



**HAL**  
open science

# L'adaptation des langues humaines aux contraintes acoustiques du territoire dans le cas des modalités de parole à distance

Julien Meyer

► **To cite this version:**

Julien Meyer. L'adaptation des langues humaines aux contraintes acoustiques du territoire dans le cas des modalités de parole à distance. Actes de la Conférence Annuelle 2022 sur l'activité scientifique du Centre d'Etudes Francoprovençales. Grammaires de l'espace et relation, Région Autonome de la Vallée d'Aoste, A paraître. hal-04389268v1

**HAL Id: hal-04389268**

**<https://hal.science/hal-04389268v1>**

Submitted on 16 Jan 2024 (v1), last revised 18 Jun 2024 (v2)

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

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



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

## **L'adaptation des langues humaines aux contraintes acoustiques du territoire dans le cas des modalités de parole à distance**

Julien Meyer

Université Grenoble Alpes, CNRS, GIPSA-Lab, Grenoble, France

julien.meyer@cnrs.fr

### **Résumé**

L'acquisition et la compréhension des langues humaines s'adaptent à différents types de variations sonores intra- ou inter-individuelles. Cette flexibilité est aussi exploitée à travers différents types/modalités de parole. La forme parlée dite 'usuelle' c'est à dire sans instruction ou forçage particulier représente le moyen d'expression le plus commun, elle est souvent nommée également 'parole modale' par les phonéticiens en référence au mode de phonation. Cependant, à plus grande distance ou dans le bruit, il est possible et fréquent de forcer sur les cordes vocales jusqu'à parfois utiliser le mode crié de la parole. Plus extrême encore, plusieurs populations dans le monde ont également développé des modes sifflés et/ou tambourinés qui transforment de manière encore plus drastique les indices phonétiques saillants de la parole ainsi que la dynamique du discours. Ces 'états modifiés de la parole' qui servent à parler à distance constituent de véritables systèmes de télécommunication langagiers adaptés au milieu de propagation. Nous verrons ici comment le cri, le sifflement et les percussions portent efficacement les messages de parole en milieu extérieur et exploitent des stratégies acoustiques différentes qui sont déjà en germe dans la parole vocale usuelle.

**Mots clés** : parole à distance, parole criée, parole sifflée, parole tambourinée, langage et écologie.

## **1. Introduction**

Le langage humain, sous sa forme acoustique, a la souplesse d'être émis de manières très différentes selon les contextes de communication. Si les interlocuteurs sont suffisamment proches, les phrases qu'ils échangent sont simplement parlées – on dit aussi prononcées en parole 'modale usuelle'. Ce type de parole - ou modalité de parole - est le plus commun et le plus confortable pour dialoguer en utilisant la voix humaine car ni les articulateurs ni les cordes vocales ne sont forcés. Les phrases peuvent aussi être prononcées en parole 'forte' ou même criée si les interlocuteurs sont situés à distance ou dans un environnement bruyant. En revanche, si la conversation doit rester discrète, elle pourra être chuchotée. Cette aptitude à adapter la modalité de la parole et la puissance du son au contexte a pour contrepartie de déformer les sons de la parole usuelle modale. Elle n'est possible que parce que nous sommes doués d'une certaine flexibilité perceptive : nous avons la capacité de reconnaître un stimulus comme appartenant à une catégorie même s'il s'écarte légèrement de la forme la plus commune dans laquelle il est utilisé. Dans le domaine de la parole, cette flexibilité correspond à la capacité de reconnaître des mots ou d'autres unités linguistiques avec de nouvelles prononciations (Bent, 2015), comme dans des dialectes ou des accents inconnus, ou lors de l'écoute de parole chantée, criée, chuchotée. C'est une compétence essentielle pour l'acquisition du langage chez l'enfant, mais aussi dans la reconnaissance de la parole dégradée, par exemple dans du bruit ; ou encore pour apprendre une deuxième langue.

Depuis longtemps les humains exploitent cette possibilité pour des usages complémentaires de la forme parlée usuelle, en particulier pour communiquer dans des situations spéciales. Il existe ainsi de nombreuses techniques traditionnelles de renforcement de la parole qui utilisent des transformations encore plus radicales que le chant ou le cri, soit pour des

performances artistiques et/ou rituelles, soit pour faire porter la parole encore plus loin. Une première catégorie utilise des instruments modificateurs de voix placés devant la bouche - des ancêtres des portes voix ou des masques d'identité vocale modernes. Ces outils traditionnels dans lesquels des textes sont parlés ou chantés sont faits de tubes en bois, de Calebasses, de coquilles de noix de coco ou de conques (voir Niles 2010). Dans ces cas, les caractéristiques acoustiques primaires de la voix sont intactes avant transformation dans un résonateur externe car les sons sont toujours produits avec les cordes vocales et articulés dans le conduit vocal. Mais il existe une autre catégorie de type de parole qui utilise des transformations sonores réalisées directement à la source : soit en sifflant les phrases, soit en les produisant et articulant avec un instrument de musique. C'est dans cette catégorie que l'on rencontre les systèmes de communication à distance anciens utilisés en extérieur, comme la parole sifflée et la parole tambourinée. Ce sont ceux-là qui vont nous intéresser dans cet article car l'environnement devient alors le support du signal de parole. Les caractéristiques écologiques du territoire ont alors un impact non négligeable sur ce type de communication langagière. Dans cet article, nous expliquons dans un premier temps pourquoi la parole usuelle modale est limitée pour la communication à distance et quelles sont les contraintes environnementales les plus communes qui justifient une adaptation en d'autres types/modalités de parole. Ensuite, nous décrivons plus précisément comment la parole s'adapte à trois modalités différentes - la parole criée, la parole sifflée et la parole tambourinée, et comment ces modalités adaptent la parole aux contraintes du terrain. Notamment, nous verrons comment les sifflements et les percussions sont adaptés à la pression exercée par les contraintes acoustiques environnementales les plus communes rencontrées en milieu rural et résistent bien mieux à la propagation en milieu extérieur. En transformant les indices phonétiques de la parole de manière aussi drastique, de véritables systèmes de télécommunication langagiers ont évolué.

## **2. Contraintes environnementales et leur impact sur la parole modale**

Au quotidien, la compréhension de la parole vocale usuelle implique presque toujours une écoute dans un bruit de fond plus ou moins fort et masquant. Les recherches sur la parole dans le bruit ont d'ailleurs montré que les signaux issus d'une source vocale incorporent de nombreuses propriétés acoustiques qui contribuent à compenser les distorsions du signal et les interférences du bruit qui sont inévitables (Assmann et Summerfield 2004, pour une revue de littérature) et, qu'en parallèle, notre système perceptif et cognitif est adapté à surmonter des dégradations importantes impliquant des distorsions et/ou des fragmentations du signal de parole (ref). Malgré ces atouts, la parole vocale usuelle perd rapidement de son efficacité quand la distance qui sépare les interlocuteurs augmente, et ceci dès les premiers mètres (Meyer et al. 2013). En effet, l'impact du bruit sur les signaux de parole issus de la voix et sur les performances d'intelligibilité augmente significativement avec l'éloignement entre l'auditeur et le locuteur. Ceci reste valable quand la perturbation se limite aux conditions les plus simples de propagation dans les bruits les plus calmes rencontrés en extérieur (c.a.d. combinant un bruit de fond naturel de bas niveau au simple effet d'atténuation d'amplitude dû à la distance, cf. Meyer et al 2015). Or, ce type de conditions est non seulement assez commun dans la vie de tous les jours mais il correspond en outre aux contraintes environnementales de base qui ont imprégné la communication parlée tout au long de l'histoire humaine et que l'on retrouve encore de manière sous-jacente en tous lieux. En effet, les bruits de fond naturels ont tous des propriétés de base communes caractérisées par une répartition non uniforme des fréquences bruitées, riche en basses fréquences car celles-ci voyagent plus loin sans être atténuées, et avec des niveaux de puissance plus faibles à mesure que les fréquences sont plus élevées.

