades (Sarakatsan parlant grec) ou semi-nomades (Aroumain ou Vlaques parlant aroumain<sup>45</sup>). Sur l'île d'Eubée, c'est une pratique propre au village d'Antia depuis plusieurs générations. Une telle situation a alimenté les débats sur l'origine des habitants de ce village. Trois explications différentes ont été données:

---

<sup>45</sup> Nous avons fait une courte enquête dans la population Aroumaine des montagnes du Nord de la Grèce qui mériterait d'être poursuivie. Cette population parle une langue romane distincte du roumain mais ayant des origines communes. Pour des raisons historiques, la langue aroumaine a du mal à se faire une place dans la société grecque actuelle, par contre elle est reconnue et défendue au niveau européen.

l'une raconte que cette pratique remonterait à l'époque des guerres Perses (Theocharis 1959 cité par Charalambakis 1994), la seconde que les habitants d'Antia seraient les descendants de prisonniers emmenés depuis Aïnos en Thrace jusqu'à Antia par l'Amiral de Venise prénommé Canale en 1409 (Karatzas 1972 cité par Xirometis et Spyridis 1989), la dernière est celle des habitants, ils affirment que leurs ancêtres auraient développé un langage secret en raison de l'arrivée de nombreuses populations étrangères dans la région (Charalambakis 1994).

Il y a eu trois publications scientifiques évoquant la langue sifflée grecque. A l'origine de cet intérêt, le réalisateur Stravos Joannou fit un documentaire en 1982 sur le village d'Antia, il y emmena le linguiste Charalambakis et le musicologue Amargianakis. Les enregistrements réalisés à cette occasion furent utilisés dans une introduction très générale qui incluait quelques spectrogrammes (Xirometis et Spyridis 1989). Les résultats de l'enquête ethnolinguistique furent publiés par Charalambakis (1994) et une analyse acoustique des voyelles fut présentée la même année par Xirometis et Spyridis (1994). Suite à cette période, les habitants d'Antia refusèrent de réaliser des enregistrements avec d'autres scientifiques<sup>46</sup> (com. pers. Charalambakis et Spyridis 2004). Nos enregistrements ne purent être obtenus que grâce à un séjour long dans la taverne locale et une collaboration avec l'école qui firent naître des relations amicales entre nous. La présentation des enregistrements des autres pays nous aida beaucoup. A partir de nos données une première analyse a été publiée de manière à expliquer les réductions phonétiques des voyelles et les articulations des consonnes du grec (Meyer 2005).

### **3.3.3.3. Le turc des montagnes de l'Est, au bord de la Mer Noire**

La langue turque fait partie de la branche Turkic de la très vaste et controversée famille Ouralo-Altaique qui s'étend de la Turquie à la Mongolie et la Chine. D'un point de vue morphologique, la langue turque a été classée parmi les langues agglutinantes ce qui signifie que les mots sont constitués par un radical (élément sémantique fondamental) suivi éventuellement d'un ou plusieurs suffixes de dérivation (éléments sémantiques secondaires modifiant le sens de la racine) et d'une désinence (porteuse d'un ensemble de valeurs permettant d'actualiser le sens du radical et de marquer ses rapports avec le reste de la phrase). Chaque désinence ne porte qu'une signification ce qui entraîne leur multiplication et donne parfois des mots très longs.

La langue sifflée est pratiquée non loin de la mer Noire, au Nord-Est du pays, à l'intérieur des terres montagneuses. Autrefois, les villages de plusieurs vallées pratiquaient cette forme de la langue. Aujourd'hui seuls quelques villages la maintiennent mais elle n'est plus beaucoup pratiquée. Elle sert encore aux bergers qui montent en été sur les hauts plateaux avec leur troupeau de moutons, là où les routes ne vont pas encore. Avec la modernisation de la vie pastorale et de la récolte du thé, des noisettes et des cerises qui font la richesse de cette région depuis des centaines d'années, elle a de moins en moins de raison d'être utilisée. Malgré cela de nombreux sifflleurs la maîtrisent encore, en particulier dans le village de Kusköy dont le nom signifie « village des oiseaux ».

---

<sup>46</sup> L'abandon d'une thèse grecque sur le sujet, commencée à partir de l'ancien corpus, en est l'illustration.

### 3.3.3.3.1. Description phonétique de la langue

Les documents sonores qui nous ont servi à réaliser ces descriptions sont issus de deux sources différentes:

- Les enregistrements réalisés dans un village voisin de Kusköy avec deux bons siffleurs<sup>47</sup>.
- Les enregistrements réalisés en 1967 par l'équipe pluridisciplinaire de Busnel dont les analyses ont par ailleurs été publiées dans une monographie dont nous citerons certains passages (Busnel et al 1970)<sup>48</sup>.

### 3.3.3.3.2. Voyelles

Le vocalisme de la langue turque est particulier. Tout d'abord il est constitué de 8 voyelles réparties comme dans le tableau suivant:

**Tableau 13 : Voyelles turques**

		<i>Antérieures (« acute »)</i>		<i>Postérieures (« grave »)</i>	
		Étirées (« Plain ») <sup>49</sup>	Arrondies (« Flat »)	Étirées (« Plain »)	Arrondies (« Flat »)
Fermées	Haute (“lowered-high”)	i [ɪ]	ü [Y]	ı [i]	u [ʊ]
Diffuses	Moyenne (« lower mid »)	e [ɛ]	ö [œ]		o [o]
	Basse <sup>50</sup>			a [a]	

Dans ce tableau les voyelles sont à la fois présentées avec leur orthographe turque moderne officielle ainsi que celle de l'A.P.I. (entre crochets). Nous avons trouvé plusieurs descriptions différentes du turc en ce qui concerne les voyelles. Celles ci varient surtout pour le « o » qui est parfois décrit comme un o ouvert ou le « ı » qui est parfois décrit comme un [u]. Les choix que nous avons faits reposent sur la langue parlée à Kusköy et dans la région de la ville voisine de Görele.

Le triangle vocalique est représenté de manière complémentaire Figure 45.

<sup>47</sup> Nous remercions Mr. Cemal Patan pour son accueil dans son village et ses voisins pour avoir accepté de siffler.

<sup>48</sup> Nous tenons à ce propos à remercier le Professeur René-Guy Busnel pour nous avoir autorisé à exploiter ces données uniques car la langue Turque était largement répandue et pratiquée à l'époque. Elle reste aujourd'hui la plus vigoureuse des langues non tonales connues à ce jour. Nous tenons également à remercier Bernard Gautheron qui a conservé ces données en bon état jusqu'à aujourd'hui et qui a pu nous les transmettre.

<sup>49</sup> Les termes mis entre parenthèses correspondent à la terminologie mise au point par Jakobson, Fant et Halle qui se réfère à une étude acoustique des voyelles (cités par Leroy 1970).

<sup>50</sup> L'opposition phonétique entre voyelles basses et voyelles moyennes est redondante d'après certains auteurs en raison de la présence de règles d'harmonie vocalique en turc.

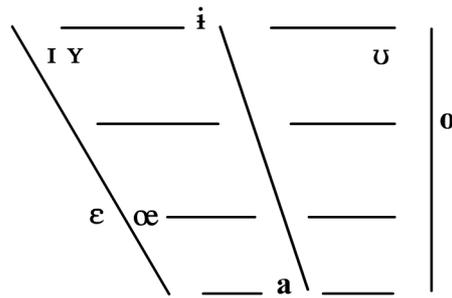


Figure 45 : Triangle vocalique du turc

## Harmonie vocalique

Les 8 voyelles suivent des règles d'harmonie vocalique propres au Turc qui fixent certains aspects de l'enchaînement des syllabes dans un même mot agglutinant.

L'harmonie vocalique est un procédé par lequel une partie des oppositions de qualité des voyelles est neutralisée par un effet d'assimilation entre une voyelle d'une syllabe et celle de la suivante. Les règles s'appliquent de gauche à droite et ainsi seules les voyelles non initiales sont concernées. Les règles sont les suivantes:

- Si la première voyelle est prononcée antérieure (i[I],e[ε],ü[Y],ö[œ]) ou postérieure (ı [i],u [u], a[a], o[o]) les voyelles subséquentes seront respectivement antérieures ou postérieures. Ceci classe les mots en deux catégories.
- Si une voyelle fermée (ou Haute) est étirée ou arrondie, la voyelle suivante sera respectivement également étirée ou arrondie. Par contre une voyelle diffuse (ou Basse) en position non initiale sera toujours étirée. La conséquence directe est que les voyelles ö et o ne pourront être qu'en première syllabe.

On peut résumer toutes les possibilités de l'harmonie vocalique turque sous la forme suivante:

*a et ı -----peuvent être suivies de ----- a et ı*  
*o et u----- peuvent être suivies de -----a et u*  
*e et i-----peuvent être suivies de ----- e et i*  
*ö et ü-----peuvent être suivies de ----- e et ü*

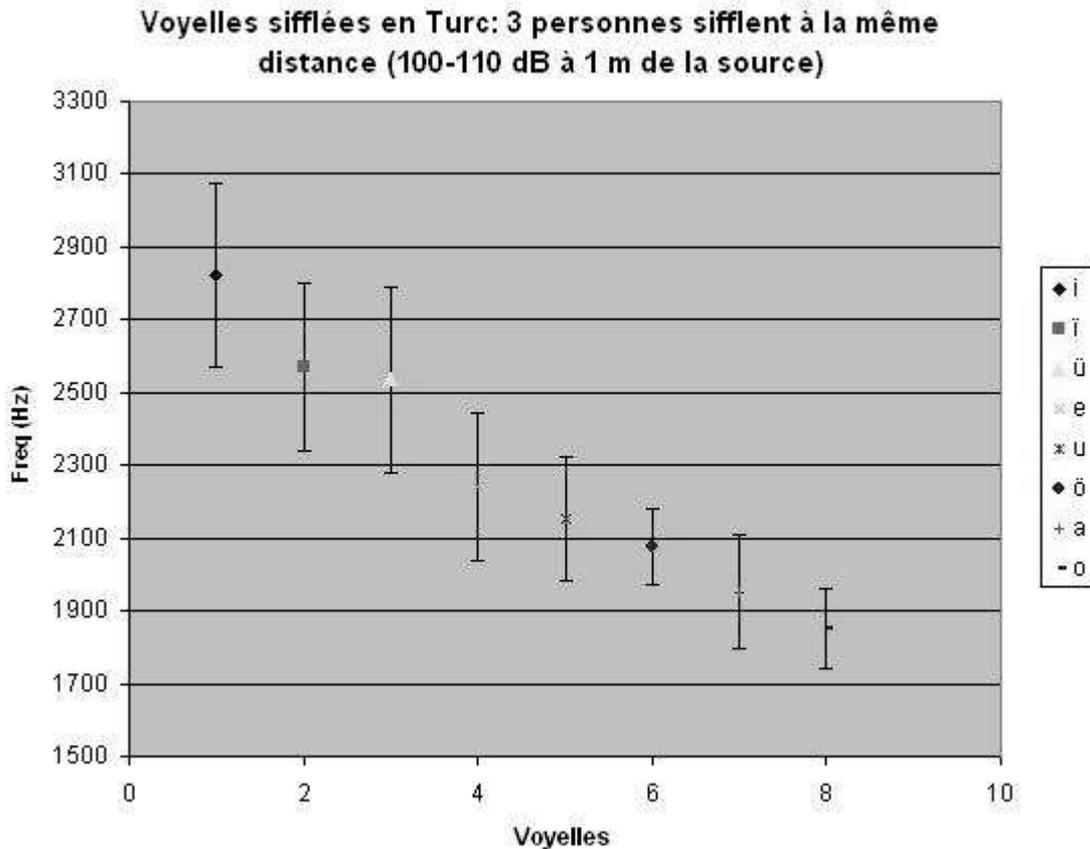
Dans toute syllabe non initiale, le système décrit se réduit à 6 voyelles. Les seules oppositions qui subsistent par ce processus sont celles entre voyelle Haute ou voyelle non Haute.

## Forme sifflée

Le sifflement des 8 voyelles du turc est réalisé en 8 bandes de fréquences comme nous pouvons le voir sur la Figure 46.. La voyelle i([I]) est relativement bien démarquée de ces deux plus proches voisines car sa

fréquence sifflée est en moyenne plus élevée. Certaines bandes de fréquences se chevauchent plus que d'autres :

-premièrement (ı([i]) et ü([Y])) qui ont des bandes de fréquences presque confondues<sup>51</sup> même si ı est en moyenne sifflée légèrement plus haut que ü.



**Figure 46 : Répartition des voyelles turques.**

**La grande variabilité de l'écart type s'explique par le fait que les données sont issues de trois personnes.**

-Deuxièmement, les bandes de fréquence des voyelles (e [ɛ], u [ʊ], ö [œ]) se chevauchent largement. La voyelle e est en moyenne plus élevée que u et ö. D'autre part ö est en moyenne moins variable et de fréquence plus basse que les deux autres. Il semble que leurs bandes de fréquences se chevauchent suffisamment pour que ces voyelles soient difficilement discernables sans le contexte lexical.

-Troisièmement, les bandes de fréquences sifflées des voyelles (a [a] et o [o]) sont sifflées de manière proche. Mais o [o] est en moyenne sifflée à une fréquence plus basse.

<sup>51</sup> Cf. partie intelligibilité où il sera montré quelles sont souvent perçues confondues également, lorsqu'elles sont tirées d'un contexte lexical.

## Diphthongues

Les voyelles du turc peuvent former 4 diphthongues qui sont ay (/aj/), ey (/ej/), oy (/oj/) et uy (/uj/).

Elles sont sifflées par un enchaînement rapide de la voyelle initiale suivie d'une montée fréquentielle vers la bande de fréquence des i. Par conséquent, ay et oy sont facilement confondues.

### Conclusion pour les voyelles

On constate également sur les données que nous présentons que le sifflement turc est relativement aigu par rapport au grec et au Silbo. D'une manière générale les contraintes de l'articulation sifflée entraînent les regroupements fréquentiels visibles sur la figure suivante :

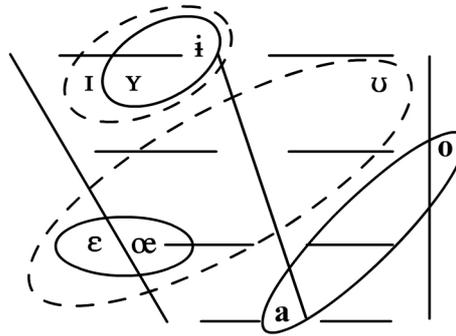


Figure 47 : Triangle vocalique du turc sifflé

Ce système constitué de nombreuses voyelles est avantageusement complété par les règles d'harmonie vocalique propres au turc.

### 3.3.3.3.3. Accentuation

La langue parlée turque possède un accent intonatif qui intervient sur les particules précédant l'expression de l'interrogation, la négation et sur l'impératif exprimant la prohibition. Cet aspect n'a jamais été abordé en ce qui concerne la langue sifflée. Parmi les phrases du corpus que nous avons examiné, plusieurs d'entre elles ont l'une ou l'autre de ces situations:

Tableau 14 : Exemple 1

Phrase en turc	Kalemin var mı
En phonétique	[kalɛmin var mɨ]
Grammaire	crayon-POSSESSIF2sg il y a INTERROGATION
En français	Est ce que tu as un crayon?

Dans cet exemple le a de Var est accentué en voix parlée, au moins en intensité. Dans les 6 prononciation sifflées de cette phrase que nous avons examiné, seule une n'est pas accentuée au niveau fréquentiel. Pour les autres, le /a/ a une valeur proche du maximum de la bande des fréquences des /a/ sifflés. L'accentuation

semble donc être réalisée également par une légère augmentation de l'amplitude comme la montre la Figure 48.

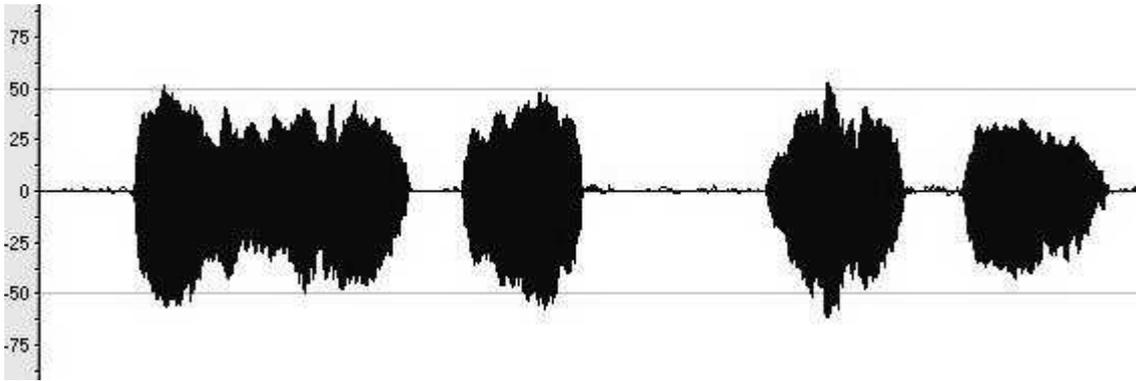


Figure 48 : L'avant dernière unité de parole est la syllabe /var/ sifflée accentuée

Voici trois autres configurations d'accentuation turque pour lesquelles le sifflement reproduit cet aspect de la langue parlée de manière assez fidèle, essentiellement à travers la fréquence. Et plus modestement à travers l'intensité

Tableau 15 : Exemple 2

Phrase en turc	Arkadasım askere gitmedi
En phonétique	[arkadaʃım askere gitmedi]
Grammaire	Ami-POSSESSIF1sg service national -DIR aller-NEG-PARF TEST-0
En français	Est-ce que mon ami n'est pas allé au service national?

Dans cet exemple le i de git est accentué dans la forme parlée et sifflée

Tableau 16 : Exemple 3

Phrase en turc	Sürücülerı karsı dagda bırakma
En phonétique	[syrydzlırı karʃi dayda bırakma]
Grammaire	Conducteur-PLU-ACC contre montagne-LOC laisser-NEG-0
En français	Ne laisse pas les conducteurs sur la montagne d'en face

Le a de rak est accentué dans la version parlée et sifflée

Tableau 17 : Exemple 4

Phrase en turc	Evin kapısını kapa
En phonétique	[evin kapısını kapa]
Grammaire	Maison-POSS3sg-ACC fermer-IMP3sg
En français	ferme la porte de la maison

Le a de ka est accentué dans la voix parlée et sifflée

D'après ces quelques données, l'accentuation est bien marquée en turc sifflé.

### 3.3.3.3.4. Consonnes

#### Version parlée

Nous présentons les consonnes sous la forme du tableau récapitulatif suivant.

Tableau 18 : phonétique des consonnes turques

	<i>Bilabiale</i>	<i>Labio dentale</i>	<i>Labio velaire</i>	<i>dentale</i>	<i>alvéolaire</i>	<i>Post alvéolaire</i>	<i>Retroflexe</i>	<i>Palatale</i>	<i>Velaire</i>	<i>Uvulaire</i>	<i>Glottale</i>
stop	p b			t̥ d̥				c ɟ	k ɡ		ʔ
implosif											
clic											
trill					r						
Tap											
flap											
Fricative sib		f v			s z	ʃ ʒ			y		h
Affriquée						t̥ʃ d̥ʒ					
Nasal	m			ŋ							
Laterale fricative											
Laterale approximant								ʎ			
approximant								j			

#### Version sifflée

Les consonnes sifflées du turc ont un comportement similaire aux consonnes sifflées du silbo et du grec, avec les mêmes classes de regroupement en fonction des formes de modulations fréquentielles. Elles sont articulées de manière extrêmement précise.

#### Nouvelles consonnes

Les consonnes nouvelles par rapport à celles que nous avons déjà examiné dans les autres langues sont:

-[ʔ] la glottale, essentiellement présente dans des mots issus de l'arabe assimilés par le turc. L'aspiration [h] est légèrement plus fréquente tout en étant aussi due souvent à des emprunts. Ces deux consonnes sont sifflées à l'intervocalique sous la forme d'une très légère modulation convexe vers le haut (voir Figure 49).<sup>52</sup>

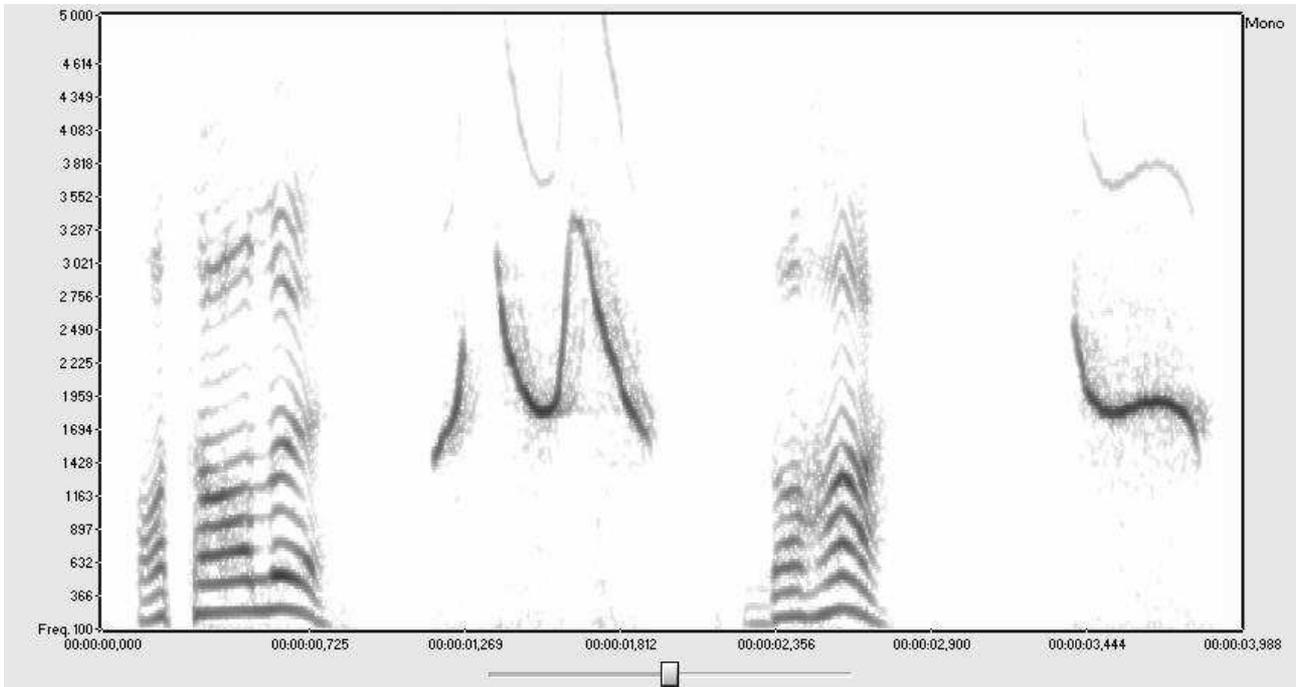


Figure 49 : mots turc « atali » et « dahar » en version parlée puis sifflée

-[ʒ] sifflée comme [j] du silbo et du grec.

-[λ] sifflée comme [l] du silbo et du grec

-[c] et [j] sont sifflées comme [ʃ] ou [t]

-[f] [v] et [ɣ] qui eux mêmes sont sifflées comme en grec et en silbo, ce sont les seules consonnes qui étaient absentes de l'étude de Leroy (1970).

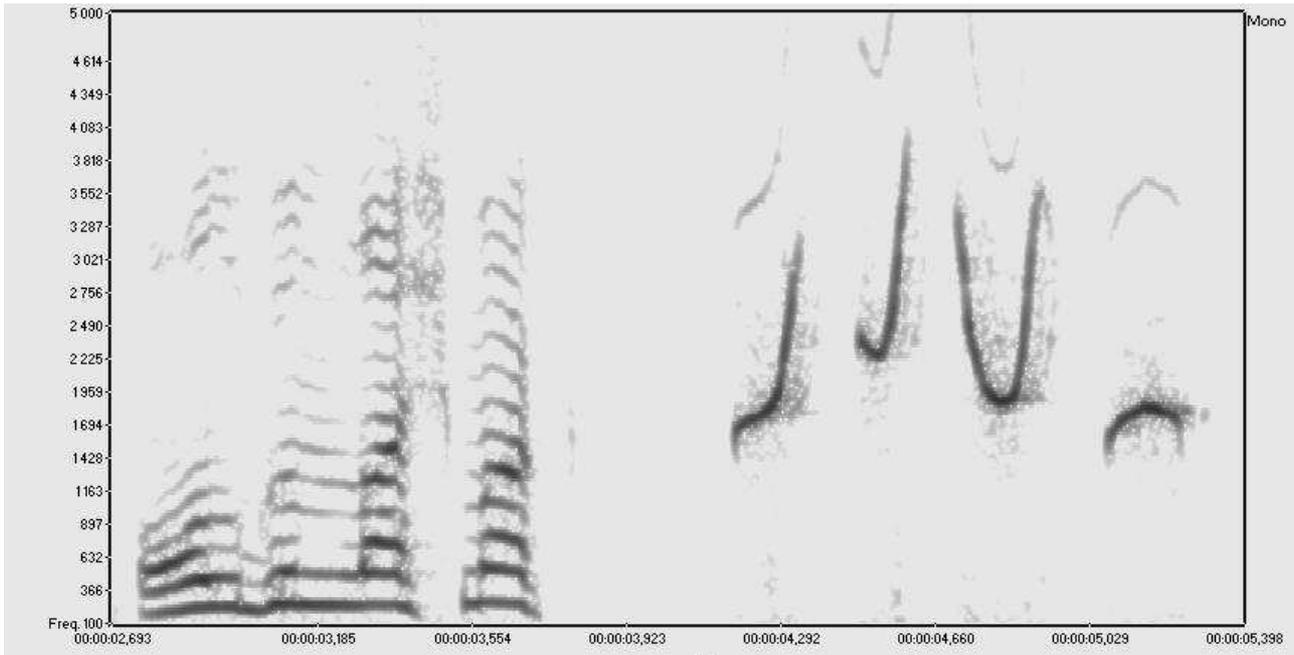
-[tʃ] [dʒ] sont sifflées comme un cluster respectivement de [t] et [ʃ] et de [d] et [ʒ].

## Clusters

En ce qui concerne les clusters de consonnes, le sifflement turc est très explicite, même sur un spectrogramme, on peut se rendre compte que les clusters reproduisent toutes les consonnes qu'ils contiennent mais les contraintes de l'articulation entraînent plusieurs phénomènes:

-Dans les clusters où se trouve une consonne qui entraîne une interruption de la modulation par un silence, les représentations phonétiques sifflées sont légèrement tronquées à ce niveau mais les caractéristiques de chaque consonne sont cumulées. On peut voir différentes illustrations de ce phénomène sur la s modulations propres à la consonne initiale (attaque du cluster) et à la consonne finale (relâchement du cluster et attaque de la voyelle qui suit) sont phonétiquement présentes).

<sup>52</sup> Comme nous le verrons dans la langue Chepang, ce comportement est différent du [h] des langues qui transposent le sifflement des glottales par leur effet sur l'intonation du noyau syllabique.



**Figure 50 : mot turc parlé /olgunlaşmak/ et sifflé /ol--/gun/--/laʃ/--/mak/ où chaque silence est un cluster de consonnes**

-Dans les clusters constitués uniquement de consonnes continues, l'effet cumulé des consonnes concernées est souvent réalisé par une seule modulation plus longue. C'est le cas de [ll] de [rl].

### 3.3.3.3.5. Historique de l'étude de la langue sifflée Turque

La langue sifflée de Turquie fut découverte par les scientifiques grâce à une campagne de construction d'école conduite dans les villages de Turquie dans les années 60 à laquelle participa une compagnie pétrolière dont les prospecteurs signalèrent l'existence d'une langue sifflée. Plusieurs articles de journaux furent publiés sur le sujet ce qui éveilla la curiosité des deux principales personnalités scientifiques intéressées par le sujet à cette époque: René-Guy Busnel et André Classe. En 1967 Busnel organisa une expédition pluridisciplinaire avec l'aide d'une fondation allemande<sup>53</sup>. Les résultats de cette étude restent non seulement les seuls publiés sur le sujet mais ils font aussi référence en matière d'étude des langues sifflées en raison des multiples modes d'analyses qui furent développés à cette occasion.

En ce qui concerne la description linguistique des voyelles et des consonnes, plusieurs approches furent envisagées: descriptives et perceptives. Nos résultats, qui reposent en grande partie sur le corpus récolté à la fin des années soixante, sont cohérents avec l'ensemble des observations de Moles (1970) et Busnel (1970). En ce qui concerne l'étude de Leroy (1970), nous avons une légère divergence de résultat pour le u et le i. Nous trouvons le i bien distingué de ses plus proches voisines fréquentielles qui sont le ɪ et le ü. Moles trouve la même tendance au niveau de la perception<sup>54</sup>. Leroy par contre décrit le i complètement confondu avec le u et le ɪ. Elle situe donc le u avec les voyelles les plus élevées alors que d'après nos données, le u est

<sup>53</sup> Wiener Green Foundation

<sup>54</sup> Voir test d'intelligibilité des non-mots dans la partie dédiée à l'intelligibilité

une des voyelles situées à des fréquences intermédiaires. Peut être que l'approche de Leroy a porté sur un siffleur utilisant le u de façon très personnelle.

### **3.3.3.4. Etude comparative des langues sifflées articulées non tonales**

D'une manière générale, les résultats que nous avons présentés démontrent que les siffleurs s'appuient en premier lieu sur des considérations articulatoires partagées avec la voix parlée. En effet, de nombreux regroupements vocaliques et consonantiques sont dus à des proximités d'articulation que l'on retrouve dans la version parlée. Du point de vue de la phonétique, les langues sifflées que nous avons considéré jusqu'à maintenant réalisent donc naturellement *une étape descriptive naturelle et instructive*. Le signal acoustique sifflé qui en résulte permet une analyse directe de ces phénomènes car il est constitué d'une bande étroite de fréquences plus facilement caractérisable que le timbre issu du spectre de la voix.

#### **3.3.3.4.1. Systèmes vocaliques**

Chaque voyelle est définie par une hauteur relative de sifflement qui correspond à un espace de résonance dans la bouche. Cet espace, certes large, est tout de même limité et fixe le cadre dans lequel des hauteurs relatives peuvent varier. Suivant la technique utilisée, et la physiologie de chaque personne le sifflement obtenu sera plus ou moins aigu.

#### **Modification avec la distance**

Nous avons mesuré que pour un même siffleur, le sifflement utilisé à 500m est en moyenne plus élevé de 150 Hz que celui à 50 m (changement de technique). Si l'on mesure les fréquences moyennes des voyelles produites par deux siffleurs (deux hommes utilisant la même technique et dont les fréquences moyennes de sifflement sont similaires lorsqu'ils sifflent à la même distance), l'un cherchant à atteindre une distance de 300m et l'autre de 1000 m, on obtient les résultats statistiques présentés pour le Silbo. Ceux-ci semblent indiquer que le siffleur parlant à un interlocuteur situé à 1000m produit des sifflements en moyenne plus aigus. Mais ce décalage se manifeste de manière graduelle dans le détail des voyelles. Plus la voyelle sifflée est aiguë plus elle est décalée vers le haut. Ainsi pour le /i/ le décalage est en moyenne de 300 Hz entre 300m et 1000 m.

#### **Tendances générales des bandes de fréquences vocaliques**

Si l'on considère ensemble les répartitions des voyelles sifflées du silbo, du turc et du grec, on remarque des tendances générales: /i/ est invariablement le plus aigu. Il est parfois approché par des voyelles ayant une articulation proche: [Y] ou [I] ou [i]. /o/ fait invariablement partie des voyelles les plus graves. Il partage souvent son intervalle de fréquence avec une autre voyelle (/u/ en silbo espagnol, /a/ en grec et en turc et même parfois en Silbo). /e/ et /a/ sont des voyelles intermédiaires, /e/ étant plus aigu que /a/. Suivant leur réalisation dans la langue parlée, leurs intervalles de fréquence respectifs se chevauchent plus ou moins. D'une manière générale, les voyelles intermédiaires partagent leur intervalle de fréquence avec les voyelles voisines. Ainsi les valeurs fréquentielles de /e/ interfèrent avec celles de /i/ et /a/ alors que les valeurs

fréquentielles de /a/ interfèrent avec celles de /o/ et /i/. D'autre part, lorsqu'il y a de nombreuses voyelles intermédiaires comme en turc, leurs valeurs vont largement se chevaucher à tel point qu'il ne semble pas qu'elles puissent être distinguées facilement autrement que grâce au contexte. La voyelle /u/ a un comportement particulier lorsqu'elle est sifflée, elle est le plus souvent associée à une voyelle intermédiaire mais en silbo elle est sifflée comme la voyelle la plus grave, à l'extrême de la bande des fréquences de la langue.

### Analyse des voyelles par paires (duplets)

Les valeurs relatives des voyelles successives par paires sont susceptibles d'intervenir en complément du positionnement fréquentiel de chaque voyelle. C'est pourquoi nous les avons mesurées pour les deux siffleurs de silbo. La répartition pour les deux siffleurs est similaire à part quelques inversions dues aux sons [o] ou [u] qui sont peu nombreuses dans notre corpus. Les résultats de ces calculs pour le siffleur Luis M. sont présentés sur la Figure 51

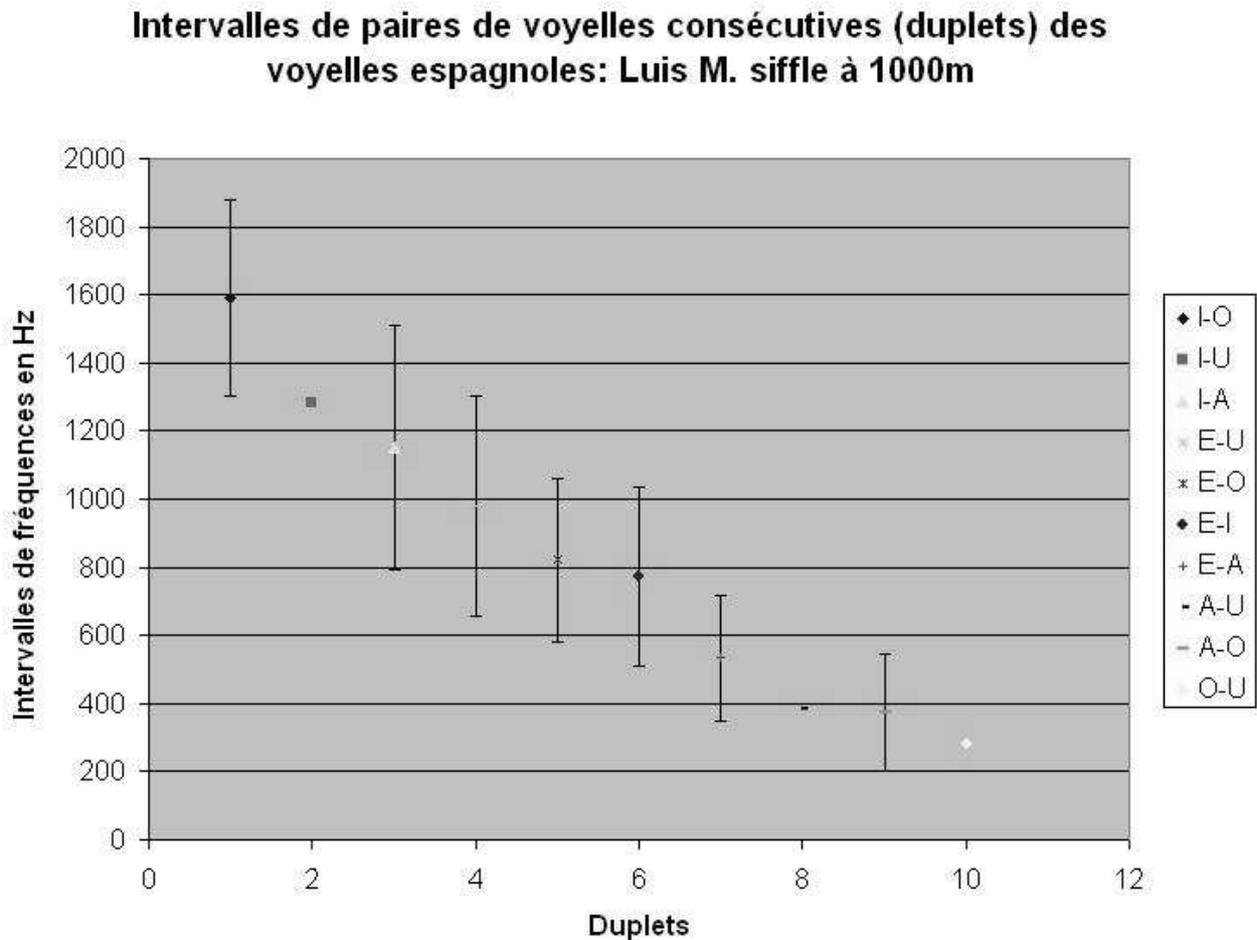


Figure 51 : Répartition des duplets du silbo espagnol

Les valeurs pour lesquelles l'écart type n'est pas représenté sont celles correspondant à un nombre trop limité d'occurrences<sup>55</sup>.

<sup>55</sup> Moins de 6 occurrences alors que les autres duplets sont mesurés sur plus de 10 valeurs.

Nous avons observé que dans de nombreux cas où les voyelles ne sont pas proches de la fréquence moyenne d'intervalle de fréquence dans lequel elles évoluent, cette imprécision est compensée par la réalisation d'intervalles relatifs précis (duplets). A contrario, lorsque la voyelle est bien placée dans son intervalle la différence relative n'est pas forcément respectée. Il semble donc qu'il y ait compensation des deux phénomènes. Cependant ce n'est pas systématique. Nous n'avons pas fait d'analyse précise des conditions qui régissent l'apparition d'une telle compensation mais des considérations lexicales ou phonotactiques doivent être en cause. Il existe aussi quelques cas exceptionnels d'assimilation des voyelles. Nous pensons que le contexte lexical permet alors de lever l'ambiguïté car c'est souvent le fait de mots très courants, le plus souvent des mots de liaison.

Ainsi l'information portée par un doublet semble dépendre du contexte mais agit indubitablement de manière significative dans la perception des voyelles. Cette information s'ajoute pour toutes les langues à la répartition fréquentielle des voyelles sifflées et à des particularités propres à la structure de chacun des idiomes (quantités, accentuation, harmonie vocalique...).

### **Influence des consonnes**

Alors que la hauteur de la voyelle influence fortement la modulation consonantique, nous observons très rarement l'influence de la modulation sur la fréquence de la voyelle en turc, grec et silbo. Le seul cas relativement récurrent est dû à une augmentation de la fréquence d'une voyelle située entre deux modulations vers le haut.

### **Quantité vocalique**

Dans des langues qui n'ont pas de distinctions de quantité vocaliques qui ont une valeur phonologique, la durée des voyelles n'a pas de valeur autre que de permettre de faciliter l'intelligibilité de la langue. Ainsi, il n'est pas rare, particulièrement à grande distance qu'une voyelle soit maintenue pendant plus d'une seconde comme nous l'avons déjà signalé. D'autre part, lorsque la voyelle finale et la voyelle initiale de deux mots consécutifs sont identiques, elles sont presque systématiquement sifflées en une seule voyelle à la fois en termes de durée et de fréquence par un effet de liaison vocalique. D'autre part, tout comme dans la version parlée, la segmentation mot à mot n'est pas toujours respectée même si les deux voyelles consécutives ne sont pas identiques ainsi dans la phrase « tiene que ir » ei de « que ir » est sifflé comme une diphtongue, de manière similaire à celle de « tiene ».

### 3.3.3.4.2. Accentuation

#### **Remarques générales**

L'accentuation est marquée en fonction de son rôle dans chaque langue. Elle s'exprime le plus souvent par un effet combiné d'augmentation de l'amplitude et de la fréquence. Il arrive également que l'accentuation soit marquée par un allongement de la voyelle sifflée, le plus souvent dans les mots pour lesquels les processus habituels d'accentuation sont perturbés (Classe donne l'exemple des mots paroxytons en Silbo).

Pour le grec, le turc et le silbo, l'augmentation de la fréquence se fait dans les limites relatives de l'intervalle de la voyelle concernées. La fréquence de la voyelle accentuée dépend des fréquences des voyelles voisines. Par conséquent, la fréquence sifflée de la voyelle accentuée n'est pas systématiquement dans la partie haute de la bande fréquentielle même si c'est le cas dans la majorité des situations. Ces éléments prosodiques, même s'ils sont limités sont amplement suffisants dans les langues non tonales concernées.

Parmi les langues que nous avons étudiées, l'accentuation est particulièrement régulièrement marquée en grec, cela reflète l'importance de l'accent en grec parlé. A la fois le turc et l'espagnol respectent également dans la version sifflée les tendances d'accentuation de la version parlée.

#### **Cas particulier du béarnais**

En ce qui concerne le béarnais<sup>56</sup>, il est très difficile d'avoir une opinion claire en raison des performances très différentes des siffleurs dont nous avons récupéré des données. Chez les plus chevronnés qui sont en général les plus âgés, les voyelles accentuées semblent se comporter comme dans les trois autres langues que nous avons décrites. Chez les siffleurs plus jeunes et maîtrisant essentiellement des phrases stéréotypées, il semble que les voyelles soient essentiellement distinguées par leur accentuation, à l'exception du /i/ qui reste toujours dans les fréquences élevées. Nous pensons que ce phénomène est intéressant pour l'analyse des processus progressifs de perte de vitalité des langues: il semble indiquer qu'à une perte de vocabulaire est associée une perte de précision dans la prononciation de certains aspects du langage, ce qui se manifeste par une dégradation des régularités acoustiques.

### 3.3.3.4.3. Systèmes consonantiques

Dans toutes les langues non tonales observées le système consonantique atteint une certaine complexité qui est limitée par la dextérité des siffleurs et par les contraintes articulatoires dues au sifflement. Nous retrouvons des comportements similaires dans toutes les langues qui sont synthétisés sur les schémas de formes de la figure.

#### **Influences des voyelles**

Les consonnes sifflées sont caractérisées par une combinaison de modulations de fréquence et d'amplitude. La fréquence consonantique modulée dépend des voyelles entre lesquelles la consonne est placée. Dans le cas d'une configuration VCV, la modulation de la consonne débute par la fréquence de la voyelle qui la

---

<sup>56</sup> voir plus de détails en Annexe **Erreur ! Source du renvoi introuvable.**

précède et elle s'achève par la fréquence de la voyelle qui la suit. A l'initiale et à la finale d'un mot, une seule voyelle influence la modulation.

## **Types de modulations simples**

Il existe trois types de modulations de fréquence de consonnes simples (c.a.d. pas clusters):

- Modulation de fréquence interrompue par un silence franc et clair (modulation d'amplitude abrupte pour les consonnes occlusives)
- Modulation de fréquence marquée par une modulation d'amplitude progressive donnant parfois lieu à un court silence, suivant la distance d'écoute (fricatives et nasales)
- Modulation continue (Trille, Tap ou Flap).

En général, les consonnes voisées respectent ces trois distinctions. Le voisement est reproduit par les meilleurs siffleurs par une atténuation de la rapidité de la modulation d'amplitude, ce qui se traduit dans les cas 1 et 2 par des transitions plus douces.

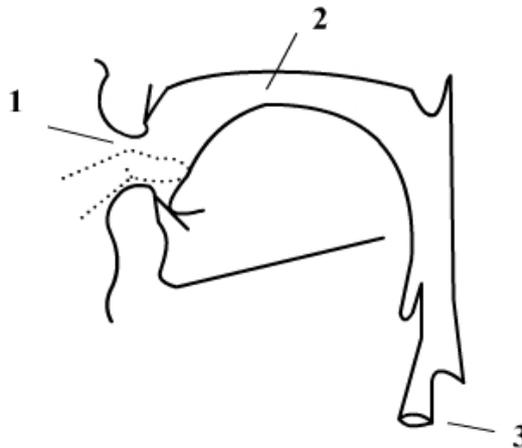
## **Modulations complexes : plusieurs consonnes**

Les transitions consonantiques complexes qui cumulent plusieurs consonnes (clusters) sont marquées par l'ensemble des consonnes qui les constituent. Cependant il existe une limitation importante: Une forme concave et une forme convexe de modulation ne peuvent être observés dans un même cluster que de part et d'autre d'une interruption silencieuse du signal. Ou retrouve les trois cas des transitions simples:

- Interruption nette avec silence
- Interruption marquée par une modulation d'amplitude plus douce
- Modulation de fréquence continue.

### **3.3.3.4.4. Origine articulatoire**

Lors du sifflement, les lèvres sont fixes, la fréquence de résonance est réglée par articulation. Pour ce type de langues non tonales, l'articulation est la plus proche possible de celle réalisée lors de la voix parlée.



**Figure 52 : rappel : articulation du sifflement**

## Fréquences non modulées : voyelles

Comme pour la voix parlée, le siffleur réalise l'articulation des voyelles avec un tractus vocal relativement ouvert c'est à dire que la langue ne vient pas obstruer la réalisation de la vibration sonore. La position de la mâchoire et de la langue vont l'orienter la résonance vers une zone ciblée du palais ou du fond de la bouche. Lors de cette gymnastique, seuls deux des trois critères de l'articulation des voyelles de la voix parlée peuvent être retenus car les lèvres sont fixes ce sont :

- la position antérieure-postérieure
- la hauteur de la langue.

L'arrondissement est perturbé par la position assez fixe des lèvres.

La voyelle /i/ est sifflée en ciblant une résonance vers l'avant du palais (zone alvéolaire peut être un peu plus en arrière), /e/ est sifflée en ciblant la zone post-alvéolaire, /œ/ est sifflée en ciblant une résonance privilégiée dans la zone palatale, /ɨ/ cible la zone vélaire et /a/ une zone au fond du palais (plutôt zone uvulaire). /o/ est réalisée encore plus en arrière et plus bas que /a/.

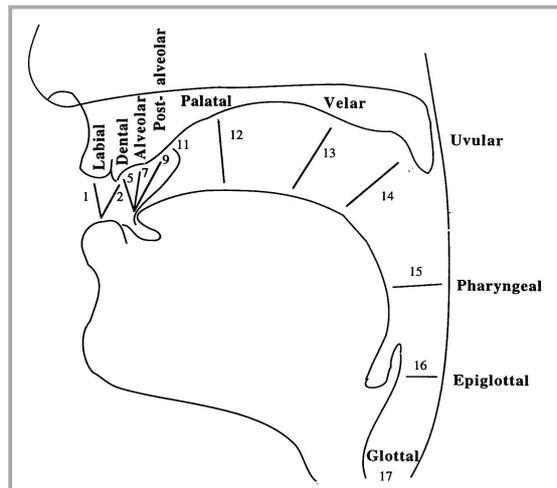
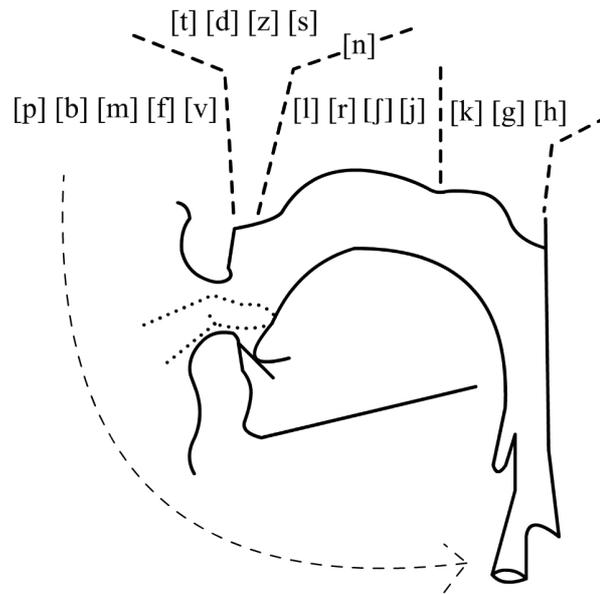


Figure 53 : zone d'articulation du palais, du pharynx et du larynx (in Ladefoged & Maddieson, 1996, p13)

## Fréquences modulées : les consonnes

La direction de la modulation est déterminée par le lieu d'articulation de la consonne sifflée. Il correspond au lieu d'explosion ou de résonance principale de chaque consonne. Cet aspect permet de distinguer plusieurs classes de consonnes qui émergent naturellement à travers la pratique des siffleurs et que l'on peut observer sur sonagramme. Sur la Figure 54 nous avons décrit les articulations sifflées les plus représentatives en fonction du lieu d'explosion ou de résonance principale de la partie supérieure de la bouche (dents, palais, haut du pharynx).



**Figure 54 : Lieux d'explosion ou de résonance principale des consonnes sifflées (l'occlusion glottale joue un grand rôle pour [p], [b],[m], [v],[f], [k],[g],[h])**

**L'articulation, qui est similaire à celle de la voix parlée explique les groupes de consonnes observés dans la forme des modulations sifflées.**

La Figure 54 explique non seulement les classes de consonnes sifflées mais aussi le fait que certaines d'entre elles oscillent entre deux catégories (c'est le cas pour [n]). D'autre part les lèvres fixes perturbent la réalisation sifflée de [f] et [v] ou de la nasale [m], ou des occlusives [p] et [b] qui, comme nous l'avons vu, sont souvent réalisées à l'aide d'une occlusion glottale.

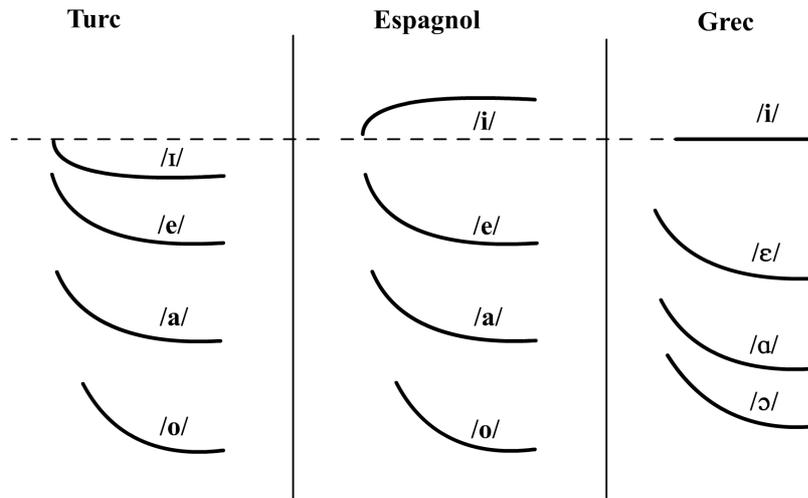
## Conclusion

Trois paramètres principaux déterminent la forme du sifflement:

- premièrement les fréquences des voyelles environnantes liées à une zone de résonance privilégiée dans la partie supérieure de la bouche,
- deuxièmement les lieux d'articulation des consonnes (entraîne souvent une modulation de fréquence (FM))
- et troisièmement le caractère continu, intermédiaire ou interrompu du sifflement (dû à différents types de modulation d'amplitude (AM)).

Ces paramètres se combinent de manière spécifique à chaque langue en fonction des caractéristiques de l'origine parlée à tel point que l'espace sifflé reflète des différences entre les langues à la fois au niveau des voyelles et des consonnes. Nous avons résumé ces différences pour le cas représentatif de la consonne /t/ combinée avec les voyelles de toutes les langues étudiées (Figure 51).

**Modulation du /t/ en fonction des voyelles**



**Figure 55 : Répartition des voyelles sifflées dans chaque langue et modulation du /t/ sifflé**

**(il arrive que la modulation de /ti/ en espagnol soit dirigée vers le haut à cause de la variabilité des fréquences vocaliques et des usages des sifflés)**

D'autre part, nous remarquons que la forme de la modulation fréquentielle est soit concave soit convexe. Tout point d'inflexion résultant d'une variation modulée largement perceptible indique le début ou la fin d'une voyelle. La réciproque n'est pas vraie ce qui a pour corollaire que toute voyelle n'est pas forcément entourée de deux points d'inflexion.