A ces propriétés sous-jacentes ou conditions de base en milieu extérieur s'ajoutent une grande diversité de contextes et d'ambiances sonores perturbateurs qui doivent fréquemment être surmontés pour permettre une bonne efficacité de communication. C'est pourquoi, avant de décrire les différentes formes d'adaptation de la parole aux conditions de communication parlée à distance, nous dressons ici une liste représentative de l'éventail des principaux effets perturbateurs sonores non artificiels rencontrés sur les territoires habités de la planète. Ceux-ci résultent toujours de l'impact conjugué de bruits de fond concurrents et de dégradations sonores supplémentaires occasionnées par des obstacles comme la végétation, les cours d'eaux ou les reliefs :

Omniprésents en milieu extérieur, les '*bruits ambiants*'/'*paysages sonores*' ruraux sont assez variables, même lorsqu'ils n'incluent pas de source sonore mécanique anthropique. Ceci explique d'ailleurs pourquoi la plupart des études sur le langage humain essayent de les éviter en enregistrant les locuteurs de près, et loin des sources sonores concurrentes. Ces bruits naturels varient en fonction de la situation géographique, du type de terrain et de la végétation qui caractérisent le milieu de propagation, et dépendent aussi des conditions météorologiques ainsi que des sources sonores environnementales concurrentes telles que les cris d'animaux (biophonie) ou les sons du milieu physique (géophonie, par exemple due aux rivières ou à la mer). Un bruit de fond fort, proche du récepteur et similaire au signal - ou 'bruit de premier plan' - crée des problèmes plus importants qu'une source de bruit calme, distante et dissemblable (Bradbury et Vehrencamp 1998). Les types de bruits de premier plan qui peuvent interférer le plus avec les communications humaines sont les communications acoustiques animales et les effets acoustiques que font le vent ou les bruits d'eau au contact de la végétation ou du terrain (cf. Busnel et Classe, 1976). Les fréquences de la parole humaine qui contribuent le plus à son intelligibilité se situent entre 1 et 4

kHz. Les bruits ambiants non animaux – dits abiotiques - les plus communs ne sont intenses qu'à des fréquences inférieures à 1 kHz et n'en perturbent que les fréquences les plus basses (Meyer et al., 2013) (exceptées les quelques sources abiotiques liées à la pluie, aux courts d'eau et au vent qui couvrent toutes les fréquences). Les humains bénéficient également de leur capacité à concentrer leur attention sur la dynamique d'un signal spécifique, sur la direction de sa source et sur sa distance (Bergman 1990 ; Kondo et al. 2017). Comme la plupart des bruits liés aux animaux sont intermittents, ceci permet d'écouter dans les « trous » et de saisir une partie du signal pour reconstruire cognitivement la partie masquée (Hoen et al. 2007 ; Varnet et al. 2012). Les arthropodes cependant produisent des signaux avec généralement des motifs réguliers, et parfois en chorus et donc provenant de nombreuses directions différentes en simultanée, résultants en un signal général quasi-stationnaire dans la gamme relativement aigüe de la voix 2,5–10 kHz. Ils peuvent donc être très gênants pour la communication vocale humaine, surtout en forêt tropicale et équatoriale où ils sont presque omniprésents malgré leurs variations saisonnières et circadiennes. D'autre part, certaines espèces d'oiseaux, de batraciens et de mammifères produisent aussi des chants qui utilisent les mêmes bandes de fréquences que la parole humaine, mais ils sont beaucoup plus intermittents, saisonniers et/ou plus spécifiques à un biotope (Luther 2008, Meyer 2020). Notamment, les mammifères et les oiseaux dont les productions sonores sont généralement beaucoup plus courtes et moins stationnaires que les insectes et consistent la plupart du temps en la répétition de motifs modulés stéréotypés.

Quant aux *effets de la distance* qui modifient le signal acoustique à mesure qu'il se propage dans le milieu, les humains y sont implicitement habitués, en particulier les habitants des territoires ruraux. Outre le phénomène physique d'atténuation sphérique

déjà cité, omniprésent et entraînant une perte d'intensité sonore proportionnelle à la distance (d'environ 6 décibels à chaque fois que la distance de propagation double), les autres principaux facteurs qui modifient fortement les sons sont ceux liés à l'orographie et à la transmission à travers la végétation (Wiley et Richards 1978, 1982 ; Blumenrath et Dabelsteen 2004) : (a) la réflexion et l'absorption sur le sol et la canopée, sachant que l'effet de sol dépend fortement de la topographie et du terrain. (b) La déformation des sons par filtrage sélectif des fréquences (par réverbération sur le feuillage et les troncs, ou par effet des turbulences atmosphériques, Wiley et Richards 1982 ; Dabelsteen et al. 1993). Les gradients de température atmosphérique jouent également un rôle mais sont difficiles à quantifier, ils dépendent de l'intensité du soleil, du vent et de l'humidité et créent des couches d'air changeant la vitesse du son dans l'air. Ce n'est que la nuit et par temps de pluie que ces gradients de température s'égalisent, ce qui explique de fortes différences d'ambiances et de comportements acoustiques entre le jour et la nuit, auxquels humains et animaux s'adaptent.