**3.3.3.4.5. Conséquences pour une comparaison avec la voix parlée**

**Précisions sur la phonétique de la voix parlée**

La littérature sur l'articulation de la voix parlée est conséquente. Tout un ensemble de mécanismes ont été décrits à propos de la position des organes articulatoires lors de la prononciation de différentes voyelles: hauteur de la langue ouverture de la bouche, forme et taille de la section du canal d'air laissé libre au niveau du pharynx ou du palais. Ces paramètres ont permis de classer les voyelles et de remarquer l'importance de certains paramètres acoustiques associés à leur spectre fréquentiel, comme par exemple les formants<sup>57</sup>. Nous présentons dans les deux paragraphes suivants une synthèse des données présentées par Stevens (1998) à propos des conséquences de la position de la langue sur les formants de la voix.

<sup>57</sup> Les voyelles de la parole sont caractérisées par leur spectre de fréquences ou timbre vocalique. Ce timbre est constitué d'harmoniques dont les zones les plus intenses en énergie sonore dessinent des regroupements que les phonéticiens ont appelé formants.

## Hauteur de la langue et conséquences sur le signal

En fonction de la position haute ou basse de la langue les voyelles sont également qualifiées de *hautes* ou *basses*.<sup>58</sup> Stevens a observé que le premier formant de la voix est plus stable et plus bas en fréquence pour les voyelles *hautes* (Stevens 1998).

## Distinction antérieure et postérieure

Les voyelles ont été également classées en fonction de la position antérieure ou postérieure de la langue lors de l'articulation: en position antérieure la langue est légèrement plus haute qu'en position postérieure. Stevens (1998) a observé qu'il y avait des conséquences acoustiques communes aux déplacements vers l'avant ou vers l'arrière de la langue: une avancée entraîne une augmentation de la fréquence maximum du second formant de la voix. Cette valeur est plus élevée pour les voyelles *hautes* que pour les voyelles *basses*. D'autre part, les voyelles antérieures sont toujours caractérisées par un vide fréquentiel large entre la zone des harmoniques définissant le formant 1 et la zone des harmoniques définissant le formant 2. Pour les voyelles postérieures le formant 2 de la voix est déplacée vers une zone proche de ses valeurs minimales et il est voisin du formant 1. Pour les voyelles ayant une position de langue intermédiaire, les harmoniques des formants 2 et 3 se combinent pour créer une proéminence spectrale plus élevée que le formant 2. Enfin, dans le cas des voyelles antérieures mais qui ne sont pas *basses*, comme le /u/, un rapprochement des formants 2 et 1 est réalisé par l'arrondissement des lèvres (Stevens 1998). Certaines des configurations d'articulation et de leurs conséquences pour le passage du flux d'air et l'énergie sonore sont présentées Figure 56).

---

<sup>58</sup> Nous mettons en italique ces appellations liées à la position de la langue pour les distinguer d'éventuelles références à une fréquence haute ou basse.

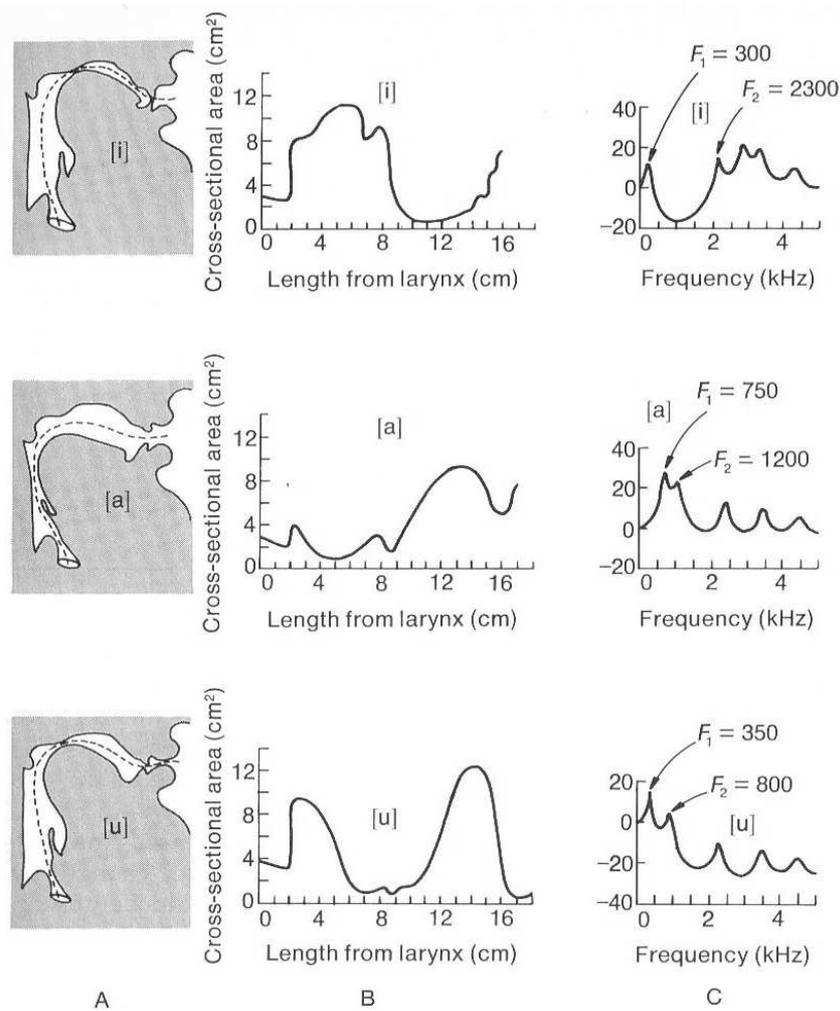


Figure 56 : Articulation et acoustique de [i], [a] et [u] en voix parlée

“Approximate midsagittal sections (A), cross sectional area functions (B), and acoustic transfer function (C) of the vocal tract”(Lieberman 1975, p.75).

## Sifflement articulé et formants

D’après notre description de l’articulation du sifflement des langues sans tons, le signal résultant n’est pas la reproduction directe d’un des formants ni même sa transposition. Il est plutôt le résultat de l’adaptation de l’ensemble des paramètres d’articulation – l’un des principaux est le locus- aux contraintes du sifflement. Ces origines articulatoires communes, ont pour conséquences certaines corrélations acoustiques. A ce propos, comme le deuxième formant de la voix est principalement le résultat de la résonance provoquée par la cavité formée entre le palais et la langue, c’est souvent lui qui sera le plus proche de l’articulation sifflée. Cependant, sa forme acoustique est bien plus diffuse qu’un sifflement. Dès lors il est difficile de tirer des conclusions définitives. La question reste ouverte pour savoir si le sifflement a un locus d’articulation plus élevé en moyenne ou si tout simplement il l’approche de manière plus précise ce qui donne l’impression que la valeur est différente de celle obtenue à partir des formants de la voix. Ces derniers ne permettent en effet qu’une estimation à partir d’une interpolation entre les concentrations de l’énergie portée par les harmoniques.

En effet, jusqu'à maintenant tous les travaux sur le sujet ont expliqué que le locus de sifflement était plus élevé que celui de la parole. Une des preuves visuelles -sur sonagramme- à l'appui de cette interprétation était que la modulation du formant de la voix issu de /ti/ ou de /te/ était vers le bas (comme schématisé pour [di] sur la Figure 57) alors que celle de /ti/ en sifflement était vers le haut.

Pourtant dans notre corpus, lorsque le maestro de Silbo Lino Rodriguez réalise un /ti/ la modulation du sifflement est également vers le bas comme schématisé sur la Figure 55 pour l'espagnol et montré sur le sonagramme Figure 58. Nous ne trouvons par conséquent pas de différence entre sa manière de parler et sa manière de siffler du point de vue des directions de modulation du /t/ et ce sur toutes les occurrences de /ti/.

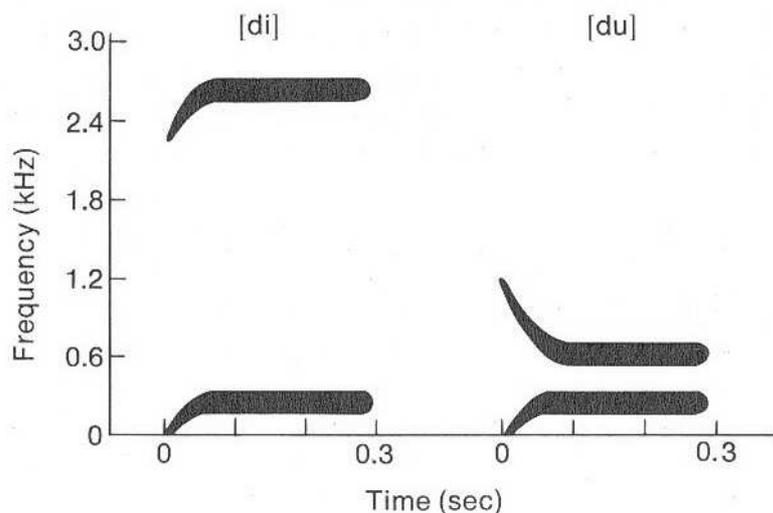


Figure 57 : Schématisation des modulations de formants 1 et 2 occasionnées par la lettre [d] pour deux types de contextes vocaliques : [i] et [u] (Lieberman 1975, p. 77)

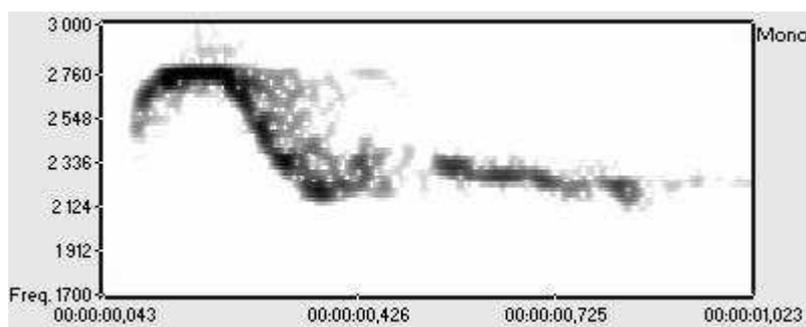


Figure 58 : Mot « tiene » sifflé par le Maestro de silbo Lino Rodriguez

Remarque : Il se peut que la différence observée chez d'autres siffleurs à propos de la voix parlée et de la voix sifflée soit due à des différences de *dialecte sifflé*<sup>59</sup> sur l'île de la Gomera (cette remarque est suggérée par les différences observées entre les langues sur la Figure 55).

<sup>59</sup> Nous avons déjà signalé ce phénomène à la Gomera qui avait aussi été observé par Classe (1963).

## Mouvements acoustiques dus aux consonnes

Malgré la similarité entre formant 2 de la voix et modulation observée chez les bons siffleurs, nous ne pouvons pas conclure que c'est un phénomène systématique. Il n'est valable de manière récurrente que pour les occlusives et le [j] dans les cas où la voix parlée est grave et prononcée lentement. Les liquides [l, r] par exemple se comportent différemment des formants 2 tout en étant issus des mêmes mouvements buccaux. Nous pensons que les modulations acoustiques sifflées sont fortement influencées par la forme et la tension du tractus vocal qui changent avec les différentes consonnes ou voyelles. Le siffleur sur ce point se comporte comme le ferait un ventriloque.

## Voyelles

Dans le cas des voyelles, le parallèle entre formant 2 et sifflement est acceptable en silbo car le /u/ sifflé se positionne à une fréquence plus basse que le /a/. Mais alors seul l'ordre décroissant de /i/ à /u/ entraîne une corrélation. Les écarts fréquentiels entre les voyelles sifflées et ceux des formants 2 ne sont pas du même ordre. Par contre, en turc, en grec, en béarnais ou également en tepehua une telle répartition commune n'existe pas car le /u/ sifflé se positionne au niveau du /e/.

Les regroupements des triangles vocaliques sifflés que nous avons tracés pour le turc et le grec, et même pour le silbo sur certains points, suggèrent plutôt que le sifflement opère une transposition synthétique des éléments les plus saillants des différentes harmoniques de la voix (et donc indirectement les formants sont concernés à différents degrés en fonction du type de voyelle). Dès lors, plusieurs descriptions des voyelles basées sur une étude corrélée de la perception et de la phonétique des voyelles parlées seront intéressantes pour avancer dans la comparaison avec les voyelles sifflées :

-d'une part les analyses des aspects perceptifs des formants décrits par Carlson et al (1970) et Bladen et Fant (1978).

-d'autre part les analyses des rapports de formant (Peterson et Barney 1952, Miller 1989).

Nous aborderons ces aspects lors des tests perceptifs des voyelles sifflées présentés § 4.3.1 dédié à l'intelligibilité des voyelles sifflées.

### **3.3.4. Langues sifflées utilisant une stratégie de transposition intermédiaire<sup>60</sup>**

Plusieurs langues sifflées adoptent une stratégie de sifflement qui ne correspond à aucun de celles décrites dans les études générales traitant des langues sifflées. Il s'agit d'un ensemble hétéroclite de langue du point de vue de la distinction tonales/non tonale. En effet, il regroupe à la fois des langues décrites comme tonales

---

<sup>60</sup> Terminologie employée dans ces paragraphes: « hauteur tonale » fait référence à l'attribut de la perception. « Ton » fait référence à une distinction phonologique de hauteur tonale au niveau du mot. L'intonation fait référence à la « hauteur fondamentale » en général. Nous utilisons cette dernière pour caractériser le noyau vocalique des langues qui n'ont pas de tons.

(surui en Amazonie Brésilienne (com. pers. De Lacerda 2004)) et des langues décrites comme non tonales (kickapoo du Nord du Mexique (Hausler 1960, Voorhis 1971)) ou à tons naissants (chepang, (Caughley, 1976)). Leur point commun est de reproduire des éléments de l'intonation de la syllabe- soit les tons comme en surui, soit l'accent intonatif et le poids des syllabes comme en kickapoo et en chepang- conjointement à des modulations de fréquences pour les consonnes. Ces dernières sont moins développées que celles des langues sans tons mais influencent parfois la hauteur fréquentielle de la voyelle de manière sensible.

Jusqu'à aujourd'hui seules deux langues ayant une forme sifflée de ce type ont été signalées: le kickapoo et le chepang mais seul, le chepang a fait l'objet d'une description et analyse détaillées. Nous avons pu nous procurer quelques enregistrements de chepang<sup>61</sup>. Nous en présenterons quelques analyses complémentaires après avoir résumé les travaux publiés sur cette langue.

D'autre part, le surui, une langue amazonienne du Rondonia (l'un des états du Brésil) reproduit à la fois les tons des voyelles et des éléments de l'articulation des consonnes. La version parlée de cette langue est encore en cours de description, notre analyse est donc une approche préliminaire réalisée à partir de listes de mots collectées lors du travail de terrain de Mariana de Lacerda et de Denny Moore. Une description plus complète sera faite en collaboration avec les équipes de linguistes du Museu Goldi de Bélem au Brésil.

### **3.3.4.1. Langue chepang du Népal**

Le chepang est une langue de la famille Tibéto-Birmane parlée au Népal, dans la région du Chitwan, entre les villes de Hetauda et de Narayanghat. La forme parlée de la langue a fait l'objet de plusieurs descriptions linguistiques. Ces études ont qualifié le chepang de langue à tons naissants « *because consonants affect the stress and pitch of the syllable and in certain instances the presence of a phonemic glottal may be manifested only by a contrastive pitch* » (Caughley 1976 in Sebeok et Umiker Sebeok, p 998).

#### **3.3.4.1.1. Histoire de l'étude de la langue sifflée**

En 1969, la présence d'une version sifflée de la langue chepang, utilisée pour faciliter les communications lors de la chasse, fut signalée par Dahal et Bandhu de la Tribhuvan University aux linguistes de la SIL Caughley et Pike (in Caughley (1976)). Cette forme de la langue permit d'approfondir la réflexion sur la nature tonale ou non de la langue. Une publication de Pike (1970) relate les questions abordées à cette occasion, à partir de la forme sifflée de la langue. Par la suite, une analyse plus complète spécifiquement sur cette version sifflée fut publiée sous le titre « *Chepang Whistled Talk* » (Caughley 1976).

---

<sup>61</sup> Nous remercions ici le linguiste Ross Caughley, car lors de notre visite au Népal, la situation politique locale ne nous a pas permis de nous rendre dans la région où vivent les Chepang.

### 3.3.4.1.2. Le système phonétique

#### Voyelles

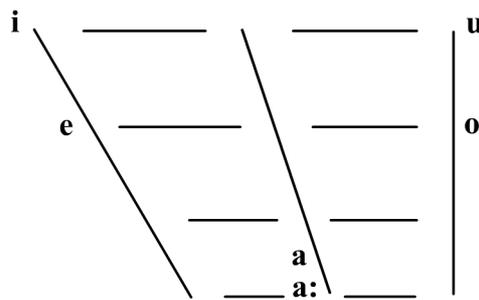
##### Forme parlée

La hauteur de la syllabe dépend des éléments qui la composent: « *The constitution of the nucleus appears to effect the height of the syllable pattern as a whole rather than its shape* » (Caughley 1976). Les caractéristiques du noyau syllabique sont résumées dans le Tableau 19.

**Tableau 19 : Voyelles chepang de la forme parlée (in Caughley 1976, p. 999)**

voyelle	antérieure	centrale	Postérieure (Back: B)
Haute (High: H)	i		u
Moyenne (Mid: M)	e	a	o
Basse (Low: L)		aa	

Ce qui donne le triangle vocalique de la Figure 59 que nous avons adapté au tableau précédent:



**Figure 59 : Triangle vocalique du chepang**

##### Forme sifflée

Caughley observe que l'effet du /i/ et du /o/ sont réguliers. Il les définit ainsi: « *generally higher average pitch with the high front vowel /i/, lower with the low back vowel /o/* ». Il remarque par ailleurs que comme pour la forme parlée, il est possible d'affecter à chaque noyau vocalique une contribution de hauteur à l'ensemble de la syllabe en fonction de deux des qualités de la voyelle: hauteur (H, M ou L) et antériorité (NB)/postériorité (B) (Tableau 20).

**Tableau 20 : Poids du noyau vocalique (NB (pour Non Back) : non postérieure et B : postérieure) (Ibid, p.1005)**

	H : 1	M : 0	L : -1/2	NB : 0	B: -1	Total poids des voyelles
i	+			+		1
e		+		+		0

	<b>H : 1</b>	<b>M : 0</b>	<b>L : -1/2</b>	<b>NB : 0</b>	<b>B: -1</b>	<b>Total poids des voyelles</b>
a		+		+		0
aa			+	+		-1/2
u	+				+	0
o		+			+	-1

Nous remarquons que la répartition des voyelles en fonction de leur contribution en hauteur est similaire à la répartition des fréquences des voyelles des langues sans tons déjà décrites (grec et turc en particulier). Elle délimite ainsi trois groupes de voyelles: /i/ qui tire les fréquences de la syllabe vers le haut. /e/, /a/ et /u/ qui sont neutres et /aa/ et /o/ qui tirent les fréquences vers le bas; l'influence de /o/ étant plus intense que celle de /aa/. Il est important de noter que la voyelle n'apporte qu'une contribution qui est susceptible de varier assez largement avec l'influence des consonnes contrairement à ce que nous avons observé sur les langues non tonales précédentes.

### 3.3.4.1.3. Consonnes

#### Version parlée

**Tableau 21 : phonétique des consonnes chepang**

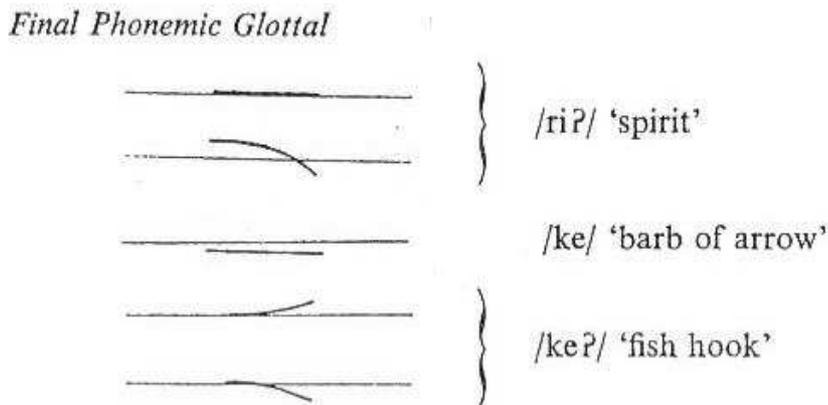
	<i>Bilabial</i>	<i>Labio dental</i>	<i>Labio velaire</i>	<i>dentale</i>	<i>alvéolaire</i>	<i>Post alvéolaire</i>	<i>Retroflexe</i>	<i>Palatale</i>	<i>Vélaire</i>	<i>Uvulaire</i>	<i>Glottale</i>
Occlusive	p b			ṭ ḍ				c ɟ	k g		ʔ
Implosive											
Click											
Trille					ɾɾ						
Tap											
Flap											
Fricative					s						h
Affriquée											
Nasale	m̩			ŋ̩ n̩					ŋ̩		
Latérale fricative											
Latéral approximante					l̩						
Approximante			w̩					j̩			

## Forme sifflée analysée par Caughley

### Effet de la consonne l'intonation de la syllabe et la modulation proche de la voyelle

Comme pour les voyelles, Caughley a développé un tableau des influences de la consonne sur la modulation locale de fréquence avec une distinction entre la position initiale et finale. «*Among the more obvious results are the effect of depressing pitch by a velar such as /k/, and raising of it by the grooved fricative /s/.* » (Ibid, p. 999). Un autre effet moins marqué mais jugé comme «*consistent* » concerne l'abaissement de la hauteur avec le voisement.

Le système de modélisation des modulations fréquentielles sifflées mis ainsi au point par Caughley prédit, de manière assez précise la réalisation des phrases sifflées telle qu'il les a perçues à l'écoute. Sur les données qu'il présente, il parvient à retrouver 216 cas sur 278 extraits (soit environ 77%). L'auteur souligne cependant que certains mouvements ne sont pas bien prévus: ainsi un /k/ en position finale de mot se manifeste souvent par une modulation montante alors que, dans la plupart de ses réalisations, il se manifeste comme une modulation descendante. Pour /ʔ/ en position finale d'un mot la modulation redescend légèrement alors que la plupart du temps elle reste montante (Figure 60).



Note: The variant forms of these no doubt reflect the corresponding variant manifestations of a FINAL phonemic glottal in speech, viz.,

- i) full glottal closure.
- ii) marked decrescendo and falling pitch.

**Figure 60 : Effet de la glottalisation sur le noyau syllabique sifflé (Ibid. p.1001)**

Les contributions des consonnes décrites par Caughley ne définissent pas les mêmes groupes de consonnes que pour les langues non tonales, ni même les mêmes effets. Ainsi /p, t, ʔ/ ont les mêmes conséquences pour le sifflement. C'est également le cas pour /m/ et /n/. Pourtant il précise qu'à l'écoute on peut les différencier clairement: «*while listening to whistle speech under good conditions it is possible even for an inexperienced person to identify some segments by these transients.* » (Ibid. p1017)

## Analyse acoustique complémentaire

### Rôle de l'intonation des voyelles

En ce qui concerne l'intonation, deux points apparaissent remarquables à notre avis:

- Le premier concerne le rôle prépondérant des voyelles. Or la manière dont elles influencent l'intonation est similaire à la manière dont les voyelles du turc, du grec et même du béarnais se répartissent en fréquence. Dans ces dernières langues, les voyelles intermédiaires (comme le /a/ et le /e/) sont toujours celles qui sont le moins clairement définies car leurs bandes de fréquences se chevauchent. Il se trouve qu'en chepang, nous avons pu vérifier que les voyelles intermédiaires /a/, /u/ et /e/ sont plus susceptibles d'interférer entre elles car elles seront plus facilement influencées par les consonnes. Ainsi /a, aa/ varie de 1241 à 1572 Hz, /e/ de 1271 Hz à 1715 Hz et /u/ de 1142 à 1563 Hz alors que /i/ reste proche de 1800 Hz.
- Nous remarquons également que les fréquences vocaliques du chepang sifflé couvrent une bande de fréquence moins large que les langues non tonales transposant le spectre vocalique. Ceci confirme que le sifflement est lié à l'intonation de la voix car la fréquence de cette dernière varie dans une bande plus limitée que ses éléments spectraux.

### Analyse des segments transitoires

Les consonnes ont, à notre avis, deux types de contribution dans la forme sifflée:

- sur l'intonation de la syllabe et la modulation proche de la voyelle étudiée par Caughley en détail,
- sur les transitoires en entraînant des modulations.

### Remarque préliminaire

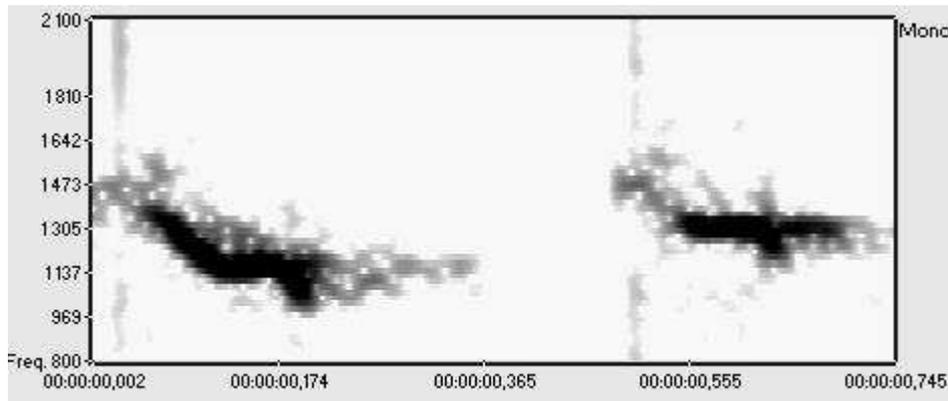
Le deuxième type de contribution consonantique explique certaines possibilités de distinctions (par exemple entre des consonnes telles que /p/ et /t/). Même si Caughley ne les a pas reproduites dans ses schémas et ses tableaux prédictifs, il a pourtant précisé que les consonnes sont articulées « *in a manner as close as possible to spoken speech* » (ibid p.1117). Nous avons cherché à compléter son analyse, car il affirmait lui-même à la fin de son article : « *more elaborate experimentation is needed to determine just what are the identifying features of whistled speech. The most that can be said at this stage is that intonation and context probably play the major rôle in identifying the content of a message while the transients and possibly some pitch variations reduce ambiguity* ». (Ibid. p1018).

## Analyse

A l'écoute des 11 phrases chepang<sup>62</sup> du corpus dont nous avons disposé /d/, /t/, /k/ et la semi voyelle /j/ sont clairement articulées comme dans la voix parlée et parfois sifflées d'une manière proche de celle du sifflement des langues tonales.

Les représentations sur sonagramme permettent de vérifier cette première impression d'écoute:

-Prenons l'exemple du /t/ qui présente une modulation qui la caractérisait déjà dans les langues non tonales. Elle est moins évidente ici car moins bien marquée sur sonagramme (Figure 61). On remarque d'autre part qu'en chepang les dentales sont souvent réalisées en séparant assez clairement la transitoire du noyau de la syllabe.



**Figure 61 : 2 extraits de chepang avec transition en t /to/ et /te/. Le sonagramme est réglé pour faire apparaître les transitoires**

-/k/ est sifflée par un stop et une transitoire. Comme les enregistrements ont été réalisés à courte distance, la transitoire est souvent présente sur le sonagramme comme pour le /t/. Il y a de fortes chances que cet aspect soit atténué à moyenne distance dans la forêt, lors d'un usage pour la chasse :

-/p/ est marquée par un stop comme dans toutes les langues que nous avons analysées,

-Les modulations observées pour /s/ par Caughley sont explicables par la proximité de réalisation sifflée avec le /t/ (Figure 62). Le /s/ sifflé est marqué de manière régulière sous la forme d'une modulation visible sur sonagramme. Il semble que ce phénomène acoustique soit dû à l'insistance des siffleurs sur l'aspect fricatif.

<sup>62</sup> Nous remercions Ross Caughley pour nous avoir prêté une partie de son corpus en raison de notre impossibilité de nous rendre dans la région.

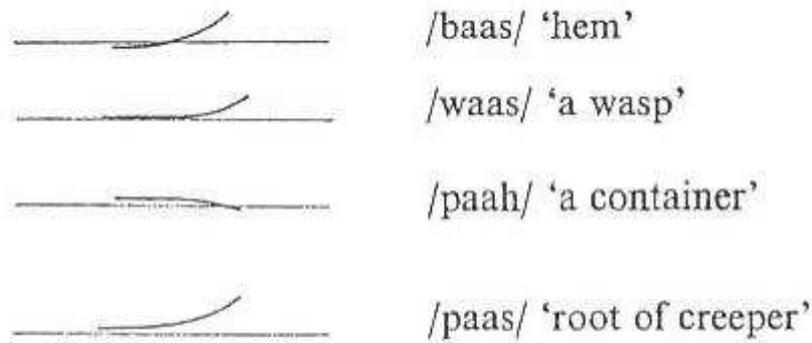


Figure 62 : schématisation de sifflements chepang : illustration du cas du /s/ (Ibid, p.1000)

Par contre, pour d'autres types de consonnes, plusieurs types de réalisation sont présentes sur les enregistrements.

-Ainsi /v/ est parfois sifflée comme une dentale parfois comme un /k/. Ceci pourrait être dû à un effet de coarticulation.

#### 3.3.4.1.4. Conclusion pour la langue chepang

Ainsi le comportement du chepang nous a permis d'observer un type de compromis nouveau de sifflement tranposant à la fois la « Hauteur Fondamentale » et le spectre dans une même syllabe. Il s'exerce à la fois sur les voyelles et les consonnes alors que dans les langues précédentes nous ne l'avions observé qu'à travers l'accentuation sur les voyelles.

Un autre aspect nouveau émerge : il s'agit d'une part de l'importance de l'influence du caractère voisé/non voisé de la consonne sur la fréquence de sifflement du noyau vocalique et d'autre part du fait que la glottale est parfois présente à la finale dans la forme sifflée à travers une modulation de fréquence vers le bas. Ces phénomènes ne s'exprimaient pas de la même manière dans la forme sifflée d'une langue sans ton comme le turc qui possède elle aussi des glottales et fait une distinction légère entre voisement et non voisement (voir mot /dahar/ de la Figure 49).

Ce qui maintient la langue chepang dans le classement des langues non tonales est le fait que la glottale à la finale se manifeste encore par un abaissement de fréquence de l'intonation dans de nombreux cas (Caughley 1976). Le sifflement reflète cet aspect et donc confirme le phénomène. Tous les enregistrements que nous avons écoutés ont été réalisés en sifflement bilabial. Il serait très intéressant de les obtenir avec la technique labiodentale que Caughley signale dans son article, car celle-ci produit assurément des sifflements d'amplitudes plus intenses, ce qui permettrait de confirmer les tendances observées avec un autre point de vue.

### 3.3.4.2. Langue surui d'Amazonie

Le surui fait partie de la famille Monda d'Amazonie. Cette langue a été très peu étudiée et est encore en cours d'analyse linguistique. Les éléments d'analyse phonétique que nous présentons pour la version parlée sont issus du travail de thèse de De Lacerda.

#### 3.3.4.2.1. Système Tonal

Le surui est une langue décrite comme ayant deux tons.

#### 3.3.4.2.2. Le système vocalique

La langue parlée surui a 20 voyelles distinctives phonologiques qui reposent sur le triangle vocalique suivant:

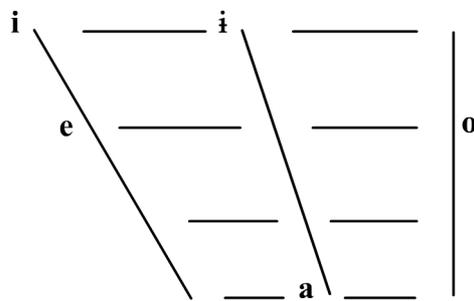


Figure 63 : Triangle vocalique du surui

Chaque voyelle de ce triangle est également présente sous 3 autres formes distinctes : nasale, longue et nasale-longue.

#### 3.3.4.2.3. Consonnes

Le suruí a 18 phonèmes consonantiques, qui sont listés dans le tableau suivant:

Tableau 22 : Tableau phonétiques des consonnes du surui (com. pers. De Lacerda 2004)

	<b>bilabiale</b>	<b>dentale</b>	<b>palatale</b>	<b>vélaire</b>	<b>glottale</b>
occlusive	p b	t d		k g	
nasale	m	n		ŋ	
tap		r			
fricative	β		ʃ		h
affriquée		tʃ (dz rare)			
approximante		l	j		

### 3.3.4.2.4. Système sifflé

Notre analyse du surui s'appuie sur une liste de mots établie par les linguistes du Museu Goldi, Bélem Brazil<sup>63</sup>.

#### Hauteurs des tons

Le surui sifflé s'appuie en partie sur les tons portés par les noyaux vocaliques de la syllabe. Il existe donc deux hauteurs distinctives de sifflements liées aux voyelles. Nous avons analysé la répartition statistique des tons de 86 mots sifflés par un même locuteur (le premier quart de la liste a été sifflé par l'informateur à une fréquence plus haute que les trois quarts suivants).

La moyenne des tons hauts et des tons bas sifflés de la deuxième partie de la liste ainsi que leur variabilité sont exprimés dans le tableau suivant :

**Tableau 23 : Répartition fréquentielle des tons sifflés en surui**

	<b>Tons hauts</b>	<b>Tons bas</b>
Fréquence moyenne	1200 Hz	1070 Hz
Largeur de la bande	1146 -1312 Hz	1004-1117 Hz

En ce qui concerne le premier quart de la liste qui est sifflé sur un autre niveau de fréquences, les tons bas atteignent une valeur moyenne de 1180Hz, ils sont associés à des tons hauts d'environ 1300 Hz.

#### Durée des tons sifflés

La quantité de la voyelle est respectée dans 90% des cas. La version sifflée s'appuie donc sur les durées distinctives ce qui réduit les probabilités de confusion.

#### Remarques à partir des mesures fréquentielles

Ces mesures nous ont permis de remarquer deux phénomènes qui méritent d'être signalés :

-La hauteur du ton ne dépend pas du type de voyelle ni même de sa nasalité ou de sa durée (en cela cette langue se comporte différemment du chepang)

-La fréquence sifflée des tons est relative. Sinon le siffleur n'aurait pas pu changer de niveau de référence de hauteur au cours de l'entretien réalisé avec De Lacerda. Comme pour toutes les autres langues, ce phénomène permet d'utiliser des registres fréquentiels différents en fonction de la distance de communication.

<sup>63</sup> Nous remercions Mariana De Lacerda et Denny Moore pour leur accueil .L'intérêt pour les versions sifflées des langues qu'ils étudient vient du fait qu'ils se servent du sifflement comme d'un outil linguistique d'aide à la description du système tonal.

## Consonnes

Le surui sifflé transpose certaines caractéristiques consonantiques. Nous ne citerons ici que les plus robustes.

### Consonnes facilement reconnaissables pour un auditeur non entraîné

-Le /t/ marque la fréquence d'une modulation ciblant le locus d'articulation du /t/ similaire à celle des langues non tonales. La modulation n'est pas aussi nette que pour le turc, le grec ou le silbo mais elle est très souvent bien marquée au point d'avoir un effet sur le noyau syllabique de manière similaire à ce qui se passe en chepang. /d/ se comporte de la même manière ainsi que /tʃ/. La quantification précise de l'effet du /t/ sur le ton est encore à réaliser et à observer dans des phrases.

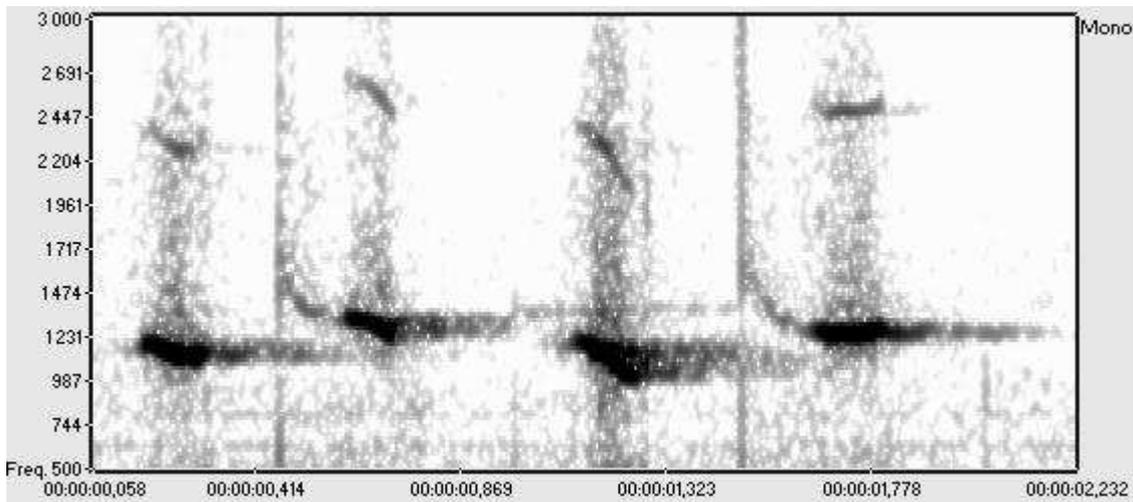


Figure 64 : Sifflement 2 fois de suite du même mot contenant le /t/

Ce sonagramme est aussi instructif pour comprendre le sifflement du/t/ dans les langues du groupe 1 (sans tons).

-Le /k/ est clairement audible dans de nombreuses occurrences. Comme on peut le voir sur la Figure 65 la transitoire apparaît, mais très légèrement. Lorsqu'elle n'est plus visible sur spectrogramme elle est encore perceptible. C'est pourquoi nous avons cherché à l'observer avec un outil de paramétrisation dédié au sifflement humain développé en fin de thèse à l'occasion d'une collaboration (Figure 66). Le /k/ est d'autre part marqué par un silence avec une légère modulation en direction de la transitoire. Il se peut que cet abaissement ne se manifeste pas par une modulation mais par une stabilisation du ton.

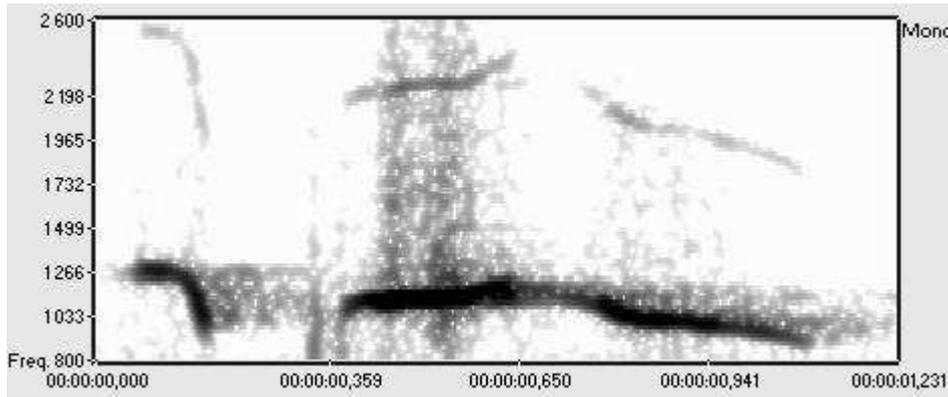


Figure 65: Mot surui /moko:wa:/ en représentation sur sonagramme

On aperçoit la transitoire qui est prononcée (le spectrogramme a dû être réglé en fonction)

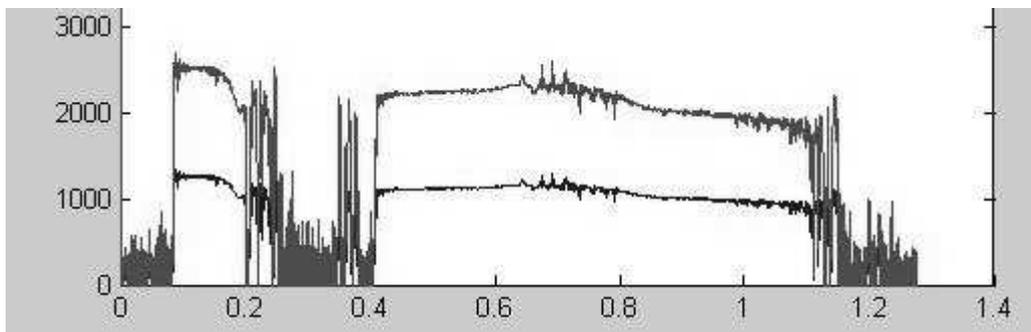


Figure 66 : Mot surui /moko:wa :/ visualisé avec un outil de paramétrisation des sifflements (fréquence fondamentale en bleu et harmonique en vert)

La transitoire est visible beaucoup plus aisément car ce programme a été développé spécifiquement pour les sifflements humains (cf. Annexe A.4.2.4)

A distance dans un environnement sylvestre la transitoire ne sera plus perceptible mais il est important de montrer qu'elle est réalisée suivant un processus d'articulation similaire à celui de la voix parlée.

-La semi voyelle /j/ est sifflée en une modulation de fréquence comme sur le spectrogramme de la Figure 67. La similarité entre le sifflement et le formant 2 est frappante.

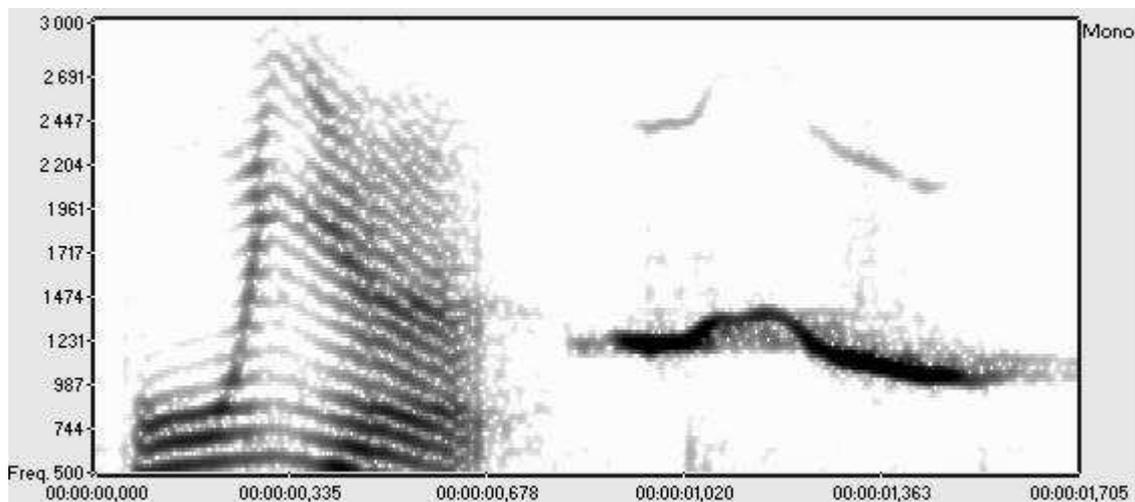


Figure 67 : Syllabe /uja/ en voix parlée et en parole sifflée

## Autres consonnes ayant parfois un comportement similaire aux langues non tonales

-/r/ est dans 80% des cas sifflée sous une forme proche du /r/ des langues non tonales mais il existe des occurrences différentes,

-/n/ est sifflé comme dans les langues non tonales. Il est très proche de /r/ et comme lui il est parfois sifflé en une continue et d'autres fois plus comme un /t/.

/n/ et /r/ sont les seules consonnes avec le /j/ et le [v] à être très souvent réalisées de manière continue.

Le comportement variable de toutes ces consonnes cache peut être un aspect de la phonologie que nous ne maîtrisons pas.

## Consonnes ayant un comportement différent des langues non tonales

/l/,/v/,/β/, /m/,/ŋ/,/ŋ/ sont sifflés de manière différente des consonnes des langues tonales. Pour certaines plusieurs types de modulations ont été observés, par exemple pour /ŋ/ qui est sifflée souvent avec la même qu'un k ou un g.

### 3.3.4.2.5. Conclusion pour le surui

Malgré le peu de données analysées le sifflement de la langue surui présente un intérêt particulier ici dans toute notre étude car il souligne qu'il existe des langues pour lesquelles le siffleur siffle le ton des voyelles, donc la hauteur fondamentale de la voix, et conjointement le spectre de la voix, pour certaines consonnes (parfois avec des similarités saisissantes avec les formants de la voix parlée comme pour le [j]). Mais nous n'observons pas cette stratégie sur toutes les consonnes. Le fait que certaines consonnes puissent être transmises soit d'une manière proche du sifflement des langues non tonales soit d'une autre manière semble indiquer à nouveau que ce qui est important dans les consonnes est avant tout l'aspect temporel.

Dans la plupart des cas, nous n'avons pas analysé l'effet d'influence des consonnes sur le centre du noyau vocalique mais uniquement sur les bords extérieurs proches des transitoires. Une analyse systématique à partir d'un corpus plus large permettrait sûrement de tirer des conclusions instructives et de les comparer avec le chepang en particulier en ce qui concerne l'effet du voisement, de l'aspiration ou de la glottalisation.

Cette première approche a été faite à partir d'un sifflement bilabial utilisé pour une très courte distance. Il serait intéressant de réaliser des enregistrements avec d'autres techniques si elles existent, en condition réelles d'usage. Ceci permettrait d'avoir un autre point de vue avec un sifflement plus intense. Une analyse avec des phrases serait du plus grand intérêt.

### **3.3.4.3. Analyse comparative des systèmes de transposition intermédiaires**

Les deux systèmes linguistiques que nous avons étudiés sont similaires pour un certain nombre de caractéristiques essentielles qui expliquent pourquoi nous les avons regroupés<sup>64</sup>.

Leur principal point commun réside dans le fait de reproduire certains des aspects articulatoires des consonnes tout en les atténuant par rapport à ce que nous avons décrit pour les langues sans ton. Ce comportement est très intéressant il offre un second point de vue articulatoire sur les consonnes et indique des phénomènes sous jacents aux modulations du turc, du grec ou du silbo. Le caractère fondamental de l'aspect temporel des consonnes est ainsi précisé.

Leur second point commun est de ne pas reproduire le timbre des voyelles, mais, soit l'intonation, soit le ton porté par la voyelle. Dans ces deux cas, le sifflement du noyau vocalique couvre un domaine de fréquence moins large que celui des voyelles des langues à tons ce qui indique que ce sont essentiellement des aspects de Hauteur Fondamentale qui sont transposés (donc liés à la vibration des cordes vocales).

Cependant de nombreux paramètres influencent la fréquence du sifflement :

- Pour les deux langues l'influence de l'articulation de la consonne semble importante. Une étude exhaustive reste à faire concernant le surui. L'explication de cet aspect permet d'aborder un nouveau type de problèmes linguistiques à travers le filtre des sifflements. Par exemple le rôle du voisement dans l'abaissement du sifflement<sup>65</sup> en chepang ou l'influence de l'aspiration et de la glottalisation en terme de hauteur en chepang et en kickapoo (Voorhis 1971). L'analyse des langues sifflées peut alors confirmer de manière précise si un processus de tonogénèse (Hombert et al, 1979) ou plus généralement de transphonologisation<sup>66</sup> (Hagège et Haudricourt 1978) est en cours.

-Par contre les deux langues se différencient au niveau de l'influence des qualités de la voyelle sur le sifflement. En effet, alors qu'en surui, le ton n'est pas marqué par le type de voyelle sifflé, dans le cas de la langue chepang, l'intonation sifflée est fortement marquée par les qualités d'antériorité/postériorité et de hauteur de la langue. Sur ce dernier point, le chepang est proche des langues sans tons du groupe 1 et le surui proche des langues à ton du groupe 3.

---

<sup>64</sup> Jusqu'à aujourd'hui, deux autres langues sifflées (qui ont été partiellement décrites au niveau linguistique), semblent faire partie de ce groupe intermédiaire de transposition. Il s'agit du kickapoo du Mexique et du wam de Papouasie Nouvelle Guinée. Nous donnons une approche succincte de la langue kickapoo en Annexe **Erreur ! Source du renvoi introuvable.**

<sup>65</sup> Dans un phénomène indiquant qu'un processus de tonogénèse (naissance de ton) est peut être en cours dans cette langue.

<sup>66</sup> Transphonologisation : phénomène lié à la nécessité de maintenir certaines oppositions alors que la langue change : « Une opposition ayant une valeur distinctive est menacée de suppression ; elle se maintient par déplacement d'un des deux termes, ou de l'opposition entière, un trait pertinent, de toute manière, à distinguer ces termes » (Hagège et Haudricourt 1978).

Ce groupe confirme donc que la forme sifflée s'adapte à la structure de la langue de manière subtile et en représente des éléments phonologiques parmi les plus saillants. Il nous rappelle qu'entre des langues sans tons ayant un accent très peu émergent et des langues ayant un grand nombre de tons, il existe tout un ensemble de possibilités de structures dynamiques linguistiques.

### **3.3.5. Langues dont la transposition sifflée cible la hauteur fondamentale de la voix**

On estime que les langues à tons représentent plus de 60 à 70 % des langues du monde. On appelle « ton » l'utilisation de la « Hauteur Fondamentale » pour véhiculer de l'information qui permet de distinguer phonologiquement les syllabes au point de changer la sémantique des mots. Il existe une grande variété de langues à tons. D'après notre enquête, les formes sifflées des langues qui ont plus de deux tons phonologiquement distinctifs reproduisent principalement en sifflement le ton et ses variations au niveau fréquentiel. Afin d'illustrer ce type de comportement, nous présentons ici deux langues sifflées tonales : le mazatèque et le hmong. Nous verrons que l'analyse de leur forme sifflée permet d'approfondir l'idiosyncrasie de leur système tonal.

#### **3.3.5.1. La langue Mazatèque : groupe ethnolinguistique chjota éna<sup>67</sup>**

La langue sifflée mazatèque est une langue de la famille Otomangue. Elle est parlée dans toute une région dont la partie montagneuse est appelée « Sierra Mazateca » ou zone haute. Elle possède de nombreuses variantes. Un rapport de l'Instituto Nacional de Anthropologia la décrit ainsi: *"La langue mazatèque appartient au groupe linguistique olmèque-otomangue, sous-groupe otomiano-mixteco, famille popoluca. La diversification de cette famille remonte à plus de 500 ans avant J. C. avec la séparation du mazatèque du chocho, de l'ixtatèque et du popoluca. La diversification interne du mazatèque est évaluée à environ 1000 ans après J. C.. La langue mazatèque est orale<sup>68</sup>, elle possède dix variantes dialectales."* (Instituto Nacional de Anthropologia 1999, p.19, traduction libre) Malgré le nombre important de variantes, le dynamisme local et les échanges ont maintenu une bonne intercompréhension par zones géographiques. L'intercompréhension est moins bonne entre la zone haute et la zone basse. La version sifflée est utilisée principalement dans la zone haute. Les hommes l'utilisent et les femmes la comprennent mais ne sifflent presque pas. *"Cette langue fonctionne de manière équivalente au langage parlé : on siffle le même nombre de syllabes avec leur tons correspondants, de cette manière il est possible d'articuler des phrases entières"* (ibid p.22, traduction libre). Notre travail porte sur deux variantes proches de la zone haute : le Mazatèque de Huautla et le Mazatèque de Eloxochitlan. Celles ci se différencient essentiellement par des habitudes différentes de vocabulaire liées au degré de contact différent avec la langue dominante espagnole.

---

<sup>67</sup> Dénomination en langue mazatèque qui signifie : « les gens de notre langue ».

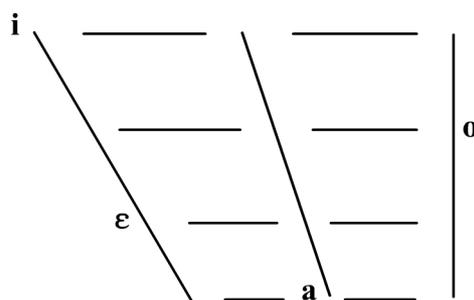
<sup>68</sup> Une pierre gravée qui se trouve à Elokochitlan possède un type d'écriture mixtèque (Ñuine), qui n'a été retrouvée à aucun autre endroit. Il n'a pas été non plus retrouvé d'écritures en hiéroglyphes.