### **3. Différentes stratégies d'adaptation pour parler à distance : crier, siffler, tambouriner**

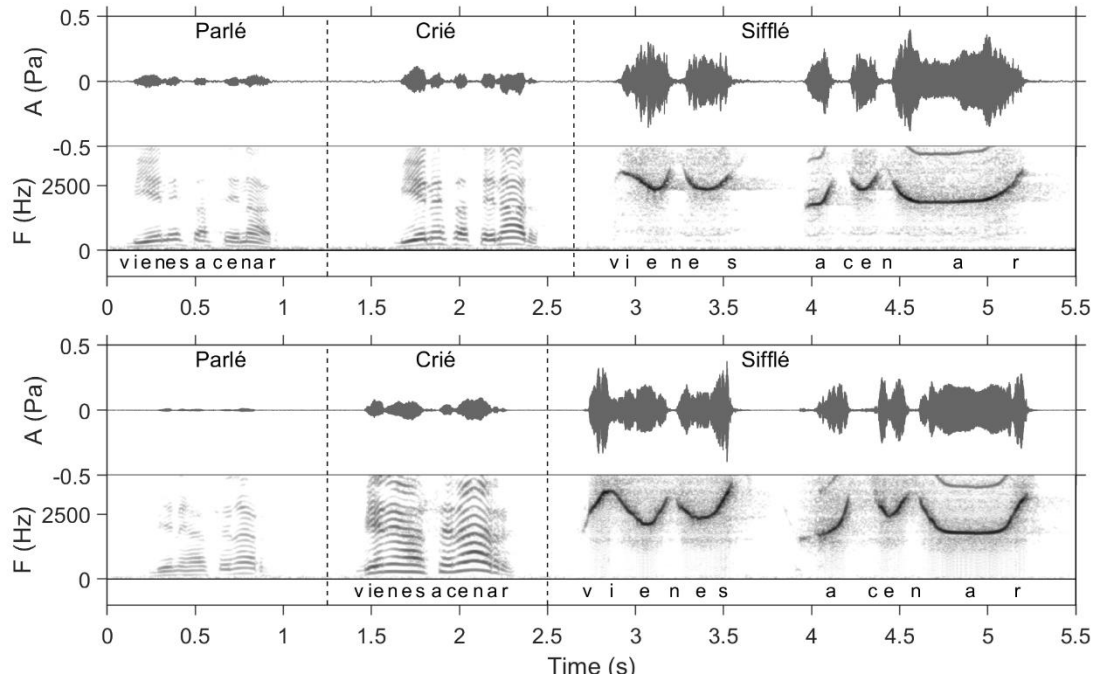
Les territoires ruraux représentent le cadre dominant de l'évolution au cours de l'histoire humaine, et l'étude des échanges langagiers dans ce type de cadre révèle qu'il peut être très utile de communiquer à distance avec toute la richesse communicative d'une langue locale, y compris à longue distance. Par exemple, pour signaler la nature d'un danger ou donner une nouvelle importante à quelqu'un situé de l'autre côté d'une rivière ou d'une vallée, ou pour informer un interlocuteur difficile à localiser car il est dans la forêt, ou chez lui. De telles situations impliquent à la fois des facteurs environnementaux et sociaux. Dans de nombreuses populations du monde, la répétition fréquente au quotidien de ce type de situations a favorisé



l'évolution des modalités de parole dédiées à la transmission à distance. Ces modalités sont de trois types principaux, avec pour base soit (1) la voix en mode crié, soit (2) le sifflement, soit (3) les percussions (Meyer 2020). La parole criée est une réponse universelle qui est donc la plus connue et étudiée mais elle n'est pas réservée aux communications à distance, alors que la parole sifflée et la parole tambourinée sont moins répandues et spécifiquement dédiées à la distance. Ainsi, ces deux derniers types de parole ont l'originalité de représenter de véritables systèmes de télécommunication anciens qui consistent à transformer certains traits linguistiques saillants des phrases d'une langue existante en des sons très différents de la voix humaine. Ces pratiques spéciales sont parfois aussi nommées « langue/langage sifflé(e)/tambouriné(e) » mais il est important de préciser que dans cet article nous ne parlons pas de langues ou langages indépendant(e)s des langues parlées<sup>1</sup> mais bien de mode ou versions sifflée/tambourinée de la parole (tout comme il existe un mode crié). Ceci veut dire par exemple que le grec, le wayãpi ou l'espagnol sifflés sont toujours du grec, du wayãpi ou de l'espagnol pour lesquels les sons de la parole vocale usuelle sont simplifiés en sifflements (Figure 1).

---

<sup>1</sup> Sont exclus de cet article les "idéogrammes lexicaux" qui sont parfois sifflés ou tambourinés mais ne ressemblent pas phonétiquement à une langue parlée puisqu'ils créent des codes parallèles indépendants (Stern 1957, Zemp & Kaufmann 1969). Sont aussi exclues les représentations secondaires de la parole telles que le code Morse (qui s'appuie sur une codification d'alphabet écrit).

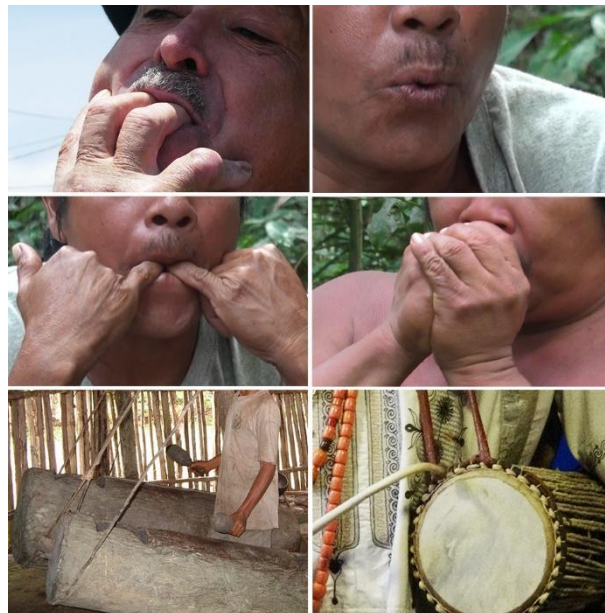


**Figure 1 :** Forme d'onde (A(Pa)) et spectrogramme (F(Hz)) montrant une phrase espagnole parlée (cible à 3m), crieée (cible à 70m), sifflée (cible à 140m) par deux locuteurs différents (en haut et en bas). La phrase en espagnol est « ¿vienes a cenar ? », prononcée [bjenes a senar] (ce qui veut dire 'Viens-tu dîner ?'). Cette phrase, ici sifflée par deux jeunes professeurs de parole sifflée est une des plus communément donnée en exemple par les siffleurs traditionnels (cf. Meyer et Manfredi 2024). La forme sifflée reflète différents aspects phonétiques de la parole modale tout en étant une simplification fréquentielle du signal acoustique parlé : on voit sur cet exemple que les voyelles sont les parties les plus stables du signal et que différentes qualités vocaliques sont sifflées à différentes hauteurs de fréquences (/i/ en moyenne supérieur à /e/ supérieur à /a/). De plus, les fréquences vocaliques sont modulées par coarticulation avec les consonnes avoisinantes et les formes fréquentielles qui en résultent reflètent les changements de taille de la cavité orale (montant dans les aigus pour les coronales par exemple).

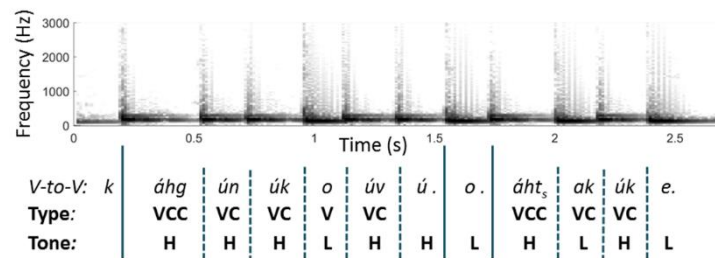
Dans le cas du sifflement, l'air est utilisé comme source sonore à la place des vibrations des cordes vocales et les mots sont articulés en sifflant, ce qui limite un peu les mouvements articulaires du conduit vocal (Figure 2). Il en résulte une mélodie qui imite des sons saillants - acoustiquement et perceptivement - des syllabes parlées (Figure 1). Sur cette base réduite, les mots et les phrases sifflés restent cependant compréhensibles, mais seulement pour les locuteurs qui sont entraînés à reconnaître leur langue sous cette forme (e.g., Moles 1970, Busnel 1970) et avec un taux d'intelligibilité moindre que la version parlée (Meyer 2015). Cette difficulté à reconnaître les sons de la parole quand ils sont profondément modifiés par une articulation spéciale, accompagnée en outre d'un changement sonore

important, existe aussi dans certains registres de chant, comme en voix soprano par exemple (Novák and Vokřál 2000, Gregg and Scherer 2006).