### 3.3.5.1.1. Système tonal et notations

Les variantes que nous avons étudiées utilisent 4 tons distinctifs au niveau phonologique. Les linguistes travaillant en Amérique centrale ont l'habitude de les noter par des chiffres après la voyelle de la syllabe comme sur l'exemple ci dessous: *Haut* : *ti1*, *Semi-Haut* : *ti2*, *Semi-Bas* : *ti3*, *Bas* : *ti4*. Notre informateur principal de la ville de Huautla de Jimenez, Juan Casimiro, est un écrivain local ayant travaillé avec les linguistes de passage. Il utilise la notation de l'écriture mazatèque enseignée aujourd'hui dans certaines écoles de la région qui marque les tons Haut et Semi-Haut avec un accent sur la voyelle et laisse Bas et Semi Bas sans accent. Lors d'un travail linguistique de description, il préfère noter les tons ainsi: *Haut*: *tí'* ; *Semi-Haut*: *tí̃* ; *Semi-Bas* : *tī̄* ; *Bas* : *tī̀* car cette notation prête moins à confusion (comm pers Casimiro 2003).

### 3.3.5.1.2. Voyelles et consonnes du Mazatèque parlé

La langue parlée à Huautla de Jimenez a fait l'objet de plusieurs descriptions linguistiques. Pike et Pike (1947) ont fourni une première approche qui récemment a été réinterprétée puis simplifiée par Golston (2002). La différence entre le système décrit par Golston et celui décrit par Pike et Pike tient à l'attribution de la glottalisation et de l'aspiration au noyau vocalique de la syllabe plutôt qu'à la consonne. D'après Golston, il n'y a pas de différence en mazatèque entre pré et post glottalisation ou pré et post aspiration ce qui a pour conséquence que le Mazatèque a une structure de syllabe simple et non hautement articulée comme Pike et Pike le soutenaient. Une telle remarque rapproche le mazatèque des autres langues de la même famille. Le système d'écriture utilisé aujourd'hui par Casimiro dans les traductions de livres qu'il fait pour étoffer la littérature disponible en mazatèque semble être influencé par celui de Pike<sup>69</sup>. Il signale en effet l'aspiration par une lettre j (par exemple « jme » noté « hme » par Pike et m̃e par Golston) et l'aspiration par un accent avant ou après la consonne à ne pas confondre avec les accents de tons placés sur les voyelles (par exemple so'nde de Casimiro noté [soʔnde] par Pike et Cowan et [soŋde] par Golston).



**Figure 68 : Triangle vocalique du mazatèque (les versions nasales, creaky (glottalisées) et aspirées des mêmes voyelles existent aussi)**

<sup>69</sup> Casimiro a été un informateur de la SIL (créé en partie par Pike) jusqu'à ce que les pressions de prosélytisme le poussent à arrêter cette collaboration (com. pers. Casimiro 2005)

Tableau 24 : phonétique des consonnes mazatèques

	<i>Bilabiale</i>	<i>Labio dentale</i>	<i>Labio palatale</i>	<i>Labio velaire</i>	<i>dentale</i>	<i>alvéolaire</i>	<i>Post alvéolaire</i>	<i>Retroflexe</i>	<i>Palatale</i>	<i>Vélaire</i>	<i>Uvulaire</i>	<i>Glottale</i>
stop						t				k		ʔ
implosif												
clic												
trille						r						
Tap												
flap												
Fricative	β					s	ʃ					h
Fricative sib												
Affricate						ts̺	tʃ̺	tʂ̺				
Affricate sib												
Nasal	m					n			ɲ			
Lateral fricative												
Lateral approximant						l						
approximant									j			

Toutes les consonnes qui marquent un stop : [t], [ts̺], [tʃ̺], [tʂ̺] et [k] ont des formes distinctives en versions aspirées, aspirées nasalisées, glottalisées nasalisées et nasalisées. Les nasales [m], [n] et [ɲ] et les *glides* [β] et [j] ont toutes des versions distinctives aspirées et glottalisées.

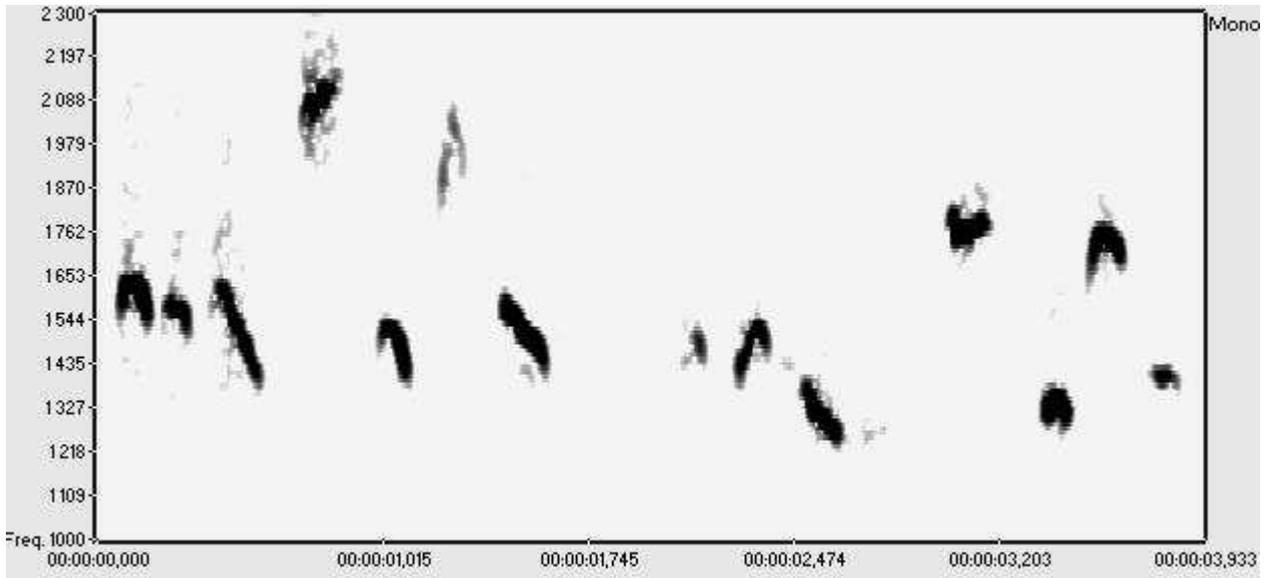
### 3.3.5.1.3. Version sifflée

#### Sifflement des tons

Notre description de la version sifflée est le résultat d'une enquête dans plusieurs villages. Nous nous sommes concentrés sur deux lieux différents : Huautla de Jiménez et Eloxochitlan de Flores Magon. Les deux variantes étudiées sont très proches et sont principalement le fait de différences de vocabulaire. De plus nos deux informateurs principaux avaient tous les deux de la famille dans la région de l'autre variante.

Pour la phrase : « Dans le monde entier les hommes parlent en sifflant » le locuteur de Huautla (Juan) a donné la traduction suivante (en écriture mazatèque actuelle) : « Nga kijnda so'nde tji'n chjota χi male b'exi » et le locuteur de Eloxochitlan (Miguel) a donné la traduction : « Nga je so'nde tji'n tcha χi b'exi ». La première traduction est plus littéraire que la seconde.

La partie commune de ces deux phrases est représentée sur le sonagramme Figure 69. Le premier siffleur est à une fréquence moyenne légèrement supérieure au second. On observe une correspondance exacte entre le nombre de syllabes en version parlée et sifflée et que certaines modulations comme celle de la troisième syllabe sont marquées dans chacune des énonciations. D'autres modulations comme celle de la fin de la phrase ne sont pas phonologiques mais dues à la position de la syllabe dans la phrase.



**Figure 69 : Partie commune des deux phrases citées dans le texte, « Nga so'n-de tji'n χi b'e-χi » (en sept unités de parole sifflée, en notation mazatèque le χ est un /ʃ/ et le j est un /h/ )**

Si nous considérons maintenant le début de la phrase de Juan Casimiro on constate que les quatre registres de tons disponibles en mazatèque y sont présents : « Ngā kījndá sō'ndē tji'n chotā »<sup>70</sup>. Sur la version sifflée Figure 70 on observe bien quatre hauteurs de tons témoins du fait que le sifflement s'appuie en grande partie sur les distinctions tonales. On remarque également deux modulations de tons vers le bas qui montrent que toute la complexité des tons de la voix est reproduite et pas seulement la hauteur de registre.

Sur tout notre corpus nous avons pu observer quatre hauteurs de sifflement relativement stables. Elles sont indépendantes des distinctions de voyelles des triangles vocaliques (voir Figure 71 : statistique de la répartition des tons de 6 longues phrases sifflées par Juan Casimiro).

<sup>70</sup> En écriture mazatèque avec notation des tons adoptée dans l'éducation scolaire mazatèque (voir p. 150).

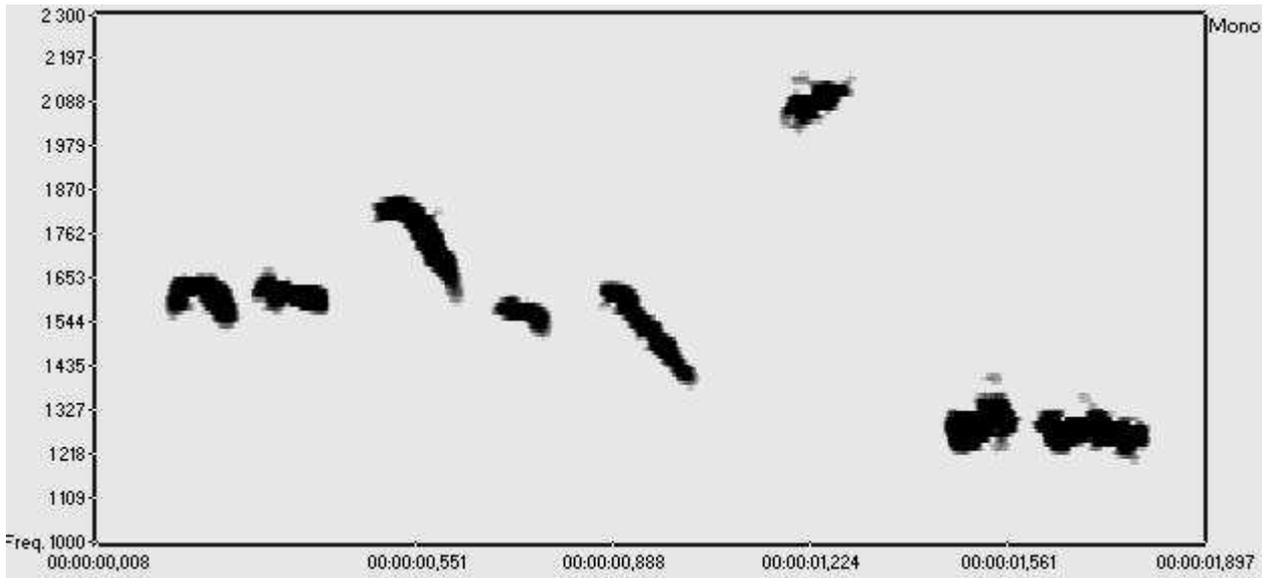


Figure 70 : Début de phrase du siffleur Juan : « Ngā kījndá sō'ndē tji'n chotā »<sup>71</sup>

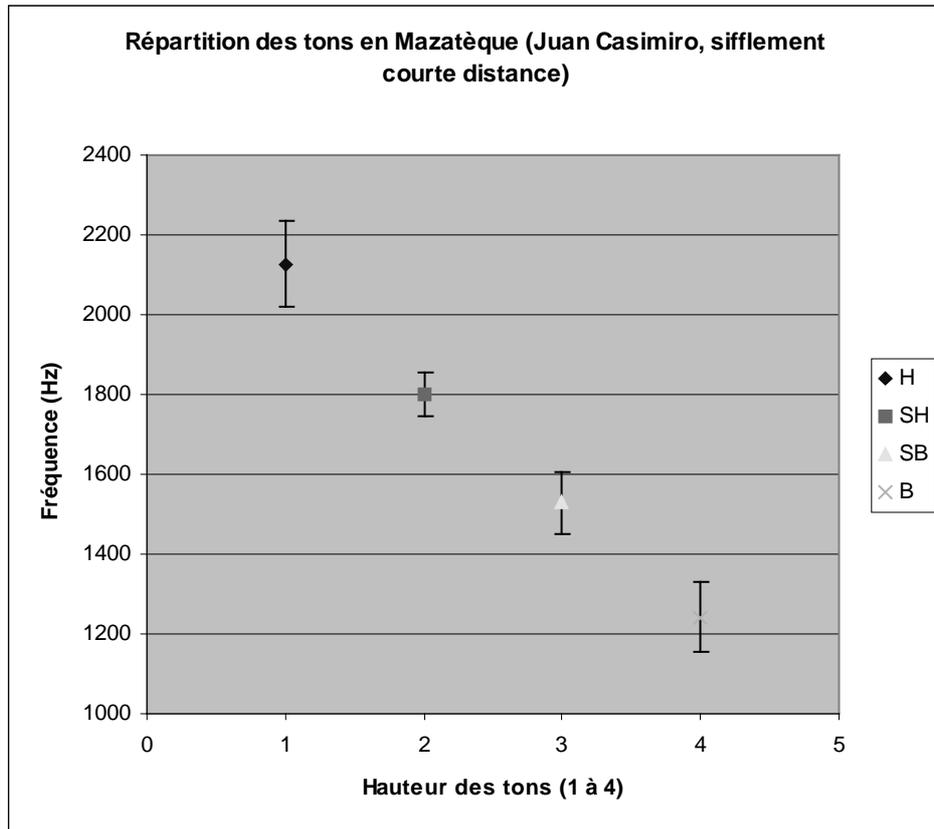


Figure 71 : Statistiques de la répartition des tons d'un siffleur utilisant une seule technique à une distance donnée

<sup>71</sup> En écriture mazatèque adoptée par l'éducation mazatèque avec notation des tons.

## Modulations de tons

### Types de modulation

Plusieurs modulations récurrentes apparaissent en sifflement, elles peuvent être de quatre types types :

-soit ce sont des contours issus de registres (ou glides internes à la syllabe mazatèque: Cowan en a signalé 6 différents dans ses exemples : 2-4, 3-4, 4-3, 1-3, 3-2 et 2-3)

-soit elles sont dues à des diphtongues où chaque voyelle porte un ton différent (aspect non mentionné par Cowan et que nous illustrons sur la Figure 72)

-soit elles sont dues à une fin de phrase ce qui fait qu'elles dépendent de l'effort du siffleur à la fin de son élocution (exemple déjà mentionné Figure 69)

-soit, et ceci est spécifique au sifflement rapide, elles sont dues à une liaison sifflée entre deux syllabes portant des tons différents.

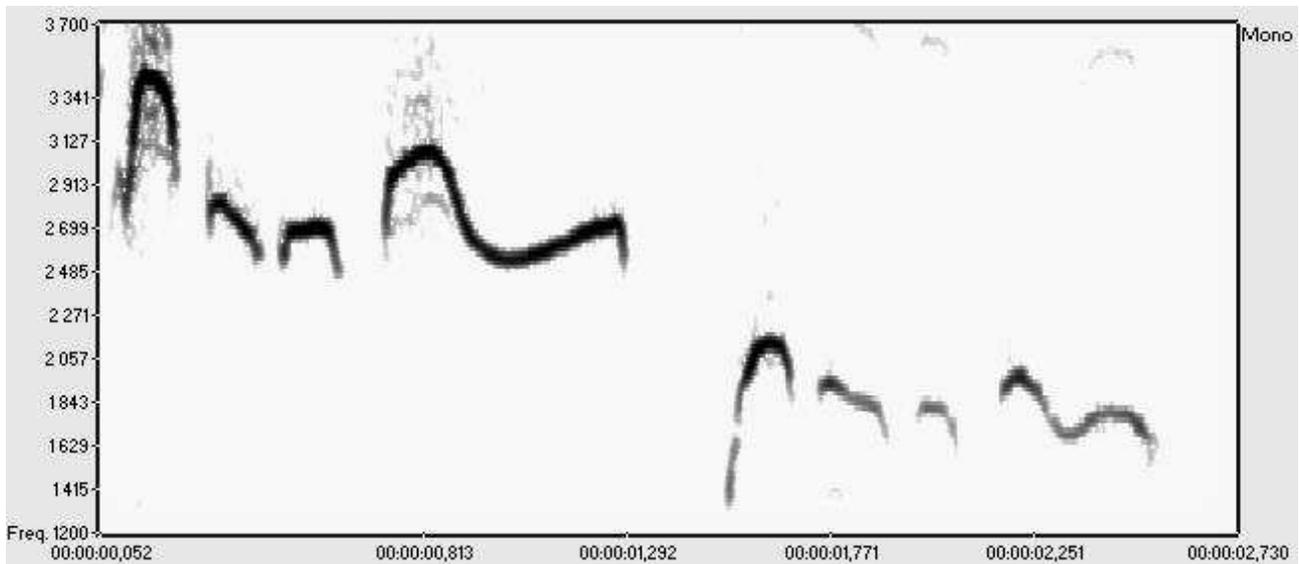


Figure 72 : Phrase “Jme' χi fin.îē”(notation mazatèque) ou /hme1 ʃi3 ti-pai2-3/ (notation phonétique)

La phrase est sifflée deux fois à longue et courte distance par la même personne. La modulation 2-3 de la dernière syllabe apparaît dans les deux cas.

### Directions des modulations

Juan Casimiro nous a communiqué son interprétation des tons sifflés (Figure 73), qui, après vérification sur nos enregistrements est tout à fait adaptée, bien que légèrement différente de celle de Cowan (1948). Sa description ne considère que les aspects relatifs des modulations, venant de la part d'un siffleur formé à la linguistique, c'est un point important<sup>72</sup>. Les modulations 2. 3. 5. et 6. correspondent à des modulations internes au syllabes. 1. et 4. correspondent à des modulations entre des syllabes mais pour que son système

<sup>72</sup> Les notations de Cowan distinguent par exemple la modulation 1-3 et 2-4 qui correspondent toutes les deux à la descente rapide signalée au point 5 de la Figure 73.

soit complet il faut considérer que le point 6 peut aussi concerner des modulations entre syllabes, sinon la modulation de la diphtongue Figure 72 n'est pas prise en compte.

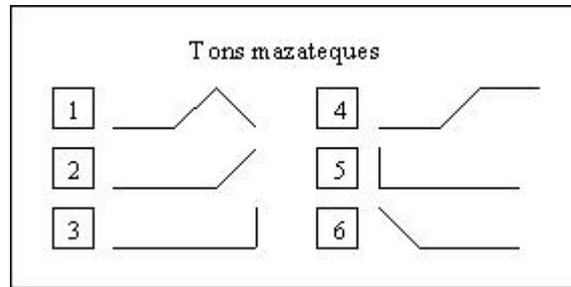


Figure 73 : Sifflement des modulations des tons mazatèques (com. pers. Casimiro 2003)

### 3.3.5.1.4. Historique de l'étude de la langue sifflée Mazatèque

Malgré le fait que la langue sifflée mazatèque soit la deuxième langue sifflée la plus connue du milieu scientifique après le silbo, il n'y a eu que deux études publiées spécifiquement sur cette forme de la langue. La première est celle réalisée par Cowan (1948), collaborateur de Pike et Pike. C'est également la première véritable étude linguistique sur une langue sifflée. Il y décrit de nombreux aspects qui sont vérifiables grâce à des exemples donnés en notation phonétique à la fin de son article. Il a également remarqué la différence entre forme sifflée et forme parlée sur les *glides* : « *In one respect only do the glides of the whistle differ from the glides of the spoken language* » (Cowan 1948, p.284). Il n'a pas fait de description détaillée sur ce point mais a remarqué que toutes les distinctions du mazatèque parlé sont présentes en sifflement.

La seconde est une étude uniquement acoustique (Busnel, 1974a). Le sonagramme, le spectre de fréquence d'un mot sifflé par plusieurs personnes à plusieurs distances a été analysé. Des aspects liés au milieu et aux techniques de production ont également été analysés.

### 3.3.5.1.5. Conclusion pour le mazatèque

La langue mazatèque nous permet d'observer la première langue sifflée tonale qui ne reproduit pas les consonnes par une production sonore. Elles sont présentes sous forme de silences dont nous n'avons pas encore étudié les détails mais il est certain qu'ils jouent un rôle non négligeable au niveau rythmique dans la phrase. Cowan a fait la remarque fondamentale « *In spite of the high probability of ambiguity, the actual instances where confusion occurs are amazingly few* ». (Cowan 1948, p.283). qui souligne que malgré l'élimination en sifflement des données fréquentielles liées aux consonnes, l'intelligibilité de la phrase reste élevée. Cet aspect est caractéristique de toutes les langues tonales sifflées qui ne s'appuient que sur la Hauteur Fondamentale.

### 3.3.5.2. Langues tonales sifflées avec une feuille

Deux « langues sifflées avec feuille » des peuples montagnards d'Asie du Sud-Est ont pu être analysées. Il s'agit du hmong et du akha. Ce sont deux langues sifflées de familles linguistiques différentes développées dans les mêmes milieux. Le hmong possède une structure tonale de 8 tons répartis sur cinq niveaux de

hauteur. Dans cette structure complexe tonale, les différences de formes des modulations jouent un rôle très important. Les linguistes pensent que dans un tel cas, les informations linguistiques ne sont plus seulement codées dans les différences relatives entre les tons ou dans leurs positions respectives mais dans un prototype de forme appelé contour tonal (Yip 2002).<sup>73</sup> « *If a language places a heavy information on tonal contrasts, level tones do not suffice* » (Yip, 2002, p27). Pour décrire un contour, plusieurs méthodes ont été appliquées dans le passé : indication des fréquences initiales et finales mais également des pentes des modulations et les maxima ou minima par lesquels elles passent. Une forme typique est également souvent utilisée.

Le akha utilise quant à lui un système de 6 tons répartis suivant trois niveaux de hauteur et possède également des contours. La durée des tons est un facteur important qui permet pratiquement de dédoubler le système tonal akha. La description du akha est fournie en Annexe **Erreur ! Source du renvoi introuvable.**

Les formes sifflées akha ou hmong n'ont jamais été décrites au niveau linguistique. Seul le système hmong a par ailleurs fait l'objet d'une analyse en ethnomusicologie (Brunet 1972) et en acoustique (Busnel et al 1989). Ces deux approches avaient permis de remarquer un parallèle saisissant avec les tons de la voix et invitaient les phonéticiens à s'y intéresser. Comme nous l'avons vu dans le chapitre précédent, de nombreuses particularités acoustiques du timbre de la *voix sifflée* de la feuille font qu'à notre avis la feuille sifflée est particulièrement bien adaptée au sifflement des langues tonales ayant de nombreux contours de tons<sup>74</sup>.

### 3.3.5.2.1. La langue sifflée avec feuille des Hmongs (Hmoob)

La langue hmong fait partie de la famille Sino-Tibétaine, Hmong Mien. Elle est originaire des montagnes du sud de la Chine, du nord du Vietnam et du Laos. Elle est constituée d'un grand nombre de dialectes locaux qui se distinguent en plusieurs groupes ethnolinguistiques principaux: parmi eux on trouve le Ghao Xong ou Hmong rouge, le Hmu ou Hmong Noir et le A Hmao parlés surtout en Chine ; on trouve également le Hmong Njua ou Hmong vert, le Hmong Daw ou Hmong Blanc dont les populations sont réparties en Chine, au Laos, au Vietnam, en Thaïlande et, depuis le 18<sup>ème</sup> siècle, également en Australie, en Guyane Française, en France métropolitaine et aux Etats Unis ou au Canada. La langue utilisée dans notre corpus est le Hmong Daw ou Hmong Blanc.

---

<sup>73</sup> Les contours de tons du hmong sont différents qualitativement des modulations internes à une syllabe observées en mazatèque. Celles ci pouvaient toujours être expliquées par le regroupement de deux registres (ou niveaux) de tons par des phénomènes cumulatifs (par exemple mot + forme conjuguée). En Hmong de telles explications n'existent plus.

<sup>74</sup> Chez les Hmongs c'est la seule forme du sifflement que nous ayons pu trouver alors que chez les Akhas, nous avons rencontré dans un village du Laos une forme sifflée sans feuille. Même si notre enquête n'est pas exhaustive en raison de la très grande zone géographique couverte par ces populations, nous pensons que ces données méritent d'être signalées car elles ont été confirmées par les associations de peuples montagnard avec lesquelles nous avons travaillé (MPCD-SEAMP pour les Akha et IMPECT pour les Hmong).

## Alphabet

Un certain nombre d'alphabets différents ont été utilisés pour écrire le hmong, le plus largement répandu aujourd'hui pour le hmong blanc est le Romanised Popular Alphabet (RPA).

## Tons

Le hmong est une des langues qui possèdent le plus de tons distinctifs au monde. En hmong Daw on compte 8 tons répartis sur cinq niveaux. En plus des niveaux de tons, le système tonal utilise des modulations de tons qui affectent une même syllabe et qu'on appelle contours<sup>75</sup>. En RPA, les tons sont signalés par l'ajout d'une consonne à la fin de chaque mot. Ceci est possible car les mots ont une structure de type CV et les rares mots se terminant par une consonne sont des emprunts (com. pers. Ly 2005).

Le système tonal est le suivant :

**Tableau 25 : Tableau des tons du hmong blanc**

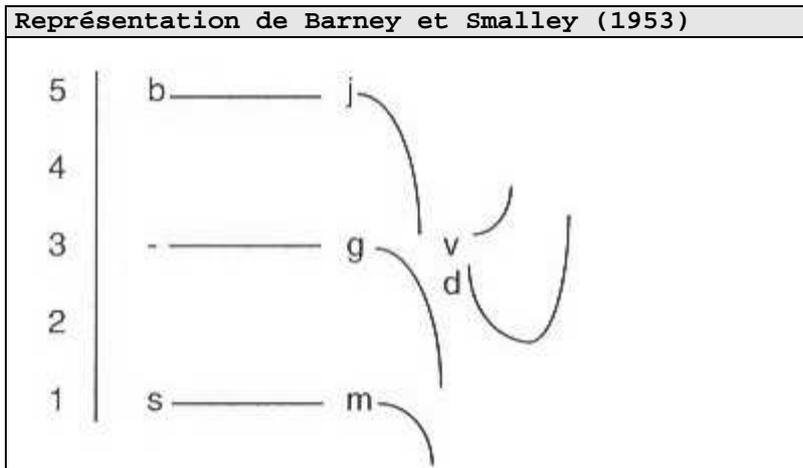
Description RPA	Notation phonétique
Il y a donc 5 niveaux et 5 contours modulés et 3 contours plats	
b : ton haut	55
j : ton haut tombant	52
v : ton moyen montant	34
Rien : ton moyen	33
s : ton mi-bas	22
g : ton mi bas avec expiration	21
m : ton bas glottalisé (légèrement descendant)	1?
un ton bas montant noté d n'a lieu que dans certains contextes : il remplace le ton m à la fin d'une phrase	

On remarque dans ce tableau que la description RPA est plus vague que la description phonétique. Ceci a une raison simple, certains contours ont lieux à plusieurs hauteurs (par exemple le ton v peut partir du niveau 3 ou du niveau 4).

<sup>75</sup> Le hmong est une des rares langues à cinq hauteurs distinctives de tons et à plusieurs formes de contours à n'avoir pas pu être réduite à un nombre de 4 niveaux (Yip 2002).

Certains auteurs ont représenté les tons du hmong grâce à un système de visualisation des contours :

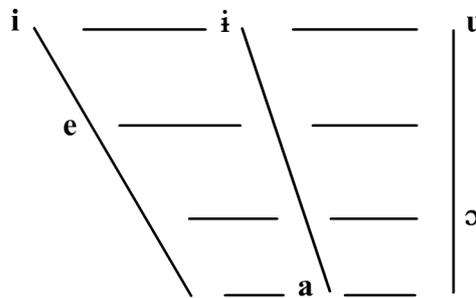
**Tableau 26 : Schéma des contours de tons d'un dialecte hmong différent de celui du RPA**



Ce type de représentation permet de visualiser la variabilité des systèmes tonals en fonction des dialectes. Le système RPA que nous utilisons dans notre étude est celui utilisé pour le hmong blanc et le hmong vert.

## Voyelles

Les voyelles sont présentées sur la figure ci-dessous :



**Figure 74 : Triangle vocalique du hmong**

(i, u et a peuvent être nasalisées et il y a 4 diphtongues : /ai/, /au/, /aɪ/, /ua/)

## Consonnes

Tableau 27 : phonétique des consonnes hmong blanc

	<i>Bilabiale</i>	<i>Labio dentale</i>	<i>Labio velaire</i>	<i>dentale</i>	<i>alvéolaire</i>	<i>Post alvéolaire</i>	<i>Retroflexe</i>	<i>Palatale</i>	<i>Velaire</i>	<i>Uvulaire</i>	<i>Glottale</i>
Occlusive	p			t				c	k	q	
Click											
Trille											
Tap											
Flap											
Fricative		f v		s		ʃ ʒ	ʂ ʐʔ				h
Affriquée				tʃ		tʃ	tʃ				
Nasale	m			n							
Laterale fricative											
Laterale approximante					l̥						
Approximante			w								

## Systeme sifflé

### Introduction

Nous décrivons ici le système sifflé par deux personnes parlant le dialecte hmong Daw. Notre description est basée sur 12 phrases de deux dialogues différents enregistrés par les soins des équipes de Busnel dans les années 80. Un certain nombre d'autres enregistrements ont pu confirmer la répartition des tons que nous avons extraits, mais ils n'ont pas participé à la description car ils n'avaient pas été traduits mot à mot. Le linguiste hmong Chô Ly a réalisé la transcription phonétique des phrases. La Figure 75 illustre le fonctionnement de ce système.



Figure 75 : Phrase hmong parlée puis sifflée « koj puas paub tshuab nplooj »

Les sifflements reproduisent les tons de la langue avec toute leur complexité. Au niveau fréquentiel, toutes les variations de la Hauteur Fondamentale issues de la voix sont transposées.

### Tons et contours de tons

Les cinq niveaux de hauteur ainsi que toutes les formes de contour de ton ont pu être retrouvés à partir des extraits sonores de feuille sifflée. Sur la figure la longue phrase « *Awm koj ab tswj ua ntej moog lov seb pos nthau los tsw tau nthau es peb mam le koj ntev lawv qab tuaj nawb* » permet de montrer la plupart des configurations rencontrée dans cette langue.

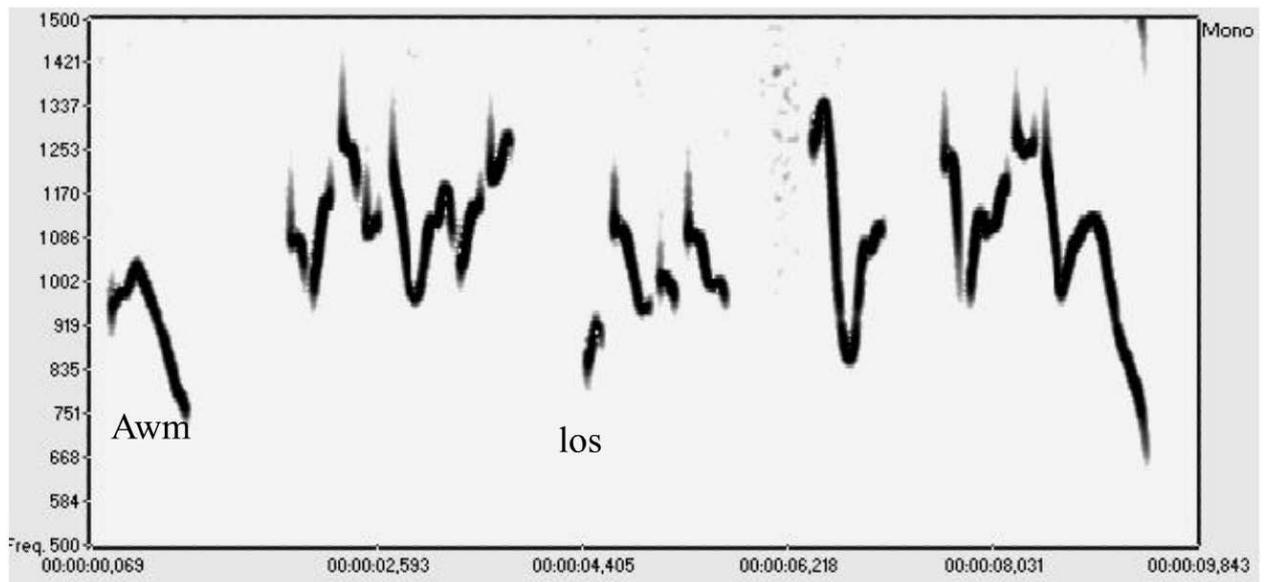


Figure 76 : Phrase hmong présentant les 5 niveaux de ton et de nombreux types de contours

Le Tableau 25 permet de lire cette phrase sifflée. La première et la dernière syllabe se terminent à un niveau très bas. Le niveau 1 correspondant en réalité à celui de la petite modulation montante située au milieu de la phrase : « los ».

Remarques :

-On remarque sur cette phrase que quand c'est possible, les modulations de deux syllabes consécutives sont liées. Ceci arrive quand 2 conditions sont réunies :

-la seconde syllabe débute par une des consonnes /l/, /m/, /n/ (labiales et nasales),

-la seconde syllabe débute dans le même groupe de respiration que la première

Ainsi la transition de milieu de phrase « nthau los » marque un silence car le siffleur marque une pause pour respirer.

-le /t/ semble bien marqué par une modulation vers le haut comme dans les langues sans ton, ce qui suggère de faire une analyse phonétique de chaque consonne.

## **Conclusion pour le hmong**

Le cas de la langue hmong nous a permis d'observer que même dans les langues qui ont la plus grande complexité tonale le sifflement produit une reproduction fidèle de la Hauteur fondamentale tout en prenant parfois en compte des qualités des consonnes. En effet, les consonnes sont sifflées soit par des silences soit par des modulations pleines quand le contexte le permet. Dès lors à la fois les rythmes intra et extra syllabique sont transposés de manière fidèle, ce qui garantit une bonne intelligibilité du hmong sifflé avec la feuille.

### **3.3.5.3. Conclusion pour les langues à tons**

Les observations que nous avons faites sur les systèmes sifflés des deux langues de ce groupe montrent des différences importantes suivant la langue concernée. Nous n'avons pas observé de réduction de tons comme pour les voyelles des langues sans tons, même pour le hmong qui a 5 niveaux de fréquences de tons, un chiffre qui entraîne des confusions dans les voyelles. Nous pensons que cela est possible en vertu de la moins grande variabilité des tons (origine dans la dynamique de la phonation) et peut être aussi grâce au mode de vibration de la feuille sifflée. De plus, contrairement aux modulations de consonnes des langues sans tons, la forme de la modulation des contours de tons à l'intérieur d'une syllabe ne varie pas en fonction du contexte. Seules les modulations liant deux syllabes peuvent varier. Ce dernier cas se présente dans chacune des langues, montrant que le sifflement prend toujours en compte quelques aspects segmentaux dans les langues tonales. De plus les formes fixes des contours peuvent être parfois réalisées identiquement à plusieurs niveaux dans une même phrase, même si la plupart du temps les niveaux de réalisation sont contraints par la phonologie de la langue. Dans tous les cas le sifflement se fait à une hauteur relative adaptée à la distance de communication et au milieu ambiant.

### 3.4. Conclusion de la typologie comparative

La description-comparaison<sup>76</sup> typologique développée dans cette partie nous a permis de faire un bilan sur l'ensemble des langues sifflées documentées à ce jour. La comparaison systématique avec la phonétique du système parlé fait émerger trois types de comportement du sifflement de la parole en fonction des options phonético-phonologique de la structure de chaque langue.

En décrivant en détails sept langues représentatives de la diversité des stratégies de sifflement linguistique et en apportant des analyses complémentaires en Annexe C nous avons pu montrer qu'une première catégorie de langues s'appuie essentiellement sur l'articulation des voyelles et des consonnes; une autre catégorie siffle l'intonation à travers le poids de la syllabe tout en articulant les consonnes; la dernière catégorie, formée uniquement de langues tonales complexes retransmet en priorité les paramètres liés aux tons. En outre, nous avons observé dans chaque groupe de langue une grande diversité.

Par conséquent, les langues sifflées effectuent naturellement une partie du travail de description typologique. Parallèlement, le sifflement développe une véritable description phonétique des caractéristiques vocales des éléments transposés: il définit des classes de consonnes, indique des similarités de timbre des voyelles et précise le comportement des tons. Tout cela dans le respect des particularités de chaque langue.

D'autre part, la hiérarchie entre les éléments de la voix établie par le sifflement n'exclue pas que le signal sifflé relate des variations dues à des dimensions « secondaires ». Ainsi l'accentuation est relativement bien marquée en grec, et joue aussi un rôle en silbo ou en turc. Le hmong quant à lui tient compte de certaines qualités des consonnes. Et l'on peut dire que la recherche d'un équilibre entre les contributions respectives de la « Hauteur Brute » et de la « Hauteur Fondamentale » est la caractéristique distinctive du second groupe de langues.

Un retour à la source d'émission (la bouche) nous a permis de préciser l'origine de certains comportements sifflés comme les regroupements en classes de phonèmes et les modulations. La fréquence de sifflement des voyelles et des tons est ajustée en leur associant une fréquence qui correspond intuitivement à une zone de résonance dans la bouche. C'est pourquoi les langues sifflées confirment une des observations expérimentales de Békésy: « *In judgment of the location of a narrow band, a most disturbing feature is the discovery that, during speech and singing, we are accustomed to localize the low-pitched sounds back in the throat and the high pitched sounds toward the front of the mouth* » (Békésy 1963, p. 599).

L'ensemble des données réunies ici montre que le simple fait de regrouper et d'organiser le comportement des langues sifflées remplit l'objectif que toute typologie se fixe. En effet, d'après Lass il s'agit « *de réduire l'ordre déconcertant d'éléments de l'univers du discours à un nombre de classes analysables* »<sup>77</sup> (Lass 1984, p122). Cette étape de la réflexion nous indique les liens qui existent entre phonétique et

---

<sup>76</sup> Description typologique des langues sifflées et comparaison systématique avec la phonétique des systèmes parlés.

<sup>77</sup> « *to reduce the bewildering array of items in the universe of discourse to a tractable number of classes* ».

phonologie. Elle souligne donc le rôle crucial joué par l'intelligibilité. Elle nous invite donc à explorer plus précisément l'influence des processus perceptifs convoqués par les siffleurs.

## **CHAPITRE 4. INTELLIGIBILITE**

## 4.1. Introduction

Un des paramètres fonctionnels principaux des langues sifflées est leur intelligibilité. Elles répondent au besoin d'assurer vite et bien une communication à distance dans des conditions où le bruit de fond de la nature n'est pas négligeable. Tous les témoins des langues sifflées ont été très étonnés de l'efficacité des systèmes qu'ils observaient. Ainsi Cowan, dans son premier papier de linguistique sur ce thème, relate: « *Un jour, Eusebio Martinez, debout devant sa maison, siffla en direction d'un homme qui était à une distance considérable. Il passait sur le chemin en contrebas pour aller y vendre des paquets de feuilles de maïs qu'il portait. L'homme répondit à Eusebio par un sifflement. L'échange fut répété plusieurs fois avec différents sifflements. Finalement l'homme fit demi-tour, revint sur ses pas et monta sur le sentier qui venait à la maison d'Eusebio. Sans dire un mot, il laissa tomber sa charge au sol. Eusebio vérifia la charge, rentra dans sa maison, ressortit avec de l'argent et le paya. Celui ci fit demi-tour et s'en alla. Pas un mot n'avait été dit. Ils avaient parlé, marchandé le prix et étaient arrivés à un accord satisfaisant pour les deux en utilisant seulement le sifflement comme mode de communication*<sup>78</sup> » (Cowan 1948, p.280, traduction libre). Le succès d'une telle communication tient à un grand nombre de paramètres: la motivation de l'émetteur et du récepteur, l'efficacité de l'encodage linguistique, la propagation du signal dans l'air, la manière dont le signal sera noyé dans le bruit de fond à son arrivée au niveau de l'oreille du récepteur, les capacités perceptives de l'auditeur et bien sûr, de manière cruciale, la connaissance du vocabulaire et du système de règles de la langue employée.

Si les langues sifflées surprennent ce sont indirectement les capacités du cerveau humain et de l'oreille qui étonnent. La principale originalité tient à l'usage d'un signal sifflé qui exploite la redondance de la voix parlée pour sélectionner un *squelette informatif* en fonction de la structure de la langue. Celui ci est opté pour répondre aux contraintes liées aux conditions d'usage qui entraînent une réduction du rapport signal sur bruit et une augmentation de la réverbération. Les mesures que nous avons faites auparavant montrent que la parole sifflée relève d'une stratégie qui est la continuité directe de celles de la voix parlée puis criée.

Comme le sifflement s'appuie sur des éléments phonétiques et phonologiques de la langue locale, les processus cognitifs qui permettent à un siffleur de comprendre le message sifflé sont donc, à notre avis, non seulement similaires à ceux qui interviennent lors de l'écoute d'une langue mais ils impliquent fortement ceux en jeu lors de l'écoute dans des conditions difficiles liées au milieu ambiant.

---

<sup>78</sup> Discussion rapportée dans une publication faite en espagnol par Cowan: A: Qu'as tu donc acheté? B: C'est une charge de maïs? A: Et où l'amène tu? B: Je l'amène à Tenango. A: Tu vas la vendre? B: Je vais la vendre. A: Combien en veux tu? Vends la moi ici. B: Ce sera deux pesos cinquante centavos pour la charge? A: Tu ne veux pas deux pesos vingt cinq centavos? C'est ce que je peux te donner. B: La où je vais la vendre, ils m'en donneront trois pesos. A: Oui mais c'est loin, que décides tu ? B: Je vais laisser ma cargaison ici. A: A la bonne heure, tu demandais beaucoup. (Cowan 1976 in Sebeok et Umiker-Sebeok p.1396, traduction libre).

La parole humaine étudiée dans des conditions idéales d'écoute ne permet pas toujours de rendre compte de ces derniers aspects. Pourtant la voix parlée est quotidiennement utilisée dans le bruit et parfois même à des distances non négligeables pour le signal phonétique qui parvient à l'auditeur. De nombreuses études sur la parole dans le bruit ont montré que le rapport signal sur bruit et la réverbération sont les principaux facteurs qui affectent les performances de reconstitution cognitive de la parole. Compte tenu de ces remarques, l'analyse des langues sifflées est susceptible de fournir un matériau d'étude très riche du fonctionnement cognitif langagier.

La performance de reconnaissance des mots par un auditeur est parfois appelée l'«*intelligibilité de la parole*». Le but de toute communication linguistique humaine est d'atteindre une intelligibilité maximale avec un effort adapté au contexte, si possible optimisé. Etudier l'intelligibilité au sens général implique l'analyse de l'influence des différents paramètres affectant le taux de reconnaissance des mots d'une phrase. Une telle approche concerne à la fois un locuteur, principalement à travers le signal qu'il émet et sa localisation dans l'espace, un milieu de transmission acoustique appelé canal (pour les langues sifflées, le canal est le plein air dans 99% des cas<sup>79</sup>) et un récepteur qui fait un prétraitement acoustique au niveau de l'oreille puis une reconnaissance et une interprétation au niveau du système central (Figure 77).

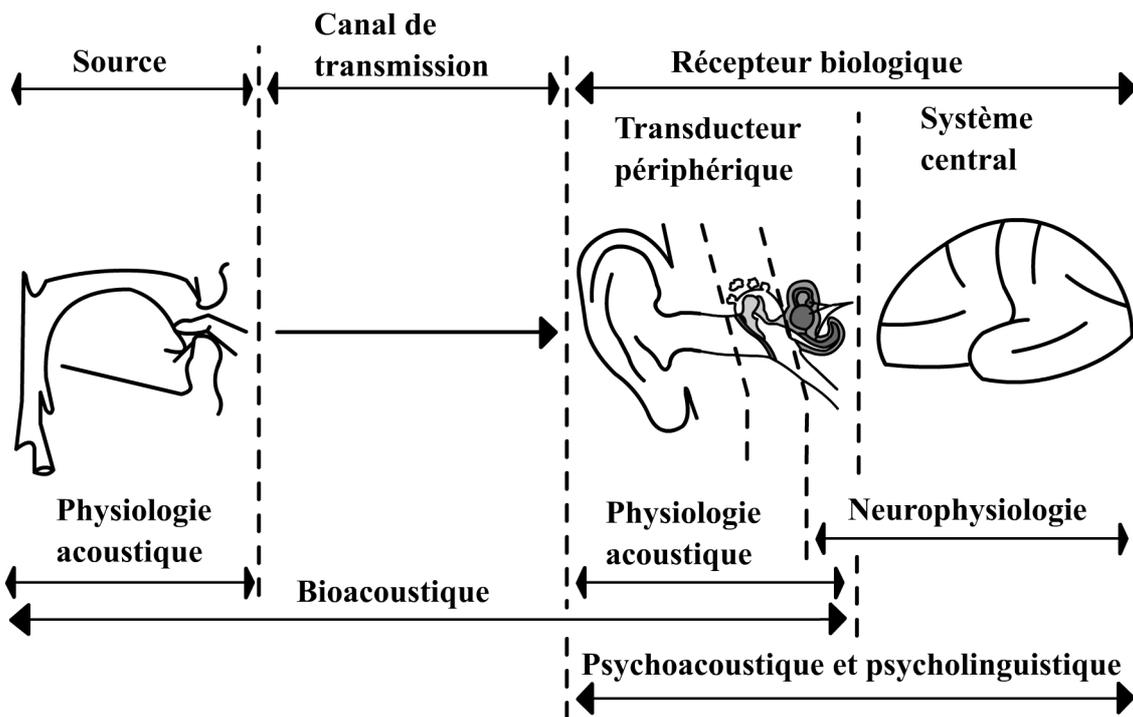


Figure 77 : Schéma des étapes de la transmission et de la reconnaissance d'un message linguistique

<sup>79</sup> Un de nos informateurs Gomero appelle souvent au téléphone son frère résidant sur l'île de Tenerife en utilisant le sifflement par boutade.

Les scientifiques travaillant sur les systèmes de télécommunication ont été les premiers à attirer l'attention sur ce problème de l'intelligibilité (Campbell, 1910, Fletcher & Steinberg 1929, Chavasse 1962). Leurs travaux sur les caractéristiques du message parlé et sur l'adaptation des moyens de transmission aux possibilités de la perception périphérique ont mis en valeur la nécessité de mesurer l'identification correcte des sons de la parole. Ils ont créé la théorie metrologique dont l'usage, entre autres, pour la mesure de la surdité est d'un considérable intérêt. Ils ont perçu la notion différentielle qui doit séparer netteté, d'intelligibilité. L'« *intelligibilité* » est la compréhension des idées, qui se mesure avec des phrases et des mots significatifs d'une langue donnée. Elle diffère de la « *netteté* » qui se réfère, elle, à des associations conventionnelles de sons ou de mots dépourvus de sens (« *logatomes* » et « *non-mots* »).

Mais, pour les ingénieurs, l'intelligibilité est exclusivement relative au courant d'information et à ses variations au travers des systèmes de transformation et de transmission (par exemple en téléphonie: par le microphone, la ligne, l'écouteur) dont la mesure doit être considérée comme constante, bien qu'elle se fasse au travers d'un récepteur humain qui ne devrait théoriquement que « mesurer » ce « courant d'informations » sans pouvoir l'influencer. Pourtant, les propriétés psychophysiologiques du récepteur sont à la base de toute mesure d'intelligibilité. Comme il n'est pas pensable de s'affranchir complètement des variations dues aux jugements subjectifs des individus et aux conditions variables de l'expérience, des protocoles particuliers ont été mis au point afin de normaliser les résultats obtenus dans les études technologiques<sup>80</sup> (pour un bilan sur le sujet voir Cartier 1989).

Dans notre cas, notre but n'est pas d'améliorer l'efficacité d'un canal de transmission (puisque cela a déjà été fait par l'évolution naturelle) mais de comprendre l'influence des différents paramètres intervenant dans la compréhension de l'information linguistique encodée dans les phrases telles qu'elles parviennent à l'oreille du siffleur. La notion d'intelligibilité au travers du récepteur humain et de son système cognitif telle que nous l'envisageons est donc, par essence, différente de celle qui se rapporte au courant d'information mesuré suivant des normes technologiques telle que définie par les ingénieurs de téléphonie.

Dans cette partie, nous verrons dans un premier temps que notre connaissance actuelle du système périphérique auditif définit un cadre dans lequel le sifflement se positionne avantageusement. L'audition testée en psychoacoustique par des sons simples éclaire également certains des choix faits par les siffleurs. A partir du cadre défini par ces résultats expérimentaux préliminaires, nous pourrions aborder les recherches qui permettent de comprendre comment l'être humain organise les sons complexes qu'il reçoit. Par exemple,

---

<sup>80</sup> Les protocoles d'évaluation de l'intelligibilité existant aujourd'hui se distinguent en méthodes directes et indirectes.

Les méthodes indirectes font appel au jugement des sujets. Les méthodes directes ont souvent été mises en place grâce à des méthodes indirectes, ce sont des méthodes de calcul des taux de reconnaissance à partir de normes (comme les normes RASTI), d'Indices d'Articulation des mots (ou I.A.) ou de modèles mathématiques. Les normes RASTI qui sont utilisées dans l'industrie restent mal adaptées à certains milieux (Tisseyre, 1998). Les indices d'articulation se basent sur l'acoustique de la voix, ils ne sont donc pas adaptés pour les langues sifflées. Les modèles de probabilité sont nombreux et les plus récents sont bien adaptés à l'écoute dans le bruit (Bronkhorst et al, 1993).

l' « *analyse de la scène auditive* » (Bregman, 1990) explique comment l'oreille humaine réalise une véritable enquête sur l'origine des sons (Risset, 1994). Un certain nombre de regroupements des attributs de la perception en résultent, ils expliquent de nombreux aspects de la perception des sons complexes comme ceux de la musique ou du langage. Ils suggèrent également d'analyser les processus de reconnaissance de la parole à plusieurs échelles. De telles démarches, dont nous signalerons quelques résultats, ont été initiées pour l'étude des langues. *L'analyse de la prosodie* est un des domaines qui a cherché à répondre à cette exigence. Nous verrons pourquoi l'on peut considérer que les langues sifflées redéfinissent la prosodie en proposant une approche tenant compte de la perception.

Après ces considérations générales, nous ferons un bilan progressif de la dynamique de l'intelligibilité des langues sifflées en partant de la perception des voyelles jusqu'à l'intelligibilité des phrases. Certaines étapes de cet exposé apporteront des éléments expérimentaux nouveaux, en particulier sur les voyelles sifflées et sur la dégradation de la structure phonétique avec la distance. Le cadre de notre étude rappellera en quoi elle peut être instructive pour l'ensemble des langues et ne doit pas être confinée à la curiosité exotique.

## **4.2. Adaptation étonnante du sifflement à l'audition humaine**

Les bioacousticiens ont observé sur de nombreuses espèces l'adaptation des systèmes de communications aux possibilités de perception acoustique. Comme chaque espèce animale, nous sommes immergés dans un monde sonore dont l'oreille ne capte qu'une partie de façon très inégale selon les fréquences, les niveaux d'amplitude.

### **4.2.1. Système périphérique**

Le récepteur humain est formé d'un ensemble: celui ci comprend un premier étage de transducteurs déjà très complexe mais à partir desquels les chercheurs ont pu faire des mesures relativement reproductibles. Le système auditif périphérique effectue le codage du signal acoustique en influx nerveux. C'est un transducteur non linéaire de sons qui comporte trois niveaux physiologiques distincts: l'oreille externe (transmission dans l'air jusqu'au tympan), l'oreille moyenne (transmission mécanique grâce aux osselets et à un différentiel de surface), l'oreille interne (transmission en milieu aqueux dans la cochlée et activation des cellules ciliées). Sur la Figure 77, nous avons appelé ce stade le canal de la physiologie acoustique. Il est relié à l'étage central par le nerf auditif. Le traitement réalisé par cette étape de la perception fait subir des transformations au signal sonore en fonction de ses propriétés. Celles-ci le préparent à l'analyse réalisée par le système central. Cela influence l'intelligibilité car la résolution du signal au niveau de l'amplitude, de la fréquence et du temps est modifiée en même temps que le rapport signal sur bruit.

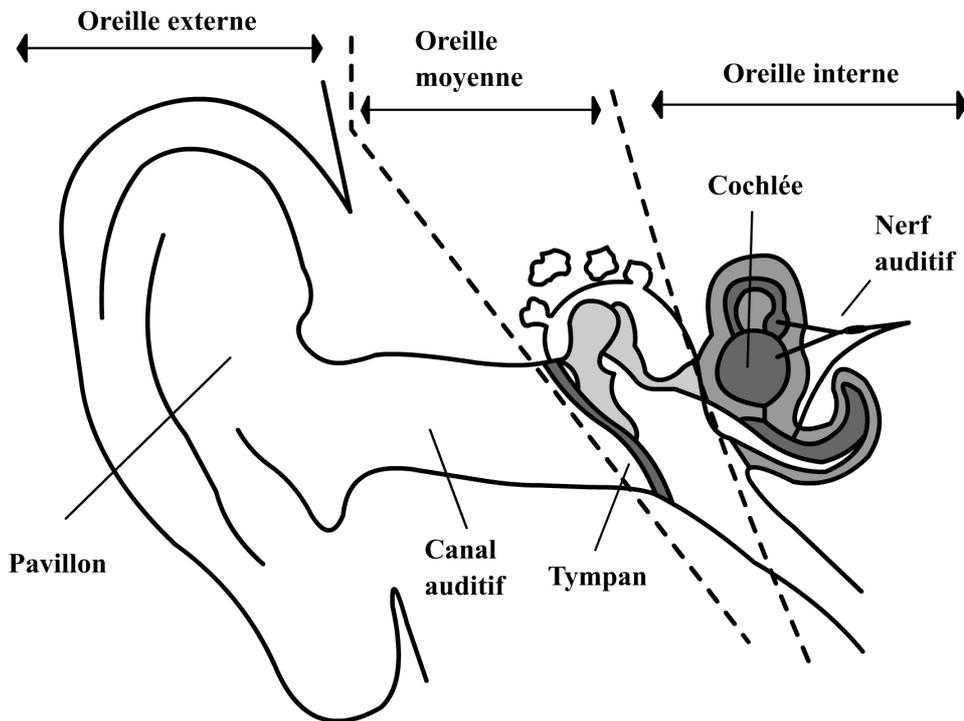


Figure 78 : Schéma du système auditif périphérique

#### 4.2.1.1. Le prétraitement réalisé par l'oreille externe: La physiologie auditive

A ce niveau de la perception auditive, la réception est facilitée par la conformation du pavillon et du canal auditif. Cet ensemble joue un rôle important dans l'audition. Des mesures réalisées en plaçant un microphone au niveau du tympan ont permis de conclure qu'au niveau acoustique ces structures physiologiques forment un filtre dont la réponse en intensité en fonction de la fréquence n'est pas linéaire (Batteau 1967, Busnel 1976 in Busnel et Classe 1976).

Figure 79 : Réponse du filtre de l'oreille externe (Busnel et Classe 1976, p. 42)

Outre les propriétés utiles pour la localisation auditive et la représentation spatiale des stimuli sonores du pavillon (Batteau 1967, Canevet 1989), l'effet du pavillon améliore le rapport signal sur bruit. Pour l'ensemble pavillon-canal, certaines bandes de fréquences sont favorisées. Ainsi, comme on peut l'observer sur Figure 79, le domaine de fréquence 1000-5000Hz est le plus réactif. Le canal auditif résonne à une fréquence d'environ 3800 Hz et crée une amplification de niveau des sons de 15 décibels.

A l'intérieur du domaine fréquentiel ainsi favorisé, l'augmentation maximale d'intensité est de 20 à 25 dB pour la bande de fréquences 1800-3500 Hz. C'est précisément le domaine de sensibilité maximum de l'audition humaine prise dans son ensemble et mesurée couramment avec la technique des audiogrammes. Ainsi, l'oreille externe prépare le signal acoustique à son traitement par les autres niveaux de l'oreille (intermédiaire et interne) et par le système central.

#### 4.2.1.1.1. Conséquences pour le langage sifflé

Le domaine de fréquences favorisé par l'oreille externe correspond à celui exploité pour encoder le langage dans les langues sifflées. L'application du filtre du pavillon et du canal auditif dans une écoute de face<sup>81</sup> à un signal sifflé en turc, enregistré à 150m, montre que la fréquence fondamentale est renforcée (rapport signal sur bruit amélioré de 7% en moyenne<sup>82</sup>). Par contre cela ne change rien pour les harmoniques qui n'émergent pas mieux du bruit ambiant (Figure 80).

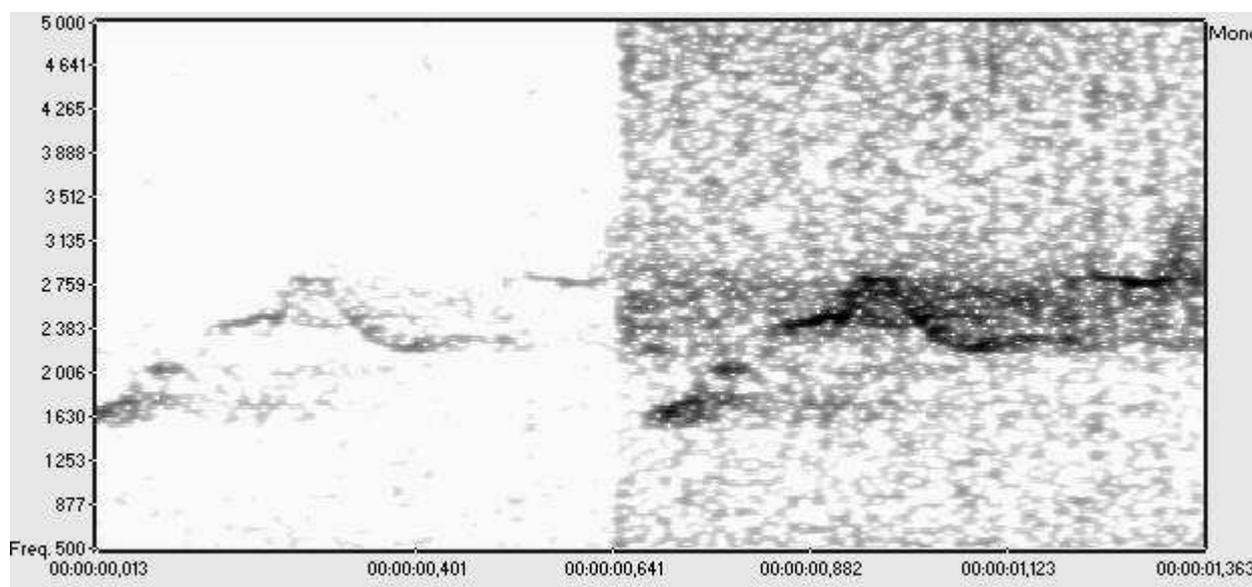


Figure 80 : Effet du filtre de l'oreille interne modélisé à partir de l'application d'un filtre artificiel à partir des valeurs de la Figure 79, position face à l'émission.

#### 4.2.1.2. Aspects importants des autres niveaux physiologiques de l'oreille

##### 4.2.1.2.1. L'oreille moyenne

La fonction de l'oreille moyenne est de permettre de dépasser l'effet de réflexion d'une grande partie de l'onde acoustique qui a habituellement lieu lors du transfert d'énergie d'un milieu à impédance faible (air) à un milieu à impédance moyenne (fluide). Un tel transfert entraînerait une perte de 30 dB en amplitude (Glatke, 1973) mais ce problème est résolu par l'oreille moyenne qui "...augmente la pression d'environ 30 dB" (Borden & Harris, 1980, p.165). Un tel résultat est obtenu par l'action combinée des osselets de l'oreille et du

<sup>81</sup> Condition d'écoute normale d'un siffleur.

<sup>82</sup> Calcul effectué sur les pics d'énergie les plus intenses de chaque voyelle.

rapport d'environ 17:1 entre la surface du tympan et celle de la fenêtre ovale (Bekesy, 1960). La fonction de transfert du gain en pression de l'oreille moyenne n'est pas uniforme mais présente un pic à 1000 Hz, relativement plat (moins de -10 dB par rapport au maximum dans le domaine des fréquences de 300 Hz à 7000 Hz) et qui décroît à plus de 20 dB d'atténuation en deçà de 100 Hz et au delà de 10000 Hz (Nedzelnitsky, 1980). Cette étape de la transduction auditive est adaptée aux domaines de fréquences à la fois de la voix parlée et de la parole sifflée.

#### 4.2.1.2.2. L'oreille interne

##### Description générale

Le point le plus important à comprendre est que l'oreille interne réalise une analyse fréquentielle du signal acoustique grâce aux bancs de filtres créés par l'ensemble des filtres nerveuses. Quelles que soient les interprétations des multiples analyses scientifiques réalisées ces dernières années concernant les possibilités de codage du domaine fréquentiel perçu -soit codage tonotopique de 200 Hz à 20000 Hz, soit codage temporel 20 Hz à 4000 Hz soit les deux - le domaine de fréquences couvert par la voix humaine et par les sifflements permettent aux deux principes d'être appliqués. Dans l'hypothèse où ils seraient tous les deux impliqués dans l'analyse fréquentielle ceci signifierait que la bande de fréquence de 200 Hz à 4 kHz serait une zone privilégiée, car elle permettrait un traitement fréquentiel suivant deux modalités ayant des propriétés différentes<sup>83</sup>.

##### Les fibres nerveuses de l'oreille interne

Les fibres nerveuses de la cochlée répondent sélectivement au son dont la fréquence est proche de leur fréquence de décharge (ou fréquence caractéristique). Elles se comportent chacune comme un filtre passe bande. Un point important tient au fait qu'en plus de coder la fréquence d'une composante du signal, elles préservent dans une certaine mesure leur information temporelle en produisant une décharge à un instant précis de l'onde qui les stimule. Ce phénomène est appelé *phase locking*.

La structure même des fibres est complexe, elle pourrait expliquer certaines des non-linéarités de la perception auditive. En effet, différents modes d'activation ont été observés:

- (i) certaines fibres ne déclenchent leur décharge que pour des variations spécifiques du signal (répétition, ou temps de montée de l'enveloppe) (Koch and Piper, 1979).
- (ii) Beaucoup de cellules nerveuses du noyau cochléaire répondent peu -ou pas du tout- à un bruit filtré dont la largeur de bande est supérieure à une valeur donnée. Cependant, lorsque la fréquence centrale d'une telle bande de fréquence varie rapidement, ces cellules nerveuses répondent vigoureusement et ont une sélectivité fréquentielle prononcée (Miller, 1979).

---

<sup>83</sup> Dans ce cas, la bande de fréquence couverte par le téléphone (400 Hz à 4000 Hz) et celle couverte par la parole sifflée sont particulièrement favorisées.

- (iii) D'autre part, LePage (1987) a fourni la preuve que la réponse fréquentielle cochléaire ainsi que sa précision dépendaient de l'intensité du stimulus.
- (iv) De plus, des interactions entre groupes de fibres introduisent des composantes fréquentielles qui sont absentes du signal acoustique ou suppriment partiellement des composantes comprises dans le signal.

## **Conclusion**

Les découvertes réalisées sur l'oreille interne démontrent que le système auditif périphérique n'est pas un transducteur passif, comme l'est un microphone. Nous retiendrons également qu'il existe différents modes d'activations des fibres neuronales suivant les propriétés dynamiques du signal et que des processus de correction ou d'augmentation de la précision de la réponse en fréquence en fonction de l'intensité du signal sont en jeu. A ce jour, aucune conclusion définitive ne peut être tirée à propos du mode de traitement des fréquences par la cochlée ou de son contrôle par des mécanismes centraux. Enfin, il apparaît clairement que les sifflements et la voix sont dans les zones les plus sensibles de cette partie dynamique de l'oreille.

### **4.2.1.2.3. Conclusion pour l'oreille périphérique**

A chaque étape de filtrage réalisé par le « canal physiologique » le signal de parole est transformé suivant de nombreux processus pour la plupart non linéaires qui améliorent le rapport signal sur bruit général en privilégiant certains domaines de fréquences. Une dynamique particulière des signaux est également favorisée. Le domaine de fréquences mis en valeur par l'effet cumulatif des trois niveaux de l'oreille périphérique (externe, moyenne et interne) va de 1000 à 4000 Hz. Cette bande de fréquence est efficace car elle se situe largement au delà de la majorité des bruits de fond de notre environnement. D'autre part, jusqu'à 4000 Hz l'oreille est en mesure de réaliser une analyse temporelle précise de la dynamique des signaux. Etant donné que cette bande de fréquence permet d'obtenir une intelligibilité de plus de 90% avec un signal de voix (Moore, 1982<sup>84</sup>) et contient l'ensemble des signaux sifflés, nous pouvons raisonnablement penser qu'elle représente une zone clef de la perception de la parole.

## **4.2.2. Psychoacoustique: une première approche du système central**

La psychoacoustique a pour objet l'étude expérimentale des relations quantitatives entre des stimuli acoustiques mesurables physiquement et les réponses de l'ensemble du système auditif de l'être humain. Elle permet donc de confirmer les hypothèses que nous avons émises au vu des résultats de physiologie auditive,

---

<sup>84</sup> Moore obtient 90% d'intelligibilité des phrases avec un signal de parole filtré dans la bande 1000-2000 Hz.

tant sur les limites de décodage<sup>85</sup> des sons simples que de certaines dynamiques. Elle permet également de tester la résolution de la perception de la hauteur, de l'intensité et de la durée.

#### **4.2.2.1. Le domaine d'audibilité: résolution de la perception auditive humaine en fonction de la fréquence et de l'intensité**

##### 4.2.2.1.1. Analyse synthétique

La courbe de "réponse" de la sensation subjective auditive d'égales intensité (lignes isosoniques: Fletcher & Munson 1933) en fonction de la fréquence donne une idée générale de la perception auditive humaine (Figure 81). Cette courbe est statistiquement normale à des excitations de fréquences pures variables pour un ensemble d'intensités différentes. Contrairement aux travaux sur l'oreille externe, il n'est plus possible d'obtenir des mesures physiques directes mais c'est le jugement de sujets testés suivant un protocole précis qui permet de tirer des conclusions sur la perception acoustique. Les tests qui ont permis d'établir les résultats de la Figure 81 ont été réalisés à partir de sons purs, c'est à dire des sons de type sinusoïdaux ayant une fréquence et une intensité donnée à chaque présentation. Les sensations de hauteur du sujet sont appelées « hauteur tonale » (HT). *« Nous dirons que la HT d'un son pur est l'attribut perceptif du son sur la base duquel il est possible de lui apparier, par ajustement de fréquence, un autre son pur différant par le niveau d'intensité. Et nous dirons aussi qu'après appariement de leur HT, les deux sons ne diffèrent que par la « sonie ». Ce qui signifie que deux sons purs quelconques ne peuvent différer que par la HT et/ou la sonie »* (Demany 1989, p. 42). Ces approches psychoacoustiques ont permis de remarquer que l'audition humaine ne peut détecter un changement que si la variation de l'excitation dépasse une certaine quantité ou *seuil différentiel*.

---

<sup>85</sup> Dans ce travail, nous considérons que le signal est encodé par l'émetteur et décodé par le récepteur. Certains chercheurs utilisent le terme encodage de manière différente puisqu'ils l'utilisent pour le passage de l'acoustique à l'influx nerveux réalisé dans l'oreille.

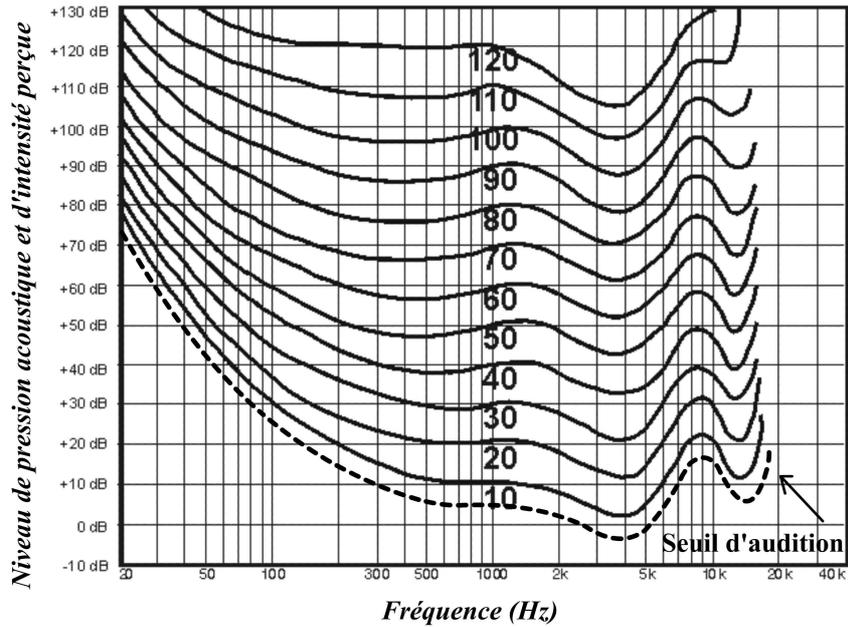


Figure 81 : Lignes isotoniques et niveau d'audition

Cette représentation graphique nous fournit plusieurs renseignements :

- Le domaine d'audibilité humain s'étend entre les deux courbes des seuils d'audibilité et de douleur.
- Une zone de fréquences prédomine s'étendant autour de 3000Hz (de 2000 à 5000 Hz) où l'être humain parvient à distinguer les plus faibles intensités.
- De part et d'autre de cette zone, la sensibilité diminue. Ainsi à 30 Hz, il faut augmenter l'intensité de 60 dB (c'est à dire multiplier l'amplitude du son par 1 million) pour que ce son de basse fréquence soit tout juste audible, comme l'était celui de 3000Hz. La sensibilité décroît également du côté des hautes fréquences et au delà de 20000 Hz, la sensation auditive ne permet plus de percevoir les sons.

### Comparaison avec quelques analyses spectrales de parole sifflée

Les données présentées ici sont issues d'un sifflement émis à 105 dB émergeant à 15 m (distance d'enregistrement) de 40-50 dB d'un bruit de fond de 30-40 dB(Figure 82 et Figure 83).

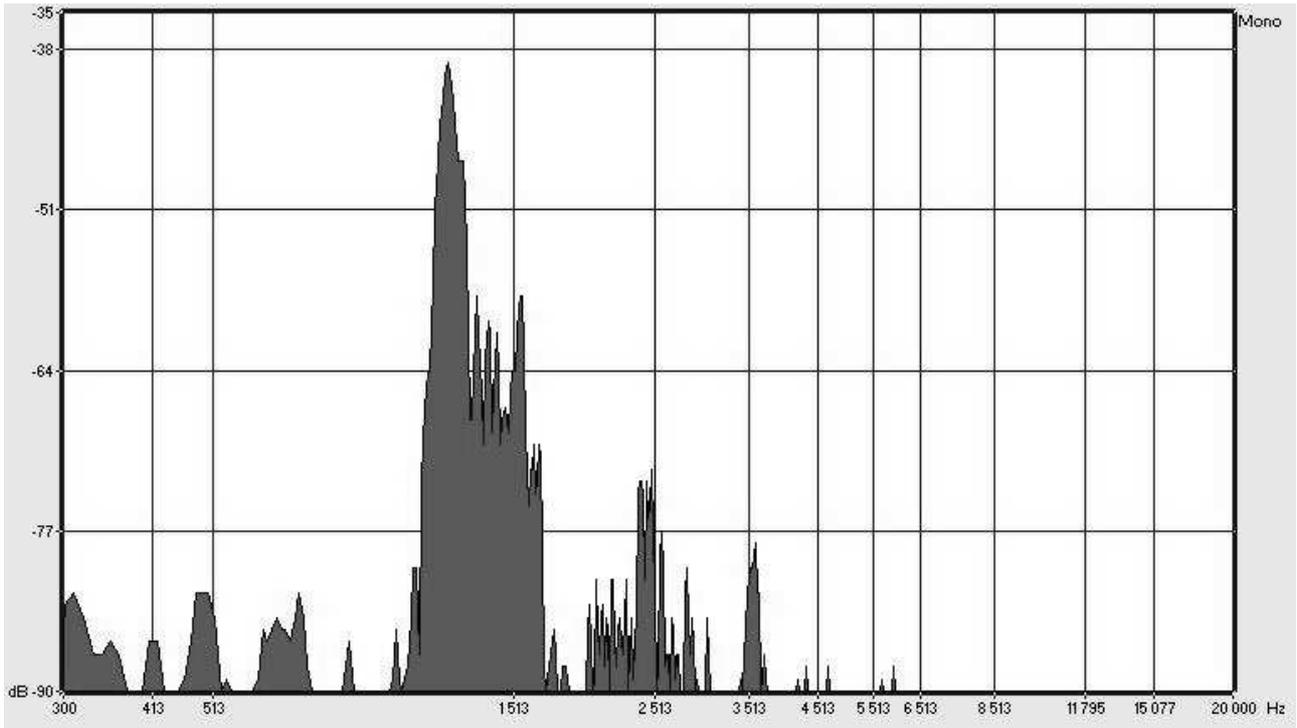


Figure 82 : Analyse fréquentielle de la voyelle «a» de la dernière syllabe du mot espagnol « montaña »

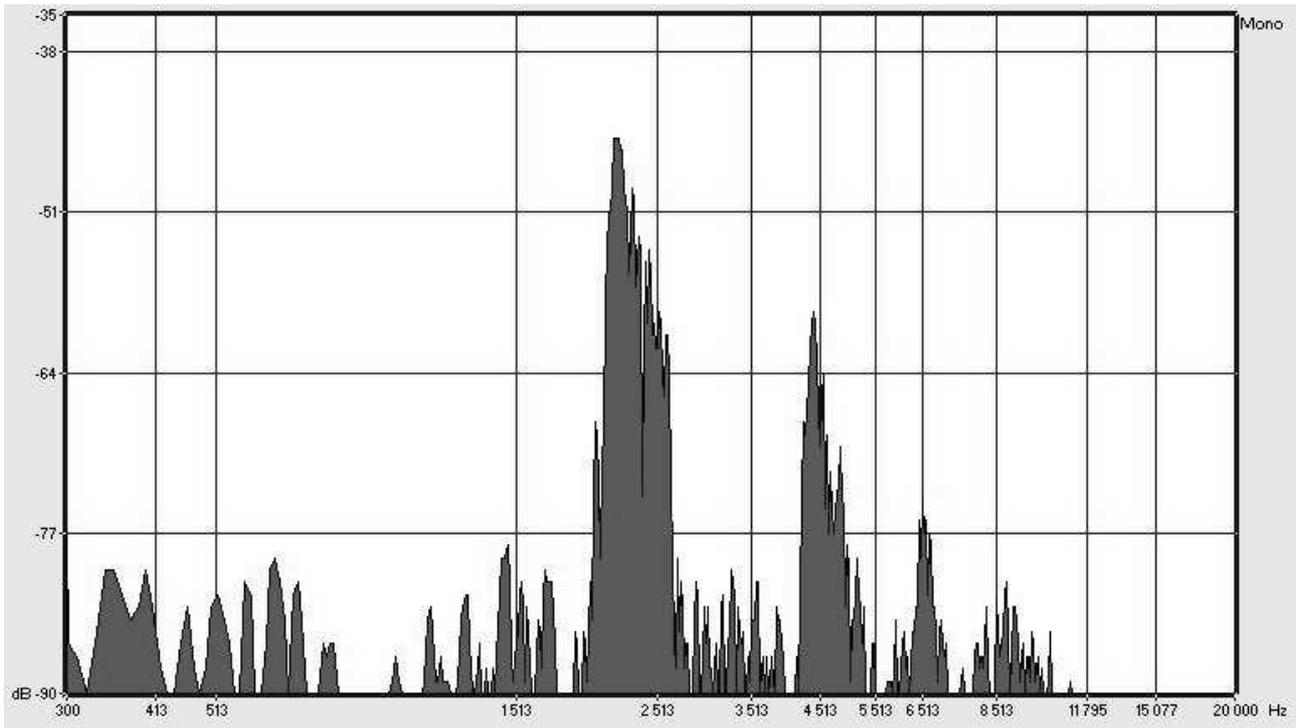


Figure 83 : Analyse spectrale de la fréquence d'un point de la modulation de la consonne «ñ» du mot espagnol « montaña »

*Remarques :*

-Les analyses spectrales de la fréquence présentées sur les Figure 82 et Figure 83 témoignent du fait que le sifflement se situe dans la zone privilégiée par la perception en fréquence (Figure 86).

-La fréquence ayant l'amplitude la plus élevée sur ces deux figures est la fréquence qui ne sera pas dégradée à moyenne et longue distance, c'est donc elle qui porte l'essentiel de l'information alors que ses harmoniques sont vite dégradées.

-A courte distance le sifflement est également dans la zone la plus précise de la perception de l'intensité.

#### 4.2.2.1.2. Discrimination en fréquence ou résolution fréquentielle

Le seuil différentiel en fréquence, plus petit intervalle perceptible entre 2 sons purs, varie beaucoup avec la fréquence. Il présente un minimum entre 1500 et 4000Hz et croît autour de ces valeurs. Le minimum est de l'ordre de 1/300e d'octave à 2000 Hz et le maximum est de l'ordre de 1/6eme d'octave à 30 Hz. Pour une même fréquence, le seuil différentiel diminue quand l'intensité augmente. Ce qui confirme que la hauteur perçue est déterminée en première approximation par sa fréquence. Au-dessus de 1000 Hz la fréquence doit être plus que doublée pour obtenir une sensation de hauteur double. Ces résultats ont été obtenus par une méthode qui consiste à rechercher la plus petite différence de fréquence que doivent présenter deux sons purs stables et successifs pour qu'ils soient juste discriminés. L'intervalle de silence entre les deux sons n'est pas critique, i.e. trop long pour empêcher de les comparer, tant qu'il est compris entre 100 et 1000 ms (Harris 1952).

#### 4.2.2.1.3. Echelles de discrimination ou de sélectivité en fréquence

##### Cadre de l'analyse

L'échelle de HT peut être graduée en *mel* (définition de Carlson et al, 1970). Un écart constant de *mel* correspond à un écart constant de hauteur perçue. Une autre échelle psychoacoustique perceptive des fréquences a été mesurée à partir de la largeur de la bande critique pour laquelle deux fréquences peuvent être discriminées (sélectivité fréquentielle). Cette bande critique variable en fonction de la fréquence est appelée *bark*. Un écart constant en *bark* va correspondre à un écart de perception entre des fréquences simultanées. Ces échelles, construites comme des approximations mathématiques de la perception fréquentielle humaine, à partir de données expérimentales, témoignent du fait que le mode de perception des fréquences par le système cognitif humain se rapproche d'une forme logarithmique<sup>86</sup>. Ces deux échelles ne sont pas acoustiquement équivalentes malgré certaines approximations (ou redéfinitions) faites dans des travaux de recherche de référence (Peterson et Barney 1952, Zwicker & Fastl, 1990). Les études psychoacoustiques de perception de la fréquence réalisées par Zwicker et Fastl rendent équivalentes la sélectivité fréquentielle (*bark*), la discrimination de fréquentielle (jugements de la fréquence : « *frequency-jnds* ») et la perception de la hauteur (*mels*) dans un souci d'uniformisation. Cependant cette simplification ne rend pas compte des différences réelles entre les différents modes d'observation de la perception de la fréquence. Cette conclusion est imposée par la considération de résultats plus récents sur les largeurs de bande des filtres auditifs perceptuels : ceux-ci ont montré que les largeurs de bandes critiques sont en général

---

<sup>86</sup> Nous verrons un peu plus loin que la réalité est plus tridimensionnelle.

surévaluées en particulier à basse fréquence et que la sélectivité fréquentielle du système auditif est mieux représentée par l'échelle ERB, échelle utilisant une méthode de détermination des largeurs de bandes fréquentielles mise au point par Patterson (1976) et qui est aujourd'hui préconisée par Rosner & Pickering (Rosner & Pickering, 1994) pour étudier la perception des voyelles.

## **Masquage et rapport Signal sur Bruit des langues sifflées**

L'échelle ERB est importante pour expliquer certains aspects perceptifs de l'émergence du signal sifflé dans le bruit. En effet d'après les données de Patterson (1976) et de Moore et Glasberg (1983) les filtres auditifs seraient des filtres à pentes exponentielles et convexes au sommet dont la largeur peut être estimée par le calcul d'un Equivalent Rectangular Bandwidth (ERB) :

$$\text{ERB} = 6,23f^2 + 93,39f + 28,52$$

Avec f en kHz et ERB en Hz.

Ainsi pour un sifflement centré sur f=2 kHz, ERB=250 Hz. Donc de 1 à 4kHz on obtient une bande ERB qui varie entre 120 et 500 Hz. Nous avons mesuré qu'à courte distance (15 m) le signal émerge du bruit de fond avec une largeur de bande de 450 Hz au maximum (Figure 82 et Figure 83). Il activera au maximum 4 filtres perceptifs de l'oreille à un instant donné. A plus grande distance, la largeur de la bande de fréquence qui émerge est plus réduite. Cependant nous avons mesuré qu'à 550m un signal qui est noyé dans le bruit de fond reste intelligible. Cela n'est possible que parce que la sélectivité de l'oreille a pour conséquence que, lorsqu'un bruit à large bande masque un son relativement pur, seules les composantes fréquentielles du bruit proches de la fréquence du signal sont effectivement masquantes pour ce dernier<sup>87</sup>. Par conséquent, il est pertinent d'observer l'émergence d'une phrase en langue turque comprenant toute l'information sifflée dans une bande de fréquence de 1500 à 3000 Hz en ne considérant que le bruit dans ce domaine (le résultat du filtrage est présenté sur les Figure 84 et Figure 85). A la limite d'intelligibilité le son émerge encore de 20 dB de cette bande avec une largeur de bande de 150 Hz. On peut donc dire que la bonne sélectivité auditive combinée aux caractéristiques du sifflement permet d'expliquer à la fois la précision de la perception et la bonne émergence du bruit.

---

<sup>87</sup> Le masquage réalisé par un son de fréquence donnée sur des sons de fréquences différentes traduit l'amplitude et l'étendue de l'excitation du système auditif (Zwicker 1970). Zwicker a su adapter l'explication des phénomènes de masquage à des sons dont le niveau varie rapidement au cours du temps comme ceux de la parole. Il a mis en évidence qu'à un niveau donné, lorsqu'un bruit vient interférer avec le signal, le masquage qu'il exerce sera d'autant plus important que sa fréquence est inférieure et son intensité supérieure au signal de parole. Dans une bande de fréquence de 250 Hz (bande d'émergence typique d'un sifflement), très peu de signaux seront concernés.

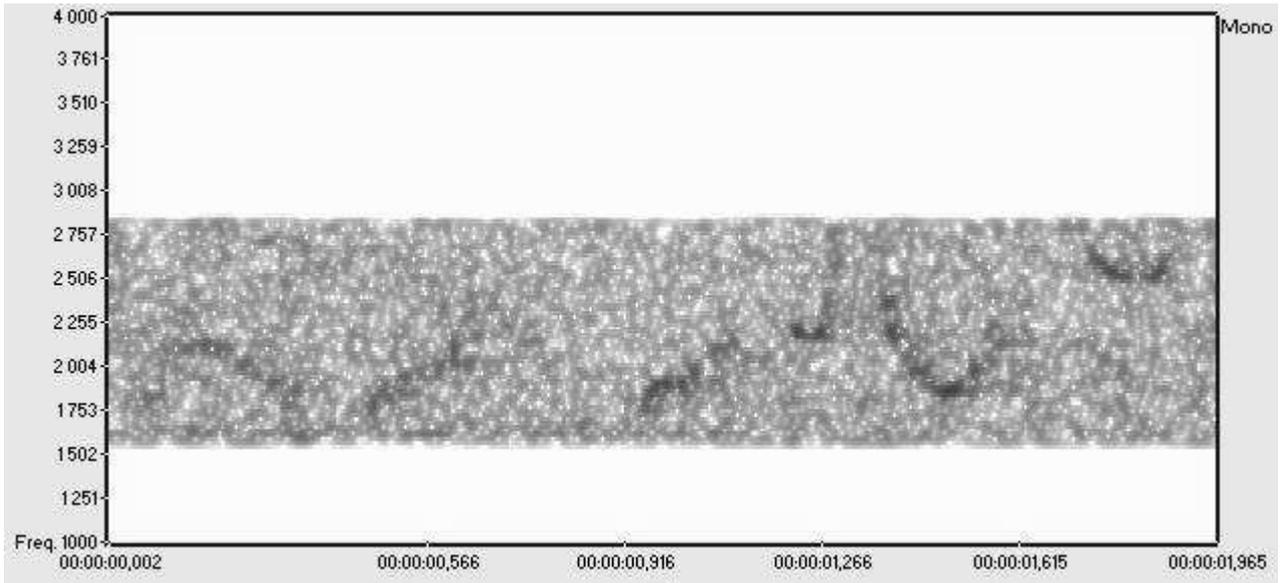


Figure 84 : Sonogramme de la phrase turque « Mehmet okulagit » à 550m élimination des bandes de fréquence en dessous de 1500 Hz et au dessus de 3000 Hz

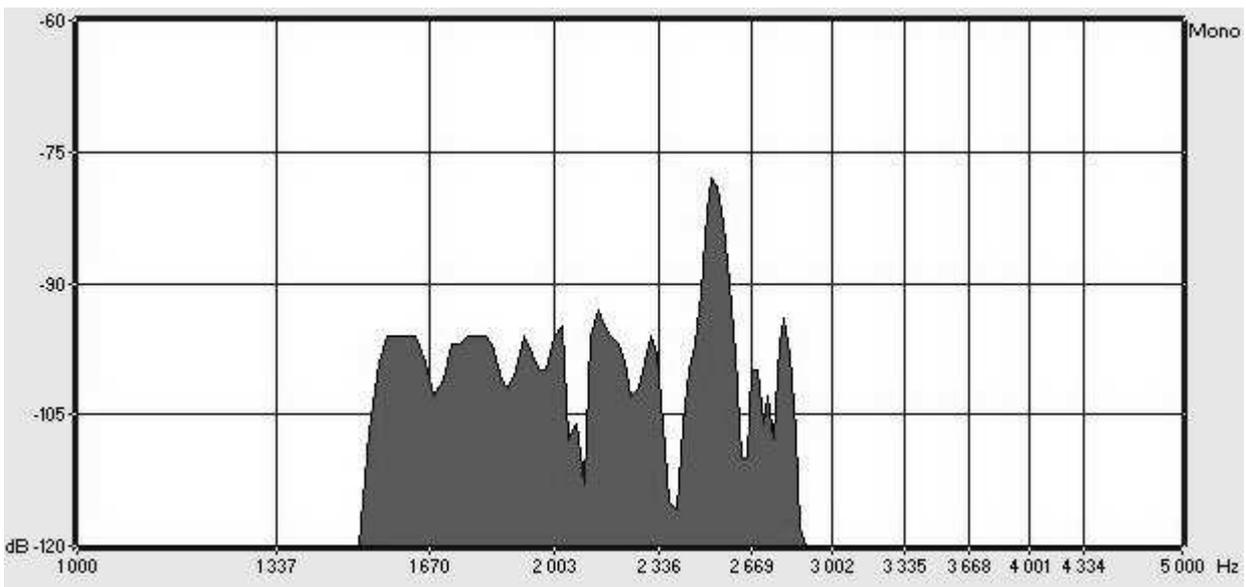


Figure 85 : Emergence en amplitude du sifflement de la phrase turque à 550m après filtrage (1500-3000 Hz)

#### 4.2.2.1.4. Discrimination en intensité

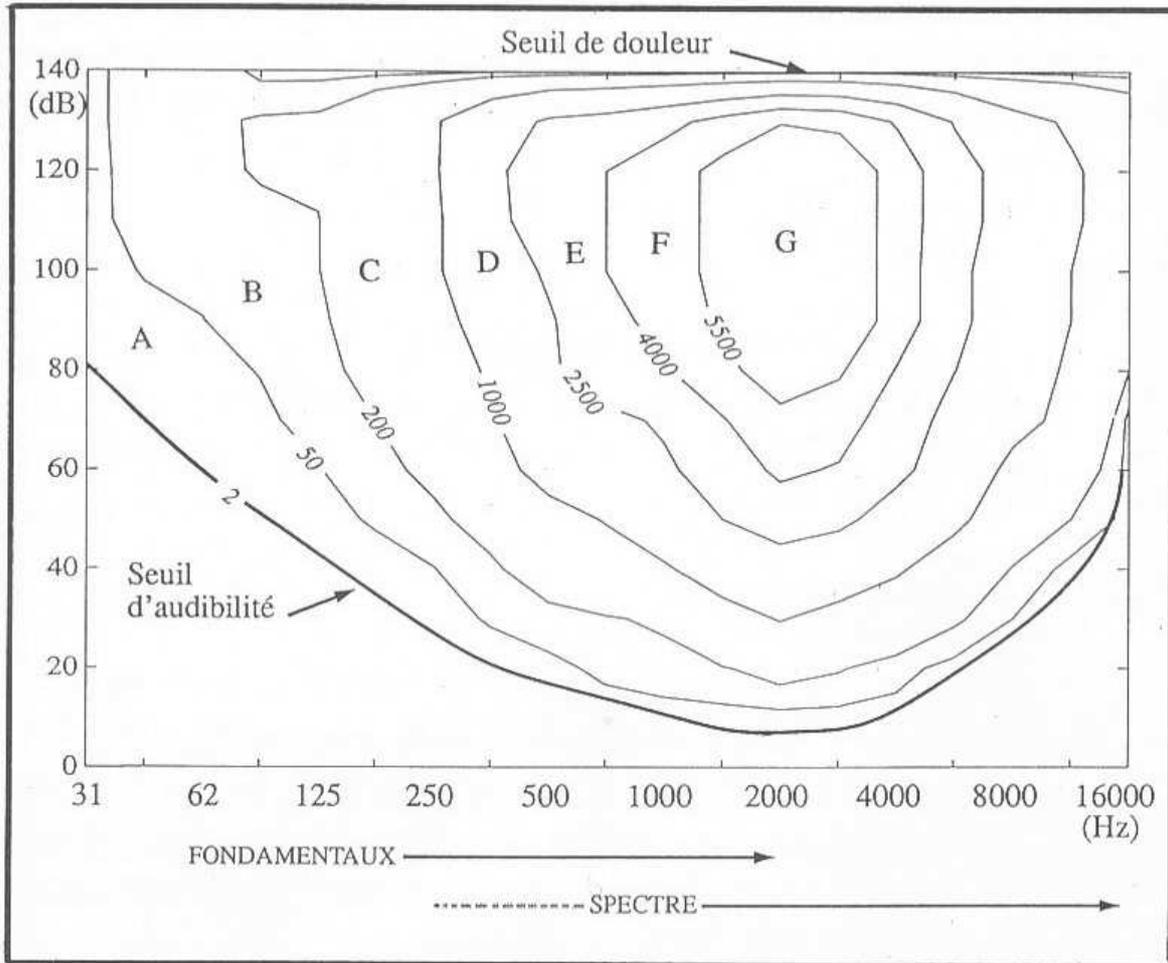
Le seuil différentiel d'intensité suit des variations similaires à la fréquence, sa valeur minimale témoignant d'une sélectivité optimale de l'oreille étant de 0,02 dB pour un son de 4000 Hz à 90 dB (Stevens et Davis, 1938 p.140). La « sonie » (intensité perçue d'un son pur) est déterminée, en première approximation par son niveau de pression acoustique. Mais, à niveau de pression acoustique égale, les sons à fréquence basse ou très élevée ont une sonie inférieure aux sons de fréquence moyenne. L'échelle des décibels qui est la plus communément utilisée pour rendre compte de la sensation d'intensité est une des plus représentatives de la perception. Elle a été établie à l'origine par les premiers ingénieurs travaillant sur le téléphone qui avaient

trouvé en elle une échelle appropriée pour relier l'amplitude sonore à la perception humaine suivant une fonction non linéaire de l'intensité ou de la pression acoustique.

Dans le but d'étudier les éléments de la phonétique, il est légitime de se demander dans quelle mesure l'échelle des décibels est appropriée pour rendre compte de la netteté de la perception des voyelles des consonnes, des tons et des contours de tons. En effet, la perception humaine de l'intensité est également complexe car elle est dépendante de la fréquence et de la durée du stimulus. Il y a trois échelles principales qui peuvent être considérées pour paramétrer l'amplitude du signal de parole. Ce sont l'échelle des pressions (Pascal), l'échelle dérivée des jugements d'intensité (intensity-jnd-rate), l'échelle de sonie et l'échelle des décibels (loi de Fechner). Dans une étude comparative réalisée par Mannell (1994), les échelles jugées comme les plus représentatives de la position des pics (maxima) d'amplitude est l'échelle des Pascals, alors que la plus adaptée à rendre compte des évolutions de l'amplitude le long de la forme spectrale est l'échelle des logarithmes de sonie et l'échelle des décibels. Un des résultats les plus intéressants de cette étude concerne le fait que ces deux dernières échelles sont les plus efficaces pour les consonnes faisant un stop (occlusives), car leur perception nécessite la prédiction de l'amplitude basse et de l'amplitude haute de ces composants phonétiques.

#### 4.2.2.1.5. Diagramme synthétique

A partir des données précédentes, on peut calculer le nombre total de sons élémentaires que le système auditif peut distinguer et étudier leur répartition dans l'aire auditive en dressant une carte de la finesse de résolution de l'oreille en fréquences et en intensité (Figure 86). Le cœur de ce schéma est une zone privilégiée située entre 70 et 120 dB dans l'intervalle de fréquence 1500-4000Hz correspondant à de meilleurs seuils d'audibilité et à une plus grande sélectivité de l'oreille (Stevens et Davis 1938).



**Figure 86 : Variation de la finesse de résolution de l'oreille humaine**

Variation de la finesse de résolution de l'oreille humaine « à partir de l'estimation du nombre de sons élémentaires perceptibles par unité de surface ( $10 \text{ dB} * \frac{1}{2}$  octave, les courbes délimitent les zones de finesse croissantes (de A à G) dont les seuils inférieurs sont respectivement 2, 50, 200, 1000, 2500, 4000 et 5500, le maximum (au cœur) étant de l'ordre de 7000 unités l'intensité (d'après Stevens et Davis, 1938 p.140 in Zenatti, 1994, p.89).

Certaines productions sonores humaines sont partiellement adaptées à cette zone de perception. Tout d'abord la parole classique, première et plus importante source d'éducation de l'oreille. Par exemple, les variations spectrales qui permettraient de distinguer les voyelles et les principales transitions phonétiques, donc de reconnaître les mots, évoluent entre 300 Hz et 4000 Hz. D'autre part, de nombreux instruments traditionnels exploitent ces domaines (flûtes, violons, ...).

En ce qui concerne les sifflements, les conditions d'émission du signal évoluent typiquement entre 80 et 110 dB. Donc, en production, les langues sifflées se situent dans les zones de meilleure perception à la fois en intensité et en fréquence (zones F et G sur la Figure 86). En raison de l'usage des langues sifflées à longue

distance, l'amplitude de perception n'a de sens qu'en terme de contrôle de la dynamique par l'émetteur<sup>88</sup> ou d'écoute à courte distance (par exemple par dessus un torrent comme en Turquie, dans une taverne comme en Grèce ou dans un marché comme au Mexique). Dans ces cas, c'est un point d'explication supplémentaire à la bonne émergence du sifflement dans le bruit.

#### **4.2.2.2. Autres aspects de la perception mis en valeur par la psychoacoustique**

##### **4.2.2.2.1. Composition de la Hauteur Tonale des sons complexes**

Parmi tous les sons dont le système auditif est capable d'extraire une sensation de hauteur tonale, la grande majorité est constituée de sons écologiques complexes et périodiques (ou quasi-périodiques). C'est le cas des voyelles de la voix et des sons de très nombreux instruments de musique. Ces sons complexes périodiques sont harmoniques<sup>89</sup>.

En général, un auditeur à qui l'on confie la tâche d'apparier un son pur à un tel son complexe, ajustera la fréquence du son pur à la fréquence du premier de tous les harmoniques (Davis et coll 1951). Cette fréquence est dite *fréquence fondamentale*. Dans la voix humaine c'est la fréquence de phonation ou de vibration des cordes vocales. La différence de qualité des timbres du son pur et du son complexe rend cette tâche plus difficile qu'entre deux sons purs.

Le timbre est un attribut perceptif lié au spectre de fréquences d'un son à tel point que deux sons complexes ayant la même fréquence fondamentale seront tout de même discriminés en fonction du contenu du spectre fréquentiel. Si l'énergie de l'un est plus grande et compacte aux fréquences élevées on dira qu'il est plus brillant et plus « *aigu* » (Demany 1989, com pers Gautheron 2005). Par conséquent, la HT d'un son complexe périodique a deux qualités distinctes: la qualité de timbre, également appelée qualité de *Hauteur Brute* (HB), (Risset, 1968) et la qualité de *Hauteur Fondamentale* (HF).

Comme la HF manifeste des ambiguïtés d'octaves, cette qualité est jugée par certains auteurs comme parente de la qualité de *chroma*. La HF est en effet plus liée à la périodicité qu'à la fréquence (Plomp 1967). Le chroma serait une qualité selon laquelle deux sons purs dont le rapport de fréquence est proportionnel à 2 et qui forment donc un multiple d'intervalle d'octave sont perçus comme similaires ou identiques (Bachem 1950). La qualité de HB est quant à elle associée à la qualité de tonie.

Demany (1989) précise que ces deux qualités peuvent être contrôlées séparément en production : « *La HB et la HF peuvent être variées de façon complètement indépendantes, en modifiant les caractéristiques physiques distinctes de ce son complexe périodique* ». Il explique aussi qu'elles sont perçues sans

---

<sup>88</sup> Ce facteur peut être particulièrement important à l'acquisition de la technique de la langue sifflée après la période d'apprentissage passif du bébé. En effet, les siffleurs de Turquie, de Grèce, de la Gomera ou du pays mazatèque expliquent avoir appris le sifflement simultanément à la version parlée, vers 2 ans.

<sup>89</sup> Leurs composantes spectrales ont des fréquences qui sont des multiples entiers successifs d'une même fréquence  $f$  ; tout autre son complexe est dit « inharmonique » comme les percussions des tambours par exemple.

apprentissage particulier par tout être humain : « *Il est permis de dire que HB et HF sont deux qualités de HT qu'évoquent spontanément les sons complexes périodiques. En général, pour la plupart des sons complexes périodiques émis dans notre environnement courant, ces deux qualités s'imposent à la conscience sans qu'un effort d'attention soit nécessaire pour les appréhender. Elles sont le produit de ce que Helmholtz appelait l'écoute « synthétique »* » (Demany 1989 p.55). Dans les sons complexes périodiques de la parole classique humaine, ces deux qualités sont exploitées pour encoder les différents éléments porteurs de sens, suivant des principes généraux mais également des règles propres à chaque langue. Par exemple, l'auditeur perçoit les tons ou l'intonation au niveau de la HF alors que le timbre des voyelles est perçu au niveau de la HB. D'une manière générale, suivant la structure des langues les deux qualités de HB et de HF sont combinées différemment pour marquer la phonologie de la langue. La grande variabilité des combinaisons possibles rend nécessaire un approfondissement de la compréhension de la perception de ces deux qualités.

## **Le spectre complexe de la Hauteur Brute (HB) dans les langues**

Pour toutes les langues, la présence d'un son complexe périodique indique l'existence du phénomène de « voisement » qui caractérise les voyelles et certaines consonnes (« consonnes voisées » utilisant les vibrations des cordes vocales). Deux voyelles ou deux consonnes voisées identiques peuvent porter des tons ou des intonations différents tout en gardant leur identité perceptuelle. De même plusieurs types de voix ayant des HF différentes encodent les mêmes voyelles ou les mêmes consonnes. Dans tous ces cas, ce sont les similarités de la répartition spectrale des harmoniques et de l'intensité qu'elles contiennent qui rapprochent perceptivement différentes réalisations. Plusieurs auteurs ont montrés qu'un auditeur parvient en effet à utiliser des indices dans les harmoniques (Hartmann et al 1986, Plomp 1965).

### **Approches phonétiques et formants**

De nombreux phonéticiens se sont intéressés à cet aspect de l'acoustique des voyelles et des consonnes voisées. Ils ont observé que les zones de densité spectrale les plus intenses sur les harmoniques forment des « régions fréquentielles de concentration d'énergie » inter-harmoniques appelés formants (Potter et Steinberg 1950). Plus la voix est grave, plus ses harmoniques sont rapprochées et donc plus les formants seront facilement visibles sur un spectrogramme. C'est pourquoi les voix d'hommes posent moins de problèmes à l'identification des formants que les voix de femmes ou d'enfants dont la HF (ou Fo) est plus haute. D'autre part, plus l'articulation est réalisée clairement syllabe par syllabe, plus les concentrations fréquentielles des formants sont graphiquement identifiables et stables.

Malgré cette différence de résolution graphique, l'identité perçue des voyelles ne décroît pas avec la hauteur du Fo et, dans certaines limites, la perception de la parole n'est pas affectée par la rapidité de production (Gay 1977)<sup>90</sup>. D'autre part, en parole continue, les formants vocaliques sont régulièrement déplacés de leur

---

<sup>90</sup> Gay (1977) a montré par exemple que sur de la parole rapide, les transitions de formants des consonnes changeaient de plus de 60 Hz pour le /p/ et de 25 Hz pour le /b/ sans affecter la reconnaissance des syllabes. Il semble donc que les modulations pertinentes pour la reconnaissance ou la détection d'éléments de la parole varient en fonction du type de parole : rapide/lente, statique/naturelle entre autres.

fréquence habituelle, ce qui n'affecte pas l'intelligibilité des phrases. De nombreux expérimentateurs y ont vu l'effet d'une assimilation contextuelle (Lindblom 1963, Stevens et House 1963). L'existence d'un grand nombre d'irrégularités formantiques et la difficulté méthodologique de leur estimation est encore une réalité aujourd'hui.

### **Hauteur Fréquentielle (HF), tons et intonation:**

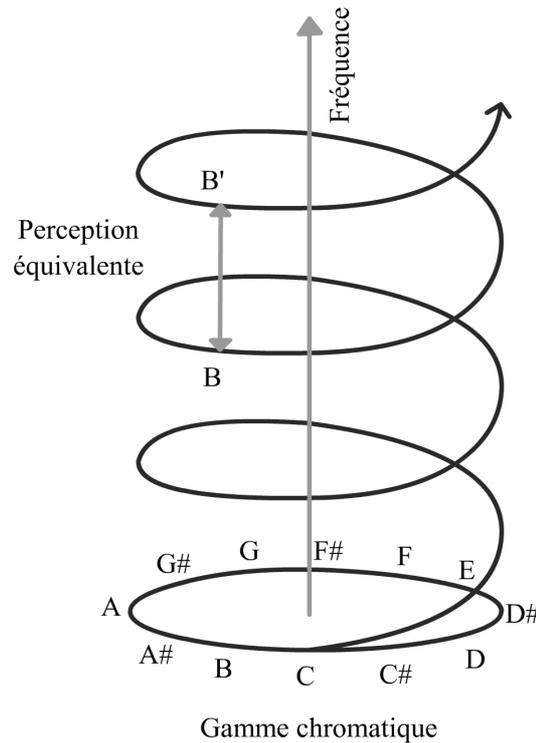
Parallèlement à l'existence de formants dans la qualité HB, les phonéticiens ont remarqué que les tons phonologiquement distinctifs caractéristiques des langues tonales sont exprimés à travers la HF. En raison de l'indépendance des qualités de HB et de HF, une même voyelle (HB) peut porter tous les tons (HF) d'une langue tonale (comme nous l'avons vu pour le mazatèque dans la partie de la typologie). Cette souplesse est également exploitée à la fois dans les langues à tons et les langues sans ton pour transmettre des informations liées à l'intonation.

### **Hauteur Brute, Hauteur Fondamentale et langues sifflées**

Les langues sifflées qui n'encodent l'information linguistique qu'à travers une seule bande de fréquence privilégient l'une des deux hauteurs mais comme nous l'avons vu, ceci ne veut pas dire que l'autre n'est pas prise en compte. Ainsi la transposition des langues sifflées est le résultat d'un compromis entre HB et HF qui permet de rendre compte des éléments les plus phonologiquement distinctifs de chaque langue afin d'optimiser l'intelligibilité du message.