Dans le cas de *la parole tambourinée*, il s'agit d'adapter la parole vocale usuelle pour permettre une communication à très longue distance à travers des rythmes et des mélodies portées par les percussions. La source du signal est un tambour (Figure 2) et la parole est codée en séquences de battements de tambour (Figure 3) qui restent compréhensibles pour les récepteurs entraînés (Stern, 1957, Sebeok et Umiker-Sebeok 1976).



**Figure 2 :** Différentes techniques de sifflement (4 photos du haut) et de tambours (2 photos du bas). Les différentes techniques sifflées servent à parler en sifflant dans des situations un peu différentes. La protrusion bilabiale (en haut à droite, montrée par un siffleur wayãpi de l'Oyapock en Amazonie) est très répandue dans le monde car c'est aussi la plus adaptable (de discret à relativement puissant). Chez les Wayãpi, elle sert pour s'appeler ou se saluer sur le fleuve, dans le village ou en forêt (voir aussi l'enregistrement Figure 5) et est aussi employée pour ne pas effrayer les animaux lors de la chasse. La technique de résonance dans les mains (au milieu à droite, par un siffleur wayãpi) est également discrète en forêt et produit un sifflement plus puissant et de fréquences plus graves que la technique bilabiale. Les deux autres techniques utilisant deux doigts insérés dans la bouche, permettent de couvrir de plus grandes distances (en haut à gauche par un siffleur de l'île de El Hierro aux Canaries, et au milieu à gauche par un siffleur Wayãpi). Pour les tambours, sont présentés ici le double tambour à troncs évidé d'Amazonie appelé Manguaré qui sert à envoyer des messages longues distances chez les Bora/Miraña de la frontière Pérou/Colombie (en bas à gauche, cf. Figure 3) ; ainsi que le 'tambour parleur' Yorùbá dùn-dún utilisé au Nigéria principalement pour réciter des proverbes et passer des messages à une audience réunie pour une occasion particulière [Photos : Julien Meyer (Tous droits réservés)]



**Figure 3** : Spectrogramme de la phrase tambourinée de Bora /káʔgúnúkòúβú ò áʔtsàkúnè/ « Je termine la cahuana (boisson à base d'amidon de manioc) ». Les lignes sous la figure fournissent la transcription phonologique, le type de combinaisons Consonne (C) et Voyelles (V) présentes dans la segmentation en intervalles 'voyelle à voyelle' (V-toV), ainsi que les types de tons (L pour Low et H pour High) (en langue bora un ton haut est transcrit avec un accent aigu un ton bas avec un accent grave). [Adapté de Seifart et al 2018].

Ces trois modalités de parole à distance représentent différents types d'adaptation acoustique avec différents niveaux d'efficacité pour résister à la dégradation sonore imposée par les contraintes acoustiques environnementales. La réponse la plus courante au besoin de communiquer à distance est aussi la plus simple : conserver l'usage de la voix en augmentant simplement la puissance du signal, c'est-à-dire crier. *La parole criée* ne change donc pas la source sonore mais renforce juste l'énergie dans les cordes vocales (et donc dans les formants des voyelles et de certaines consonnes) pour optimiser le rapport signal sur bruit au niveau du récepteur du signal. Le locuteur ajuste sa voix en fonction des contraintes du territoire le séparant de l'auditeur en augmentant tacitement les niveaux d'amplitude, ce qui s'apparente à l'effet Lombard (Lombard 1911). En parallèle il y a une tendance à réduire le débit de parole (et donc un allongement des syllabes) et à augmenter la fréquence de vibration des cordes vocales avec l'accroissement de la distance (Meyer 2008 ; Cheyne et al. 2009 ; Fux 2012, Meyer et al 2018). En raison des limites physiologiques des cordes vocales, qui fatiguent rapidement à l'effort, la plupart des dialogues criés sont brefs et ne permettent pas d'atteindre de très grandes distances, hormis dans des configurations topographiques particulières, comme des vallées entourées de falaises qui guident les ondes (il arrive d'ailleurs que des toponymes et des témoignages locaux évoquent ces lieux propices à la communication criée).

*La parole sifflée*, quant à elle, est une généralisation à toute la phrase des avantages des sifflements. Ceux-ci sont parmi les signaux les plus simples et les plus puissants qui peuvent être produits sans outils par le tractus vocal et par le corps humain en général. La simplification que la parole sifflée opère par rapport à la parole modale opère essentiellement au niveau des fréquences. Il s'agit d'utiliser la bande étroite d'une simple fréquence sifflée qui évolue par modulations dans une gamme correspondant aux meilleures performances de discrimination sonore de l'oreille humaine (de 1 à 4 kHz). Les fréquences sifflées sont bien adaptées à l'environnement puisqu'elles sont plus aiguës que la plupart des bruits ambiants les plus communs et qu'elles correspondent largement à la fenêtre acoustique environnementale pour laquelle la réverbération change le moins avec la distance en présence de végétation (1 à 3 kHz, cf. Padgham 2004). Enfin, *la parole tambourinée* transforme toute la phrase en sons explosifs/percussifs puissants et riches en basses fréquences ce qui garantit une bonne propagation. Les signaux tambourinés exploitent en effet les propriétés bio-acoustiques naturelles des percussions pour une propagation optimale du son dans les environnements naturels (Rossing 2000). Les basses fréquences sont les ondes sonores qui sont le moins perturbées par les obstacles en raison de leur grande longueur d'onde et qui ne sont donc pas bloquées par la végétation (Bass et al 1990, Padgham 2004). Par conséquent, lorsqu'elles codent des hauteurs spécifiques, celles-ci seront conservées facilement. De plus, leurs amplitudes élevées et courtes ainsi que la bande de fréquence étroite codant la hauteur utile réduisent davantage les effets de masquage du bruit (Zwicker 1982, Green 1985). La simplification sonore induite par les tambours est encore plus drastique que celle de la parole sifflée car elle opère une réduction à la fois en fréquence, en amplitude, mais aussi en durée puisque ce type de son a une résonance fortement limitée. Elle repose notamment sur une

découpe rythmique des phrases qui reproduit les intervalles de temps entre les voyelles d'un même mot (Meyer et al 2012, Seifart et al 2018, Dentel et Meyer 2020, cf. Figure 3).

Ainsi, à partir de la stratégie polyvalente de résistance aux bruits de proximité qui caractérise la forme parlée usuelle, les langues ont su développer trois autres types de stratégies plus ciblées et efficaces quand l'impact de la distance impose plus de contraintes environnementales liées aux territoires. En effet, l'étude de la forme parlée en milieu extérieur a permis de mettre en évidence que les stratégies acoustiques *criée*, *sifflée* et *tambourinée* sont déjà en germe dans les différents phonèmes de la parole usuelle en observant leurs différentes performances de résistance à la dégradation lors de tests d'intelligibilité et en corrélant ces résultats à des caractéristiques phonétiques. Il a ainsi été montré récemment que les phonèmes qui résistent le mieux aux dégradations dues à la distance en parole modale sont, d'une part, les voyelles et, d'autre part, 3 types de consonnes (cf. Meyer et al 2013, 2015). Parmi ces types de consonnes on trouve, dans l'échelle de résistance, d'abord les fricatives fortes en énergies (/s/, /ʃ/, apparentées au sifflements), puis les consonnes avec formants bien marqués (/l/, /n/, qui avec les voyelles sont les phonèmes les plus apparentés à la stratégie développée en parole *criée*) et enfin les plosives avec forte explosion (par exemple /k/, /t/, apparentés aux percussions).

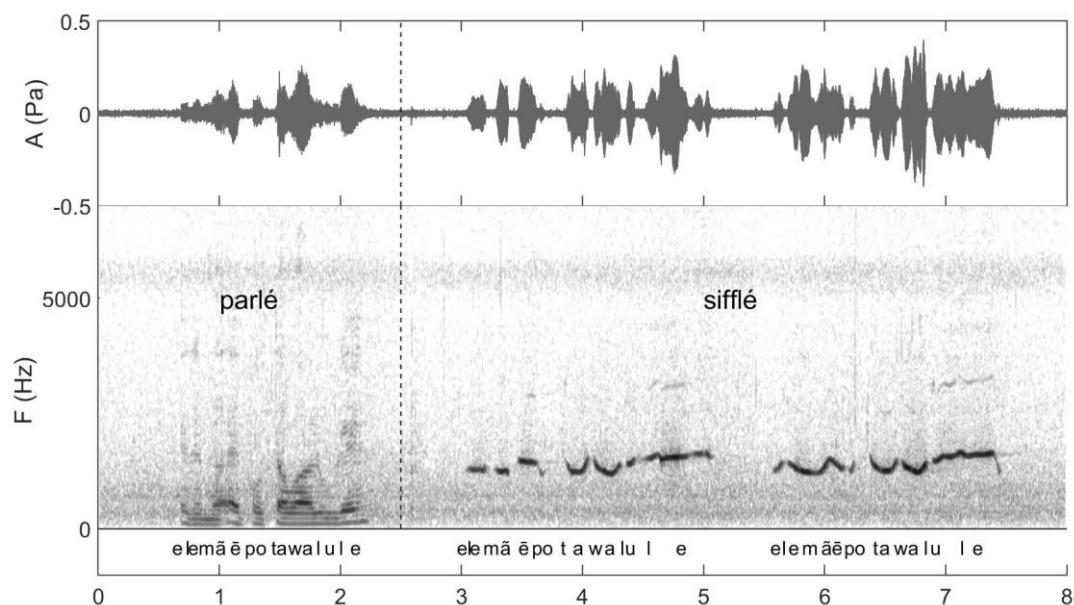
D'autre part, les différentes stratégies identifiées sont en outre adaptées à différentes portées de communication : (1) la voix humaine ne peut transmettre des informations que jusqu'à quelques centaines de mètres sous la forme *criée*. (2) La parole *sifflée* permet d'aller beaucoup plus loin : il n'est pas rare d'employer des dialogues *sifflés* à 1 voire 2 kilomètres lorsque le terrain est propice (Busnel et Classe 1976, Meyer 2015, Meyer et Diaz-Reyes 2017). (3) Les systèmes de communication utilisant les tambours peuvent encore étendre cette portée d'un facteur 10 - atteignant jusqu'à 10 km ou plus dans certains cas (Thiesen 1969, Seifart et al 2018).

Finalement, on observe que plus la distance à atteindre est grande, plus l'encodage acoustique est réduit pour concentrer l'énergie en sons puissants qui se propagent bien et contrastent efficacement avec le bruit de fond ambiant en extérieur. Cette réduction de la complexité du signal s'accompagne d'une adaptation de la structure des messages allant vers plus de simplicité pour faciliter la compréhension. Cette simplification peut passer soit par des phrases plus courtes, soit au contraire par des phrases plus longues fortement contextualisées (avec une redondance informative des textes qui facilite l'intelligibilité). Par exemple, *la parole criée* utilisée à distance s'appuie sur des phrases plus courtes et plus simples que la parole usuelle modale. Et comme nous l'avons dit plus haut, les dialogues criés sont de toutes façons en général assez courts en raison de la fatigue vocale. *Les phrases sifflées* ont également tendance à être simples avec des énoncés à proposition unique et des réparations conversationnelles également simples (Sicoli 2016, Diaz 2008). Celles-ci permettent de tenir des discussions plus longues qu'en parole criée car le sifflement évite la fatigue vocale. Par contre, quand les conversations ont lieu à grande distance, le temps de transmission du signal entre les interlocuteurs limite la dynamique d'échange des dialogues. En effet, il faut par exemple attendre environ 3 secondes pour recevoir un message à 1 km<sup>2</sup>. De plus, pour ajouter du contexte il est fréquent de cadrer un échange sifflé en commençant la discussion par un appel « phatique », comme *aaaa* ou *oye* en espagnol canarien, *oooo* en béarnais pyrénéen ou simplement un sifflement montant en tamazight marocain (Busnel et Classe 1976, Rialland 2005, Diaz 2008, Meyer et al 2021). Une autre méthode servant à faciliter la compréhension parole sifflée est de remplacer un mot par un synonyme ayant une structure syllabique plus forte et des consonnes plus distinctives car constituées de transitions plus rapides, par ex. le mot espagnol *abrigo* « manteau » est souvent remplacé par *chaqueta* « veste » (Diaz 2008). De plus, la compensation des réductions acoustiques peut s'opérer de différentes manières

---

<sup>2</sup> La vitesse du son dans l'air étant d'environ 340 m/s

visant à renforcer la redondance des informations (cf. Jakobson 1960, Shannon 1948). En parole sifflée il est commun de répéter la phrase. Dans les populations vivant en forêt dense, notamment, les siffleurs s'adaptent souvent à une situation d'urgence en produisant d'emblée deux répétitions de chaque phrase, surtout quand ils ne voient pas leur interlocuteur. Il en est ainsi chez les wayãpi et les gavião du Rondônia en Amazonie (Figure 4, et Meyer 2015). Ceci a plusieurs avantages : garantir de capter l'attention avec la première version, assurer l'écoute de l'ensemble de la phrase au moins une fois avec la seconde version (Figure 4).



**Figure 4** : Forme d'onde (en haut) et spectrogramme (en bas) montrant une même phrase wayãpi parlée (de 0 à 2.5 sec.) puis sifflée deux fois (2.5 à 5 sec., puis de 5.3 à 7.5 sec.). /elemãẽpo tawalu le/ signifiant « As-tu vu une tortue d'eau? », où /tawalu/ est la tortue d'eau (pour écouter cette phrase en version sifflée se rendre au repère de temps 01:09 de <https://vimeo.com/810960718>). On observe ici les mêmes caractéristiques linguistiques que pour la Figure 3, avec cette fois-ci une double répétition de la phrase sifflée. La seconde répétition est plus précise, en particulier dans les premières syllabes. On observe aussi un allongement et une plus forte intensité de la dernière syllabe /le/, lui conférant un effet emphatique. Cet enregistrement est réalisé sur une pirogue située sur la rivière, non loin de la berge, et l'on observe que les répétitions sifflées se placent en fréquence juste au-dessus des fréquences basses du bruit de fond ambiant de plus forte intensité (zone sombre dans la partie basse de la figure) et bien au-dessus de la bande stationnaire des insectes située un peu au-dessus de 5000 Hz. Par contre, la version parlée, elle, est plus masquée par le bruit ambiant.