#### **Particularités de la HF et perception des langues sifflées**

Risset (2000) souligne que les effets d'ambiguïté d'octave cachent parfois que notre perception de la HF est en fait mieux représentée sous forme d'une spirale (Figure 87). Il a obtenu une illusion sonore en acoustique musicale qui illustre ce phénomène : un son qui semble descendre en hauteur peut conduire à une hauteur perçue finale bien plus élevée que celle du départ. « *La notation en spirale rend compte de la similarité des notes à intervalles d'octave. Les gammes, descentes ou montées indéfinies correspondent à la dégénérescence de la spirale en un cercle lorsqu'il y a extrême ambiguïté d'octave* » (com. pers. Risset 2005).



**Figure 87 : Perception de la hauteur fondamentale (schéma inspiré de Shepard, 1965)**

Par apprentissage et imprégnation culturelle la spirale peut être affaissée en un cercle, comme dans le cas de la musique occidentale tonale. Deux notes de fréquences différentes mais de position équivalente sur le cercle chromatique seront perçues similaires (Par exemple B et B' toutes deux perçues comme un « si »).

Il se peut que l'affaissement de la spirale par imprégnation culturelle puisse expliquer pourquoi les transpositions sifflées de langues tonales évoluent dans une bande de fréquence limitée à un octave pour une même phrase (la même remarque est valable pour les formes parlées). Une rupture de cette limite entraînerait des possibilités de confusion entre les hauteurs perçues des tons comme dans le cas de l'illusion provoquée par Risset. Cette remarque n'est pas limitée aux langues tonales puisqu'en turc et en grec, la même limite de un octave semble s'appliquer également. Par contre, comme nous l'avons vu, ce facteur n'a apparemment pas la même importance lors de la perception des voyelles sifflées du silbo car ces dernières évoluent couramment sur plus d'un octave dans une même phrase. Il se peut que les Gomero soient culturellement moins marqués par l'octave. Ce serait un aspect intéressant à tester avec ces sifflés.

#### 4.2.2.2.2. Perception temporelle

L'analyse des indices acoustiques de la parole, en particulier liés aux consonnes, permet d'observer que la durée est un paramètre important à la fois pour la différenciation d'évènements proches et pour la détection de la fréquence des indices phonétiques. Green (1985) a distingué deux classes principales de phénomènes auditifs temporels, l'intégration temporelle et l'acuité temporelle.

## Intégration temporelle

Les études sur l'intégration temporelle cherchent à trouver la longueur des intervalles pour lesquels le système auditif intègre l'information acoustique de différents éléments. En effet, le phénomène d'intégration temporelle, qui est valable sur des durées allant jusqu'à 200 ms est à la fois lié à la possibilité de détecter des silences (*gap detection*) et au phénomène de masquage non simultané (ou *masquage temporel*). Pour attribuer une hauteur précise à un son pur il faut non seulement qu'il ait un certain niveau de pression acoustique appelé seuil de perception tonale mais également une durée d'au moins 10 ms. En dessous de cette limite, il tend à être perçu comme une transitoire ou un clic auquel une hauteur peut difficilement être attribuée.

## Acuité temporelle

Les études sur l'acuité temporelle analysent la rapidité de réponse du système auditif à des événements acoustiques brefs. Les durées minima de séparation de deux événements successifs sont mesurées. Ainsi les seuils de détection de silence (*gap detection thresholds*) ont des valeurs uniformes de 6 à 8ms sur toute la bande de fréquence de perception humaine. Par contre les seuils différentiels de durée de deux sons varient: ils sont de 50 ms pour des sons purs de 1 s (sons ayant des caractéristiques temporelles proches des voyelles) et de 2 à 3 ms pour des sons purs de 5 à 20 ms (sons ayant des caractéristiques proches des consonnes).

## Précisions sur le masquage temporel

D'autre part, en ce qui concerne le phénomène de masquage temporel, une différence entre l'effet de masquage « antérieur » (*rétroactif*) et celui de masquage « postérieur » (*proactif*) montre que la réponse temporelle du système auditif est asymétrique avec une réaction plus rapide aux attaques de stimuli (*stimulus onsets*) (perception des variations de 20 dB en 10 ms) qu'aux relâchements (*offsets*) (perception des variations de 20 dB en 20 ms). Le rôle fondamental de l'entité *Consonne-Voyelle (CV)* observé par les linguistes dans la plupart des langues est ici en partie justifié.

La résolution temporelle moyenne liée au masquage est de 20 à 30 ms. Avec l'apprentissage, ces valeurs sont améliorées. Elles varient également en fonction du contexte sonore (Figure 88).



**Figure 88 : 3 formes d'excitations temporelles provoquées par des sons purs de durées différentes (Zwicker 1982)**

Pour chaque événement on observe que l'attaque engendre un masquage rétroactif plus court que le masquage proactif. C'est pourquoi un silence de 120 ms pris entre 2 excitations de 400 ms aura la même durée perceptive qu'un son de 30ms.

## **Durée et perception de l'intensité**

Pour des durées sonores inférieures à 1 seconde, qui caractérisent la plupart des éléments de la parole, la sonie (intensité subjective) augmente avec la durée (Botte 1989). D'après Pedersen et al (1977), la relation entre la sonie et la durée est une fonction exponentielle dont la constante de temps est 80 ms et qui s'approche de sa valeur limite à 180 ms.

## **Conclusion pour la perception temporelle et conséquences pour l'analyse des langues sifflées**

Si l'on tient compte de tous ces facteurs perceptifs temporels on conclut que la résolution temporelle phonétiquement significative sera entre 10 ms et 30 ms et sera la même pour à peu près toutes les fréquences. L'acuité temporelle varie en fonction de la durée des sons impliqués. Jusqu'à ce jour, les adaptations non linéaires de la réponse du nerf auditif et l'amélioration des performances par apprentissage ont compliqué les tentatives de détermination d'un modèle temporel des réalisations phonétiquement pertinentes.

D'autre part, le temps de résolution temporelle du système auditif est de l'ordre de 6 à 8 ms pour des signaux sinusoïdaux dont le signal sifflé est un parent proche. Cette précision accrue justifie l'intérêt de faire une étude des régularités temporelles encodées dans ces systèmes linguistiques particuliers. Dans cette perspective, les résultats que nous avons présentés dans les paragraphes précédents permettront de préciser la validité des mesures réalisées.

### **4.2.2.3. Conclusion sur la psychoacoustique et sifflements**

Le système auditif réagit comme un banc de filtres passe-bande se chevauchant et dont la largeur de bande croît avec la fréquence. Une telle organisation favorise la perception relative des amplitudes et des fréquences. Les sifflements se situent dans une zone où la largeur du filtre auditif varie entre 120 et 500 Hz. Ils n'activent que peu de filtres auditifs à un instant donné, car d'après nos mesures, la largeur de bande sonore qui émerge du bruit est d'environ 400 Hz à courte distance et 150 Hz à 550m. Cette précision a pour conséquence de réduire les possibilités de masquage et c'est une raison pour laquelle un sifflement à une bonne résistance au bruit.

De plus l'acuité temporelle de l'audition est sensiblement la même à toutes les fréquences, mais elle varie en fonction de la durée des événements sonores. Elle est plus précise pour les sons sinusoïdaux et peut être améliorée par apprentissage. Enfin, la zone pour laquelle l'audibilité et la sélectivité de l'audition humaine sont les plus performantes se situe entre 1 et 4 kHz et 75-120 dB. Le sifflement se situe là encore, dans les zones de perception fréquentielle les plus efficaces de l'oreille. Si l'on considère le signal sifflé au niveau de l'émetteur, la zone de perception la plus précise en termes d'intensité est également concernée.

On peut donc dire que l'observation de la réaction du système auditif à des sons purs ou des éléments simples extraits de la parole est un premier pas permettant déjà de préciser le protocole d'analyse du signal de parole et de faire émerger l'adaptation spectaculaire des langues sifflées à l'audition.

### 4.2.3. Structuration perceptive du flux de parole

Nous avons vu que la perception des sons est organisée suivant plusieurs dimensions ou attributs de la perception (intensité, hauteur, durée principalement). Cependant le signal de parole naturelle n'est pas perçu comme un ensemble d'entités distinctes à chaque instant, mais plutôt comme un flux organisé à plusieurs échelles (segments-syllabe-mots-phrases) et entrecoupé par des discontinuités. C'est pourquoi l'analyse d'éléments segmentaux ou tonaux considérés comme isolés risque d'exclure un grand nombre de mécanismes mis en oeuvre au quotidien lors de la production et du traitement perceptif de la parole.

Deux domaines de la recherche actuelle, a priori différents, adoptent une perspective globale multiéchelle sur le problème de la perception du continuum sonore d'une phrase : il s'agit d'une part l'« *analyse de la scène auditive* » (Bregman, 1990) qui a permis d'expliquer comment notre cerveau scanne l'espace auditif et en tire des informations sur l'origine des sons, d'autre part de l'*analyse de la prosodie* développée en linguistique pour comprendre la mélodie et le rythme du langage. L'usage à distance des langues sifflées a dirigé notre étude vers le premier domaine et l'usage d'une bande unique de fréquence pour la parole nous a suggéré un rapprochement avec la prosodie. Il s'est avéré que les deux domaines éclairent de façon complémentaire les relations perceptives fondamentales qui forment la base de l'intelligibilité des langues sifflées.

#### 4.2.3.1. Attributs de la perception et « Analyse de la scène auditive »

##### 4.2.3.1.1. Organisation de la scène auditive: origine

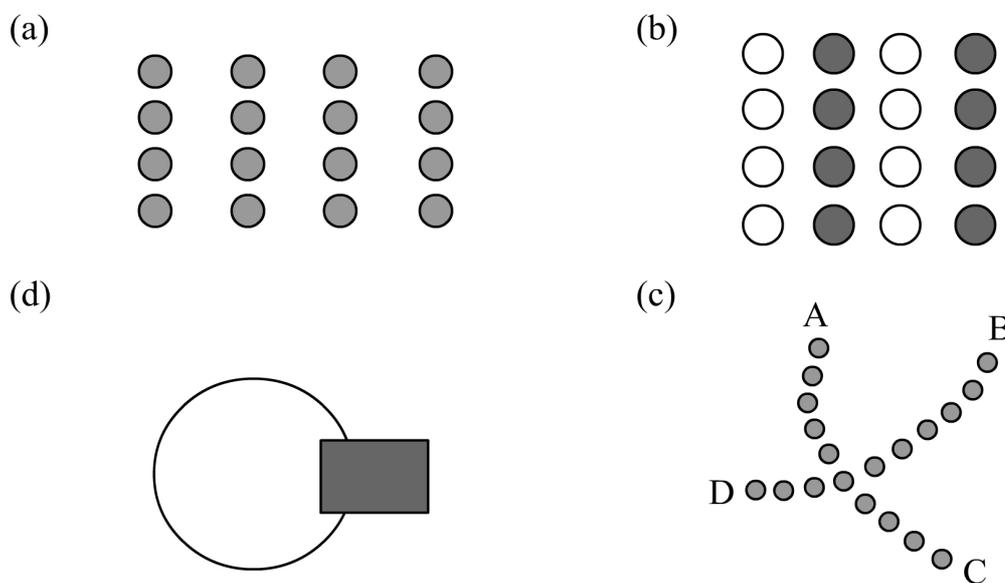
A l'écoute d'une phrase dans un environnement bruyant, notre expérience est bien différente de la réception d'un ensemble de sons dont les fréquences et les amplitudes varient dans le temps. Nous structurons plutôt le monde sonore en entités cohérentes que l'on peut (i) détecter, (ii) séparer, (iii) localiser, et si possible (iv) identifier. La notion d'intelligibilité d'une forme sonore s'appuie sur l'existence d'aptitudes perceptives auditives permettant de réaliser ces tâches vitales au quotidien. « *L'oreille a développé ses capacités dans un monde où les sons sont presque tous d'origine mécanique: elle procède à des enquêtes, elle essaye de débusquer le mode de production physique qui est -ou pourrait être - à l'origine du son [...] elle effectue des inférences complexes à partir d'indices subtils pour évaluer la direction de la source sonore, sa distance et l'intensité émise à sa source. L'audition organise le complexe sonore qui lui parvient, elle sépare ou regroupe les éléments constitutants, elle y distingue des « voix » ou des « images » des sources sonores différentes* » (Risset 1994, p. 103).

Les modes de groupements sont d'autant plus importants pour la description des langues et de tout phénomène sonore qu'ils soulignent que ce ne sont pas directement les paramètres physiques du signal qui importent pour le récepteur mais la relation qu'ils vont avoir lorsqu'il les perçoit. Cet aspect fondamental est exploité par les langues humaines et sans lui les langues sifflées ne pourraient pas transposer une langue.

#### 4.2.3.1.2. Lois de regroupements et images acoustiques

##### Introduction

Dès lors, un des défis majeurs est de comprendre comment les caractères physiques se manifestent sous la forme de relations perceptuelles. Selon Helmholtz (1821-1894), l'observateur met en oeuvre des lois perceptives qui donnent naissance à l'interprétation la plus efficace de l'environnement. Ce type de raisonnement peut expliquer pourquoi les philosophes de la Gestalt ont élaboré des lois fondamentales d'organisation de la perception (en particulier Wertheimer). Ce sont les lois dites (a) de proximité, (b) de similitude, (c) de continuité et (d) de clôture, qui sont présentées sous leur forme appliquée à la vision sur la Figure 89.



**Figure 89: Illustration des lois d'organisation perceptive énoncées par les psychologues gestaltistes**  
(a) proximité, (b) similitude, (c) continuité, (d) clôture (reproduction de Deutsch 1982)

Ces lois s'appliquent également à l'audition : (a), (b) et (c) favorisent les groupements, (d) la segmentation (Deutsch 1982).

D'après Bregman (1990), les séparations et les regroupements sont les processus élémentaires à la base de la faculté d'« analyse de la scène auditive », dont la fonction est de faciliter l'extraction d'informations sur l'origine des signaux dans un monde sonore complexe. Cette tâche est nécessaire parce que des sons indépendants issus de nombreuses sources arrivent à chaque instant à l'oreille. L'auditeur n'a accès qu'à un mélange continu d'ondes. Comme Helmholtz le faisait remarquer, il y a 150 ans, notre oreille est presque dans la même situation qu'un oeil qui regarderait un point précis de la surface de l'eau, et observant l'ensemble des perturbations à sa surface, devrait analyser l'origine des ondes qui la composent.

Or, comme l'emplacement et les caractéristiques physiques des objets changent relativement lentement par rapport à la célérité des vibrations des sons, deux événements sonores à la fois proches dans le temps<sup>91</sup> et dont d'autres dimensions perceptives sont communes (hauteur, intensité<sup>92</sup>, timbre...) ont de bonnes raisons d'avoir été produits par la même source (Bregman 1990). La localisation de la position de la source est également souvent un critère important<sup>93</sup>.

## Image acoustique et flux

D'après Bregman et McAdams (1979), l'analyse de scène réalisée permet de dresser une *image acoustique* stable de chaque source sonore de l'environnement.

### Dimension verticale

Suivant une dimension « verticale » de cette image, les éléments spectraux, jugés comme issus de la même source sonore sont fusionnés en un événement sonore complexe (phénomène de *groupement simultané* comme par exemple pour le timbre de la voix). Cette catégorie de regroupements peut être issue de plusieurs types de relations:

- (1) Coordination par synchronisation spectrale : le timbre en est l'illustration grâce à la cohérence temporelle de son attaque, qui participe à son établissement (Bregman & Pinker 1978). La structure consonne-voyelle exploite ce phénomène en variant les attaques mais en établissant toujours au final une cohérence temporelle. Il n'est donc pas étonnant que les langues sifflées sans tons traduisent ce type d'événements par une continuité de modulation entre la consonne et la voyelle<sup>94</sup>.
- (2) Coordination par microvariations communes : Chowning a mis en évidence que l'audition s'appuie sur le destin commun de composants fréquentiels différents: elle associe les composants qui ont des microvariations synchrones (par exemple des composantes affectées de « vibrato » similaires, ou alors des composantes se déplaçant ensemble dans l'espace). Dans ces cas, l'enveloppe spectrale relate les changements de dynamique de sa source sans être affectée par la propagation et l'atténuation des sons. La perception humaine de la perspective auditive exploite cette particularité (Chowning 2000).

---

<sup>91</sup> À l'échelle de la perception de l'oreille.

<sup>92</sup> Évolution de l'enveloppe spectrale entre autres.

<sup>93</sup> Elle est mesurable par le système auditif, d'une part grâce aux différences temporelles (ITD) et d'intensité (IID) qui existent pour chaque onde qui parvient successivement à une oreille puis à l'autre, d'autre part au niveau du pavillon entre les ondes directes et les ondes réfléchies sur les différentes cavités de l'oreille externe (Batteau, 1967).

<sup>94</sup> Principalement dans les langues utilisant une transposition s'appuyant sur ces segments. Notre typologie montre que lorsque leur rôle informatif décroît, cette modulation perd progressivement de l'importance en sifflements.

## Dimension horizontale

Au niveau « horizontal » de la métaphore de l'image auditive, ce sont des flux temporels continus qui sont formés. La continuité peut émerger de plusieurs indicateurs perceptuels:

- (1) la proximité fréquentielle et la proximité temporelle ont été les plus étudiées (Van Noorden 1975),
- (2) la similitude de la qualité spectrale des sons permet également des regroupements en flux (Wessel 1979). Ce dernier phénomène est particulièrement important puisqu'il souligne la possibilité d'associer des événements sonores en fonction de leur timbre: il est à la base de l'intelligibilité des voyelles d'une langue.

## Conséquences pour la fission d'évènements

La création de flux perceptifs a des conséquences importantes au niveau des relations entre les événements acoustiques qui les constituent: par exemple les coordinations temporelles entre deux événements faisant partie de deux flux différents sont impossibles (Bregman et Campbell, 1971). Cette dernière propriété est constamment utilisée dans l'élocution d'une langue.

Par opposition à la fusion en flux, il existe donc des phénomènes de ségrégation de flux ou fission. Ils apparaissent surtout quand les éléments à l'origine de la séquence perçue sont notablement différents en fréquence, en spectre, en temps de montée de l'enveloppe ou en localisation spatiale.

### 4.2.3.1.3. Conclusion

Les mécanismes de groupement interagissent entre eux de manière complexe: chaque image auditive possède des attributs perceptifs qui ne sont pas le résultat simple de la somme des événements perceptifs qui la constituent. A chaque instant, l'image possède une dynamique qui dépend du contexte, ce qui entraîne des évolutions où des transferts sont effectués entre attributs perceptifs (Bregman et Pinker, 1978). Les caractéristiques de cette dynamique vont permettre d'encoder des formes perceptives. Pour cela, un facteur important est la possibilité d'encodage, suivant une dimension donnée, en présence de changements suivant d'autres dimensions (indépendance par rapport aux autres dimensions). Le facteur temporel a donc une grande robustesse pour l'encodage de formes. On dira également qu'une dimension donnée est un bon candidat pour véhiculer des formes perceptives auditives si un grand nombre de possibilités de configurations peuvent être codées suivant cette dimension. Par exemple, nos capacités d'encodage et de perception des fréquences relatives vont permettre le développement de lignes mélodiques variées suivant la dimension fréquentielle. Les langues sifflées exploitent tous ces aspects.

A travers de telles relations la perception auditive bouleverse parfois la structure acoustique d'origine. Ces phénomènes ont des implications directes sur la perception de la parole.

## 4.2.3.2. Prosodie et langues sifflées

### 4.2.3.2.1. Introduction : l'étude de la parole continue en linguistique

La notion de prosodie est très ancienne en linguistique, puisqu'elle fut utilisée dans la philologie classique puis trouva un nouvel essor dans la description comparative des langues indo-européennes. Elle fut finalement intégrée dans les diverses tentatives de théorisation des écoles de linguistes du XX<sup>ème</sup> siècle (Jakobson, 1931; Pike, 1954, ...). Son usage expérimental fait souvent l'objet d'une définition assez vague comparée aux autres concepts de la linguistique car les paramètres physiques étudiés cachent la méthodologie d'étude que sous-entend cette notion.

Sur ce point, l'étymologie du terme « prosodie » fait référence à un lien entre musique et parole: « ôdê », veut dire chant en ancien grec ou plus précisément chant accompagné d'une musique instrumentale. Il a ensuite été utilisé pour désigner la « science de la versification » ou *métrique*, qui gouverne la voix humaine lorsqu'elle est en train de lire de la poésie. D'après l'encyclopédie Universalis la définition du terme est : « *prosodia: accent, quantité dans la prononciation* » ou « *Règles concernant les rapports de quantité, d'intensité, entre les temps de la mesure et les syllabes de la musique vocale* ». L'analyse de la prosodie suppose donc un point de vue synthétique, elle invite à observer les paramètres physiques du signal et leurs attributs perceptifs à différentes échelles (microprosodie à l'échelle du phonème, macroprosodie à l'échelle de la phrase). Elle s'appuie à la fois sur les résultats de la phonétique et de la phonologie pour considérer des phénomènes relatifs qui ne sont pas liés à un phonème particulier mais plutôt à l'enchaînement des segments entre eux (éléments suprasegmentaux). C'est pourquoi les notions d'accentuation, de mélodie, de durée relative et de rythme sont étudiées en linguistique dans le cadre de l'analyse de la prosodie.

### 4.2.3.2.2. Les composantes de la prosodie

La méthode utilisée pour observer ces éléments consiste à considérer les variations d'intensité et de hauteur (fréquence et amplitude perçues) ainsi que les durées successives des segments syllabiques (Calliope 1989). La complexité des phénomènes observés explique la difficulté qu'ont eu les chercheurs à cerner les dimensions acoustiques les plus pertinentes à analyser. Suivant une approximation courante, ces éléments sont essentiellement abordés à travers l'étude conjointe des variations de la fréquence de phonation et de la courbe d'amplitude du signal de parole. C'est pourquoi ces deux composantes physiques du signal de parole sont les composantes principales de l'analyse de la prosodie :

- L'amplitude représente l'énergie sonore du signal à chaque instant,
- La fréquence fondamentale (F0) et correspond à la fréquence de vibration des cordes vocales (par exemple, en Français, le F0 augmente à la fin d'une phrase interrogative).

De nombreuses études considèrent que l'analyse du Fo seul est représentative de la prosodie d'une langue.

Trois autres composantes d'ordre prélexicales sont parfois utilisées (Ramus 1999):

- La structure syllabique dont l'importance varie en fonction de la phonologie de la langue,

- La structure Consonne/Voyelles (CV) qui est un facteur commun à toutes les langues ;
- La structure phonémique qui est liée au fait que certains types de sons n'existent pas dans toutes les langues : on rencontrera des sons glottalisés en turc ou en chepang, mais pas en grec. Par ailleurs des sons que deux langues ont en commun peuvent tout de même servir à les discriminer, s'ils ont des distributions très différentes dans les deux langues.

Ce regroupement en cinq composantes n'est rarement utilisé à l'heure actuelle en linguistique mais nous verrons que les langues sifflées soulignent le rôle important des composantes prélexicales.

### **Information portée par ces composantes prosodiques**

L'information de ces composantes se situe à trois niveaux différents :

- L'intonation, ou l'enchaînement des tons dans les langues tonales, sont fournis par la fréquence fondamentale et l'amplitude,
- Le rythme: il provient de plusieurs paramètres comme la fréquence fondamentale, l'amplitude, la structure syllabique<sup>95</sup> et la structure CV.
- La phonotactique est donnée par l'enchaînement des phonèmes.

#### **4.2.3.2.3. Langues sifflées : une approche pragmatique et naturelle de la prosodie**

##### **Un regard alternatif sur la notion de prosodie**

L'analyse de la prosodie a donc développé une manière propre à la linguistique d'analyser le flux sonore dans une parole continue. Cependant, en raison de sa méthodologie reposant avant toute chose sur le Fo, une approche prosodique d'une langue sifflée fera émerger des différences avec le même type d'analyse réalisé sur la version parlée de la même langue. Ceci sera particulièrement visible sur les langues sans tons et apparaîtra également dans les langues à tons qui utilisent quelques éléments spectraux différents du Fo comme base du sifflement. En effet les langues sifflées considèrent naturellement que la HB peut porter des éléments à valeur prosodique. La « *prosodie interne* » des éléments spectraux des consonnes et des voyelles n'est donc pas à négliger, du point de vue de la perception et de la production des siffleurs.

Par conséquent, s'il est vrai que chaque langue sifflée propose une description prosodique d'une langue à partir d'une seule bande étroite de fréquence, ses moyens de l'obtenir et les critères que les siffleurs retiennent sont différents de l'approche développée par la tradition scientifique: alors que l'analyse de la prosodie s'appuie systématiquement sur la fréquence fondamentale, la description naturelle des langues sifflées, quant à elle, compose avec les différents éléments de la voix soulignant ainsi que les composantes « prosodiques » les plus pertinentes pour l'intelligibilité de chaque langue diffèrent en fonction de la structure de la langue.

---

<sup>95</sup> Les langues humaines ont une rythmicité plus ou moins accentuelle ou syllabique ce qui fait que suivant la langue la syllabe jouera un rôle différent dans le rythme.

#### 4.2.3.2.4. Collusion entre l'analyse de la prosodie et l'analyse de la scène auditive

La prosodie est connue comme un facteur améliorant l'intelligibilité de la parole. Pour ce faire, elle s'appuie sur tout un ensemble de paramètres qui ont leur origine dans la faculté d'analyse de la scène auditive. Nous en détaillons ici les points les plus importants.

##### **Traquage de la voix d'un locuteur**

Les propriétés perceptives qui permettent de reconnaître les variations prosodiques sont soumises à l'identification préalable d'un flux de parole issu d'une seule et même source: le locuteur. Cette tâche est réalisée en grande partie par l'analyse de la continuité de Fo. La cohérence des variations du Fo permet à un auditeur de continuer à décoder une voix particulière, même pendant les périodes non voisées ou de silence lors desquelles le signal acoustique est souvent en compétition avec celui d'autres voix ayant des fréquences fondamentales différentes (Nootheboom et al, 1978).

##### **Continuités spectrales intra mot**

Tous les autres indices permettant d'identifier des sons cohérents viendront renforcer l'analyse de la prosodie. Ainsi, à l'échelle du mot ou de la phrase l'analyse des continuités de Fo est complétée par une analyse des continuités des concentrations énergétiques portées par les harmoniques (Dorman et al, 1985). A ce niveau, le rôle de la synchronisation des attaques lors de la fusion des enveloppes spectrales a été étudiée spécifiquement pour les sons de la parole. Par exemple, Darwin (1981) a montré que des sujets ayant écouté de manière décalée dans le temps, d'abord les harmoniques constituant le premier formant d'une voyelle auxquelles ont été ensuite ajoutées les harmoniques du deuxième formant, ont souvent identifié individuellement ces deux entités, ce qui a modifié la perception de la voyelle d'origine. Une telle identification séparée de groupes d'harmoniques est bien plus difficile si les deux groupes débent simultanément. Les consonnes servent donc en partie à synchroniser les éléments spectraux de la voyelle. Une conséquence importante de ce phénomène est de lier fortement la consonne et la voyelle en un même flux auditif. La réalité perceptive de la syllabe en est renforcée (d'autant plus que la perception de la synchronie n'implique pas une identification consciente)<sup>96</sup>. Cela se traduit dans les langues sifflées sans ton par une modulation continue et unique de la voyelle avec la consonne.

---

<sup>96</sup> Bien que les phonèmes soient parfois les plus petites unités distinctives permettant de distinguer deux mots, ils ne sont pas perçus séparément: les temps de réaction des sujets lors d'expériences de psycholinguistique testant la perception de phonèmes cibles le montrent : ils sont plus courts pour des syllabes ou des mots que pour des phonèmes ce qui suppose un traitement à plus haut niveau pour les phonèmes (Savin et Bever, 1970; Segui 1988).

## Continuité temporelle dans une phrase

### Analyse spectrale: évolution temporelle de l'énergie sonore

D'après plusieurs auteurs, le système auditif serait capable d'estimer la cohérence de l'évolution de l'énergie au cours du temps dans un même flux sonore (Botte 1989). Ce facteur est un des éléments clef de la perception de la continuité temporelle de la prosodie d'une phrase. En effet, les variations d'intensité permettent de déterminer la position de l'accent lexical propre à de nombreuses langues, ce qui conditionne en partie le rythme d'une langue<sup>97</sup>. En général, l'enveloppe d'amplitude joue un rôle non négligeable de marqueur de contrastes distinctifs qui donnent une cohérence rythmique d'ensemble au signal. C'est pourquoi un bruit blanc modulé en amplitude par l'enveloppe d'un signal de parole peut être perçu comme un mot (Katz et Berry 1971). Or, Chowning a montré que des indices sonores ayant une intensité et une répartition spectrale cohérentes dans le temps avaient toutes les chances d'être issus de la même source et donc d'être groupés ensemble. Il a aussi observé que l'évaluation de la distance, grâce à la comparaison de l'intensité du signal direct par rapport aux signaux réverbérés, renforce l'estimation de cette cohérence et améliore donc la perception de l'évolution temporelle de l'enveloppe d'amplitude du signal<sup>98</sup>. Ces dimensions psychoacoustiques sont la base de la perception de la perspective auditive: « *auditory perspective is not a metaphor in relation to visual perspective, but rather a phenomenon that seems to follow general laws of spatial perception. It is dependent upon loudness (subjective !) whose physical correlates we have seen to include spectral information and distance cue, in addition to intensity* » (Chowning 2000, p.5).

Dans le cas des langues sifflées la tâche de détection de la cohérence temporelle par l'auditeur est simplifiée par l'usage d'une seule bande de fréquence dont la cohérence spectrale peut être évaluée grâce à sa largeur de bande dans la partie qui émerge du bruit.

### Durées

Au niveau des phrases qui forment des séquences sonores longues, tous les phénomènes décrits précédemment favorisent la résistance de l'intelligibilité à la dégradation du signal. Huggins (1975) a réalisé des tests sur de la parole segmentée temporellement. Il a montré deux phénomènes révélateurs sur l'organisation temporelle de la phrase: (a) la parole interrompue par des silences de durée fixe (200 ms) atteint un taux d'intelligibilité de 90% si les intervalles de parole restants sont supérieurs à une durée de 175 ms. Ce qui correspond à la durée moyenne d'une syllabe. (b) D'autre part, si les intervalles de parole sont de 60 ms, l'intelligibilité atteint 90% pour des intervalles de silence inférieurs à 70 ms. Ce qui correspond à une durée tout juste supérieure à la durée moyenne d'une consonne de la voix parlée. Dans certaines situations de

---

<sup>97</sup> La rythmicité d'une langue est en partie portée par l'accent. Ce facteur se combine avec le rôle de la syllabe pour définir une rythmicité plus ou moins accentuelle ou syllabique en fonction de la phonologie de chaque langue (Ramus 1999).

<sup>98</sup> Les expériences réalisées par Chowning étaient à courte distance (50m) dans un milieu intérieur, la conservation de l'énergie issue des réverbérations est moins complète pour les langues sifflées en raison de la déperdition d'une grande partie de l'énergie sonore par réflexion.

dégradation intermédiaire, il a également montré que les sujets sont capables de discriminer certaines transitions ou certaines voyelles mais la parole reste inintelligible. Ceci confirme qu'une certaine continuité rythmique est nécessaire à l'intelligibilité de la parole et que les consonnes y jouent un rôle clef. Au niveau de la transition consonantique, cette continuité est rompue pour une durée de silence supérieure à 100 ms (Nooteboom, 1978).

### **4.2.3.3. Conclusion**

L'analyse de la scène auditive et l'analyse de la prosodie nous ont permis de comprendre la base des relations qu'entretiennent les attributs de la perception dans le cadre d'un continuum de parole. Les points communs que nous avons dégagés entre ces deux domaines de recherche pour les besoins de l'analyse du sifflement linguistique s'appliquent à tout phénomène de langage. Une comparaison avec les résultats des travaux de recherche réalisés par ailleurs sur la parole dans le bruit est maintenant possible : ceux-ci concluent que les paramètres les plus utilisés par les personnes en condition d'écoute difficiles (rapport signal sur bruit faible, et/ou effet cocktail party, et/ou réverbération en salle) sont les suivants : (a) l'amplification sélective et l'amélioration du rapport signal sur bruit au niveau périphérique, (b) la séparation spatiale du locuteur et la perception de la perspective auditive (liées à la localisation des sons), (c) la dépendance temporelle du son ciblé (liée à la continuité rythmique), (d) l'identification et la poursuite des attributs de hauteur de la voix du locuteur (avec un suivi dynamique grâce à la perception de la continuité soit du  $F_0$  soit des concentrations spectrales d'énergie) (Bronkhorst 2000). On retrouve les éléments les plus saillants du signal de parole à l'échelle de la phrase qui sont mis en évidence par l'analyse de la prosodie et qui sont expliqués par l'analyse de la scène auditive. Toutes les recherches sur la parole considérée comme un flux sonore convergent donc vers les mêmes conclusions. Il n'est pas étonnant que les langues sifflées mettent particulièrement en valeur ceux de ces facteurs clefs de l'intelligibilité de la parole qui correspondent le mieux à la réalité acoustique d'un sifflement.

Un autre point a pu être souligné : les langues sifflées apportent des éléments de réflexion pratiques pour l'analyse de la prosodie puisqu'elles privilégient les éléments porteurs des relations les plus marquantes de la structure linguistique plutôt que seulement le  $F_0$ . Ce type d'approche est perceptivement justifié, en effet, chaque être humain est capable de détecter l'information la plus saillante dans le signal, selon la définition de Hombert & Maddieson (1998), à savoir les traits segmentaux et suprasegmentaux qui sont non seulement les plus identifiables du point de vue acoustique, mais aussi les plus discriminatoires des langues. En quelques sorte, les langues sifflées proposent une nouvelle *méthodologie naturelle d'analyse de la prosodie*, pertinente car elle repose sur plusieurs générations de siffleurs et donc de cerveaux entraînés depuis l'enfance à une pratique héritée d'un processus évolutif. Dans ces conditions, on peut penser que l'analyse du rôle prosodique du timbre a un grand avenir. En effet, en termes de hauteur le timbre a une réalité perceptive (Carlson et al 1970, Bladon et Fant 1978) dont les conséquences prosodiques valent le détour pour les langues non tonales (proéminence de l'accent par exemple). Si ce n'était pas le cas les siffleurs de Turquie, de Grèce et de la Gomera ne pourraient pas en faire un élément clef de leur transposition.

### **4.3. Intelligibilité de la parole sifflée: analyse progressive**

Nous allons maintenant présenter l'ensemble des processus impliqués dans l'intelligibilité des langues sifflées tels qu'ils sont connus à l'heure actuelle et sans en éluder la complexité. Nous nous limitons aux langues sifflées sans ton car ce sont celles que nous avons analysées le plus en détail. Nous avons organisé cette troisième partie du chapitre en une série progressive de sections. Dans un premier temps nous présentons les résultats d'une analyse de psycholinguistique portant sur l'identification de voyelles de Silbo par des sujets Français qui ne connaissaient rien au sujet des langues sifflées. Ensuite nous rappellerons les performances de reconnaissance sur des non-mots des sujets testés expérimentalement en Turquie dans les années 70. Puis nous passerons à l'analyse de la perception liée au contexte lexical, au niveau des mots et de leur contenu. Ce n'est qu'après ce stade que nous rentrerons plus avant dans la réalité rencontrée par les siffleurs avec l'analyse de l'intelligibilité des phrases. Enfin nous décrirons les résultats préliminaires d'une analyse des conséquences phonétiques de la dégradation naturelle de la parole sifflée à plusieurs distances.

#### **4.3.1. Expérience de perception des voyelles sifflées par des sujets ignorant tout des langues sifflées**

##### **4.3.1.1. Principes généraux**

Afin d'approfondir la compréhension de la perception des voyelles sifflées et leurs liens avec les voyelles articulées de la voix parlée, nous avons mis au point deux variantes d'une même expérience permettant de tester des sujets ne connaissant rien au phénomène des langues sifflées. Notre objectif était en effet de comprendre si ces personnes étaient en mesure de réaliser les mêmes catégorisations de voyelles que des siffleurs à la simple écoute de voyelles sifflées. Pour cela ils devaient effectuer une tâche simple et intuitive pour laquelle à aucun moment ils n'ont pu s'appuyer sur une quelconque correction, traduction ou explication de la répartition réelle des fréquences des voyelles sifflées.

##### **4.3.1.2. Méthode et corpus**

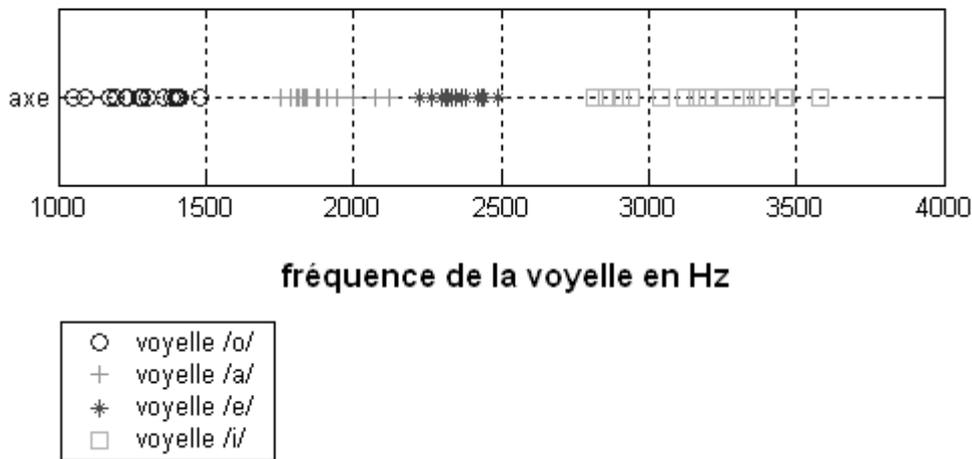
###### **4.3.1.2.1. Choix de la langue sifflée et choix de la langue maternelle des sujets testés.**

Nous avons choisi de réaliser cette expérience sur des sujets de langue maternelle Française. Les voyelles sélectionnées étaient /i/, /e/, /a/, /o/ de la langue sifflée espagnole de la Gomera. En effet, ces voyelles existent également en français avec des réalisations similaires à l'espagnol. Une autre raison de ce choix tient au fait que ces quatre voyelles sifflées (ou des réalisations très proches) ont la même répartition fréquentielle à la fois en grec, en turc. Etant donnée la structure de la langue française, on peut raisonnablement penser que si une langue sifflée se développait, ces voyelles se répartiraient de la même manière.

#### 4.3.1.2.2. Matériel sonore et stimuli

Le matériel sonore utilisé comprenait 84 voyelles de la langue sifflée espagnole de la Gomera (Silbo). Elles ont toutes été extraites de l'enregistrement de 20 longues phrases, sifflées relativement lentement en une seule session par une même personne dans les mêmes conditions (distance à atteindre, bruit, technique utilisée). Les 84 voyelles (21 /i/, 21 /e/, 21 /a/ et 21 /o/) ont été choisies suivant des critères statistiques basés sur notre analyse du Silbo. Ainsi, nous avons exclu les voyelles situées à la finale d'une phrase car elles sont souvent marquées d'un abaissement de la puissance du sifflement et nous avons décidé de ne retenir que des voyelles situées dans un intervalle de confiance à 5% de la moyenne des fréquences de chaque bande vocalique. De cette manière les bandes de fréquences des voyelles de l'expérience ne se chevauchaient pas, tout en restant relativement larges.

##### Répartition des fréquences des voyelles sifflées de l'expérience



**Figure 90** : Les voyelles jouées lors de l'expérience ont été choisies parmi celles de Luis (voir Figure 34)

D'autre part, nous avons décidé de ne pas traiter acoustiquement les sons par des filtrages de réduction de bruit afin de rester dans une situation proche des conditions d'écoute d'un siffleur.

#### Variante

Afin de tester l'effet du contexte acoustique de la phrase sifflée sur les sujets, un deuxième corpus a été établi comme suit: pour chaque voyelle du premier corpus nous avons conservé le contexte sifflé précédant son occurrence (2 à 3 secondes). Un corpus de 84 phrases sifflées se terminant par la voyelle dont nous voulions tester la perception a donc été constitué. Sur les 84 sons, 20 ont été réservés à la phase d'apprentissage et 64 à la phase de test.

#### 4.3.1.2.3. Type de tâche proposée:

Pour chaque variante, la tâche principale de l'expérience était la suivante: après l'écoute de chaque son, le sujet désigne la voyelle qu'il estime la plus proche de la voyelle sifflée qu'il vient d'entendre en cliquant sur un des 4 boutons « a », « é », « i », « o ». La tâche était donc du type « choix forcé » parmi quatre solutions.

#### 4.3.1.3. Réalisation de l'expérience:

##### 4.3.1.3.1. Programmation

Le programme graphique de l'expérience a été réalisé grâce au logiciel *Flash 5* permettant l'utilisation du langage de programmation de dessin vectoriel *Actionscript*. De cette manière il a été possible de contrôler les évènements déclenchés par l'utilisateur<sup>99</sup>, d'organiser les listes de présentation suivant un tirage aléatoire non récurrent et d'enregistrer les données dans des fichiers exploitables par un programme de traitement réalisé sous *Matlab*.

##### 4.3.1.3.2. Sujets testés

Chaque variante de l'expérience a été passée par 20 personnes situées dans une tranche d'âge de 19 - 29 ans. Les personnes ayant passé la variante 1 (voyelle seules) étaient différentes de celles ayant passé la variante 2 (phrases). Un audiogramme a été effectué afin de vérifier qu'ils avaient une bonne audition. Les sujets ne devaient pas avoir été en contact avec une langue sifflée et n'avoir jamais entendu d'explication sur le système de répartition des voyelles des langues sifflées sans tons.

##### 4.3.1.3.3. Déroulement de l'expérience:

Chaque variante comprenait trois étapes:

- Première étape: Questionnaire puis explication succincte de l'expérience suivie de l'écoute d'une phrase sifflée permettant au sujet de se familiariser avec les sonorités du sifflement articulé tout en comprenant qu'il s'agit bien d'une langue naturelle (mais la phrase n'a pas été traduite afin de ne pas donner d'indications sur les voyelles).

---

<sup>99</sup> Par exemple certains biais d'usage comme la répétition de clics successifs. Les boutons de contrôles non utiles dans l'interface pour la tâche en cours étaient également désactivés.



## Voyelles sifflées

**Prénom :**

**Etes-vous musicien ?**  **Oui**  **Non**

**Si oui, de quel instrument jouez-vous ?**

Luis est un siffleur d'une île des Canaries.  
Il parle l'espagnol en sifflant pour communiquer à distance dans les montagnes.  
Il copie en sifflement les consonnes et les voyelles.  
Les voyelles "i, é, a, o" de l'espagnol de cette île sont comme celles du français.

**Ecouter un exemple de phrase**

**suite** 

Figure 91 : Questionnaire et familiarisation avec l'expérience

-Deuxième étape: Apprentissage. Le principal objectif de cette étape était de permettre au sujet de se familiariser avec la tâche de test et avec le système vocalique sifflé.

La tâche à effectuer était donc identique à celle du test, seul le nombre de sons présentés changeait. L'étape d'apprentissage comportait 20 sons (5 /i/, 5 /e/, 5 /a/, 5 /o/) se succédant selon un *ordre préétabli* permettant au sujet d'entendre toutes les combinaisons possibles de voyelles successives différentes.<sup>100</sup>

Lors de l'apprentissage de la première variante de l'expérience (voyelles seules) la session de 20 voyelles a été présentée 2 fois de suite de manière à rapprocher le taux d'exposition sonore avant le test de celui de la deuxième variante.

<sup>100</sup> Ce critère a été retenu car un tirage aléatoire sur le nombre réduit de sons de l'apprentissage présentait parfois tous les sons correspondant à une voyelle donnée dans la première partie de l'écoute. Le risque était que les sujets n'aient pas le temps de se familiariser avec le système vocalique sifflé dans son ensemble. Nous ne voulions pas que les sujets confrontés à ces listes aient l'impression de faire toujours la même réponse au début de l'expérience.

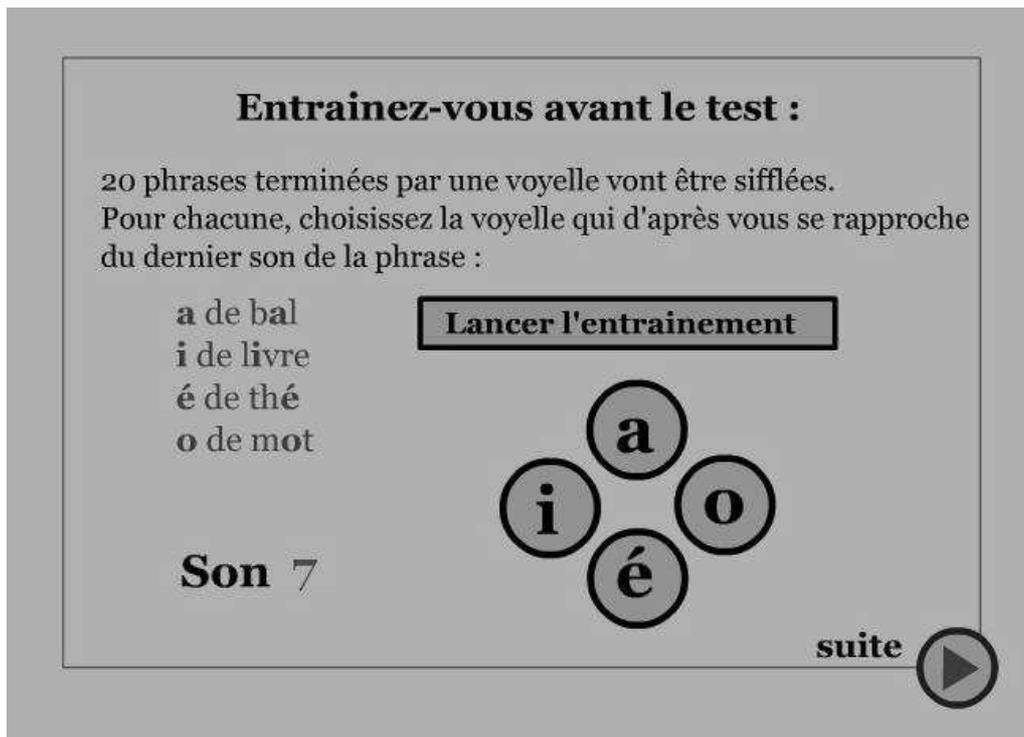


Figure 92 : Capture d'écran du déroulement d'une phase d'apprentissage de la variante 2 de l'expérience

-Troisième étape: Test

Le test proposait la même tâche que pour l'apprentissage, avec 64 voyelles sifflées (respectivement 64 phrases sifflées se terminant par des voyelles variante 2): 16 /i/, 16 /e/, 16 /a/, 16 /o/ se succédant suivant une liste choisie par tirage aléatoire. Quand le test est terminé, le bouton « Quit » permet d'enregistrer le fichier résultat (Figure 93).

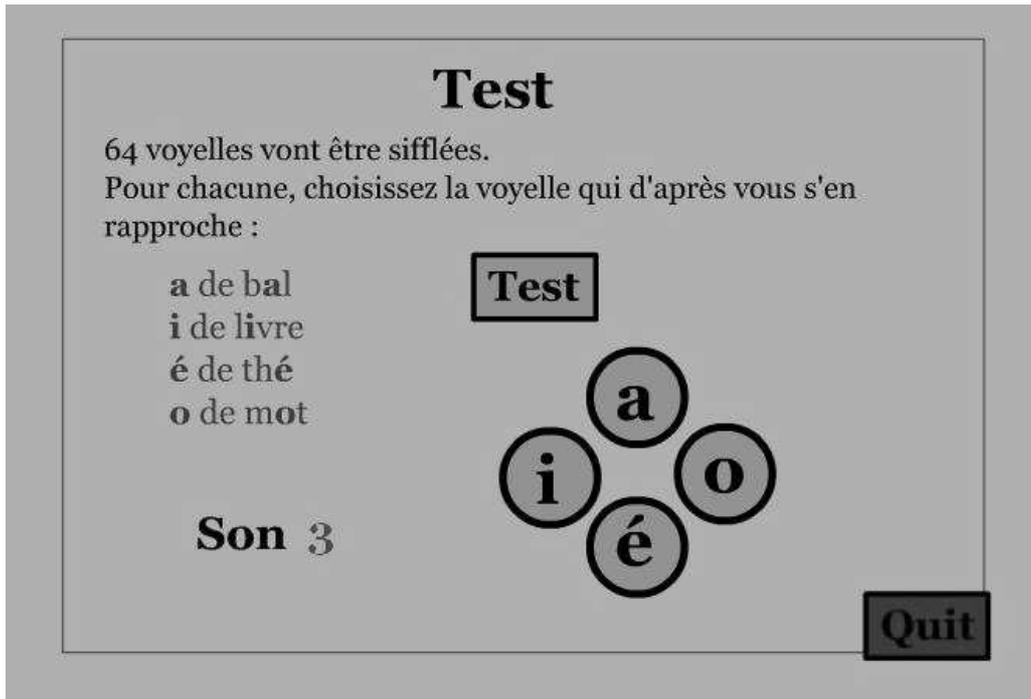


Figure 93 : Capture d'écran du déroulement de la phase de test de la variante 1

#### 4.3.1.4. Résultats

##### 4.3.1.4.1. Généralités sur le traitement des données

Les données collectées par l'interface graphique permettaient de connaître la liste des sons joués et celle des réponses. Ces données ont été analysées pour chaque individu puis ont été mises en commun. Un pré traitement a été effectué sur 5 listes avec *Excell*, puis le traitement de l'ensemble des données a donné lieu au développement d'un programme spécifique sous *Matlab* permettant d'extraire les réponses sous la forme de matrices de résultats et de les représenter sous différentes formes graphiques en réintégrant parfois des informations comme par exemple la répartition en fréquences des voyelles de la Figure 90.

##### 4.3.1.4.2. Résultats de l'expérience de perception des voyelles isolées (variante 1)

#### Tendance générale

##### Réponses justes

Le taux de réussite moyen est de 55%. Il correspond au nombre de réponses justes. Compte tenu du protocole expérimental et de la tâche à effectuer ces résultats sont largement au-dessus de la chance qui est à 25 %. Mais les taux moyens de bonnes réponses varient en fonction des voyelles.

Tableau 28 : Taux de réussite moyen de bonne réponses sur 20 sujets

	/o/	/a/	/e/	/i/
Réponses justes en %	50.63	44.06	46.88	78.44

### Ensemble des réponses

Si l'on considère maintenant l'ensemble des réponses on observe, outre les bonnes réponses, que la majorité des confusions peuvent être qualifiée de *logiques* en ce sens qu'une voyelle est en général confondue avec ses voisines fréquentielles dans 83% des cas de confusion.

Tableau 29 : Matrice de confusion de l'ensemble des réponses des 20 sujets (résultats en %).

Voyelles jouées	Voyelles répondues			
	/o/	/a/	/e/	/i/
/o/	50.63	40.31	7.50	1.56
/a/	13.44	44.06	31.56	10.94
/e/	5.94	22.19	46.88	25.00
/i/	0.00	4.38	17.19	78.44

La représentation graphique de la matrice sous forme de surface ou sur un plan avec un code couleur progressif adapté illustre visuellement ces résultats (Figure 94).