*En tambours*, il est commun d'utiliser des périphrases pour remplacer un mot court.

Carrington (1949) donne l'exemple de 'la lune' remplacé en langue kele par une expression



plus longue et moins ambiguë à décoder 'la lune regarde vers le bas sur la terre'. Le type de tambours le plus fréquent dans les communications à longue distance sont de grands troncs évidés (Figure 2, en bas à gauche), car ils produisent les signaux les plus puissants (Figure 3). Leurs messages sont structurés de manière séquentielle avec des formules générales stéréotypées qui encadrent des formules plus libres portant de l'information nouvelle. Ce sont alors des messages dédiés à l'organisation à distance de la vie publique d'une communauté. Les messages sont en général à sens unique et peuvent couvrir un large territoire s'ils sont relayés de village en village (cf. Thiesen 1969, Pepper 1950), mais il peut arriver qu'une réponse soit émise, surtout si le destinataire est en possession d'un tambour, ce qui arrive en particulier si le correspondant est d'un autre village (Thiesen 1969, Pepper 1950, 1954, Carrington 1949, et Izikovitz 1935. Ce dernier auteur donne l'exemple frappant d'une communication entre un tambour en kit embarqué sur une pirogue en Amazonie et un gros tambour de village). Le modèle séquentiel de phrase qui cadre le contenu central de ce type de message prend la forme *ouverture>nom du destinataire>objet du message>contenu du message>fermeture*, comme cela a été décrit par exemple pour les langues africaines *kele* (Carrington 1949), *bulu* (Alexandre 1967), *banda linda* (Cloarec-Heiss 1999) et la langue amazonienne *bora* (Seifart et al 2018). La séquence '*objet du message*' spécifie par une phrase stéréotypée s'il s'agit d'une demande, d'un hommage, d'une information, d'une invitation (y compris en s'adressant à un village voisin). Par contre, les tambours à peau qui servent aussi fréquemment à imiter la parole sont plutôt utilisés pour s'adresser à moins grande distance à une audience lors de réunions ou d'événements festifs (Figure 2, en bas à droite). Les messages qu'ils envoient sont souvent basés sur de larges corpus mémorisés qui expriment des formes d'arts verbaux tels que des paroles de chansons, des proverbes, des poèmes (voir par exemple Rouget 1964, Kawada 1985, 2014, Locke et Agbeli 1987). Les différentes revues d'articles sur le thème de la parole tambourinée (Stern 1957, Sebeok et

Umiker-Sebeok 1976, James 2021, Meyer et Manfredi 2024) attestent de ces différences et de la diversité des langues concernées, principalement en Afrique, Asie, Océanie et Amérique.

#### **4. Quelques particularités de la parole sifflée et de la parole tambourinée**

##### **4.1 Quels éléments linguistiques sont transmis à distance ?**

Les redondances phonétiques inhérentes aux signaux de la parole usuelle vocale permettent de se passer de certaines caractéristiques acoustiques et d'en renforcer d'autres sans trop affecter la compréhension, ce qui permet aussi de transformer le signal de parole en sifflements ou en percussions et ainsi de l'adapter aux contraintes environnementales exacerbées par la distance. Nous allons voir maintenant plus en détails comment ces deux types d'adaptation sélectionnent les indices phonétiques saillants des signaux de parole de référence.

Nous avons déjà vu que la réduction drastique en fréquence que *la parole sifflée* a la particularité d'opérer fait qu'elle ne repose plus que sur une simple ligne mélodique modulée alors que la voix et ses résonances engendrent deux niveaux de fréquence séparables en perception pour coder les mots : le pitch<sup>3</sup> et le timbre. De plus, les corrélats acoustiques correspondants à ces deux dimensions fréquentielles sont contrôlés plutôt indépendamment en production dans le tractus vocal humain : le pitch est l'attribut perceptuel du F0 des cordes vocales ; le timbre vocal est l'attribut perceptuel de l'ensemble du spectre de fréquences qui caractérise la qualité des voyelles et des consonnes (Fant 1970, Risset 1968, Stevens 1998). Les contraintes articulatoires de la parole sifflée sont fortes car pour maintenir un courant d'air sifflant lors de la prononciation des mots, la source sonore est limitée à l'avant du conduit oral. Cette limitation impose un choix entre le pitch et le timbre comme cible principale dans la transformation voix-sifflement. Et ce choix va s'opérer en fonction du rôle du pitch dans la

---

<sup>3</sup> Aussi parfois appelée en acoustique musicale la 'hauteur tonale' et qui est l'attribut perceptif de la fréquence fondamentale

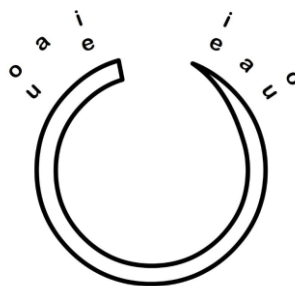
structure de la langue concernée. Ainsi, pour une langue 'tonale', les siffleurs transposent en priorité le pitch vocal (un exemple est la langue Gavião, cf. Moore et Meyer 2014), sinon la cible principale est le timbre vocal (un exemple est l'espagnol, voir Figure 1).

En ce qui concerne *la parole tambourinée*, les tambours à fente taillés à partir de troncs évidés sont les plus utilisés à longue distance et produisent un riche spectre de fréquence mais ne manipulent en général que le pitch des percussions pour coder la parole (voir Figures 2, 3). Ce pitch dépend de l'épaisseur du bois à l'endroit où la frappe est réalisée. Les valeurs que prend ce pitch imitent les changements relatifs du pitch vocal dans les langues tonales, mais certains maîtres tambours experts peuvent aussi parfois y exprimer l'influence à la fois du pitch et du timbre issus de la voix (voir l'exemple de Schneider 1952 sur la Figure 5). Si la langue n'est pas tonale, c'est le timbre de la qualité vocalique qui va être imité, mais les exemples de parole tambourinée dans de telles langues sont rares (e.g. Snyders et Haudricourt 1968, Schefold 1973). Les tambours à peau utilisés à plus courte distance pour une audience spéciale (dont un spécimen est présenté Figure 1, en bas à droite) peuvent émettre et contrôler une plus grande variété de sons. Dans ce cas, le codage réalisé pour les langues tonales passe souvent par une imitation du pitch vocal (soit avec le pitch du tambour comme pour le *dùn-dùn* Yorùbá (Akínbò 2019, 2021; Euba 1990), soit avec le timbre du tambour comme dans le cas du *Gùn-gbè* (Rouget 1964). Mais il existe des cas plus complexes où c'est une combinaison du pitch et du timbre vocal qui est imitée comme dans le cas des tambours *bàtá* Yorùbá (González and Oludare 2022). Enfin, dans les rares langues non-tonales jouées sur tambour à peau, le timbre vocal est imité par une correspondance avec différents types de frappes sur le tambour à peau (tambour wolof nommé Sabar, Winter 2014)

Le choix de la hauteur tonale (pitch) par rapport au timbre comme cible fréquentielle principale imitée lors de la transformation voix-sifflement ou voix-tambour est donc dépendante de limites acoustiques des instruments, mais aussi d'autres aspects comme (i) le

rôle de l'intonation dans la langue (par exemple, les langues « tonales » déploient clairement la hauteur du pitch pour le sens lexical plus que les langues « non tonales » ); et enfin (ii) le type de forme vocale qui est imité (chanté ou parlé) puisque que la parole chantée s'appuie particulièrement sur le pitch à travers les mélodies musicales.

Les paramètres d'amplitude et de durée sont aussi largement imités par *la parole sifflée*, en s'approchant de l'enveloppe d'amplitude de la dimension fréquentielle imitée, avec toutefois de nombreuses réductions et un tempo plus lent lié à une sur-articulation. Pour *la parole tambourinée*, les niveaux d'amplitude des coups de tambours à fente utilisés à grandes distances ne représentent pas des variations d'amplitude potentiellement informatives de la version parlée (Seifart et al. 2018). Là encore, les tambours à peau des arts verbaux permettent plus de distinctions qui sont parfois même exploitées pour distinguer différentes voyelles (cf. González and Oludare 2022, Meyer et Manfredi 2024). C'est au niveau rythmique que les tambours sont les plus intéressants car leurs sons rapides mettent en évidence certains indices rythmiques de la parole comme indiqué plus haut (voir aussi Figure 3).



**Figure 5** : Schéma d'un tambour à fente Douala (Schneider 1952, 676) utilisé pour produire à la fois un pitch soit plus grave (bord épais à gauche) soit plus aigu (bord plus fin à droite). Au-dessus de chaque bord du tambour sont représentés les emplacements où le tambourineur le plus expérimenté a frappé le tronc pour distinguer également

la qualité de trois différents types de voyelles (la hauteur moyenne perçue du timbre des voyelles était ainsi retransmise).

#### ***4.2 L'adaptation particulière de la parole sifflée à deux types de milieux écologiques***

La parole sifflée n'est pas qu'un système de télécommunication, elle a aussi l'originalité d'être associée à des milieux écologiques spécifiques dont elle épouse les contraintes acoustiques (cf. un exemple Figure 4 pour une phrase wayãpi sur le fleuve Oyapock). Les nombreuses populations humaines qui ont développé une forme sifflée de leur langue – plus de 80 ont été identifiées - vivent principalement dans des forêts denses et/ou dans des zones montagneuses escarpées. Ces milieux écologiques ont en effet la particularité commune d'isoler rapidement les individus et donc de favoriser les occasions de parler à distance. Or, le sifflement est adapté aux contraintes acoustiques de ces deux types de milieu et évite aux interlocuteurs de parcourir de longs trajets pour se parler (par exemple, à travers la forêt, d'un côté à l'autre d'une vallée, d'un côté à l'autre d'un fleuve). En général, c'est dans les milieux ouverts, comme sur un fleuve ou une montagne sans forêt, que les communications peuvent atteindre les distances les plus grandes car le guide d'onde des vallées comme le miroir de l'eau sur les rivières et les lacs créent des conditions favorables à la propagation acoustique, mais aussi parce que les occasions de se voir et donc converser sont plus nombreuses. Les propriétés acoustiques qui font que la parole sifflée est bien adaptée aux milieux ouverts de montagne (comme l'Atlas Marocain) et aux milieux fermés de forêts denses (comme l'Amazonie, milieu dits 'fermés' en raison de la couverture végétale) ont été décrites précédemment. Nous ajouterons seulement qu'en forêt les locuteurs soulignent également qu'elle est utile pour se fondre dans le paysage sonore pendant la chasse ou la pêche, et se parler/coordonner sans effrayer les animaux (de nombreuses espèces animales utilisent le sifflement).

Dans ces deux milieux, la parole sifflée possède un lien particulier avec certaines activités de subsistance traditionnelles : elle permet aux individus d'un groupe (famille, village) de

coordonner des activités telles que la chasse, la pêche, les cultures dans les vallées des montagnes (en terrasses), ou la garde du bétail par les bergers. Le maintien de ces activités conduit souvent au maintien de la parole sifflée. D'une manière plus générale, la parole sifflée facilite l'organisation de la vie quotidienne et la transmission de certains aspects de la tradition orale. En effet, elle permet de transmettre des nouvelles, de signaler un danger ou une urgence. Elle est utilisée pour parler la nuit, pour parler en secret en présence d'étrangers (contre des milices en temps de guerre, comme aux Canaries sous Franco ou au Béarn et en Papouasie durant la seconde guerre mondiale ; mais aussi plus simplement entre amoureux en utilisant parfois des codes, par exemple) ; enfin elle est parfois utilisée sous forme poétique ou ritualisée pour accompagner des moments marquants de la vie, comme c'est le cas des tambours à peau s'adressant à une audience (naissance, séduction, mariage, décès, fêtes) (Meyer 2021).

On observe que la parole sifflée est ancrée au sein de communautés relativement peu peuplées et habitant des territoires isolés, comme c'est le cas des villages wayãpi en Amazonie ou berbères dans les montagnes de l'Atlas, ou encore dans des îles (Eubée en Grèce, Canaries en Espagne ou Saint-Lawrence en Alaska). Cette pratique subit une perte de vitalité générale en raison principalement de changements de modes de vie lors desquels des activités rurales traditionnelles disparaissent (Meyer 2021). Comme la parole sifflée est issue d'une combinaison de facteurs liés à la fois au milieu écologique et au milieu social traditionnel, elle ne survit que là où les deux sont bien préservés. En Amazonie, par exemple, une étude sociolinguistique menée chez les Gavião du Rondônia a montré qu'elle ne survit que dans les villages situés au milieu de la forêt alors que presque plus personne ne l'apprend dans les villages des zones déforestées (Meyer 2015).

## **5. Conclusion**

Les modifications acoustiques et linguistiques qu'opèrent les formes criée, sifflée et tambourinée de la parole ont pour but de porter la parole à distance en extérieur et sont donc

adaptées aux différentes contraintes environnementales qu'exercent les territoires sur les communications humaines en milieu rural. Elles représentent toutes les trois des stratégies acoustiques différentes qui exploitent et renforcent de manière ciblée des propriétés déjà présentes en germe dans la forme parlée usuelle de référence. Plus la distance à atteindre est grande, plus les indices phonétiques sont réduits pour concentrer l'énergie dans des sons puissants favorables à la propagation sonore et qui contrastent avec le bruit de fond. Cette réduction est compensée par une adaptation de la structure des messages permettant de faciliter la compréhension, ce qui passe parfois par des choix guidés par des spécificités structurelles de chaque langue. Ces types de parole s'adaptent donc en production et en perception à la fois aux contraintes acoustiques du milieu, aux contraintes cognitives du traitement de l'information et à la structure linguistique des langues qu'elles expriment. Leur usage ne survit que dans les lieux où l'environnement écologique et l'environnement social traditionnel sont encore bien conservés, et donc dans des territoires isolés, loin du milieu urbain et de sa sphère d'influence. Ceci est particulièrement clair dans le cas de la parole sifflée dont l'usage s'est développé presque exclusivement en lien avec les écosystèmes de montagne et de forêt dense.