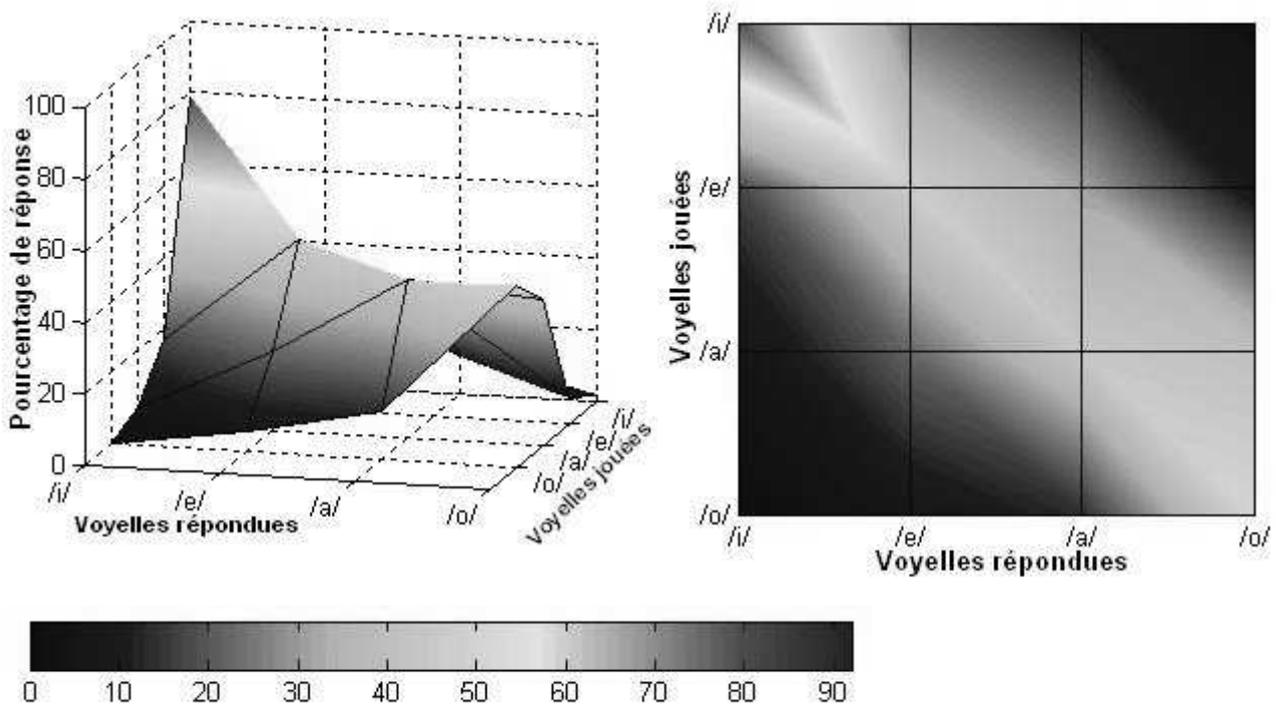


Figure 94 : Performances d'identification des voyelles sifflées seules

Afin de préciser l'influence des fréquences de chaque voyelle sifflée jouée sur les réponses des sujets nous avons également représenté les résultats en fonction de la répartition fréquentielle des voyelles de l'expérience (Figure 95). Sur cette nouvelle figure, nous avons représenté également les courbes de tendances des réponses des sujets. Elles permettent de retrouver une projection indirecte (car moyennée par la régression linéaire au second degré) de la forme en selle de cheval de la surface.

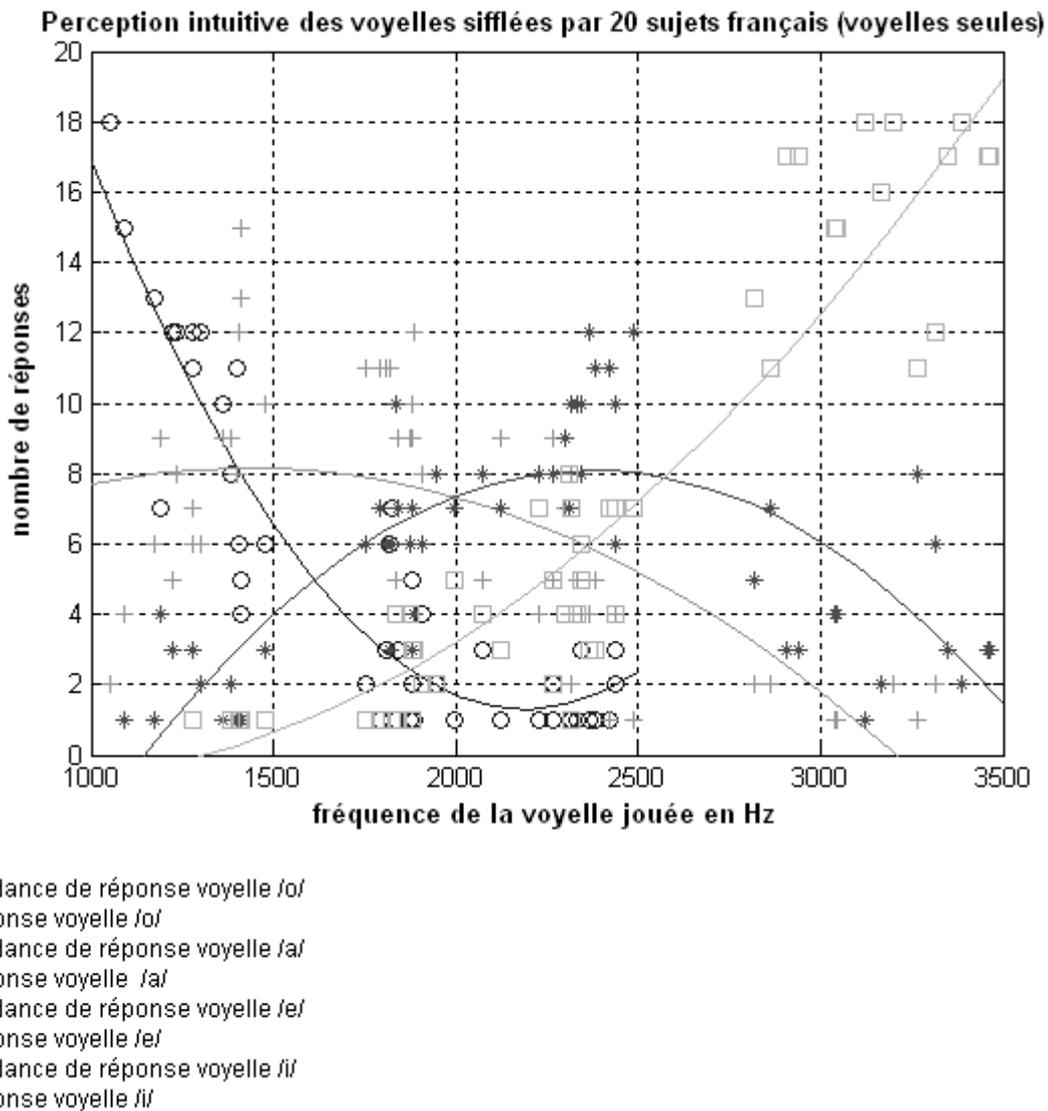


Figure 95 : Répartition des réponses en fonction des voyelles sifflées

## Différences individuelles et confusions fréquentes

### Considérations générales

Les taux de réussites et de confusion distinguent les sujets (les détails des résultats individuels sont disponibles en *Annexe Erreur ! Source du renvoi introuvable.*). En effet, deux d'entre eux se détachent positivement de la moyenne avec un taux de 73,5 %. Un groupe de 6 personnes est au dessus de 40 bonnes réponses sur 64 (62,5%). 4 autres personnes en sont très proches (taux supérieur à 37 sur 64 (58%)). Ce qui veut dire que la moitié des personnes réussit bien le test. Les 10 autres personnes ont toutes un taux de réussite supérieur à 37%. Les 4 moins performantes d'entre elles se situent entre 37 et 40% les 6 autres réussissent mieux le test puisqu'ils obtiennent des performances entre 45 et 54%.

### Détails en fonction des voyelles

En général, les moins performants ont tout de même une matrice de confusion logique. Leur taux relativement faible est bien souvent dû à une confusion systématique entre deux voyelles sifflées voisines au niveau des fréquences. La variabilité des performances des sujets est donc grande suivant le type de voyelle :

- Pour le /i/ la grande majorité des sujets ont de très bons taux de réussite puisque 16 d'entre eux ont un score supérieur à 12 sur 16 et 2 d'entre eux ont 100% de réussite. Le moins performant a un score de 9 sur 16 (56%).
- Pour le /o/, 6 personnes un taux d'identification supérieur à 10 sur 16 (62,5%). Toutes les autres personnes prennent souvent le /o/ pour un /a/.
- Le /a/ est la lettre la moins bien identifiée par les sujets car elle est souvent prise pour un /e/ et assez fréquemment prise pour un /o/.
- Le /e/ est confondu à part égales avec ses voisins sifflés /a/ et /i/

Les basses performances pour le /a/ et le /e/ s'expliquent par le fait qu'il ont chacun deux voisins perceptifs en termes de hauteur ce qui multiplie les possibilités de confusion par rapport au voyelles plus isolées /i/ et /o/. Malgré cela les sujets les plus performants arrivent très bien à les catégoriser comme des voyelles différentes, uniquement à partir de leur fréquence sifflée.

Les confusions les plus fréquentes sont donc les suivantes : le /o/ est souvent pris pour un /a/, le /a/ et le /e/ sont souvent pris réciproquement l'un pour l'autre. Cette dernière confusion s'explique en partie par la proximité des deux intervalles de fréquence de sifflement de ces deux types de voyelles.

### Différences entre musiciens et non musiciens

Parmi les personnes ayant passé le test 6 étaient musiciennes. Les résultats de ce groupe se distinguent significativement de ceux du groupe des non musiciens ( $F(1,18)=6,71, p<0,018$ ). On peut donc dire que les musiciens réussissent mieux la tâche que les non musiciens. Cette différence est clairement visible sur les Figure 96 et Figure 97 (voir aussi les courbes de tendance en fonction de la répartition en fréquence des voyelles jouées en *Annexe Erreur ! Source du renvoi introuvable.*).

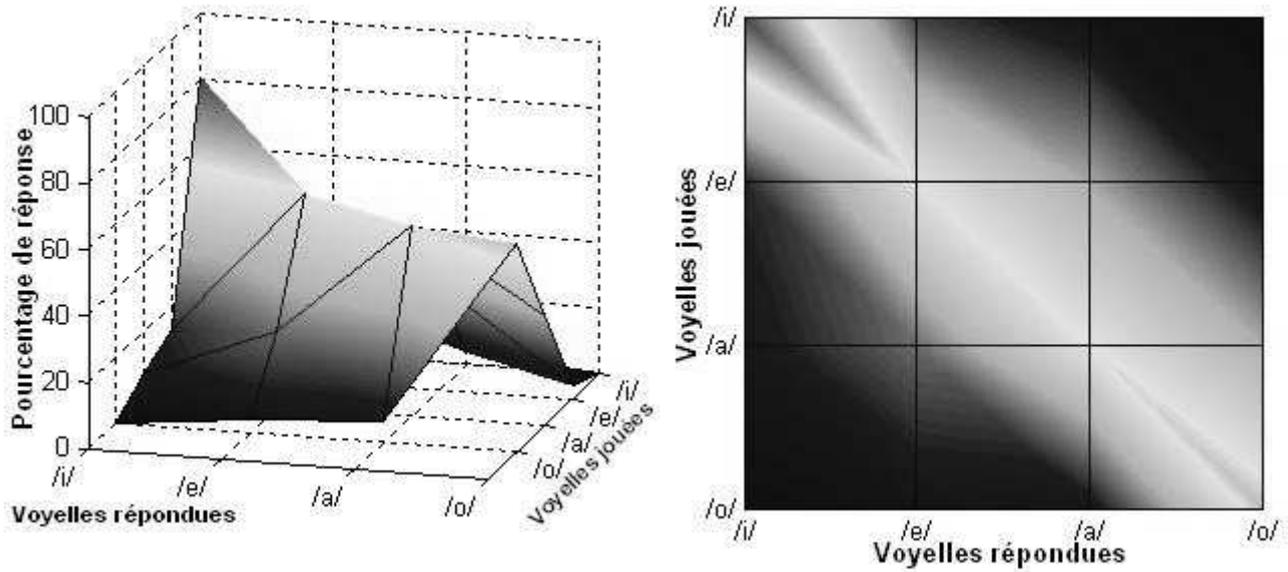


Figure 96 : Performances d'identification des voyelles sifflées seules pour les musiciens

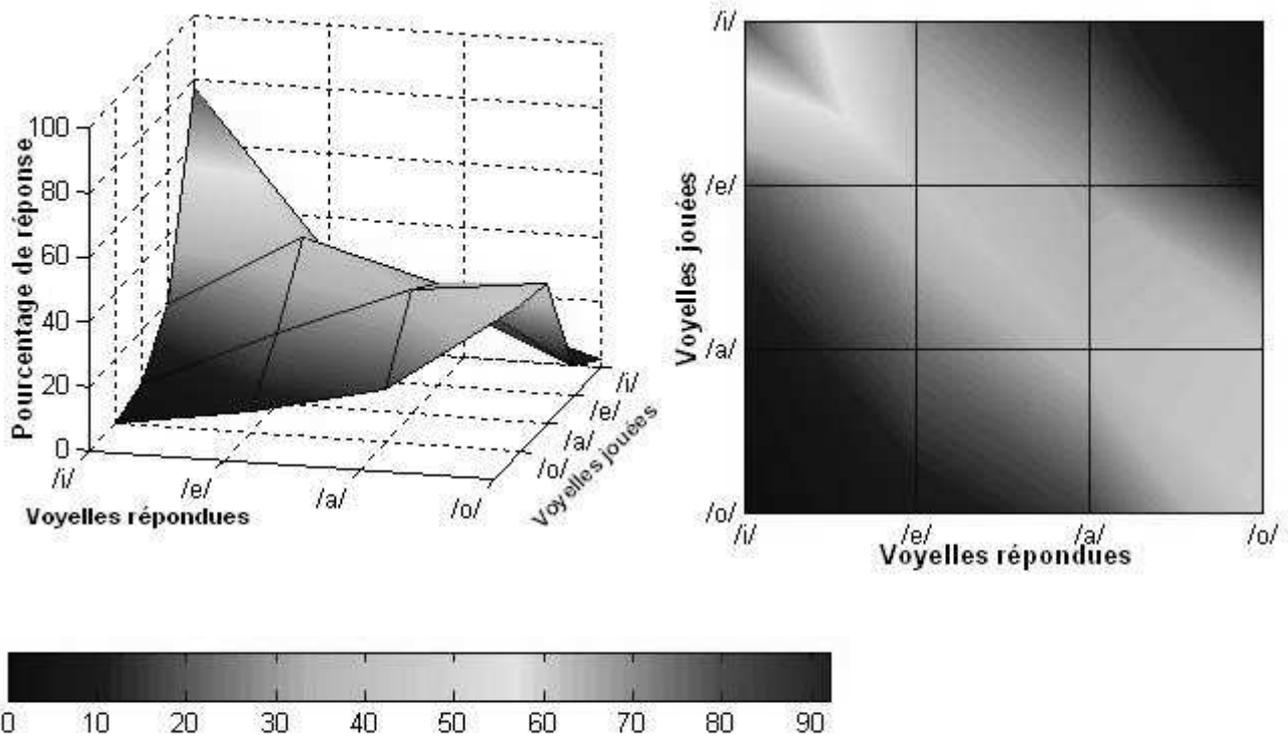


Figure 97 : Performances d'identification des voyelles sifflées seules pour les non musiciens

## Conclusion

Les tendances de bonnes réponses et les tendances de confusions montrent que les sujets réussissent en général bien la tâche. Tous ces indices soutiennent le fait que les sujets français catégorisent les voyelles sifflées « a », « é », « i », « o » comme les siffleurs de la Gomera.

Un biais expérimental a pu être observé, qui tient au fait de présenter des voyelles isolées de tout contexte sonore hormis celui de la voyelle précédemment écoutée. En effet, certains sujets répercutent des confusions sur la réponse suivante. Par exemple si à l'écoute d'un /e/ sifflé ils avaient répondu /a/ et que la voyelle jouée juste après était un /a/ il ont eu tendance à répondre /o/. Par conséquent, on observe un effet de cascade entre confusions logiques qui s'arrête quand un saut de fréquence conséquent intervient. Ce phénomène fausse les résultats des taux de réussite tout en confirmant que des sujets non-siffleurs étagent bien perceptivement les voyelles sur une répartition qui dépend de la fréquence.

Dans ces conditions, il n'est pas surprenant de constater que les musiciens réussissent mieux la tâche que les non-musiciens car ils sont plus habitués à associer les hauteurs perçues de manière isolée à une référence sonore culturellement marquée.

### 4.3.1.4.3. Résultats pour l'expérience de perception des phrases se terminant par une voyelle (variante 2)

Cette expérience a été mise au point afin de vérifier l'effet du contexte sur la perception des voyelles sifflées par des français. En particulier nous avons émis l'hypothèse qu'en nous rapprochant des conditions d'écoute des siffleurs -qui ne perçoivent pas des voyelles seules mais des voyelles sifflées intégrées à un flux sonore sifflé- nous pourrions observer une suppression ou au moins une atténuation de l'effet de répercussion en cascade des confusions logiques. Les voyelles présentées dans cette variante de l'expérience étaient donc précédées de leur contexte sifflé.

### Tendance générale

On observe les mêmes tendances générales que lors de la variante 1 mais avec des performances de réussite à la tâche d'identification de 60,2%. Les voyelles sifflées /o/ et /i/ sont encore mieux identifiées que pour la variante 1 (respectivement 73,13% et 87,81%) alors que les voyelles /a/ et /e/ sont légèrement moins bien identifiées. Ces résultats apparaissent statistiquement et visuellement sur la matrice de confusion du Tableau 30 et sur les représentations graphiques en surface et planes de la Figure 98.

**Tableau 30 : Matrice de confusion de l'identification par 20 sujets français de voyelles sifflées avec contexte**

	Voyelles répondues			
Voyelles jouées	/o/	/a/	/e/	/i/
/o/	73.13	23.13	2.81	0.94
/a/	10.94	39.06	39.38	10.63
/e/	5.00	19.38	40.94	34.69
/i/	0.31	1.56	10.31	87.81

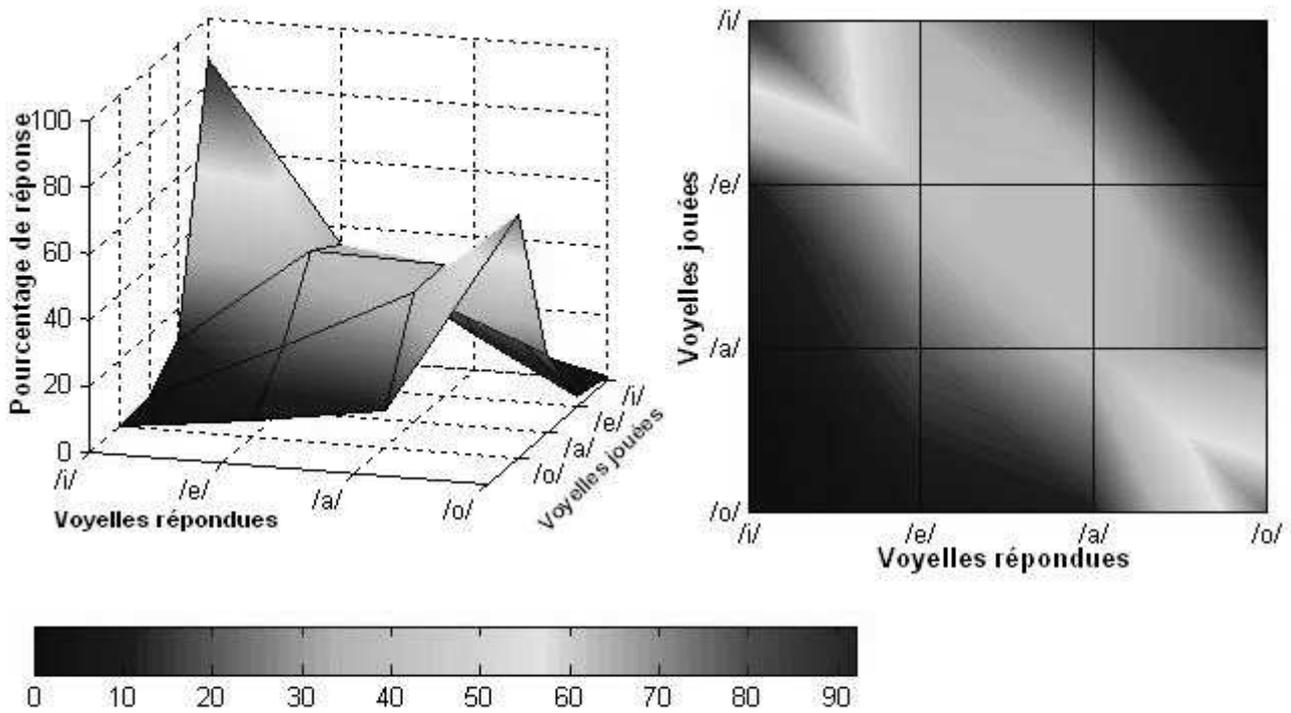


Figure 98 : Performance d'identification des voyelles sifflées avec contexte

Une représentation tenant compte des fréquences des voyelles jouées a été également produite :

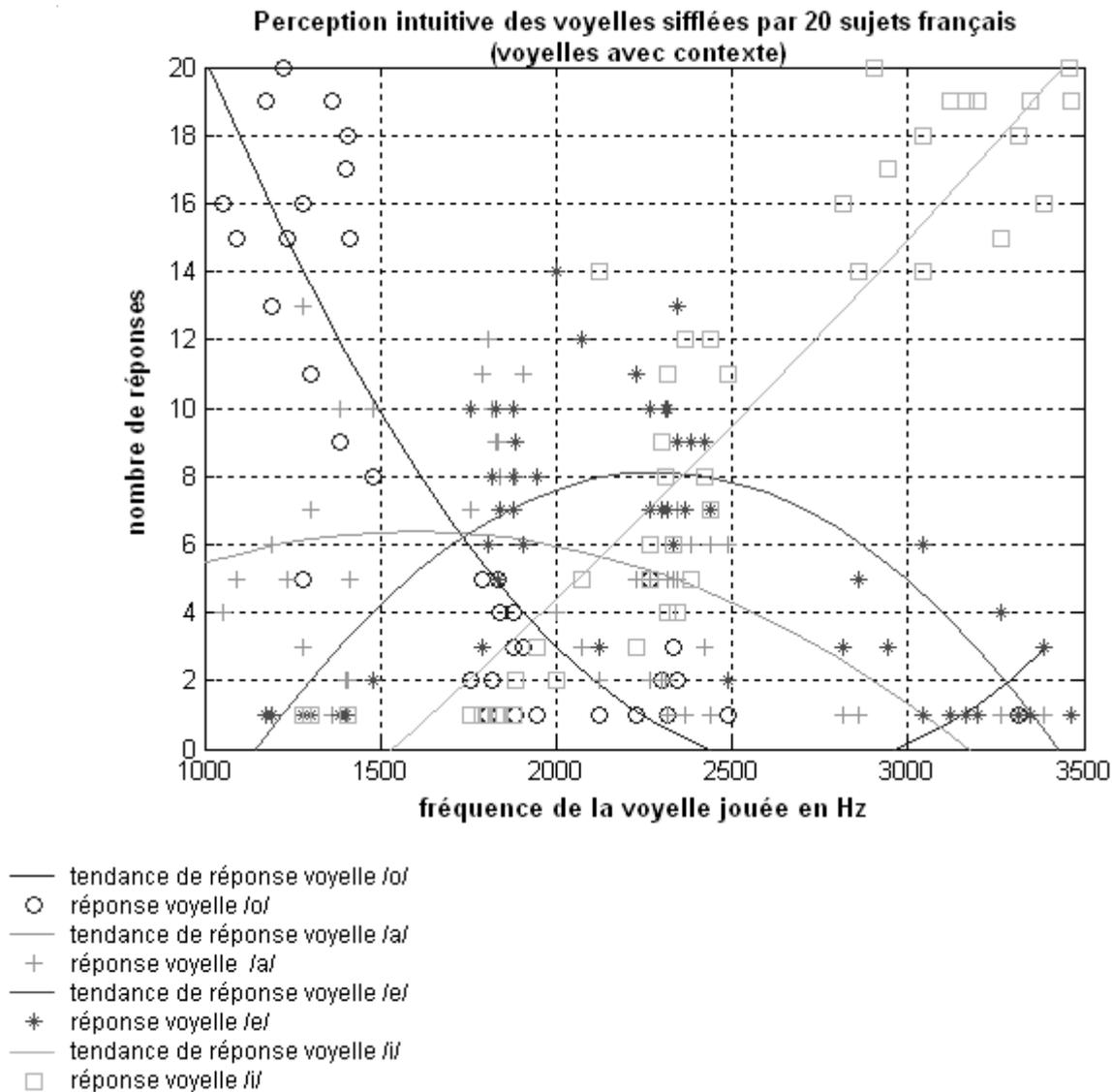


Figure 99 : Répartition des réponses en fonction des fréquences des voyelles sifflées

## Différences individuelles et confusions

8 personnes ont un score de réussite supérieur à 40 bonnes réponses sur 64 (62,5%). 3 d'entre elles ont un score plus élevé que 47 sur 64 (73%). Le meilleur sujet, un saxophoniste ayant appris les notes suivant une technique particulière de repérage des sons par leur localisation dans son corps obtient un score de 75%. A l'opposé, le sujet le moins performant en identification a un score de 46%, ce qui est plus élevé que dans la variante 1.

En ce qui concerne les confusions, l'amélioration des scores de réussite sur le /o/ montrent qu'il est bien moins systématiquement pris pour un /a/, par contre le /a/ est toujours très souvent confondu avec /e/ et réciproquement. Enfin le /e/ est très souvent pris pour un /i/ avec des différences suivant les sujets. C'est souvent au niveau de l'identification du /a/ et du /e/ que se sont faites les différences entre les sujets les plus performants et les moins performants.

Une confusion du /i/ avec un /o/ qui fait remonter notre courbe de tendance des réponses sur les /o/ pourrait être due à l'effet d'ambiguïtés d'octaves sur les sifflements. L'autre explication plus simple est que le sujet n'a pas été attentif lors de la présentation de cette voyelle.

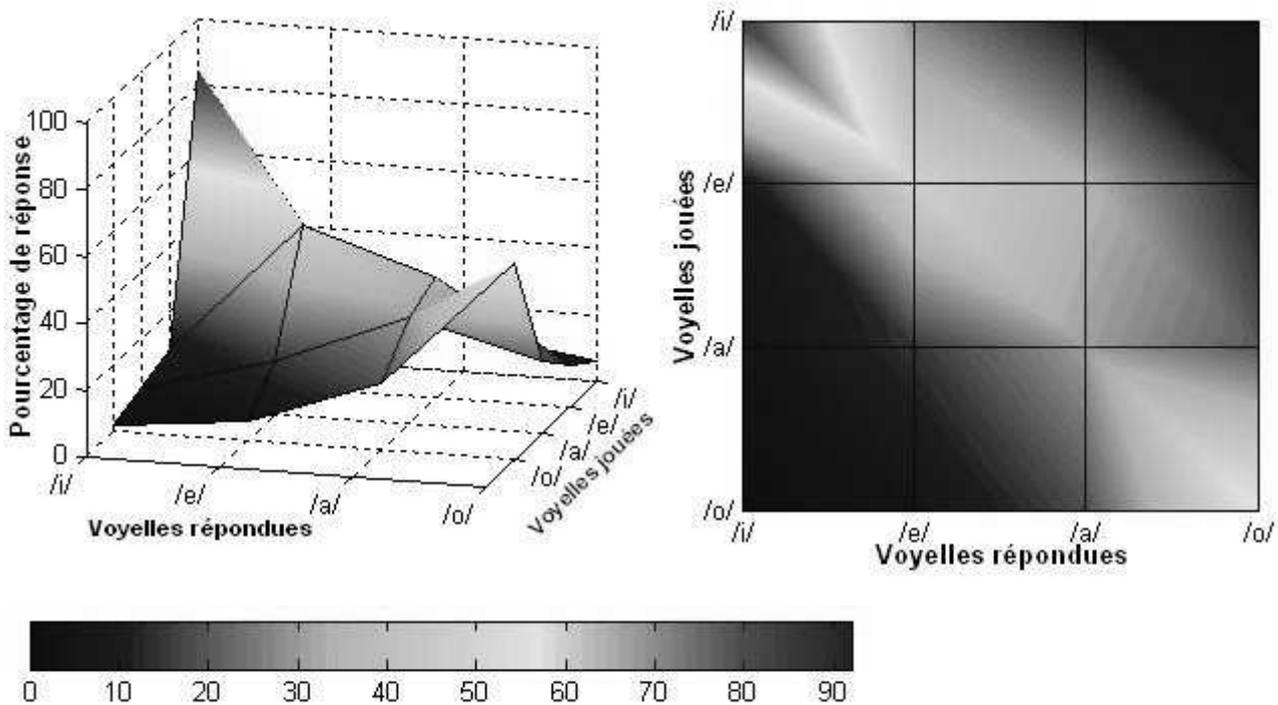
### Différence entre musiciens et non musiciens

A nouveau dans ce cas il y avait 6 musiciens parmi les sujets (et ce malgré le fait que les sujets ayant passé la variante 2 ne soient pas les mêmes que ceux de la variante 1). Un calcul d'analyse de variance similaire à celui réalisé lors de la variante 1 de l'expérience a montré que cette fois ci les résultats des musiciens n'étaient pas significativement différents de ceux des non musiciens ( $F(1,18)=6,71$ , NS). L'effet de contexte semble avoir rapproché les non musiciens des musiciens car les résultats de ces derniers ne sont pas moins élevés que pour la variante 1 (l'ensemble des représentations graphiques de ces résultats est en annexe **Erreur ! Source du renvoi introuvable.**).

### Effet d'apprentissage

En raison de l'élimination des confusions favorisées dans la variante 1 par la présentation de voyelles isolées successives, il est pertinent de tester l'effet d'apprentissage dans cette variante.

	Voyelles répondues				
Voyelles	/o/	/a/	/e/	/i/	
jouées	/o/	59.00	22.00	13.00	6.00
	/a/	20.00	32.00	36.00	12.00
	/e/	6.00	16.00	49.00	29.00
	/i/	2.00	5.00	9.00	84.00



**Figure 100 : Performance d'identification des voyelles sifflées avec contexte en phase d'apprentissage**

On remarque que dès l'apprentissage les tendances observées dans le test sont affirmées. Etant donné le peu de voyelles de la phase d'apprentissage, ceci suggère que les sujets s'appuient sur des catégorisations déjà présentes dans leurs habitudes linguistiques.

### **4.3.1.5. Discussion générale**

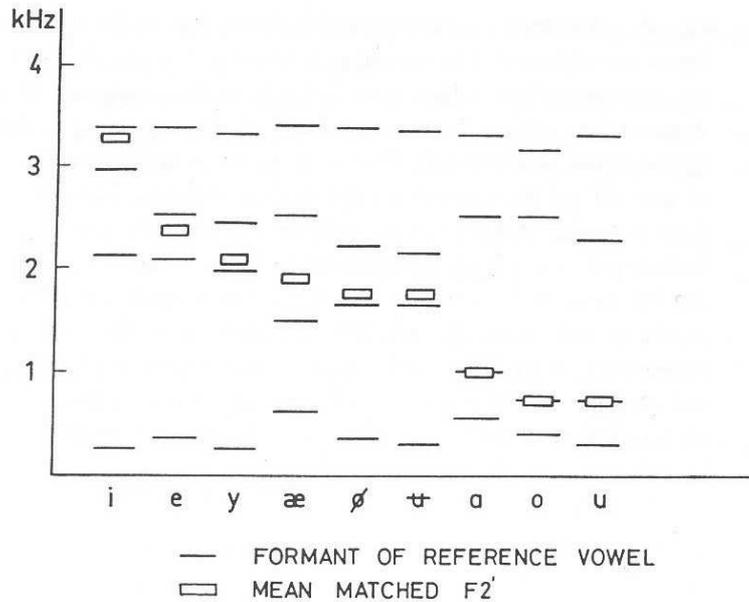
#### **4.3.1.5.1. Première interprétation**

Les résultats des deux variantes de cette expérience d'identification montrent que des sujets ignorants tout de la répartition des voyelles sifflées du Silbo espagnol et des langues sifflées en général mais ayant dans leur langue parlée le même type de voyelles que celles de l'espagnol arrivent à catégoriser de manière intuitive des voyelles sifflées du Silbo qu'ils écoutent pour la première fois. La répartition de leurs réponses est similaire à la représentation cognitive des siffleurs. Le fait que cette tendance soit déjà présente à l'apprentissage et que ce dernier ait été relativement court semble indiquer que la représentation perceptuelle des voyelles de manière étagée en fonction de la fréquence pré-existe chez les sujets testés. De plus, ces résultats semblent confirmer que les siffleurs s'appuient sur une réalité perceptive présente dans la version parlée pour transposer les voyelles en hauteurs sifflées.

#### **4.3.1.5.2. Positionnement de nos résultats dans le cadre de la phonétique et de la perception des voyelles**

Le mécanisme de la perception des voyelles a été testé par de nombreux expérimentateurs grâce à des tâches d'identification, de discrimination ou de *matching* (*qui consiste à appairer*) de stimuli. Celles qui obtiennent les résultats qui se rapprochent le plus de nos observations sont les expériences menées à partir d'un « *effective upper formant* » (Bladon et Fant, 1978 p.1) appelé F2'. F2' est la dérivée du F2 à des degrés variables de manière à tenir compte de la contribution des formants plus élevés en fréquence. Ce formant est donc considéré comme l'intégration perceptuelle de l'ensemble des formants élevés (au dessus du formant 1(F1)). Sur cette base, différentes équipes de recherche ont demandé à des sujets d'appairer des voyelles issues du suédois puis d'autres langues à une approximation synthétique à deux formants constituée du F1 et du F2' où le F2' était ajustable par les sujets eux-mêmes (Carlson et al., 1970). La justification d'une telle approche tient dans la remarque de Bladon et Fant : "*The notion that a vowel quality can be satisfactorily approximated by an acoustically-derived representation in two dimensions has held a long standing attraction for phoneticians, in the hope that such acoustic data might provide correlates for the articulatory dimensions of tongue height and tongue retraction which form the basis of conventional phonetic vowel quality diagram. More recently, perceptual studies have supported the view that a two formant model of a vowel is also a valid representation at some level (not necessarily peripheral) of auditory processing.*" (Bladon et Fant, 1978, p.1). En effet, d'une part, le travail de Plomp et al (1975) avait montré que l'identification des voyelles du hollandais issues d'une représentation spectrale extraite par une analyse en composantes principales (PCA pour Principal Component Analysis) correspondait de manière très proche aux scores d'identification des mêmes voyelles dont seulement le F1 et le F2 étaient disponibles (plan logF1

versus logF2 utilisé en phonétique). D'autre part – et c'est l'aspect qui nous intéresse le plus- l'expérience préalable de Carlson et al (1970) revenait à examiner les conséquences perceptuelles de la manipulation de l'espace entre deux prééminences spectrales dans des voyelles synthétiques. Les résultats montraient que pour des formants F1 et F2 relativement proches (de 3 à 3,5 Bark<sup>101</sup>) les sujets plaçaient F2' au niveau de F2 réalisant ainsi une intégration sur le pic de densité spectrale le plus intense (voyelles /a/, /o/, /u/ Figure 101). Carlson et al (1970) avaient également remarqué que pour des voyelles où les formants F2 et F3 étaient proches, les sujets plaçaient le F2' entre ces deux pics (typiquement voyelle /e/ Figure 101). Enfin le /i/ du suédois était un cas à part dans leur expérience puisque les sujets plaçaient le F2' entre F3 et F4.



**Figure 101 : Résultats de l'expérience de Carlson et al (1970) (in Stevens, 1998, p. 240)**

Plus tard, Chistovitch et son équipe (Chistovitch et Lublinskaya 1979 ; Chistovitch et al 1979 ; Chistovitch, 1985) ont montré que ce phénomène *d'intégration perceptuelle* apparaît à partir d'une proximité de 3,5 bark entre deux formants, quels qu'ils soient.

Bladon et Fant (1978) firent la même expérience que Carlson et al (1970) mais à partir d'un corpus de voyelles cardinales<sup>102</sup>. Leurs résultats sont similaires, ils sont présentés Figure 102<sup>103</sup>.

<sup>101</sup> 300 à 400 hz environ.

<sup>102</sup> Voyelles de l'A.P.I. (Association Phonétique Internationale)

<sup>103</sup> Dans cette étude les auteurs ont perfectionné une formule déjà établie en 1970, modifiée en 1975 (Carlson, Fant et Granström, 1975) permettant d'évaluer la valeur du F2' par un calcul, ce qui explique la présence d'un "F2' calculated" sur la Figure 102. Nous ne nous intéressons pas ici à cet aspect de leur approche.

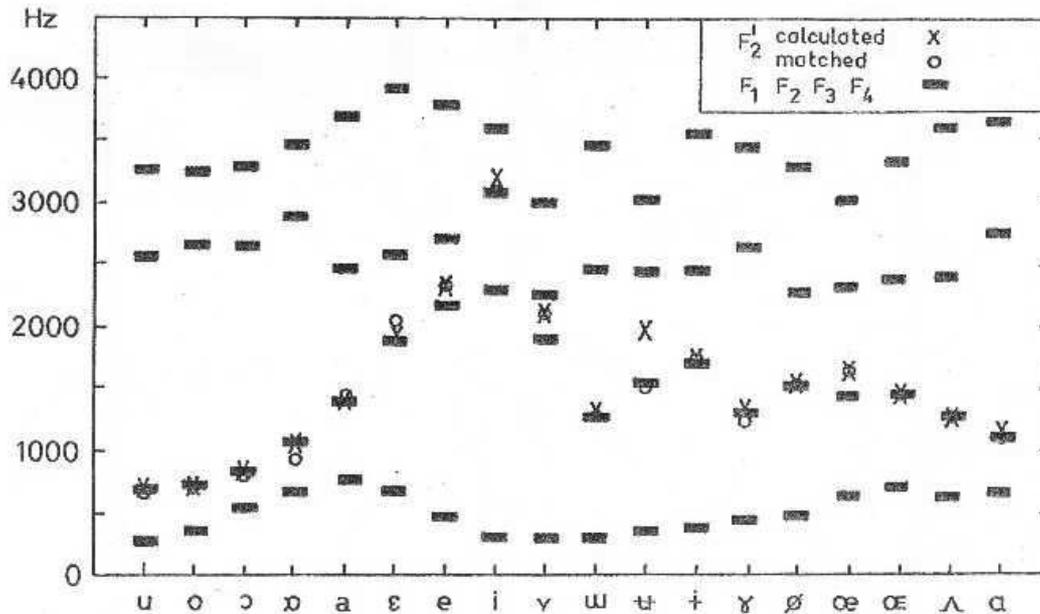


Figure 102 : Présentation séquentielle des voyelles entremes de formants F1, F2, F3, F4 mesurés et de F2' calculés et perceptuellement appariés (*matched*) (Bladon et Fant, 1978 p.9)

### Timbre, F2' et voyelles sifflées

La répartition du *effective formant* F2' nous semble particulièrement adaptée à une comparaison avec la stratégie de transposition des voyelles en langues non tonales sifflées. En effet on retrouve la distinction claire entre /i/ et les autres voyelles signalées dans la partie typologie pour le tuc, le grec et l'espagnol sifflé, mais également le groupement des voyelles postérieures et des voyelles centrales dans deux catégories différentes. De plus, elle fournit une explication plus plausible que celle d'une transposition du F2 seul.

En effet, en soulignant le rôle de l'intégration perceptuelle de deux formants de la parole en un seul F2' ces recherches mettent en évidence l'importance des zones de forte compacité dans la perception du timbre des voyelles et, d'après Stevens, le rôle de l'espacement critique mis en évidence par les expériences que nous avons citées, a des conséquences générales importantes pour la classification des voyelles : « *some aspects of the auditory system undergoes a qualitative change when the spacing between two spectral prominences becomes less than a critical value of 3,5 bark* » (Stevens 1998, p.241). Il a illustré son propos en situant l'emplacement des F2' sur une analyse spectrale de quelques voyelles (Figure 103). On peut y voir que la largeur de la bande de fréquence des deux formants les plus proches joue un rôle important. A l'intérieur de ce regroupement de fréquences, la profondeur et la largeur des vallées spectrales entre les formants sont relativement limitées. Par contre, les vallées spectrales sont plus profondes et larges autour. Ces conditions facilitent l'émergence de propriétés perceptives particulières liées à ces qualités du timbre des voyelles.

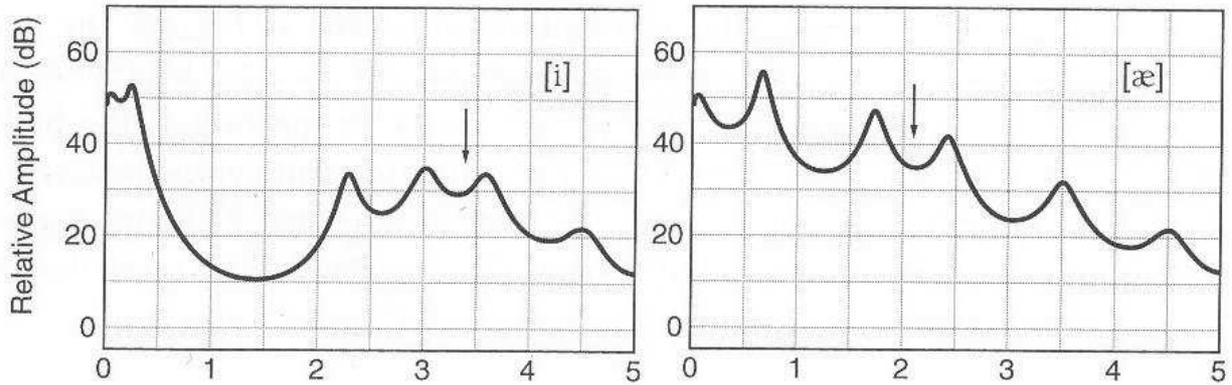


Figure 103 : Localisation du F2' (flèche) sur l'analyse spectrale de 2 types de voyelles (Stevens, 1998. p. 289)

Le /i/ est marqué par une concentration de fréquences élevées, son timbre peut être qualifié de brillant ou « aigu », le /e/ est intermédiaire (comme le [æ] de la Figure 103) et le /a/ ou /o/ ont des timbres parfois qualifiés de « grave ». Ces termes empruntés à l'acoustique musicale correspondent bien à la réalité des sifflements. Le lien fait sans grande difficulté par les sujets de notre expérience entre les voyelles sifflées et les quatre voyelles types du français nous fait penser qu'ils ont utilisé ce paramètre du prototype de voyelle qu'ils avaient en mémoire.

#### Remarques complémentaires :

Il semble que le même type de résultats ne puisse pas être obtenu avec des sujets de langue maternelle tonale, en effet, lorsque nous avons fait passer l'expérience à un étudiant arrivant du Mali et de langue maternelle Bambara, nous avons constaté qu'il n'avait pas du tout la même logique de réponse, malgré le fait qu'il ait fait tout son cursus scolaire en langue française. Il se peut qu'il ait donné la priorité à la Hauteur Fondamentale des voyelles types du français alors que les sujets Français donnaient la priorité à la Hauteur Brute marquée par la compacité du timbre. Ce type de différence de stratégie n'est pas visible dans la version parlée car les deux qualités de Hauteur perçue sont alors présentes dans la voix. Il permettrait de montrer qu'en fonction de sa culture le sujet écoute prioritairement des zones fréquentielles différentes de la voyelle. Son écoute serait donc culturellement orientée. Pour vérifier cette hypothèse il faut concevoir des expériences complémentaires.

#### 4.3.1.6. Conclusion

Les résultats de l'expérience de psycholinguistique que nous avons présentés dans cette section ont été l'occasion de mener une réflexion sur la perception du timbre des voyelles car nous avons trouvé que des sujets français arrivent à avoir, sans exposition préalable, des résultats d'identification de 4 voyelles sifflées du Silbo espagnol montrant une catégorisation des voyelles sifflées similaire à celle des siffleurs eux-mêmes. Les langues sifflées définissent un bon modèle d'analyse pour poursuivre les recherches sur les regroupements perceptifs de formants. A partir de résultats issus de recherches de référence en phonétique nous avons développé une argumentation soutenant que la catégorisation des voyelles sifflées de nos sujets

s'est appuyée sur une référence implicite aux bandes de fréquences harmoniques les plus compactes et intenses des voyelles de référence du français parlé.

Comme les siffleurs transposent la voix parlée espagnole pour siffler ces voyelles et qu'une pratique linguistique est un équilibre entre perception et production, il se peut qu'ils utilisent aussi cet indice de compacité du timbre lors de l'écoute. Actuellement, nous ne pouvons pas conclure catégoriquement en ce sens. Pour pouvoir l'affirmer de manière plus sûre, il faudrait faire des tests psycholinguistiques complémentaires avec des siffleurs de plusieurs pays. Mais si c'était le cas, les langues sifflées comme le grec ou le turc qui se limitent souvent à un octave entre la fréquence du /o/ et celle du /i/ permettraient de préciser certains aspects de la réduction phonétique vocalique. En effet, le F2' varie sur plus de 2 octaves, il y aurait donc une concentration fréquentielle de l'information d'origine.

### **4.3.2. Analyse de la perception de non-mots et de logatomes par des siffleurs**

#### **4.3.2.1. Expérience de perception des logatomes turcs**

Moles (1970) a étudié la perception de 600 logatomes<sup>104</sup> turcs sifflés par 5 siffleurs. Avant d'analyser les résultats de cette expérience il est nécessaire de rappeler que les habitants de cette zone isolée, comme tous les siffleurs en général, sont très peu familiarisés avec ce type d'exercice de reconnaissance d'éléments sans sémantique. Tous les individus testés étaient illettrés (com. pers. Busnel 2005) et donc pour eux le découpage en lettres indépendantes est une abstraction qui n'a pas de sens. C'est pourquoi, même si la segmentation en syllabes ne leur pose pas de problème technique, le fait même de devoir siffler des éléments n'ayant pas de signification leur paraît incongru et tout à fait inutile.

#### **Performances de reconnaissance des logatomes**

Le tableau de résultats disponible dans l'étude publiée par Moles (1970) et que nous reproduisons ci-dessous concerne la perception de logatomes monosyllabes de la forme Consonne-Voyelle. Or ce type de présentation n'est pas celui permettant d'obtenir les meilleurs taux de reconnaissance en raison du fait que la consonne perd certains aspects temporels à l'initiale. L'auteur remarque à ce propos que d'autres types de combinaisons ont été essayées comme les ordres VC et V1CV1. Il signale que l'ordre VC permet d'obtenir les mêmes scores que l'ordre CV, mais que l'ordre V1CV1 permet d'augmenter les performances de reconnaissance de 21% à 32%

---

<sup>104</sup> non- mots suivant des règles mises au point pour les besoins de la téléphonie

Tableau 31 : Résultats des scores de performance de reconnaissance de logatomes (Moles 1970, p. 85)

Siffleurs (% moyennes)	Adil (22%)	Bebek	Ali (25%)	Osman (20%)	Sadik (22%)
Adil (16,6%)	–	16,25	22	12,5	
Bebek	22,5	–			
Ali (20%)	19		–	17,50	24
Osman (22%)	23,75		21,25	–	21
Sadik (31%)			32,5	30	–

D'autre part, les résultats qu'il présente (Tableau 31) permettent de faire émerger les aptitudes des émetteurs (moyennes par lignes) et des récepteurs (moyennes par colonnes) et de conclure qu'il y a des différences notables entre individus. Moles a également observé, lors d'une autre expérience sur des logatomes que chaque récepteur possédait des combinaisons favorites de syllabes et qu'il reconstruisait ses mots ou ses phrases en s'appuyant sur des éléments qu'il avait véritablement eu le temps de percevoir. A ces derniers, il ajoutait ses éléments préférentiels. Il cite par exemple « *les transformations systématiques de « gel » en « çer » par Saik et la prééminence de « çok » chez Osman* » (Moles 1970, p.98).

## Perception des voyelles

### Matrice de confusion

Les tests de netteté des logatomes réalisés par Moles étaient particulièrement adaptés à l'étude de la perception des voyelles. Une analyse a été fournie à ce propos. Les résultats ont été présentés sous la forme d'une matrice de confusion que nous reproduisons (Tableau 32). L'auteur a réalisé l'analyse d'un plus grand nombre de matrices non publiées : « *Quelques remarques générales émergent de celles-ci : D'abord la relative stabilité du i chez certains siffleurs, ensuite les transformations sélectives que le e donne en ö, le ay en oy, et le u en ö* » (Moles, 1970, p.100).

**Tableau 32 : Matrice de confusion (couple Sadik-Osman) performances de reconnaissance des voyelles et des diphtongues sifflées turques.**

« les matrices de confusion sont faites en reportant sur les bords verticaux et horizontaux de la matrice, par exemple, (respectivement) la série des phonèmes émis et la série des phonèmes reçus » (Moles 1970 p.100)

	i	E	A	Ü	I	U	Ö	AY	EY	OY	UY	O
i	9											
E		4					7					1
A			2	2		1	1			1		6
Ü				3	1		6					4
I					3	1		1				
U	3			1	2	4	1					
Ö		1			1	1	4	1				4
AY		1					1	5	1	6		
EY	1	1							3	1		1
OY			1					1		2	2	1
UY	4	5			1		1		1		1	
O							1					4

### Comparaison avec notre analyse de la répartition des voyelles

Toutes les confusions citées ci dessus sont cohérentes avec les résultats statistiques que nous avons obtenus dans la la partie Typologie à propos des voyelles turques..

A partir de l'unique matrice à notre disposition qui concerne ce couple émetteur-récepteur, nous pouvons également remarquer un certain nombre d'autres tendances :

-Une forte confusion de « a » et « o ».

-Le uy est fortement confondu avec soit i soit e.

-Une seule confusion n'est pas très cohérente avec nos données il s'agit de la confusion de ü en ö (valeur de 6 items) alors que « ö » n'est pas confondu avec « ü » mais plutôt avec « o », ce qui est cohérent avec nos résultats. Nous avons cherché à comprendre pourquoi en observant les autres confusions :

Les autres voyelles prises pour un « ö » sont « e » et « o » lui-même. Ces deux confusions sont également cohérentes avec nos résultats. De son côté « ü » est légèrement confondu avec « a » et « u » et lui-même. Ces dernières confusions bien que moins marquées que celles de « ö » semblent indiquer que la différence entre nos données statistiques et les résultats de Moles viennent de la fréquence de sifflement du « ü ». Comme il existe des différences d'habitude d'élocution entre les siffleurs, il se peut que la fréquence de sifflement du « ü » de Sadik soit plus basse que celles des 3 siffleurs à partir desquels nous avons fait nos calculs.

## Conclusion

Les siffleurs turcs obtiennent des taux d'identification de 20 à 50% de logatomes issus de leur langue (performances variant en fonction du type de logatome). Des matrices de confusion de la perception des voyelles ont été extraites de ces résultats. Elles confirment que des sifflements proches en fréquences entraînent des confusions perceptives entre les voyelles correspondantes.

### 4.3.3. Intelligibilité des mots sifflés

Pour l'analyse de l'intelligibilité des mots sifflés nous nous appuyerons de nouveau sur les études de Busnel (1970) et Moles (1970) réalisées en Turquie. Busnel a mené une étude détaillée sur l'intelligibilité des mots et Moles a analysé l'impact de la fréquence des mots en constituant un début de dictionnaire sifflé.

#### 4.3.3.1. Résultats d'une expérience de reconnaissance de mots sifflés

Dans le Tableau 33 nous avons regroupé les données issues de deux tableaux de résultats d'une même expérience d'intelligibilité des mots réalisée par Busnel (1970). Le taux élevé de reconnaissance des mots émis avec la voix parlée donne une bonne référence de la difficulté de la tâche qui puisait les mots dans une liste favorisant parfois les confusions (voir dernière ligne du tableau pour des précisions sur les ambiguïtés intentionnelles du corpus)

**Tableau 33 : Résultats de reconnaissance des mots parlés et sifflés issus d'une liste favorisant l'ambiguïté (tableau que nous avons constitué à partir des données de différents tableaux de Busnel (1970), p. 46 et 49).**

Nom du receveur	Voix Parlée		Voix sifflée	
	Nombre de mots écoutés	Réponse positive Réponse négative	Nombre de mots écoutés	Réponse positive Réponse négative
Ali Cirit (m)	50	44 6	45	31 14
Osman Cirit (m)	47	45 2	45	28 17
Sadik Bebek (m)	46	44 2	46	35 11
Adil Cindik (m)	45	47 0	45	30 15
Mehmet Kosek (m)	47	47 0	45	32 13
Récapitulation globale	235	227 (96,6%) 10 (3,4%)	226	156 (69%) 70 (31%)
Exemple de mots de la liste:	Bana, sana, ormana, osmana, kapici, yapici, bakalim, yakalim, parali, yarali, yuzonbir, yuzonbin			

Les taux de reconnaissance des mots sifflés isolés restent relativement élevés (69%) mais on observe immédiatement une différence avec les 96,6 % de réussite de la voix parlée qui confirme qu'à courte distance et sans contexte lexical cette dernière reste la plus performante.

### **4.3.3.2. Eclairage par analyse du contexte lexical**

#### **4.3.3.2.1. Cadre de l'étude**

Miller et al (1951) ont trouvé que les performances d'intelligibilité ne dépendaient pas uniquement des données inhérentes au matériel sonore utilisé mais également à d'autres facteurs comme la connaissance du sujet traité dans la phrase, ou la connaissance a priori du contenu des stimulus dans lequel l'expérimentateur puise. Leurs résultats montrent que les taux de reconnaissance de mots dépendent fortement de la taille de l'échantillon de test. Ainsi, ils ont obtenu des taux 60% plus élevés pour un matériel sonore constitué de 8 mots stimuli par rapport à un autre constitué de 256 mots stimuli. Ils en ont déduit que les auditeurs avaient besoin de moins d'information acoustique dans le cas d'un petit nombre d'alternatives.

Afin d'approfondir l'analyse de l'interaction entre les performances d'intelligibilité et ce type d'informations contextuelles, plusieurs types d'approches générales sont possibles. Elles visent à quantifier l'effet de l'étendue de l'échantillon testé:

-La première et la plus souvent utilisée repose sur la constitution d'un dictionnaire fréquentiel des mots, des syllabes ou des phonèmes de la langue. Ces paramètres sont importants car ils indiquent d'une part l'homogénéité du vocabulaire et des indices phonémiques disponibles, d'autre part, ils permettent de « *prendre la température du langage. C'est à dire une notion statistique liée à l'effort intellectuel requis pour utiliser des mots de plus en plus spécifiques donc de plus en plus rares.* »(Moles 1970. p. 73)

-La deuxième approche qui peut compléter la première, tient compte de l'enchaînement relatif des mots et des phonèmes dans une phrase. Elle s'appuie sur la théorie de l'information (Shannon 1948). Elle utilise l'estimation de l'entropie linguistique<sup>105</sup> (Shannon 1951) d'un ensemble d'éléments pour quantifier l'effet du contexte.

-Enfin, d'autres approches statistiques se sont développées sur la base de ces deux premières. Ces modèles permettent de quantifier les effets du contexte à partir des relations obtenues entre les performances d'intelligibilité des sujets et les éléments constitutifs (phonèmes ou mots) du signal de parole hors contexte ou en contexte (Boothroyd 1978, Bosman 1989). Ils s'appuient sur des algorithmes statistiques développés pour la reconnaissance des mots (Boothroyd 1968).

---

<sup>105</sup> L'entropie linguistique, qui est une mesure de la densité d'information dans un flot de parole, décroît quand le nombre d'éléments d'une phrase augmente. De telles estimations sont obtenues en analysant les réponses des sujets confrontés à la reconstitution de phrases dont ils n'ont qu'une présentation partielle (Taylor 1953).

#### 4.3.3.2.2. Application aux langues sifflées

Seule la première des trois approches a été appliquée aux langues sifflées. Nous présentons le résultat de cette analyse du dictionnaire fréquentiel dans les paragraphes suivants.

#### **De rares analyses du dictionnaire des langues sifflées**

Notre souci de recréer des ambiances d'usage traditionnel de la langue sifflée fait que notre corpus est un bon révélateur des centres d'intérêts des siffleurs mais le peu de temps passé dans chaque lieu (deux mois au maximum) explique pourquoi il n'est pas suffisant une étude statistique des fréquences d'usage des mots. C'est pourquoi, nous nous appuyons ci-dessous sur les résultats d'autres études.

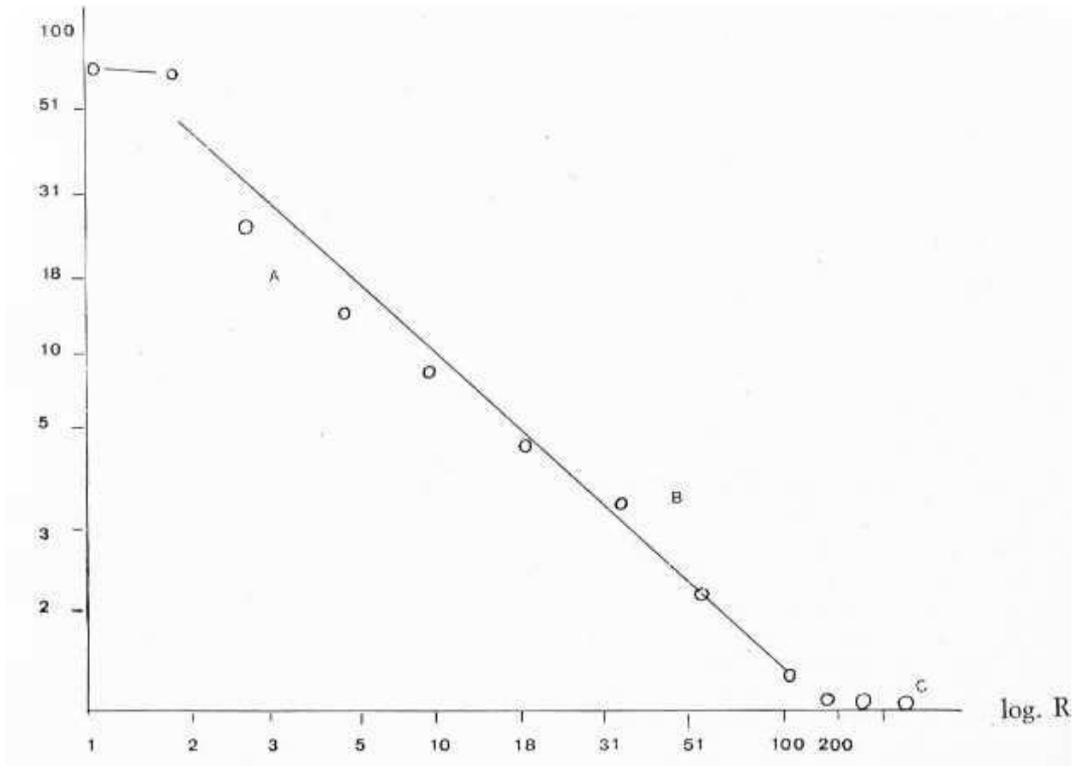
#### Constitution de dictionnaire

La tâche de constitution d'un dictionnaire est facilitée par le fait que les siffleurs utilisent un vocabulaire sifflé correspondant aux mots du quotidien. En outre, la grammaire de la parole sifflée n'est pas différente de celle de la langue locale, il n'y a pas de syntaxe spécifique. Le dictionnaire d'un langage sifflé n'a été évalué qu'en deux lieux: dans la région de Kuskoy en Turquie (Moles 1970) et dans l'île de la Gomera aux Canaries. Les estimations varient suivant les sources mais l'étendue lexicale du sifflement porte au moins sur 2000 mots. Aux Canaries une dernière estimation réalisée par les siffleurs dans le cadre de la revitalisation de la langue recense plus de 4000 mots très courants. Tout autre mot de la langue est toutefois susceptible d'être transposé en sifflements et intégré à ce vocabulaire au bout d'un certain temps d'usage. C'est par exemple le cas du mot turc signifiant « camion » qui n'était presque pas utilisé dans les années 70 lors de l'expédition organisée par Busnel et qui était d'un usage courant lors de notre passage.

#### Analyse du dictionnaire

Pour les raisons exposées plus haut une seule étude approfondie a été menée jusqu'à aujourd'hui pour évaluer l'impact des effets de contexte dans une langue sifflée. Ces résultats sont à notre avis révélateurs de l'ensemble des situations même si l'auteur de l'étude, Moles, indique qu'elle n'est qu'une approche du problème en raison du peu de temps passé sur place à collecter des données et du corpus de mots sifflés limité au 120 mots sifflés les plus courants.

*« La liste ainsi établie qui émergeait d'un corpus d'environ 1000 mots, comportait environ 120 types qui ont été sifflés sur ordre accompagnés de leur traduction en turc parlé[...] Nous avons enfin étudié la répartition des fréquences, tirées de ce dictionnaire.[...] Du point de vue linguistique, l'un des avantages du dictionnaire fréquentiel est de permettre d'établir, même de façon très approximative, une courbe de Zipf, dont la pente exprime une certaine « température du langage », grandeur strictement statistique et, donc, de la comparer avec d'autres courbes analogues, en particulier, au sujet de la langue parlée correspondante »(ibid., p84).*



**Figure 104 : Courbe de Zipf d'un corpus de 120 mots sifflés turcs avec en abscisse le Log du rang du mot dans le dictionnaire et en ordonnée le log de l'ordre d'apparition (Moles 1970, p. 113)**

Moles analyse cette courbe ainsi : « L'examen de cette courbe révèle en dehors d'erreurs locales au départ qui, avec un corpus aussi restreint ne méritent aucune interprétation, trois zones dans le vocabulaire: la zone A, de pente 1/1,2 environ, n'est pas très éloignée de la pente relative au langage parlé, elle l'épouse en quelque sorte. Dès le vingtième mot déjà, la courbe change de pente révélant une nouvelle zone B du dictionnaire, celle qui correspondrait à un effort de « pensée sifflée » ; enfin la zone C (queue de la courbe) [...] qui correspond à des mots utilisés une à deux fois ». (Moles 1970, p.113) D'après l'auteur il se peut aussi que ces derniers mots soient liés à des imprécisions de la traduction.

La constitution d'un dictionnaire fréquentiel passe par l'analyse de l'étendue du vocabulaire testé et de ses domaines d'application en prenant en compte le milieu d'évolution des sujets dont on estime les performances d'intelligibilité. Les mots les plus communément sifflés sont liés aux activités quotidiennes des siffleurs. Ce sont eux qui correspondent à la zone B de la courbe de Zipf de la Figure 104. Cet aspect est important dans la forme sifflée car le vocabulaire d'usage le plus courant est celui utilisé quand une personne est isolée, ce qui arrive surtout en dehors du village. Dans bien d'autres contextes modernes, cette tendance de la spécification du vocabulaire en fonction du métier est tangible dans la forme parlée: c'est par exemple le cas dans les métiers de la recherche.

#### 4.3.3.2.3. Conclusion

Le taux d'intelligibilité des mots isolés atteint environ 70% d'après des études menées en Turquie dans les années 60-70. Ces pourcentages sont relativement stables d'un auditeur-siffleur à l'autre. Ils représentent une augmentation de plus de 20 à 30% par rapport aux logatomes (du type Voyelle-Consonne-Voyelle). Cette différence tient en priorité à la reconnaissance de la sémantique qui repose sur l'ensemble des connaissances des sujets testés. Le vocabulaire courant en sifflement a été évalué entre 2000 et 4000 mots à la Gomera et en Turquie. Il dispose de plusieurs couches dont certaines sont plus spécifiques à l'usage sifflé.

Dans les mêmes conditions d'écoute, les mots en voix parlée sont identifiés à 96%. Mais comme la langue sifflée est adaptée, avant tout, à des conditions de communication dans le bruit de fond de la nature, il est intéressant de comparer ces mesures d'intelligibilité à celles de la voix dans le bruit. En effet, la langue sifflée serait justifiée, pour l'intelligibilité des mots parlés, qu'à partir du moment où le bruit ambiant entraîne une identification inférieure à 70 %. D'après les données de Bronkhorst et al (1993) une telle valeur est atteinte pour la voix dans des conditions d'écoute avec un rapport S/B de 3 dB alors qu'un taux d'intelligibilité des mots de 90% est obtenu pour une écoute avec un rapport S/B de 10 dB.<sup>106</sup>

---

<sup>106</sup> D'après nos mesures réalisées sur le terrain et pour des phrases (§ 2.3.6.2), la voix parlée inférieure à 80 dB porte moins loin que 40-50 m dans un bruit de fond d'environ 40 dB. C'est donc avant cette limite que le sifflement devient plus efficace que la voix parlée.

### 4.3.4. Intelligibilité des phrases sifflées (courte et longue distance)

#### 4.3.4.1. Performances usuelles

En considérant l'entité de la phrase nous atteignons la réalité de la structuration des dialogues pratiqués par les siffleurs. Tous les éléments favorisant l'intelligibilité s'intègrent à ce niveau de reconnaissance de la parole.

Sensation sonore	Perception phonétique	Perception des mots	Perception des phrases
Audibilité	Netteté	Intelligibilité	
			Compréhension

En grèce par exemple nous avons mesuré un taux d'intelligibilité de 95% à 150 m sur 10 phrases échangées spontanément dans une conversation vive et sur le sujet peu commun de la visite d'étrangers. Les phrases enregistrées à cette occasion sont caractéristiques d'une conversation courante. Elles sont assez courtes mais contiennent jusqu'à neuf mots. L'analyse de leur structure n'a montré aucune différence avec la syntaxe habituelle du langage parlé. Les 5% d'intelligibilité manquant de l'interlocutrice (Mrs. Kula) peuvent être attribués à la prise de contact à distance. Mais la même conversation a été perçue avec un taux d'intelligibilité de 100% par la fille d'un des siffleurs effectuant, depuis une distance de 10 m, en direct, la traduction du dialogue en grec parlé.

Comme nous l'avons vu, les communications ne sont pas stéréotypées car le vocabulaire courant est aussi étendu que la voix parlée utilisée dans le même type de contexte (plus de 2000 mots courants). Les arrangements syntaxiques des phrases qui suivent les règles de la langue locale réduisent la probabilité de confusion, marquent la mélodie, tout en donnant une consistance interne aux propos.

Donc, grâce au phénomène de la phrase, l'intelligibilité la plus complète est atteinte, comparable à celle de la voix parlée -quand la pratique de la langue sifflée est encore quasi-quotidienne- et ce, d'après les pratiques observées, jusqu'à des distances pouvant aller jusqu'à un à deux kilomètres suivant le milieu écologique. Ce dernier point est une spécificité sur laquelle nous allons nous attarder car ses conséquences pour la phonétique n'ont jamais été mesurées.

#### 4.3.4.2. Non-linéarité de l'intelligibilité à distance

Une expérience pilote, réalisée en pays Mazatèque en 2003 et que nous décrivons en détails en Annexe **Erreur ! Source du renvoi introuvable.** nous a permis d'observer que le phénomène d'intelligibilité de la parole sifflée a une évolution non linéaire avec la distance (comme pour la voix). Cela signifie qu'à deux distances d'écoute différentes, la même émission permet toujours au siffleur de comprendre le message avec les mêmes performances à condition de ne pas dépasser une zone limite tampon où la perte d'intelligibilité est rapide. Les éléments du signal sonore parvenant à l'auditeur sont *suffisants* pour atteindre un taux élevé d'intelligibilité des phrases, malgré la dégradation qu'ils ont subit depuis leur émission à cause des effets de la propagation et du bruit de fond.

Busnel et Classe (1976) avaient également, de manière succincte, observé le phénomène non linéaire de l'intelligibilité des langues sifflées puisqu'ils écrivaient à l'époque *"We recorded a whistled signal at the source and simultaneously, from a point 2000 m away but in visual contact, in a flat country and on a windless day. Analysis by the sonograph revealed an interesting fact: a significant portion of the frequency band had disappeared. Nevertheless the signal remains intelligible, for what remains of the informational skeleton amply suffices for a mental reconstruction of the message."* (Busnel et Classe, 1976).

Dès lors les langues sifflées présentent un cas extrême de la capacité du cerveau humain à reconstruire un message linguistique à partir d'indices acoustiques parcellaires. Du point de vue de l'analyse de l'intelligibilité du langage, le cadre expérimental offert par ce phénomène naturel est extrêmement prometteur. En effet, à des distances même moyennes (permettant un taux d'intelligibilité élevé), on peut s'attendre à ce que le signal parvenant à l'auditeur soit particulièrement réduit puisque qu'il est le résultat de la propagation d'un style de parole qui opère lui-même une concentration et une sélection physique de l'information linguistique d'une phrase dans la bande de fréquence du sifflement.

### **4.3.4.3. Expérimentation en conditions de distance**

#### **4.3.4.3.1. Introduction**

Afin d'étudier la répartition des indices acoustiques qu'un individu perçoit à grande distance et ses conséquences phonétiques, nous avons mis au point une expérience en plein air dont le protocole nous a permis d'enregistrer exactement les mêmes phrases sifflées à plusieurs distances pour plusieurs langues dont nous avons étudié la phonologie et la phonétique de manière approfondie. Il est important de connaître leur forme dégradée qui parvient à des auditeurs placés à différentes distances car nous pourrions alors comprendre non seulement quels sont les indices réduits qui suffisent à l'auditeur pour décoder le signal de parole, mais aussi nous observerions l'évolution de la dégradation du signal émis en fonction de la distance.

Dans la partie suivante nous donc développons la présentation de cette expérience qui porte sur plusieurs phrases de notre corpus de turc, de grec et de silbo. Puis nous faisons une première analyse des enregistrements obtenus.

#### **4.3.4.3.2. Préparation**

##### **Choix du lieu**

Nous avons choisi de réaliser notre expérience en milieu montagneux avec une vallée dégagée et un environnement semi boisé, comme c'est le cas dans des milieux permettant des communications jusqu'à 2 km. Notre but était en effet de faire des mesures sur des distances intermédiaires (100, 150 et 300 mètres) et une distance longue (550 m) dans un milieu qui n'était pas trop réverbérant. Après différentes investigations dans les Alpes et une expérience pilote dans le Vercors, nous avons opté pour la vallée présentée sur la Photo 27 qui forme un guide d'onde de qualité intermédiaire.



**Photo 27 : Vallée cadre de l'expérimentation à distance**

## **Equipe expérimentatrice**

Une équipe de travail de deux personnes a été constituée<sup>107</sup>. L'un des expérimentateurs était chargé de la diffusion des sons et l'autre des enregistrements. Les deux ayant évalué auparavant les distances avec un mètre de maçon de 50m.

## **Corpus et stimuli**

### **Choix des langues**

Trois langues ont été sélectionnées. Il s'agit des langues non tonales dont nous avons analysé la version sifflée le plus en détails : le turc, le grec et l'espagnol. Le turc a été diffusé en version sifflée et parlée (limite voix criée). Une partie du corpus de turc a l'avantage de présenter la même phrase énoncée trois fois de suite par 2 personnes, à la fois en version parlée et sifflée, ce qui donne plusieurs points de vue sur la même réalité linguistique et permet donc une meilleure interprétation des résultats.

---

<sup>107</sup> Je tiens ici à remercier Rémi Saudax pour sa collaboration.

## Choix des phrases

Parmi tout notre corpus, les phrases énoncées par les meilleurs siffleurs dans les conditions d'enregistrement les plus optimales et réalisées à quelques mètres de la bouche ont été sélectionnés. En tout nous avons diffusé et enregistré à chaque distance : 10 phrases de turc, 5 phrases de grec, 4 phrases d'espagnol.<sup>108</sup>

## Préparation des stimuli

La version parlée du turc a été étalonnée en amplitude sur la version sifflée afin de pouvoir réaliser une comparaison liée aux bandes de fréquences. Cela entraîne que, compte tenu des niveaux d'émission sonore, une telle aisance dans l'énonciation ne peut être obtenue naturellement (97 dB en moyenne avec des maxima à 104 dB), mais l'analyse de la voix criée qui a été réalisée par la même occasion nous a permis de le remarquer.

## Matériel de diffusion et d'enregistrement

Le matériel utilisé est présenté en Annexe G.

### 4.3.4.3.3. Déroulement de l'expérience

Le protocole expérimental conçu afin de gérer le déroulement de l'expérience consistait tout d'abord en la mesure des distances. Puis au réglage des niveaux d'émission pour chaque langue. Puis la diffusion et l'enregistrement ont été réalisés en faisant attention à ce que le bruit du vent soit minimal et qu'il n'y ait aucun bruit industriel (tracteur, avion) . La Figure 105 résume les configurations d'enregistrement (150, 300 m, 550 m). On peut voir qu'une variante a été introduite à 100m : à cette distance l'enregistrement a eu lieu dans deux positions du haut-parleur (face et 45°) afin de mesurer d'éventuels effets sur l'amplitude de l'onde directe.

---

<sup>108</sup> Nous avons également diffusé 4 phrases de Français énoncées par un Gomero. Mais ces données ne sont pas traitées ici

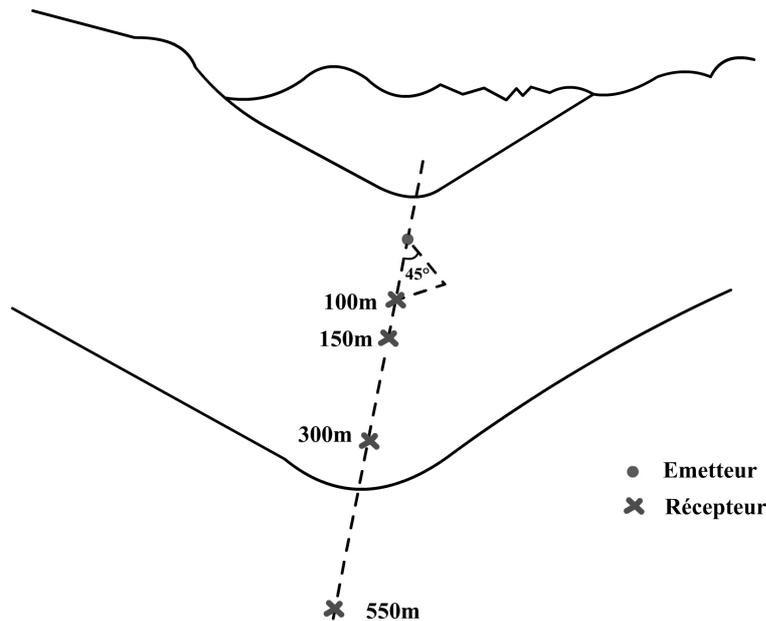


Figure 105 : Configurations de l'expérience

### Contrôle du niveau d'émission

Le son était contrôlé uniquement à partir de l'amplificateur. Un niveau de réglage a été opté pour chaque langue en fonction des conditions d'enregistrement et de la technique utilisée par les siffleurs de manière à être le plus proche possible des conditions réelles d'émission.

Tableau 34 : Niveaux d'émission des phrases pour chaque langue

	grec	espagnol	turc
Niveau moyen (dB)	89	86	97
Maximum de niveau (dB)	101	97	104

### Contrôle de l'intelligibilité

La limite d'intelligibilité n'a été atteinte que pour la langue espagnole qui est celle qui avait été émise aux niveaux de puissance les plus bas (par le siffleur et donc par nous). Pour les autres langues nous sommes restés bien en deçà de cette limite car nous n'avons pas dépassé la distance à laquelle le sifflement original était destiné. L'intelligibilité a été contrôlée dans le casque afin de correspondre à l'enregistrement réalisé. Nous avons estimé que nous étions qualifiés pour juger de l'intelligibilité des phrases car nous avons appris la technique de sifflement à la Gomera et que nous parlons l'espagnol couramment.

#### 4.3.4.3.4. Résultats et première analyse

##### Dégradation progressive 150, 300, 550m

Exemple du turc

a) *Résultat des enregistrements*

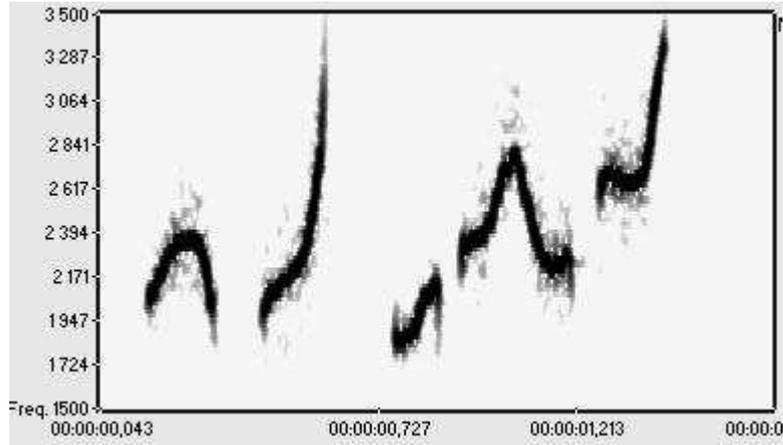


Figure 106 : enregistrement à la source de la phrase turque « mehmet okulgıt »

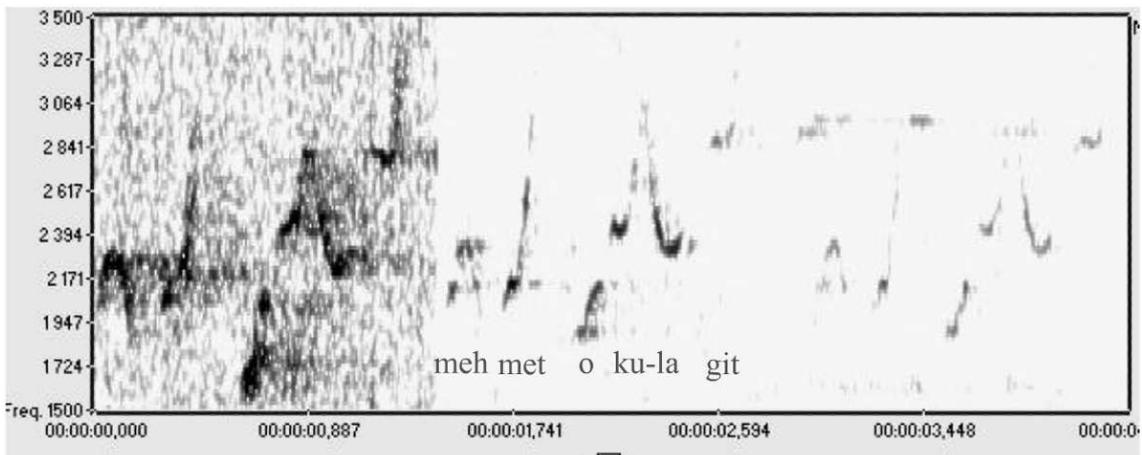
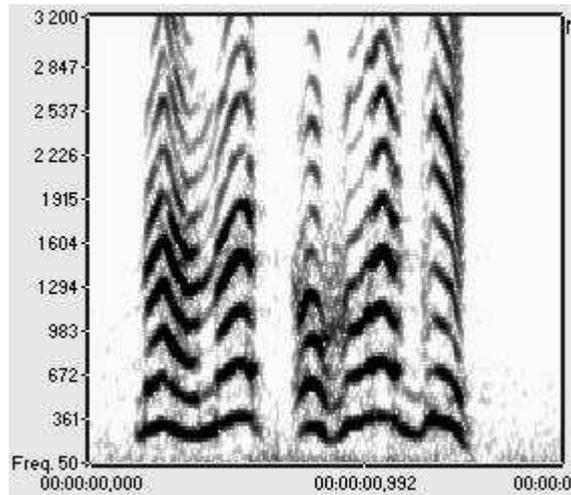


Figure 107 : Même phrase enregistrée à 150, 300 et 550 m, même réglage du spectrogramme aux 3 distances

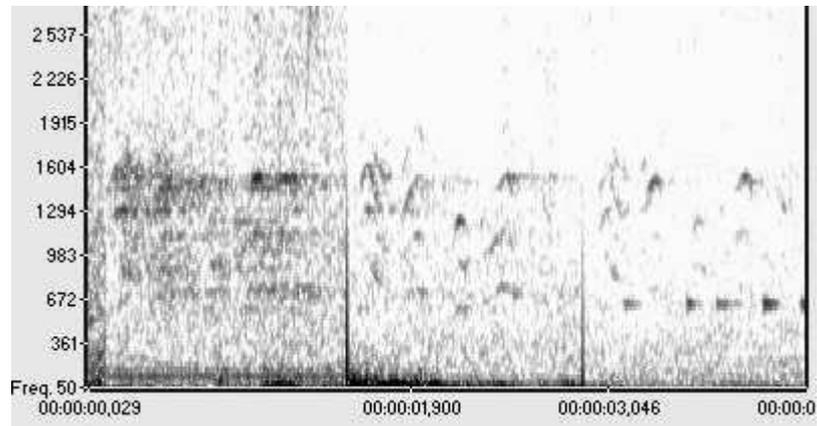
b) *Commentaires*

On observe une dégradation progressive de l'émergence du signal par rapport au bruit. L'intelligibilité est toujours complète à 550 m même si elle requiert de l'attention. En ce qui concerne le bruit de fond, il était différent en nature aux 3 distances mais toujours entre 40 et 50 dB. A 150m, le bruit était le plus élevé de toutes les distances d'enregistrement en grande partie à cause de la forte réverbération de cette partie du champ qui renvoyait le bruit des feuilles des sommets des arbres et celui de la présence des grillons.

c) *Dégradation de la voix turque*



**Figure 108 :** « Mehmet okulagit » version parlée (limite criée) à la source.



**Figure 109 :** « Mehmet okulagit » version parlée enregistrée à 150, 300, 550 m (à 550 m les fréquences visibles vers 672 Hz sont issues du tintement d'une cloche de vache)

A 550 m la voix diffusée était à la limite de l'intelligibilité. Nous avons déjà précisé qu'aucun être humain ne peut garder l'aisance d'élocution au niveau auquel nous avons diffusé cette voix. Il ne s'agit donc pas de la reproduction de conditions réelles de communication. Par contre nous pouvons remarquer que les différents milieux d'enregistrement ont eu un impact sur le signal de la voix en le déformant ou en le dégradant dans de grandes proportions. Malgré cela l'intelligibilité est toujours complète à 150 m (forte réverbération) et 300 m. Les éléments portés par la plus forte amplitude subsistent. Nous n'avons pas encore produit d'analyse phonétique détaillée comparée mais l'on remarque qu'à grande distance ce sont les fréquences situées entre 1000 et 1900 Hz qui sont le mieux transmises à travers le filtre du milieu ambiant.

## Cas de l'espagnol

### a) Résultats d'enregistrement

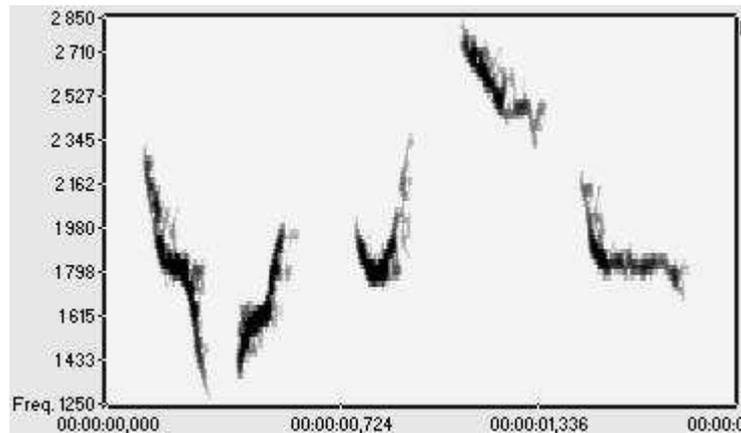


Figure 110 : Enregistrement à la source de l'extraire de phrase espagnole : « la mon-ta-ña-ta »

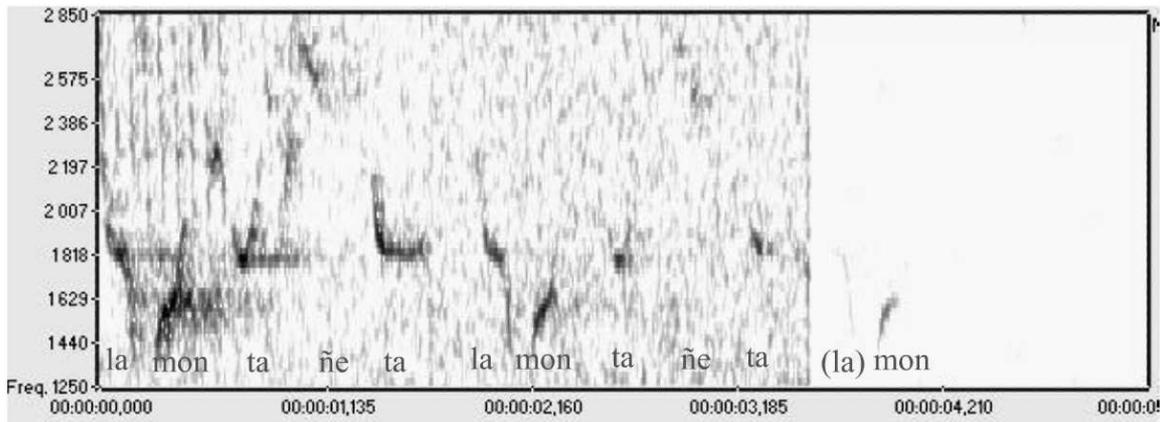


Figure 111 : Même extraire de phrase « la montaña » enregistré à 150m, 300 m et 550 m. Réglages identiques du spectrogramme aux 3 distances.

### b) Commentaires

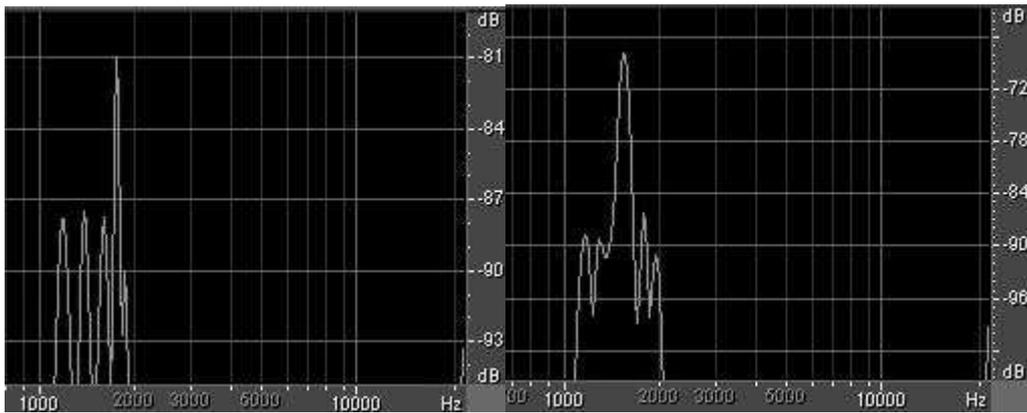
On observe la même dégradation progressive que pour le turc, mais comme le signal est émis à un niveau moins élevé, il émerge difficilement à 550 m.

Le signal est encore clairement intelligible à 300 m mais il n'est plus intelligible à 550 m. Il est pourtant net pour certaines syllabes. Mais la dégradation est trop grande pour reconstituer le signal. Nous avons par exemple confondu « silbo » avec « amigo ».

### Remarques sur le bruit

Dans tous les cas nous avons observé une réduction progressive de la réverbération du signal dans le milieu de l'expérience. Le signal à 300 m est plus agréable à l'écoute qu'à 150 m en raison de la forte réverbération à cette dernière distance. De plus, parfois, des modifications locales apparaissent à 300 m mais ne sont pas présentes à 550 m. Ceci est dû à des conditions de vent ou de réverbération différentes.

Mais le point le plus intéressant de cette étude est la possibilité qu'elle nous donne d'encadrer la limite de netteté des syllabes. En effet des cas tels que celui présenté dans la configuration 550m de la Figure 111 (syllabes « la » et « mon ») nous ont permis de conclure que le signal, pour être détecté et reconnu devait émerger d'environ 20 dB de la bande des sifflements. Si dans un mot, plusieurs syllabes ont des parties émergentes à 20 dB et quelques unes à 6 dB, il est éventuellement possible d'en reconstruire le sens mentalement, tout dépend de leur répartition. A ce niveau de difficulté, la connaissance du sujet joue beaucoup. L'exemple de la Figure 111 est typiquement représentatif de ce type de situation puisque le signal émerge de 6 dB de la bande des sifflements pour « la », et de 20 dB pour « mon » (Figure 112).



**Figure 112 : Emergence de syllabes sifflées du bruit de fond dans la bande 1200-2000 Hz, « la »(S/B=6dB) et « mon » (S/B=20dB) (voir Figure 111)**

Une émergence de 6 dB permet de détecter le signal mais pas de s'en servir de base pour une projection mentale permettant l'intelligibilité. Il ne peut que confirmer une impression donnée par un signal émergant à plus de 20 dB.

Remarques:

- les valeurs que nous donnons en décibels ici sont pertinentes pour une écoute monaurale car l'analyse spectrale ne nous permet pas de faire une évaluation de l'émergence avec écoute binaurale. En général celle ci améliore la perception de + 6 dB. Comme nos enregistrements ont été réalisés en stéréo nous en avons bénéficié dans l'écoute au casque (casque englobant les oreilles).
- D'autres considérations de l'écoute en contexte semblent jouer pour la préhension du cerveau sur les sons environnants, comme l'évaluation de la distance de la source (voir la remarque finale de l'expérience en pays Mazatèque en Annexe **Erreur ! Source du renvoi introuvable.**)

## Effet de l'angle à 45°

### Résultat type

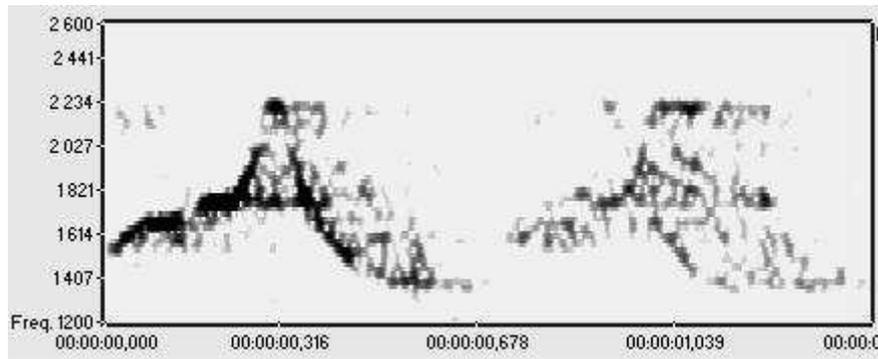


Figure 113 : Mot grec enregistré à 100m dans deux positions différentes de la source (face et 45° par rapport au récepteur).

### Commentaires

La zone d'expérimentation à 100 m était déjà une zone de forte réverbération, avec le changement d'orientation de 45° cet effet est beaucoup amplifié (Figure 113). Malgré la forte déformation du signal par réverbération et la moindre intensité de l'onde directe dans le deuxième enregistrement, les phrases sont toujours clairement audibles.

## Effet du chant des oiseaux

Lors de l'écoute nous avons également été perturbés par les sifflements d'oiseaux situés dans le même domaine de fréquence que le signal de parole sifflée. Surtout lorsqu'ils venaient de la même direction. Malgré cela, l'intelligibilité restait possible. Mais la dynamique très différente du signal des oiseaux de cette vallée par rapport à celle des sifflements humains explique que l'intelligibilité n'est que partiellement affectée. Les merles de la Gomera, dont le sifflement est très proche de celui des humains, engendrent plus de confusions chez les siffleurs en créant les conditions typiques de l'*effet cocktail party*.

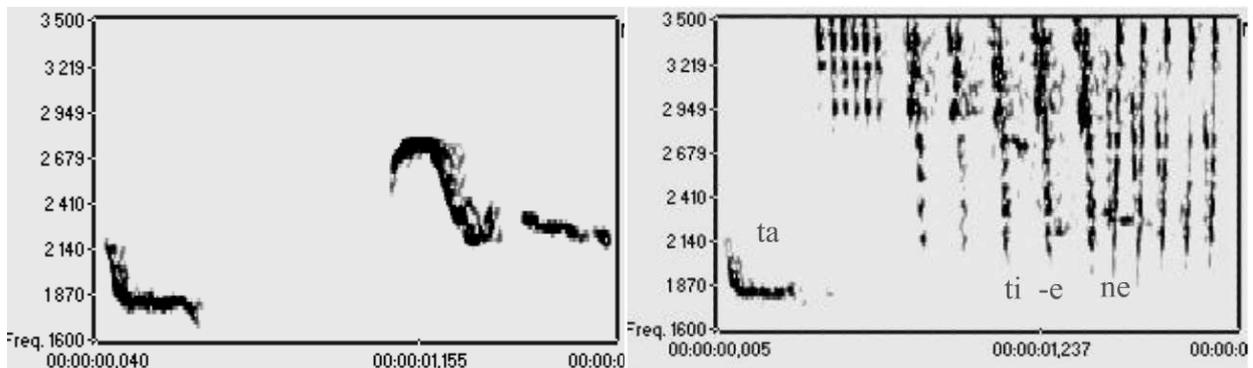


Figure 114 : extrait espagnol sifflé d'origine et à 150 m, « -ta tiene ». Le mot « tiene » est visible à travers la grille du chant d'oiseau

#### 4.3.4.3.5. Première approche phonétique

L'analyse détaillée des structures phonétiques qui subsistent à moyenne et grande distance permet de constater que les signaux entraînant une bonne intelligibilité des phrases ont une répartition qui dépend des points de plus forte énergie du signal d'origine.

-Nous avons remarqué que les voyelles apparaissent le plus clairement : elles portent la plus grande partie de l'énergie sonore de la phrase. Elles sont toujours modulées, ce qui semble indiquer la direction de variation de la (des) consonne(s) qui l'entoure(nt).

-Les éléments de consonnes qui subsistent ne sont pas uniquement ceux proches du noyau de la syllabe, ils concernent aussi parfois des points intermédiaires de la modulation de fréquence, ce qui permet d'en saisir clairement la direction et les aspects temporels liés à l'articulation. Le /t/ est par exemple souvent marqué par un silence intermédiaire qui indique une baisse puis une remontée en amplitude.

#### 4.3.4.3.6. Discussion

Il est important de considérer les problèmes d'émergence du bruit de fond que nous avons souligné grâce à l'espagnol. Dans tous les cas que nous avons observé, la parole sifflée est encore intelligible lorsque les points d'intensité les plus élevés de la syllabe émergent de 20 dB du bruit de fond de la bande des sifflements. Dans le cas particulier de la Figure 111, nous avons trouvé que la netteté des syllabes disparaît entre 6 dB et 20 dB d'émergence S/B de la même bande de fréquence des sifflements. Or, au § 2.3.7.1 nous avons montré que la dynamique intra-sifflements de la parole sifflée est en général de 20 dB (en mesure sur analyse spectrale). Cela veut dire que les éléments les moins intenses de la syllabe « mon » (Figure 111, 550 ont une émergence (S/B bande sifflée) de 0dB. Compte tenu de nos mesures il y aurait bien une dégradation des éléments perçus par l'auditeur. Pour la même Figure à 300 m on observe de nombreux cas de syllabes pour lesquelles des parties importantes du signal d'origine sont en dessous de la limite des 6 dB de rapport (S/B (Bande des sifflements)). Il y a donc souvent reconstruction cognitive de la parole pour les éléments perçus.

### Facteurs généraux d'intelligibilité de la phrase

#### Parole dans le bruit

De nombreuses recherches sur la parole dans le bruit donnent une première justification des taux élevés d'intelligibilité obtenus par les siffleurs grâce aux phrases même après dégradation. En effet, la reconnaissance de la parole dans le bruit ou en milieux contraignants ne nécessite pas la perception de tous les phonèmes, de toutes les syllabes et de tous les mots car elle s'appuie en grande partie sur l'ensemble des connaissances apprises par le sujet au cours de sa vie (Warren et Warren 1970, Warren 1970<sup>109</sup>). Les éléments manquants peuvent être facilement comblés grâce aux contraintes phonologiques et lexicales, et

---

<sup>109</sup> voir en Annexe **Erreur ! Source du renvoi introuvable.** pour une explication détaillée de l'expérience des Warren.

aux relations syntaxiques et sémantiques qui gouvernent leur enchaînement linguistique. Bosman (1989) a montré que les effets de ces indices sur l'intelligibilité de la parole se révèlent être extrêmement élevés lorsqu'elle est présentée dans des conditions d'écoute éloignées des conditions idéales. Il a par exemple mesuré les taux de reconnaissance de la parole de plusieurs types de corpus dans des conditions de bruits contraignants. Il a obtenu des niveaux d'intelligibilité jusqu'à 55 % plus élevés pour des mots insérés dans des phrases par rapport à des mots isolés (ce qui indique un effet combiné de la syntaxe et de la sémantique), et des taux jusqu'à 80% plus élevés par rapport à des non-mots (effet additionnel des contraintes lexicales). Il a également constaté que la reconnaissance des syllabes était favorisée par le fait qu'elles ne soient pas prononcées de manière isolée mais intégrée dans un continuum de parole. Ce dernier aspect souligne que la parole contient des indices de coarticulation, d'allophonie et de durée qui fournissent des informations sur la position et l'identité des syllabes et, par extension, des phonèmes.

### Rôle clef des modulations

Nous avons remarqué que même à grande distance, les modulations sifflées, qui sont si présentes et rapidement apparentes en sifflements proches, semblent jouer un rôle primordial à la fois au niveau du noyau de la syllabe et au niveau de la constriction ou de l'explosion des consonnes.

#### *a) Modulations et noyau de la syllabe*

Un aspect intéressant pour l'étude des langues sifflées tient au fait qu'il a été montré qu'une modulation du type vibrato appliquée à l'ensemble des harmoniques d'un son de la voix permet de créer une plus grande prééminence (McAdams 1984) et cohérence (Darwin et Sandell 1995, Darwin et al 1994) qu'un son non modulé. Marin et McAdams (1991) ont précisés les conditions importantes dans lesquelles une modulation favorise la prééminence d'une voyelle face à d'autres voyelles non modulées simultanées: "*We conclude that coherent, sub-audio frequency modulation on a harmonic sound source contributes to its segregation from other concurrent sounds if the modulation width is large enough*". (Marin et McAdams, p. 350). Des études perceptives récentes de Zeng et al (2005) mettent également en évidence le rôle joué par les modulations cohérentes de fréquence pour la ségrégation et l'association des éléments de parole dans des environnements de parole complexes comme un cocktail party ou un milieu bruyant. Cet aspect du problème nous donne un élément d'explication supplémentaire sur l'efficacité de la parole sifflée dans le bruit. En effet, les voyelles sifflées sont toujours affectées par une modulation, au moins à l'initiale ou à la finale du noyau de la syllabe et nous avons montré que leur bande de fréquence est assez large pour activer plusieurs filtres auditifs simultanément.

De plus, Darwin (1981) a analysé l'effet de la cohérence de modulation de l'ensemble des harmoniques d'une consonne d'attaque d'une syllabe. Il a souligné l'importance des modulations à la fois de fréquence et d'amplitude qui donnent une plus grande valeur perceptive à l'ensemble CV qu'à un phonème seul. Dans les langues sifflées non tonales du groupe 1 de la Typologie que nous avons établie, la modulation reste visible à l'attaque de la consonne, même à distance et reflète ainsi l'importance de ces cohérences en appuyant l'effet perceptif transposé.

b) Double rôle des combinaisons de modulation AM (modulation d'amplitude) et FM (modulation de fréquence)

Plusieurs facteurs ont jusqu'ici souligné l'importance des modulations dans l'organisation générale de la phrase que ce soit au niveau de la continuité temporelle observée grâce aux modulations d'amplitude (AM) en prosodie ou au niveau de la continuité fréquentielle grâce aux modulations de fréquence (FM).

Les études de Zeng et al (2005) ont à nouveau été intéressantes pour faire avancer notre compréhension à ce niveau. En effet, ces auteurs soulignent que l'avantage de la FM dépend du type de AM qu'ils ont imposé artificiellement et simultanément. Pour eux c'est le signe que *la phase* joue un rôle non négligeable dans l'intelligibilité de la parole car c'est le type de combinaison de AM et de FM qui va être important pour l'auditeur, en particulier dans le bruit : « *phase information may not be needed in simple listening task but is critically needed in challenging tasks, such as speech recognition with a competing voice* » (Zeng et al 2005, p.2298).

Or, la phase a reçu relativement peu d'attention de la part des études sur la parole. Pour comprendre et analyser les modulations des langues sifflées il convient donc de comprendre le rôle de la phase dans la perception du langage. On sait aujourd'hui que la modification de la phase ne change pas la qualité phonétique des éléments considérés à court terme (Carlson et al, 1979). Mais une des autres rares études sur le rôle de la phase dans la parole a été réalisée par Oppenheim et al (1979), elle a montré qu'un signal de parole avec un spectre d'amplitude maintenu à l'unité alors que le spectre de phase reste intact produit un type de parole appelé "*phase-only*" ayant une qualité non naturelle car bruitée mais permettant de garder un haut degré d'intelligibilité. Les spectrogrammes de ce type de parole montrent que la structure formantique est maintenue. Une telle manipulation du signal peut être comparée à l'utilisation de la phase pour la reconstruction d'images. Ces auteurs reconnaissent que la phase représentée dans des spectres étudiés sur des temps courts joue un rôle insignifiant car ce type d'observation spectrale modélise la parole comme un signal quasi-stationnaire pour lequel les composantes fréquentielles sont assimilées à des éléments stationnaires. D'un autre côté, ils remarquent que lors de l'observation du spectre de phase sur le long terme, la parole n'est plus modélisée comme un signal quasi stationnaire, elle est un signal pour lequel les composants fréquentiels changent dans le temps de manière dynamique. Ils en concluent que pour la parole, comme c'est le cas pour les images, l'information de la phase préserve la « situation relative des éléments » ("*location of events*") en signalant les bords des syllabes et tous les autres événements à variation rapide ("*lines edges and other narrow events*") (ibid. p534). En d'autres termes, les spectres tenant compte de la phase sur une longue durée encodent l'emplacement des changements les plus importants dans le signal qui, dans la parole, se traduisent par des changements dans l'amplitude des composants fréquentiels. Par conséquent on comprend que les types de combinaisons AM-FM jouent un rôle important dans les langues sifflées *au niveau de la phrase*. De surcroît, ce phénomène perceptif nous permet de comprendre pourquoi le sifflement devient si efficace sur les phrases alors qu'il est encore moyennement efficace sur les mots.

Ces résultats mettent donc en évidence que les changements fréquentiels s'accompagnant de changements d'amplitude rapides (associés à des ouvertures, des fermetures, des constrictions des couplages ou des

découplages des résonateurs du tracus vocal) jouent un rôle important d'indices acoustiques de localisation relative pour la perception. Ceci est cohérent avec les conclusions de l'analyse de la scène auditive.

L'analyse de la phase de la parole révèle également certaines limites des modes d'observation spectrale utilisant l'analyse de Fourier. Ils confirment la nécessité d'utiliser des observations sur le long terme (large bande) pour les éléments temporels. Dès lors, il nous est possible d'identifier les phonèmes pour lesquels l'analyse de Fourier à bande étroite entraîne un rendu fortement incomplet des changements de valeurs portés par la phase. Ce sont par exemple les consonnes occlusives, les affriquées (provoquées par l'ouverture et la fermeture des cavités de la bouche), les consonnes nasales (qui sont produites par un couplage et un découplage des cavités nasales et par l'ouverture et la fermeture des cavités orales), et les latérales /l/ ou /r/ (produites par un couplage et un découplage de deux cavités orales parallèles). Les langues sifflées confirment que ces phonèmes entraînent des variations rapides de AM ou de FM (voir partie Typologie, Chapitre 3)

## **Ouverture sur le rythme**

On remarque sur nos enregistrements réalisés à distance que les phrases enregistrées à longue distance sont parfois très limitées à une modulation du noyau de la syllabe. C'est par exemple le cas de l'expression « la montaña » à 300 m (Figure 107), avec seulement les voyelles pour éléments de base et de légères modulations comme un tambour parleur peut le faire (Essien 2000)<sup>110</sup>. Cet aspect a le mérite de nous rappeler l'importance du rythme dans la phrase sifflée. L'intensité jouant le rôle de marqueur rythmique soulignant un événement fréquentiel. Nous n'avons pas fait de mesures temporelles précises sur le signal sifflé jusqu'ici car nous attendions les résultats de nos enregistrements à distance. Certains aspects de la dégradation du signal à distance permettent effectivement de préciser le choix des éléments temporels à mesurer. La forme dégradée des modulations de consonnes observées sur sonagrammes suggère que tous les éléments visibles à courte distance ne jouent pas le même rôle dans le rythme et que ce ne sont pas seulement les pics d'amplitude qui sont à prendre en compte mais plutôt des zones temporelles dans chaque syllabe qui participent à identifier les consonnes qui sont marquées par la fréquence des voyelles sur lesquelles elles s'appuient.

---

<sup>110</sup> Essien a étudié le « tambour parlant » Yoruba ou « tambour bata » qui est un tambour d'aisselle biface portable (de la forme d'un sablier) dont les peaux de percussion peuvent être tendues par un jeu de courroies pressées sous l'aisselle en même temps que l'on frappe la peau, ceci permet de faire de légères modulations des sons.

#### 4.3.4.3.7. Conclusion et perspectives

Nous avons engagé ici une recherche qui se révèle très riche d'enseignements sur la compréhension de l'intelligibilité des langues sifflées. Le respect des conditions naturelles de dégradation du signal, bien que contraignant d'un point de vue expérimental est le point clef et déclencheur de nombreuses réflexions relativement nouvelles sur l'intelligibilité de la parole.

Nous avons donc commencé à explorer quelques pistes de recherche à la faveur d'une première analyse des enregistrements réalisés :

- Nous avons ainsi souligné la nécessité de considérer le rôle des combinaisons AM-FM en lien avec leur signification phonétique.
- Nous avons pu identifier des paramètres importants à considérer lors de la mesure du rythme d'une parole sifflée.
- Nous avons souligné l'importance de regarder le signal avec une méthodologie adaptée aux types de variations des différents indices acoustiques portés par les éléments segmentaux de la parole. Cette observation doit être réalisée à la fois à court terme et à long terme. C'est pour cela que nous avons commencé à mettre au point, avec l'aide de chercheurs spécialisés en traitement du signal, des outils adaptés au sifflement. Nous avons cherché des outils mathématiques permettant une résolution spectro temporelle plus stable que l'analyse de Fourier. Deux voies ont été ouvertes dans ce sens en fin de thèse. Nous les présentons succinctement en fin d'Annexe A. Il s'agit d'une part d'une approche par l'intermédiaire des ondelettes de Morlet, et d'autre part, d'une approche par l'intermédiaire d'une technique de paramétrisation du signal qui a l'avantage de s'adapter aux caractéristiques du sifflement tout en étant rapide et souple (cf. **Erreur ! Source du renvoi introuvable.**). Nous n'avons pas cependant terminé la phase d'optimisation de ces outils. De plus, dans la perspective d'une comparaison avec le rythme de la voix parlée, il est nécessaire de faire un choix parmi les différentes méthodes de mesures temporelles proposées par les phonéticiens de la voix. Nous en sommes au stade de bilan des techniques afin de choisir celle qui est la plus cohérente avec la réalité des langues sifflées et des données perceptives de la psychoacoustique.

## 4.4. Conclusion pour l'intelligibilité

Pour étudier l'intelligibilité des langues sifflées nous avons décidé de procéder par étapes en nous appuyant sur des connaissances de différents domaines. De cette manière nous avons pu souligner que toutes les connaissances actuelles en physiologie acoustique de l'oreille et en psychoacoustique convergent pour soutenir que l'audition humaine privilégie la bande de fréquences des sifflements mais aussi sa dynamique (en termes de masquage et de résolution temporelle notamment). Des considérations perceptives sont aussi particulièrement adaptées pour mieux comprendre l'équilibre perception-production utilisé par les langues sifflées pour reproduire les voyelles ou les tons. En effet, leurs stratégies de transposition de la voix choisissent entre deux qualités différentes de la hauteur: la Hauteur Fondamentale (HF : tons et intonation) et la Hauteur Brute (HB : timbre). De plus, nous avons pu faire un lien entre le fait que la HF entraîne parfois des ambiguïtés d'octave et l'habitude des siffleurs de maintenir leur signal de parole dans un intervalle d'un octave (pour une même phrase et ce dans toutes les cultures rencontrées sauf une<sup>111</sup>).

Comme la parole n'est pas statique et dispose de nombreux paramètres relatifs, nous avons cherché à éclairer notre compréhension de la parole continue sifflée avec les résultats de l'analyse de la prosodie et avec ceux de l'*analyse de la scène auditive*. L'évocation de ces deux domaines de recherche a été l'occasion, d'une part de décrire comment l'audition humaine sépare et regroupe les sons de son entourage en des flux perceptifs, d'autre part de remarquer que les langues sifflées font une proposition alternative et phonologiquement pertinente de méthodologie d'étude de la prosodie du langage. Nous avons montré que les relations entre les attributs de la perception permettent de comprendre de nombreux usages phonétiques et phonologiques et se révèlent des points clefs d'une bonne intelligibilité dans le bruit. Ceci est valable à la fois pour la forme parlée et pour la forme sifflée d'une langue.

Sur ces bases, nous avons ensuite pu présenter et concevoir des expériences testant l'intelligibilité des langues sifflées non tonales du groupe 1 de notre typologie. Lors de l'expérience de psychoacoustique sur les voyelles sifflées, il est apparu que les Français identifient les voyelles sifflées /i/, /e/, /a/, /o/ avec la même logique que les siffleurs espagnols. Il est intéressant de noter que les voyelles de l'expérience étant sur plus d'un octave, il semble que les Français, comme les Espagnols soient peu perturbés par l'ambiguïté d'octave pour l'identification des voyelles. Ces résultats nous ont suggéré un rapprochement avec ceux des expériences sur l'*effective formant* de Carlson et al (1970). La pratique sifflée semble être un modèle adéquat pour étudier la perception du timbre des voyelles. Elle nous a permis de confirmer qu'un des principaux éléments de perception de la voyelle tient à la compacité de la répartition de ses harmoniques fréquentielles et même de ses formants les mieux soulignés en amplitude.

Nous avons ensuite montré que les taux de netteté des logatomes, et d'intelligibilité des mots et des phrases étaient respectivement de l'ordre de 20-50%, 70%, 90-100% à courte distance. Nous avons remarqué que l'intelligibilité reste souvent élevée à grande distance. Nous avons donc conçu une expérience pour observer

---

<sup>111</sup> Il semble que les langues sifflées les plus marquées par ce phénomène soient celles qui transposent les tons de la voix (mazatèque, hmong, surui) ou bien celles qui ont un grand nombre de voyelles (turc)

de manière contrôlée la dégradation du signal de parole sifflée. Notre analyse des résultats qui n'est que partielle actuellement nous a permis de comprendre l'importance des modulations dans le bruit et de trouver des limites cohérentes de netteté des syllabes. Nous avons enfin retenu que la phase et le mode d'observation des éléments segmentaux et suprasegmentaux de la parole devaient être pris en compte pour analyser l'intelligibilité de la phrase et pour mesurer les paramètres temporels de la parole.

## **CONCLUSION : BILAN ET PERSPECTIVES**

Les langues sifflées étudiées dans ce travail représentent assez fidèlement la diversité des langues sifflées du monde qui elles-mêmes couvrent un très large éventail des familles linguistiques répertoriées jusqu'ici. Elles nous ont permis de confirmer que la parole sifflée est dans chaque population le résultat de l'adaptation de l'intelligence productive et perceptive humaine à un milieu biologique (conséquences acoustiques) et social (conséquences linguistiques). Pour cela nous avons partagé la vie des siffleurs de plusieurs cultures et ceux-ci nous ont fait l'honneur de nous accepter parmi eux. Ils nous ont aidé à concevoir une méthodologie adaptée à l'étude actuelle des langues locales.

Le premier point à retenir de notre étude est que la voix classique, qui est à l'origine de la parole sifflée laisse dans le sifflement l'emprunte de la manière dont elle est exploitée par les locuteurs d'une langue. Ceci est vrai à la fois en ce qui concerne les éléments clefs de la structure de la langue et en ce qui concerne les stratégies employées couramment par les êtres humains pour que leur voix porte loin. Il est possible de le montrer avec les outils de la linguistique et de la bioacoustique. Notre contribution sur ces deux points a été de montrer qu'il existe une continuité entre la voix parlée et la parole sifflée qui se fait sur plusieurs échelles.

Pour la communication à distance et dans le bruit, la parole sifflée est la suite logique de la voix criée. En effet, l'élévation de la voix par effet Lombard, l'augmentation de la fréquence et de la durée d'élocution poussent rapidement les cordes vocales à leurs limites et il est nécessaire de s'en passer. Le sifflement a été la solution retenue par de nombreuses populations. De cette manière la parole peut rester durablement dans la zone de perception privilégiée de l'oreille humaine, en dehors des zones où le bruit ambiant de la nature est le plus intense. De plus, le sifflement permet d'utiliser le langage humain avec des propriétés qui en font un véritable système de télécommunication: l'amplitude de la dynamique du langage est condensée, la bande de fréquence est réduite. Ces propriétés permettent d'améliorer la netteté du signal à distance en optimisant le rapport signal sur bruit.

Suite à la réduction de la bande de fréquence, les siffleurs doivent choisir de transposer en priorité soit la Hauteur Fondamentale, soit la Hauteur Brute de leur langue qui correspondent respectivement à la fréquence de phonation des cordes vocales et à l'ensemble du spectre de la voix. Ce choix hiérarchique semble préexister dans la version parlée des langues humaines et c'est pour cela qu'il n'affecte pas trop l'intelligibilité de la parole sifflée. Par exemple, c'est lui qui différencie les langues tonales des langues non tonales et il est bien connu que les langues sifflées s'adaptent à ces

structures linguistiques. Mais les sciences du langage ont montré également que la distinction tonale/non tonale est parfois difficile à faire pour certaines langues qui donnent un rôle important à la fois à l'accent et au timbre de la voyelle. C'est pourquoi, en élargissant le nombre de langues sifflées étudiées, nous avons montré qu'il existe un groupe intermédiaire de langues qui se caractérisent par la recherche d'un équilibre entre la Hauteur Brute et la Hauteur Fondamentale dans leur forme sifflée. Notre description typologique a donc fait émerger trois stratégies types de transposition qui dépendent de la structure phonologique de la langue d'origine. A l'intérieur de chaque groupe de langues ainsi définis, nous avons observé une grande diversité d'adaptations qui respectent les propriétés particulières de chaque langue. Nous avons conclu que les langues sifflées effectuent naturellement une partie du travail de description typologique qui consiste par définition à faire émerger des caractères généraux communs à des langues différentes. En outre, nous avons observé que le sifflement développe également une véritable description phonétique des caractéristiques vocales des éléments transposés: il définit des classes de consonnes, indique des similarités de timbre des voyelles et précise le comportement des tons.

Le deuxième point important à retenir est que tous ces aspects ont été développés dans les langues sifflées pour des raisons fonctionnelles visant à faciliter l'intelligibilité de l'interlocuteur. Nous avons montré que leur profil phonétique est avant tout destiné à la communication dans des conditions difficiles. C'est pourquoi les modulations de fréquence et d'amplitude qui les caractérisent rendent compte des points les plus importants mis en valeur par les domaines d'étude de l'analyse de la scène auditive, de la prosodie et de la parole dans le bruit. Il s'ensuit que les langues sifflées représentent des modèles pertinents d'étude des liens entre la perception et la production du langage, particulièrement en contextes contraignants d'écoute. Nous avons identifié plusieurs conséquences à cette réalité :

1. Les langues sifflées définissent une méthodologie pertinente d'approche de la prosodie : elle est à géométrie variable suivant les propriétés phonologiques qui organisent les indices acoustiques et les relations perceptives les plus saillants d'une langue.
2. Il est possible de tester des personnes ne connaissant rien des langues sifflées pour montrer que ces dernières exploitent certains aspects clefs des indices acoustiques de toute langue. Du même coup on peut réinterpréter les données d'autres expériences perceptives de la parole, et cela à la lumière de la stratégie sifflée choisie. Nous avons développé cette possibilité lors d'un test psycholinguistique de perception des voyelles sifflées espagnoles par des sujets Français qui a montré que ces derniers catégorisent les fréquences des voyelles sifflées comme les siffleurs eux mêmes. Nous avons conclu qu'un des éléments primordiaux de la perception des voyelles tient à la répartition de l'énergie sonore dans le spectre. La perception du timbre de la voix semble privilégier les zones les plus compactes et intenses du spectre fréquentiel de la parole. Suivant le type de voyelle, ce ne sont pas les mêmes formants qui seront concernés.

3. On observe pour les langues sifflées, comme c'est le cas pour la perception de la voix parlée, une augmentation des taux d'identification lorsqu'on passe des logatomes (netteté de 20-50%) aux mots (intelligibilité de 70%) et enfin aux phrases (90% d'intelligibilité). Lors de l'identification des phrases, les modulations de AM et FM, qui sont utilisées de manière extensive par les langues sifflées, sont des atouts à la fois pour donner une cohérence d'ensemble à la phrase et pour permettre d'identifier les éléments segmentaux dans le bruit.
4. Les éléments du signal qui parviennent à l'oreille du récepteur à grande distance sont *suffisants* pour atteindre un bon taux d'intelligibilité dans les situations usuelles de pratique des siffleurs, malgré la dégradation évidente du signal. Comme pour la voix parlée et la voix criée, la perte d'intelligibilité n'est pas un phénomène qui varie linéairement avec la distance. Une émergence *suffisante* du sifflement par rapport au bruit de fond contenu dans la bande de fréquence des sifflements est nécessaire pour permettre l'intelligibilité de la langue.

Ces formes de la langue qui peuvent s'appliquer à tout le vocabulaire et permettent un taux d'intelligibilité des phrases de plus de 90 % représentent donc à notre avis *une émergence informative de la langue*. Le sifflement reçu au niveau de l'oreille est alors la partie visible d'un *iceberg linguistique et acoustique* dont la partie immergée est dans le cerveau des acteurs du dialogue. Malgré tous ces atouts les langues sifflées sont toutes en situation de perte de vitalité, les siffleurs sont de moins en moins nombreux à être très compétents. La mort des langues sifflées emporte le regard alternatif qu'elles permettent de porter sur le phénomène du langage qui a encore beaucoup de choses à dire tant au niveau culturel que scientifique.

## Perspectives

Tout d'abord, nous avons commencé à montrer que ces langages se prêtent bien à la paramétrisation d'un signal ce qui laisse entrevoir des possibilités d'analyser de manière détaillée le rythme de la parole tout en permettant un parallèle avec les langages tambourinés qui soulignent de manière inharmonique ce même aspect. Les particularités du signal sifflé permettent l'usage des techniques et des logiciels d'analyse sonore développés en bioacoustique pour étudier les sifflements des animaux (oiseaux et dauphins principalement). Dès lors, une comparaison est également possible avec les productions sifflées animales ce qui permet de mieux cerner l'idiosyncrasie du sifflement humain (Busnel, 1966 ; Meyer, 2004). De plus, comme les langues sifflées ont un lien étroit avec le milieu naturel et réalisent une *optimisation bioacoustique du signal*, elles permettent de confronter la parole humaine à l'« *hypothèse de l'adaptation acoustique* » qui dit que la structure générale d'un signal de communication animal sera différente suivant les milieux d'habitat ou les niches écologiques (Mathevon et al, 2004).

Ces formes des langues peuvent également servir actuellement d'indicateur humain fiable pour évaluer la vitalité d'un biotope et donc à la fois de la biodiversité et des connaissances traditionnelles des populations

locales qui sont aujourd'hui essentielles si l'on veut comprendre ces milieux à protéger pour les générations à venir (Nettle et Romaine, 2003). L'approche que nous avons initiée sur l'oreille humaine, sur la dégradation de la parole sifflée béarnaise, sur le transfert de la technique à une nouvelle langue, sur l'analyse des profils sociolinguistiques de siffleurs et sur la comparaison typologique des langues sont autant de points qui peuvent participer à une réflexion sur l'évolution du langage humain, avec bien souvent des enseignements pour la phonétique (Demolin et Hombert, 1999). Nos données pourraient donc être complétées et analysées dans une perspective d'anthropologie cognitive. Ce développement semble souhaitable car les langues sifflées se sont développées dans des situations d'usages motivées par la survie d'un petit groupe d'individus, parfois dans des conditions très proches de nos ancêtres chasseurs-cueilleurs. A ce propos les langues sifflées ont le mérite de souligner le rôle fondamental que la contrainte du bruit ambiant a assurément joué dès l'origine du développement du langage humain.

De plus, la langue sifflée est *un style de parole* faisant partie intégrante du phénomène linguistique d'une langue locale. Dans ce cadre, les résultats d'une première étude de neurosciences utilisant la technique de résonance magnétique fonctionnelle (Carreiras & al 2005) ne sont pas étonnants. Ils ont montré que les aires neurologiques de la production (Broca) et de la perception (Wernicke) du langage sont activées chez des siffleurs entraînés, mais pas chez des sujets témoins, lors de l'écoute de phrases de Silbo : « *Our results show that the temporal regions of the left hemisphere that are usually associated with spoken-language functions are engaged during the processing of Silbo in experienced Silbadores.* » (Carreiras et al, 2005, p.31). Cela confirme l'intérêt de notre souci de toujours relier une langue sifflée à la voix parlée d'origine. Nos résultats suggèrent d'envisager une interprétation des données de cette équipe de recherche en comparaison avec celles d'études sur les bases neurales de l'écoute de la parole dans le bruit (Wong P. C.M. et al, 2005).

Enfin, dans le cas des langues sifflées comme dans la voix parlée, de nombreux éléments paralinguistiques sont liés à l'intelligibilité de la parole et relèvent aussi de la reconnaissance de formes auditives (reconnaissance du locuteur, de son sexe, de sa santé, de son état émotionnel). Nous l'avons constaté à plusieurs reprises avec nos informateurs siffleurs. Ce dernier aspect a été abordé de manière systématique dans deux études sur les langues sifflées par Busnel (1970) et Busnel et Classe (1976). Ils ont identifié dix facteurs paralinguistiques de la voix parlée et trois ont été testés sur les siffleurs : le sexe, l'âge et l'identité du locuteur qui semblent être les plus résistants, même si des performances inférieures à la voix sont obtenues. De telles recherches méritent un approfondissement suivant la méthodologie déjà mise en place. Une extension aux autres facteurs comme la valence émotionnelle nous paraît prometteuse car les siffleurs nous ont montré que c'est un paramètre qu'ils perçoivent facilement.

Le langage est souvent présenté comme porteur d'une redondance d'information afin de lutter contre le bruit. La particularité des langues sifflées de ce point de vue est de définir un cadre dans lequel cette redondance est réduite de manière ciblée, c'est leur intérêt principal pour tous ces domaines d'étude car les paramètres importants de la parole humaine y sont sélectivement soulignés de manière évocatrice.

## BIBLIOGRAPHIE

- Anderson S. (1978). Tone features. In V. Fromkin (ed.) *Tone a linguistic survey*. New York: Academic press, 133-176.
- Arripe R. (1985). *Les siffleurs d'Aas*. Imprimerie de la monnaie. Pau.
- Aubin T. (2004). Penguins and their noisy world. *Anais da Academia Brasileira de Ciências*, 76, 279-283.
- Barney G. L., Smalley W.A. (1953). *Third Report on Meo: Orthography and Grammar*. (unpublished)
- Basden G. T. (1966). *Among the Ibos of Nigeria*. Frank Cass & Co.Ltd.
- Batteau D. W. (1967). The role of the pinna in human localization. *Proceedings of the Royal Society London*, 168 (series B), 158-180
- Bausani A. (1974). *Le lingue inventate*. Ubaldini Editore : Roma.
- Beaudet J. M. (1997). *Souffles d'Amazonie, les orchestres tule des Wayãpi*. Nanterre, Société d'Ethnologie.
- von Békésy G. (1960). *Experiments in Hearing*, New York: Wiley
- von Békésy G. (1963). Hearing and complex sounds. *Journal of the Acoustical Society of America* 35, 588-601.
- Bert M. (2001). *Rencontre de langues et francisation : l'exemple du Pilat*. Thèse de doctorat, Université Lyon 2, Lyon.
- Bladon R. A. W. et Fant G. (1978). A two-formant model and the cardinal vowels. *STL-QPSR*, 1-1, 1-12.
- Blauert J. (1997). *Spatial hearing: the psychophysics of human sound localization*. Revised Edition. Cambridge, MA: MIT Press.
- Boilès-Lafayette C. (1973). Les chants instrumentaux des Tepehuas, un exemple de transmission musicale de signification. In *Musique en jeu* 12, 81-99.
- Boothroyd A. (1968). Statistical theory of the speech discrimination score. *Journal of the Acoustical Society of America*, 43, 362-367.
- Boothroyd A. (1978). Speech perception and sensorineural hearing loss. In *Audition management of Hearing impaired Children*. M. Ross et T.G. Giolas Ed. University Park, Baltimore.
- Borden G. J. & Harris K. S. (1980) *Speech Science Primer: Physiology, Acoustics and Perception of Speech*. Baltimore: Williams and Wilkins.

- Bosman A. J. (1989). *Speech perception by the hearing impaired*. Doctoral thesis, University of Utrecht, Utrecht, The Netherlands.
- Botte M. C. (1989). Perception de l'intensité sonore. In M. C. Botte, G. Canevet, L. Demany, C. Sorin. *Psychoacoustique et perception auditive*. Paris Inserm. Editions médicales internationales, 13-41.
- Bregman A. S. (1990). *Auditory scene analysis. The perceptual organization of sound*. Cambridge, Massachusetts: MIT Press.
- Bregman A. S. et Campbell J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 89, 244-249.
- Bregman A. S. et Pinker S. (1978). Auditory streaming and the building of timbre. *Canadian Journal of Psychology*, 32, 19-31
- Bronkhorst A. W. (2000). The cocktail party phenomenon: a review of research on speech intelligibility in multiple talker conditions. *Acustica*, 86, 117-128
- Bronkhorst A. W., Bosman A. J., Smoorenburg G. F. (1993). A model for context effects in speech recognition. *Journal of the Acoustical Society of America*, 93, 499-509.
- Brunet J. (1972). *Laos, musique du Nord*. Collection Musiques du monde. Galloway records GB 600531.
- Brusis T. (1973). Über die phonetische Struktur der Pfeifsprache "Silbo Gomero" dargestellt an sonographischen Untersuchungen. *Zeitschrift für Laryngologie*, 52, 292-300.
- Busnel R-G. (1964). *Documents sur une langue sifflée pyrénéenne*. SFRS. Paris.
- Busnel R-G. (1966). Information in the human whistled languages and sea mammal whistling. In K.S. Norris (Ed.), *Whales, Dolphins and porpoises*. Berkeley and Los Angeles: University of California Press, 544-565.
- Busnel R-G. (1967). *Documents sur une langue sifflée turque*. SFRS, Paris.
- Busnel R-G. (1970). Recherches expérimentales sur la langue sifflée de Kusköy. *Revue de Phonétique Appliquée* 14/15: 41-57.
- Busnel R-G (1974a). Bio-acoustique de la langue sifflée mazathèque. *Revue d'acoustique*, 29, 94-100.
- Busnel R-G (1974b). L'intelligibilité et certaines données psychophysiologiques sous-estimées. *Proceedings of the 8th International Congress on Acoustics*, London, 123-148.
- Busnel R-G. et Classe A. (1976). *Whistled languages*. Springer-Verlag. Berlin.
- Busnel R-G. et Siegfried J. R. (1990). *Parole, langages et langues sifflées*. SFRS, Paris.
- Busnel R-G, Moles A. and Vallencien B. (1962a). Sur l'aspect phonétique d'une langue sifflée dans les Pyrénées françaises. *Proceedings of the International Congress of Phonetical Science*, Helsinki. The Hage : Mouton, 533-546.
- Busnel R-G, Moles A. et Gibert M. (1962b). Un cas de langue sifflée dans les pyrénées françaises. *Logos*, 5, 76-91.

- Busnel R-G., Alcuri G., Gautheron B., Rialland A. (1989). Sur quelques aspects physiques de la langue à ton sifflée du peuple H'mong. *Cahiers de l'Asie du Sud-Est*, 26, 39-52.
- Calliope. (1989). *La parole et son traitement automatique*. Masson, Paris
- Canevet G. (1989). Audition binaurale et localisation auditive. In M. C. Botte, G. Canevet, L. Demany, C. Sorin. *Psychoacoustique et perception auditive*. Paris Inserm. Editions médicales internationales, 83-122.
- Campbell G. A. (1910). *Telephonic Intelligibility*. *Philosophical Magazine*, 19,6, 152-159.
- Carlson R., Fant, G. et Granström B. (1975). Two formant models, pitch and vowel perception. In Fant, G. and Taham, M.A.A. (Eds), *Auditory analysis and Perception of Speech*, London: Academic Press, 55-62.
- Carlson R., Granström B. et Fant, G. (1970). Some studies concerning perception of isolated vowels, *STL-QPSR 2-3/1970*, 19-35.
- Carlyon R. P. (1991). Discriminating between coherent and incoherent frequency modulation of complex tones. *Journal of the Acoustical Society of America*, 89, 329-340.
- Carreiras M., Lopez J., Rivero F., Corina D. (2005). Linguistic perception: neural processing of a whistled language. *Nature*, 433, 31-32
- Carrington J. F. (1949). *The talking drums of Africa*. The Carey Kingstate Press. London.
- Cartier M. (1989). Evaluation de la qualité de la parole codée et synthétisée. In *La parole et son traitement automatique*, Calliope (Ed), Masson, Paris.
- Chaix R., Austerlitz F., Khegay T., Jacquesson S., Hammer M.F., Heyer E., Quintana-Murci L. (2004). The genetic or mythical ancestry of descent groups: lessons from the Y chromosome. *American Journal of Human Genetics*, 15, 6, 1113-1116.
- Charalambakis C. (1994). A case of whistled speech from Greece, *Current Theory in linguistics*, 117, 389-396.
- Chavasse P. (1962). La notion d'intelligibilité. *L'audioprothésiste français*, 3.
- Cherry E. L. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, 25, 925-979.
- Chistovitch, L. A. (1985). Central auditory processing of peripheral vowel spectra. *Journal of the Acoustical Society of America*, 77, 789-805.
- Chistovitch, L. A. and Lublinskaja, V. V. (1979). The center of gravity effect in vowel spectra and critical distance between the formants: psychoacoustical study of the perception of vowel like stimuli. *Hearing research*, 1, 185-195.
- Chistovitch L. A. Sheikin, R. L. and Lublinskaja, V. V. (1979). Centres of gravity and spectral peaks as the determinants of vowel quality. In *frontiers of speech communication research* (ed. Lindblom and S. Öhman). New York: Academic Press, 143-157.
- Chowning J. M. (2000). Digital sound synthesis, acoustics and perception: a rich intersection. *Proceedings of the COST G-6 Conference ont Digital Audio Effects*, 1-6.
- Classe A. (1956). Phonetics of the Silbo Gomero. *Archivum linguisticum*, 9, 44-61.

- Classe A. (1957). The whistled language of La Gomera. *Scientific American*, 196, 111-124.
- Classe A. (1963). Les langues sifflées, squelettes informatifs du langage. In A.A. Moles A. et B. Vallancien (eds.) *Communications et langages*. Gauthier-Villars, Paris, 129-139.
- Cloarrec-Heiss (1997). Langue naturelle, langage tambouriné: un encodage économique (banda linda de Centrafrique). In *Diversité des langues et représentations cognitives*, 136-149.
- Cowan G. (1948). Mazateco whistle speech. *Language* 24, 280-286.
- Cowan G. (1976). Whistled Tepehua. In Sebeok T. and Umiker-Sebeok D.J. (Eds). *Speech surrogates: drum and whistle systems*, Mouton, Paris, 1400-1409.
- Darwin C. J. (1976). The perception of Speech. In *Handbook of Perception* (Vol. VII), EC Carterette, M. Friedman (Eds); Academic Press, 175-226.
- Darwin C. J. (1981). Perceptual grouping of speech components differing in fundamental frequency and onset time. *QJ Experimental Psychology*.33 A, 185-207.
- Darwin C. J. (1984). Perceiving vowels in the presence of another sound: constraints on formant perception. *Journal of the Acoustical Society of America*, 76, 1636-1647.
- Darwin C. J. et Sandell G. J. (1995). Absence of effect of coherent frequency modulation on grouping a mistuned harmonic with a vowel. *Journal of the Acoustical Society of America*, 97, 3135-3138.
- Darwin C. J., Ciocca V. et Sandell G. R (1994). Effects of frequency and amplitude modulation on the pitch of a complex tone with a mistuned harmonic. *Journal of the Acoustical Society of America*, 95, 2631-2636.
- Deacon A. B.(1976). Malekula. .In Sebeok T. and Umiker-Sebeok D.J. (Eds). *Speech surrogates: drum and whistle systems*, Mouton, Paris, 1199-1223.
- Demany (1989) La perception de la hauteur tonale. In M. C. Botte, G. Canevet, L. Demany, C. Sorin. *Psychoacoustique et perception auditive*. Paris Inserm Editions médicales internationales: 43-81.
- Demolin, D. et Hombert, J. M. (1999). The phonetics of the origin and evolution of speech, *Special issue of Evolution of Human Communication*, 3.
- Descola P. (2003). *Les lances du crépuscule, relation Haute Amazonie*. Collection Terres Humaines. Plon. Paris.
- Deutsch D. (1982). Grouping mechanism in music. In D. Deutsch, ed., *The psychology of music*, Academic Press, New York, 99-134.
- Diara (2002). *Oerzar: Life cycle of Zaqnyiq peoples, animals and plants*. Akha archaic text manual, 1.MPCD-SEAMP, Chiang mai, Thailand.
- Dimou A. L. & Dommergues J. Y. (2004). L'harmonie entre parole chantée et parole lue: comparaison des durées syllabiques dans un chant traditionnel grec. In. *Proceedings of Journées d'Etudes de la Parole*, 177-180.
- Dodane C. (2003). *La langue en harmonie, influences de la formation musicale sur l'apprentissage précoce d'une langue étrangère*. Thèse d'Etat. Université de Besançon.

- Dorian N. (1977). *Investigating Obsolescence. Studies in Language Contraction and Death*. Cambridge, Cambridge University Press.
- Dorman M. F., Cutting J., Raphael L. J. (1975). Perception of temporal order in vowel sequences with and without formant transitions. *Journal of Experimental Psychology: Human Perceptive Performances*, 2:121-129
- Dreher J. J. et O'Neill J. J. (1957). Effects of ambient noise on speaker intelligibility for words and phrases. *Journal of the Acoustical Society of America*, 29, 12, 1320-1323.
- Dugast I. (1976). Le langage tambouriné ou sifflé chez les Banen. In Sebeok T. & Umiker-Sebeok D.J. (Ed). *Speech surrogates: drum and whistle systems*. La Haye and Paris, Mouton, 708-741.
- Eboué F. (1935). La clef musicale des langages tambourinés et sifflés. *Bulletin du Comité d'étude historiques et scientifiques de l'Afrique occidentale française*, 18, 353-360.
- Encyclopedia Universalis (2002). *Base de donnée universalis en ligne* : <http://www.universalis-edu.com>.
- Esling J. H., Clayards, J., Edmondson, J. A., Fuyuan, Q., Harris, J. G. (1998). Quantification of pharyngeal articulations using measurements from laryngoscopic images, In *ICSLP-1998*, paper 0618.
- Essien A.J. (2000) *Contribution à la recherche sur la perception des tons du yoruba : évidences expérimentales à partir des tambours, des signaux de la parole et de la synthèse*. Thèse de Doctorat non publiée, Université Paris III.
- Everett D. L. (1986). Pirahã. In *Handbook of Amazonian Languages I*. Edited by D. Derbyshire and G. Pullum. The Hague: Mouton de Gruyter, 200-226.
- Fant G. (1970). *Acoustic theory of vowel production*. Mouton, s'Gravenhage, The Netherlands.
- Fletcher H. et Steinberg J-C. (1929). Articulation testing methods. *BSTJ*, 8, : 806-853
- Fletcher, H. et Munson, W.A. (1933). Loudness, its definition, measurement and calculation. *Journal of the Acoustical Society of America*, 5, 82-108.
- Fonagy Y. et Madgics K. (1963). Emotional patterns in intonation and music. In *Zeitschrift für Phonetik, Sprach wissenschaft und Kommunikations Forschung*, 16(1-3), 293-326.
- Fournel T. (2002). *Polyphonies Ari*. Collection Ocora/Radio France. Paris
- French et Steinberg J-C. (1947) Factors governing speech intelligibility. *Journal of the Acoustical Society of America*, 19, 90-119.
- Gardner M. B. (1969) Distance estimation of 0° or apparent 0° oriented speech signals in anechoic space. *Journal of the Acoustical Society of America*, 46, 47-53.
- Gay T. (1978). Effect of speaking rate on vowel formant movements. *Journal of the Acoustical Society of America*. 63, 223-230.
- Germain J. (1984). *Peuples de la forêt de Guinée*, Académie des sciences d'Outre-mer.
- Glatke, T. J. (1973). Elements of auditory physiology, In F.D. Minifie, T.J. Hixon & F. Williams (eds.) *Normal Aspects of Speech, Hearing and Language*. Englewood Cliffs: Prentice-Hall
- Gold B. Morgan N. (2000). *Speech and audio signal processing*. New York: Wiley.

- Golston C. et Wolfgang K. (1998). Mazatec onsets and nuclei. *International journal of American linguistics*, 64, 311-317.
- Green D. M. (1985) Temporal factors in psychoacoustics, In A. Michelsen (ed) *Time Resolution in Auditory Systems*, Springer-Verlag, Berlin, 122-140
- Grinevald C. (1997). Language Contact and Language Degeneration. In Florian Coulmas ed. *Handbook of sociolinguistics*, Oxford, Blackwell, 257-270
- Grinevald C. (2001). Encounters at the brink : linguistic field work among speakers of endangered languages. In *lectures on endangered languages 2, Kyoto ELPR*, 285-313.
- Grinevald C. (2003). Speakers and documentation of endangered languages. In *Langage documentation and description*, 11, Austin (Ed.), Londres, ELDP, SOAS.
- Hagège C. et Haudricourt A. (1978). *La phonologie panchronique*, Presses Universitaires de France, Paris.
- Haggard M. (1985). Temporal patterning in speech: The implications of temporal resolution and signal-processing, In A. Michelsen (ed) *Time Resolution in Auditory Systems*, Springer-Verlag, Berlin, 215-237
- Handley T. D. and Steer J. (1949). *Journal of speech and Hearing disorders*. 14, 363-368.
- Harris J. D. (1963). Loudness discrimination, *Journal .Speech Hearing Disorders Mon. Suppl.* 11, 1-63
- Harris J. D. (1952). The decline of pitch discrimination with time. *Journal of Experimental Psychology*, 43, 96-99.
- Hartmann W. M., McAdams S., Gerzso A., Boulez P., (1986). Discrimination of spectral density. *Journal of the Acoustical Society of America*, 79: 1915-1925.
- Hasler J. A. (1960). El lenguaje silbado. In Sebeok T. and Umiker-Sebeok D.J. (Eds), (1976) *Speech surrogates: drum and whistle systems*, Mouton, Paris, 1412-1425
- Hombert J. M. Ohala J. J., Ewan W. G.(1979). Phonetic explanations for the development of tones, *Language* 55, 37-58.
- Hombert J. M. et Maddieson I. (1998). A linguistic approach to language recognition. *UCLA working papers in phonetics*, 96, 106-118.
- Hubbard B. B. (1996). *Ondes et ondelettes, la saga d'un outil mathématique*. Ed. Pour la science. Belin.
- Huggins A. W. F. (1975). Temporally segmented speech. *Perception Psychophysics* 18, 149-157
- Huttar G. (1996) sum. Whistled speech. *Linguist List* 7.1166, Sun Aug 18 1996: Internet.
- Instituto Nacional de Antropología e Historia (1999). La gente de nuestra lengua. El grupo etnolingüístico chjota éña (mazatecos). INI (Ed.) *Configuraciones étnicas en Oaxaca perspectivas etnográficas para las autonomías*, 19-23.
- Jakobson R. (1931). Die betonung und ihre Rolle in der Wort- und Syntagmaphonologie. *Travaux du cercle linguistique de Prague IV*, reprinted in *Selected Writings I*. The Hague Mouton and co, 115-136.

- Johansson A. T. (2004). *Parametric modeling of cetacean calls*. PhD Thesis. University of Southampton, 2004.
- Junzo K. (1998). *La voix, étude d'ethno-linguistique comparative*. Ed. EHESS, Paris.
- Kersalé P. (1997). *Musique et chants des Hmong*. Collections peuples/VDE-Gallo.
- Klatt D. H. et Stefanski R. A. (1974). How does a mynah bird imitate human speech? *Journal of the Acoustical Society of America*, 55, 822-832.
- Koch R. et Piper H. M. (1979). Time segmentation in central analysis of complex signals, in Creutzfeldt et al, 1979, 128-132.
- Krauss, M. E. (1975). St. Lawrence Island Eskimo Phonology and Orthography. *Linguistics* 152, 39-72.
- Labouret H. (1923). Langage tambouriné et sifflé. *Bulletin du comité d'études historiques et scientifiques de l'Afrique Occidentale Française*, 6, 120-158.
- Lacerda M. (2004). Phonétique et phonologie du surui. Mémoire de DEA. Université libre de Bruxelles.
- Ladefoged P. et Maddieson I. (1996). *The sounds of the World's Languages*. Oxford: Blackwell.
- Lajard M. (1891). Le langage sifflé des canaries. *Bulletin de la société d'Anthropologie*, Paris, 469-483.
- Laoye H. H. (1976). Yoruba drums. In Sebeok T. & Umiker-Sebeok D. J. (Eds). *Speech surrogates: drum and whistle systems*. La Haye and Paris, Mouton, 705-707.
- Lass R. (1984). *Phonology: an introduction to basic concepts*. Cambridge: Cambridge University Press, 75-147.
- Lehiste I. (1970). *Suprasegmentals*. Cambridge: MIT Press.
- LePage E. L. (1987). A spatial template for the shape of tuning curves in the mammalian cochlea *Journal of the Acoustical Society of America*, 82, 155-164
- Leroy C. (1970). Étude de phonétique comparative de la langue turque sifflée et parlée. *Revue de Phonétique Appliquée* 14/15, 119-161.
- Lieberman P.(1975). *On the origins of language, an introduction to the evolution of human speech*. New York: Macmillan Publishing.
- Lindblom B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35, 1773-1781.
- Maddieson I. (1972). Tone system typology and distinctive features. Proceedings of the 7th International Congress of the Phonetic Science, ed. A. Rigault, R. Charbonneau. The Hague: Mouton. 957-61
- Mannell R. H. (1994). *The perceptual and auditory implications of parametric scaling in synthetic speech*, Ph D Dissertation, Macquarie University.
- Mathevon N., Aubin T., Dabelsteen T., Vieilliard J. M. E. (2004). Are communication activities shaped by environmental constraints in reverberating and absorbing habitats? *Anais da Academia Brasileira de Ciências*, 76, 259-263.

- Marin C. M. H. et McAdams S. (1991). Segregation of concurrent sounds II, Effect of spectral envelope tracing, frequency modulation coherence, and frequency modulation width. *Journal of the Acoustical Society of Acoustical*. 89, 341-351.
- Marler P. (1955). Characteristics of some animal calls. *Nature*, 176, 6-8.
- McAdams S. (1984). *Spectral fusion, spectral parsing and the formation of auditory images*, unpublished Ph.D. dissertation, Stanford University.
- Meyer J. (2004). Bioacoustics of human whistled languages: an alternative approach to the cognitive processes of language. *Anais da Academia Brasileira de Ciências*, 76, 2, 406-412.
- Meyer J. (2005). Whistled speech: a natural phonetic description of languages adapted to human perception and to the natural environment. *Proceedings of Interspeech 2005*, Lisboa, 49-52.
- Meyer J. et Gautheron B. (2005). Whistled speech and whistled languages. In *Encyclopedia of Language and Linguistics 2nd Edition*. Elsevier.
- Miller G. A., Heise G. A., Lichten W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, 41: 329-335.
- Miller J. D. (1989). Auditory perceptual interpretation of the vowel. *Journal of the Acoustical Society of Acoustical*. 85, 2114-2133.
- Moles A. (1970). Etude sociolinguistique de la langue sifflée de Kuskoy. *Revue de Phonétique Appliquée*, 14/15, 78-118.
- Moore D. (1998). Tonal system of the Gavião Language of Rondônia, Brazil, in Tupian Perspective. *Proceedings of cross-linguistic studies of Tonal Phenomena, ILCAA*. Tokyo.
- Moore B. C. J. et Glasberg B. R. (1983). Suggested formulae for calculating auditory filter bandwidths and excitation patterns, *Journal of the Acoustical Society of America*, 74, 750-753.
- Moreau M. L. (1997). La communication sifflée chez les Diola (Casamance Sénégal). *DiversCité Langues*, 2, édition internet : <http://www.quebec.ca/diverscite>.
- Moser et Dreher J. J. (1955). Effects of training on listeners in Intelligibility Studies. *Journal of the Acoustical Society of America*, 27: 1213-1219
- Moser et Dreher J. J. (1955). Phonemic confusion vectors. *Journal of the Acoustical Society of America*, 27: 874-880
- Nedzelinsky V. (1980). Sound pressures in the basal turn of the cat cochlea. *Journal of the Acoustical Society of America*, 68, 1676-1689
- Nekitel O. M. (1992). A perceptual analysis of Abu/-Wam whistled speech. *The language Game: Papers in Memory of Donald C. Laycock*. Ed. Tom Dutton, Malcom Ross, & Darrell Tryon, 299-311.
- Nettle D. et Romaine S. (1998). *Les langues, ces voix qui s'effacent*. Autrement Frontières.
- Nketia J. H. K. (1976). Surrogate languages of Africa. Sebeok T. & Umiker-Sebeok D.J. (Eds). *Speech surrogates: drum and whistle systems*. La Haye and Paris, Mouton, 825-864.
- Nooteboom S. G., Brox J. L., De Rooij J. J. (1978). Contribution of prosody to speech perception. In *Studies in the perception of language*. B. Flores d'Arcais ed., 75-108.

- Ohala J. J. (1991). The integration of phonetics and phonology. *12ème Congrès International des Sciences Phonétiques*, Aix-en-Provence, France, 1/5, 2-16.
- Oppenheim A. V. (1969). Speech analysis-synthesis system based on homomorphic filtering, *Journal of the Acoustical Society of America*, 45, 458-465.
- Oppenheim A. V., et Lim J.S. (1981). The importance of phase in signals. *Proceeding of IEEE* 69, 529-541.
- Oppenheim, A. V., Lim, J. S., Kopec, G. et Pohlig, S. C. (1979). Phase in speech and pictures, *Proceeding of IEEE, ICASSP 1979*, 632-637.
- Padgham M. (2004). Reverberation and frequency attenuation in forests - implications for acoustic communication in animals. *Journal of the Acoustical Society of America*, 115, 1, 402-410.
- Patel A. D., Peretz I., Tramo M., Labreque R. (1998). Processing prosodic and musical patterns: a neuropsychological investigation. *Brain and language* 61, pp 123-144
- Patel, A. D., Iversen, J.R., & Rosenberg, J.C. (2004). Comparing rhythm and melody in speech and music: The case of English and French. *Journal of the Acoustical Society of America*, 116:2645
- Patterson, R. D. (1976) Auditory filter shapes derived with noise stimuli, *Journal of the Acoustical Society of America*, 59, 640-654
- Pedersen O. J., Lyregaard P. E., Poulsen T. E. (1977). *The round robin test on evaluation of loudness level of impulse noise (report 22)*. Copenhagen, Technical University of Denmark, Acoustic Laboratory.
- Peretz I., Gagnon L., Hébert S., Macoir J. (2004). Singing in the brain: Insights from cognitive neuropsychology. *Music Perception*, 21, 3, 373-390.
- Peterson, K. E. et Barney H. I. (1952). Control methods used in the study of the vowels. *Journal of the Acoustical Society of Acoustical*. 36, 89-97.
- Picard F. (1991) Chine: le xiao, ou souffle sonorisé. Cahier de musiques traditionnelles. *Atelier d'ethnomusicologie Genève*:17-26.
- Pike K. L. (1948). *Tone Languages*. University of Michigan.
- Pike K. L. (1954, 1955 et 1960). *Language in Relation to a Unified Theory of the Structure of Human Behavior, Parts I, II, III*. Glendale, California: Summer Institute of Linguistics. 2nd Ed. The Hague: Mouton & co, 1967.
- Pike K. L. (1970). The role of nuclei of feet in the analysis of tone in Tibeto-Burman Languages of Nepal. *Prosodic Feature Analysis*, P. R. Léon, G. Faure and A. Rigault eds.
- Pike, K. L. (1945). *The Intonation of American English*. Ann Arbor, Michigan: University of Michigan Press.
- Pike, K. L. et Pike, E. (1947). Immediate constituents of Mazatec syllables. *IJAL* 13, 78-91.
- Plomp R. (1965). Detectability Threshold for combination tones. *Journal of the Acoustical Society of America* 37: 1110-1123.
- Plomp R. (1967). Pitch of complex tones. *Journal of the Acoustical Society of America*, 41,1526-1533.

- Plomp R. (1975). Auditory analysis and timbre perception. In Fant, G. and Tatham, M. A. A. (Eds), *Auditory analysis and Perception of Speech*, academic Press, London, 7-22.
- Potter R. K. and Steinberg J. C. (1950). Toward the specification of speech. *Journal of the Acoustical Society of America*, 807-820.
- Ramus, F. (1999). *Rythme des langues et acquisition du langage*. Thèse d'état. EHESS.
- Rialland A. (2003). A New Perspective on Silbo Gomero, dans Solé M.J., Recasens D. & Romero J., eds., *Proceedings of the 15th ICPHS conference*, Barcelona, 2131-2134.
- Risset J. C. (1968). Sur certains aspects fonctionnels de l'audition, *Annales des Télécommunications*, 23, 91-120.
- Risset J. C. (1994). Quelques aspects du timbre dans la musique contemporaine. In A. Zenatti Ed. *Psychologie de la musique* PUF. Paris. 87-114.
- Risset J. C. (2000). Perception of musical sound: simulacra and illusions. In Tsumotu Nakada (ed.), *Integrated Human Brain Science: Theory, Method, Application* (Music), Elsevier, 279-289.
- Ritzenthaler R. E. et Peterson F. A. (1954). Courtship Whistling of the Mexican Kickapoo Indians. *American Anthropologist*. 56: 1088-1089.
- Rosner B. S. et Pickering L. B. (1994). *Vowel Perception and Production*. Oxford science publications. Oxford.
- Sapir E. (1921). *Language. Introduction to the Study of Speech*. New York: Harcourt.
- Savin H. et Bever T.G. (1970). The non perceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behaviour*. 3, 295-302 .
- Schneider M.(1976). Zur trommelsprache der duala. In Sebeok T. & Umiker-Sebeok D.J. (Eds). *Speech surrogates: drum and whistle systems*. La Haye and Paris, Mouton, 669-678
- Sebeok T. et Umiker-Sebeok D.J. (1976). *Speech Surrogates: Drum and Whistle Systems*. La Haye and Paris, Mouton
- Segui J. (1989). L'accès au lexique : données expérimentales et modèles. In *La parole et son traitement automatique*. Callioppe (Ed.). Masson, Paris, 215-231.
- Shannon C. E. (1951). Prediction and entropy of printed english. *The Bell System Technical Journal*, 30, 50-64
- Shannon C.E.(1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379-423.
- Shannon R. V., Zeng F. G., Kamath V., Wygonsky J., Ekelid M. (1995). Speech recognition with primarily temporal cues. *Science* 270, 303- 304.
- Shepard R. N. (1968) Approximation to uniform gradients of generalization by monotone transformation of scale. In D. I. Moskosky (Ed.) *Stimulus generalization*. Stanford University press, Stanford, 343-390.
- Silverman D. (2004). Pitch discrimination during breathy versus modal phonation. In Local J., Ogden R. and Temple R. Ed. *Phonetic interpretation papers in Laboratory Phonology*, VI, Cambridge: Cambridge University Press, 293-321.

- Sorin C. (1989). Perception de la parole continue. In M. C. Botte, G. Canevet, L. Demany, C. Sorin. *Psychoacoustique et perception auditive*. Paris Inserm Editions médicales internationales: 123-139.
- Stern T. (1957). Drum and whistle languages: an analysis of speech surrogates. *American Anthropologist*, 59, 487-506.
- Stevens K. N. (1998). *Acoustic Phonetics*. Cambridge: MIT Press.
- Stevens K. N. et House A. S. (1963). Perturbation of vowel articulations by consonantal context: an acoustic study. *Journal of Speech and Hearing Res.* 6: 111-128.
- Stevens S. S. et Davis H. (1938). *Hearing: Its Psychology and Physiology*. New York: Wiley.
- Taylor W. L. (1953). Close procedure: a new tool of measuring readability. *Journalism Q.* 30: 415-433.
- Tisseyre A. (1998), Moulinier A, Rouard Y., (1998) Intelligibility in various rooms : comparing its assessment by (RA)STI measurement with a direct measurement procedure. *Applied Acoustics*, 53, 1-3, 179-191
- Thierry E. (2002). *Les langages sifflés*. Mémoire de DEA Philologie et linguistique. EPHE.
- Trujillo R. (1978). *El Silbo Gomero: analisis linguistico*. Ed. I. Canaria-I. Andres Bello, Tenerife.
- Unesco (2003). *Language Vitality and Endangerment*. Unesco Intangible Cultural Heritage Unit's Ad Hoc Expert Group on Endangered Languages. Unesco. Paris.
- Van Norden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. Ph.D Thesis, Eindhoven University of Psychology, Eindhoven.
- Van Tassel D. J., Soli S. D., Kirby V. M., Widin G. P. (1987). Speech waveform envelope cues for consonant recognition. *Journal of the Acoustical Society of America*, 82, 1152-1161.
- Verfaillie V. (2003). *Effets audionumériques adaptatifs : théorie, mise en œuvre et usage en création musicale numérique*. Thèse de doctorat. Université Aix-Marseille II.
- Voorhis P.H. (1971). Notes on the Kickapoo whistle speech. *International Journal of American Linguistics*, 37/4.
- Warren R. M. (1970). Restoration of missing speech sounds. *Science*, 167, 392-393.
- Warren R. M. et Warren R. P. (1970). Auditory illusions and confusions. *Scientific American*, 223, 30-36.
- Wedekind K. (1981). A Six-Tone language in Ethiopia: Tonal Analysis of Benč4non4. *Journal of Ethiopian studies*, April 1981, 129-156.
- Wessell D. L. (1979). Timbre space as a musical control structure. *Computer Music Journal*, 3, 2, 42-52.
- Wilken G. C. (1979). Whistle speech in Tlaxcala (Mexico), *Anthropos* 74, 881-888.
- Wong P. C. M, Lee K. M., Parrish T. B. (2005). Neural Bases of Listening Speech in Noise. *Proceedings of the Interspeech 2005 Conference*, Lisboa, 1745-1748.
- Xian-Ming Y. (2002). The talking musical instruments of the province of Yunnan. *Ethnomusicology research seminar*. London: Goldsmith University.

- Xirometis N. et Spyridis C. (1989) A whistling language in the village Antias in the greek island of Evia, *Glossologia* 7-8: 219-224
- Xirometis N. et Spyridis H. C. (1994). An acoustical approach to the vowels of the village Antias in the Greek Island of Evia, *Acustica*, 5, 425-516.
- Yip M. (2002). *Tone*. Cambridge : Cambridge University Press.
- Zemp H. (1996). *Les voix du monde, une anthologie des expressions vocales*. Collection Musée de l'homme. Chant du Monde/Harmonia Mundi.
- Zenatti A. (1994). *Psychologie de la musique*. PUF. Paris. 87-114.
- Zeng F. G, Nie K, Stickney G. S., Kong Y. Vongphoe M., Bhargave A., Wei C., Cao K. (2005). Speech recognition with amplitude and frequency modulation. *Proceedings of the National Academy of Sciences*, 102, 7, 2293-2298.
- Zwicker E. et Fastl H. (1990). *Psychoacoustics: Facts and Models*, Springer-Verlag: Berlin
- Zwicker E. (1982). *Psychoakustik*. Springer-Verlag: Berlin