## **Bibliographie**

- Akínbò Samuel (2021). The language of *gangan*, a Yorùbá talking drum. *Frontiers in Communication* 6, 650382.
- Akínbò Samuel (2019). Representation of Yorùbá Tones by a Talking Drum : An Acoustic Analysis. *Linguistique et Langues Africaines* 5, 11–23.
- Alexandre Pierre (1967). *Langues & langage en Afrique noire*. Payot, Paris.
- Assmann Peter, Summerfield Quentin (2004) Perception of Speech under adverse conditions. In *Speech Processing in the Auditory System*, ed 1, S. Greenberg et al, eds. Springer, pp 231–308.
- Bass Henry E., Sutherland Louis C., Zuckerwar Allan J. (1990). Atmospheric absorption of sound : Update. *J. Acoust. Soc. Am.* 88, 2019–2021.
- Bregman Albert (1990). *Auditory scene analysis: the perceptual organization of sound*. The MIT Press, Cambridge
- Blumenrath Sandra H., Dabelsteen Torben (2004). Degradation of great tit *Parus major* song before

- and after foliation : Implications for vocal communication in deciduous forests. *Behavior* 8, 935–958.
- Busnel René-Guy, Classe André (1976). *Whistled Languages*. Berlin : Springer-Verlag
- Busnel René-Guy (1970). Recherches expérimentales sur la langue sifflée de Kusköy. *Rev. Phon. Appl.* 14–15 : 41–57
- Busnel René-Guy, Moles Abraham, and Vallencien Bernard (1962). Sur l'aspect phonétique d'une langue sifflée dans les Pyrénées françaises. *Proc. Int. Congress of Phon. Sci.* (Helsinki, Finland), 533–546.
- Busnel René-Guy, Gustavo Alcuri, Gautheron Bernard, Rialland Annie (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.
- Bradbury Jack B., Vehrencamp Sandra L. (1998). *Principles of Animal Communication*. Sunderland, MA : Sinauer Associates, Inc.
- Carrington John F. (1949). *Talking Drums of Africa*. London : Carey-Kingstate Press
- Cheyne Harold A., Kalgaonkar Kaustubh, Clements Mark A., Zurek Patrick (2009). Talker-to-listener distance effects on speech production and perception. *Journal of the Acoustical Society of America*, 126, 2052–2060.
- Cloarec-Heiss France (1999). From natural language to drum language : An economical encoding procedure in Banda-Linda, Central African Republic. In *Language diversity and cognitive representations* (eds C Fuchs, S Robert), pp. 145–157. Amsterdam, Philadelphia : John Benjamins Publishing Company.
- Dabelsteen Torben, Larsen Ole N., Pedersen Simon B. (1993). Habitat-induced degradation of sound signals : quantifying the effects of communication sounds and bird location on blur ratio, excess attenuation and signal-tonoise ratio. *Journal of the Acoustical Society of America*, 93, 2206–2220.
- Dentel, Laure, and Meyer, Julien (2020). How Rhythm and Timbre Encode Mooré Language in Bendré Drummed Speech. *Proceedings of Interspeech 2020- 21st Annual Conference of the International Speech Communication Association* (Germany : DBLP), 1898–1902. doi: 10.21437/Interspeech.2020- 2532
- Díaz Reyes David (2008). *El lenguaje silbado en la Isla de El Hierro*, Santa Cruz de Tenerife, Excmo, Cabildo Insular de El Hierro.
- Euba Akin (1990). *Yorùbá Drumming, the dùn-dún tradition*. Breitinger, Bayreuth.
- Franich Kathryn, Ngemzué Ange (2021). Feeling the beat in an African tone language; rhythmic mapping between language and music. *Frontiers in Communication* 6, article 653747.
- Fux Thibaud (2012). *Vers un système indiquant la distance d'un locuteur par transformation de sa voix*. Ph.D. Dissertation. Grenoble : Université de Grenoble.
- González Mariano, Oludare Olupemi (2022). The Speech Surrogacy Systems of the Yoruba Dùndún and Bàtá Drums. On the Interface Between Organology and Phonology. *Front. Commun.* 6:652542. doi: 10.3389/fcomm.2021.652542
- Gregg Jean W., Scherer Ronald C. (2006). Vowel intelligibility in classical singing. *Journal of Voice* 20(2), 198–210.
- Hoen Michel, Meunier Fanny, Grataloup Claire, Pellegrino François, Grimault Nicolas (2007) Phonetic and lexical interferences in informational masking during speech-in-speech comprehension,



*Speech Communication*, 49 : 905–916.

Izikovitz Karl G. (1935). "Idiophones" in *Musical and Other Sound Instruments of the South American Indians: a Comparative Ethnographical Study*, ed. K. G. Izikovitz (Göteborg : Elanders Boktryckeri Aktiebolag), 7–33.

James Lucas (2021). Systems of Communication: Aspects of Culture and Structure in Speech Surrogates. *Front. Commun.* 6:653268. Doi : 10.3389/fcomm.2021.653268

Kawada, Junzo (1985). *Textes historiques oraux des Mossi méridionaux (Burkina-Faso)*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa (ILCAA).

Kawada Junzo (2014). Notes on the Drum Language of the Ancient Mossi Kingdoms (Burkina Faso). *Japanese Rev. Cult. Anthropol.* 15, 199–215.

Kawada Junzo, Jun-Ichi Oda (1996). *A study of a drum language as a communication system in an African society*. Tokyo Institute for the Study of Languages and Cultures of Asia and Africa (ILCAA).

Kondo Hirohito M., van Loon Anouk M., Kawahara Juni-Ichiro, Moore Brian CJ. (2017). Auditory and visual scene analysis: an overview. *Phil. Trans.R. Soc. B* 372: 20160099. doi.org/10.1098/rstb.2016.0099

(2) (PDF) Auditory and visual scene analysis: An overview. Available from: [https://www.researchgate.net/publication/312034897\\_Auditory\\_and\\_visual\\_scene\\_analysis\\_An\\_overview](https://www.researchgate.net/publication/312034897_Auditory_and_visual_scene_analysis_An_overview) [accessed Sep 18 2023].

Locke David, Agbeli Godwin (1981). *Drum language in Adzogbo*. Black Perspect. Music 9, 25–50.

Lombard Emile (1911). Le signe de l'élévation de la voix. *Annales des maladies de l'oreille, du larynx, du nez et du pharynx*, 37, 101–119.

Luther David A. (2008). *The Evolution of Communication in a Complex Acoustic Environment*. PhD dissertation. Chapel Hill University.2307/1214103

Meyer Julien (2015). *Whistled Languages : A Worldwide Inquiry About Whistled Speech*. Berlin : Springer.

Meyer, Julien (2020) Coding Human Languages for Long-Range Communication in Natural Ecological Environments : Shouting, Whistling, and Drumming. In : Aubin T., Mathevon N. (eds) *Coding Strategies in Vertebrate Acoustic Communication. Animal Signals and Communication*, vol 7. pp.91-113, Springer, Cham. [https://doi.org/10.1007/978-3-030-39200-0\\_4](https://doi.org/10.1007/978-3-030-39200-0_4)

Meyer, Julien (2021). Environmental and linguistic typology of whistled languages. *Annual Review of Linguistics*, 7, 493-510.

Meyer Julien, Diaz David (2017). Geolingüística de los lenguajes silbados del mundo, con un enfoque en el español silbado. *Géolinguistique*, 17, 99-124

Meyer Julien, Manfredi Victor (2024). Auxiliary speech by vocal tract modulation and musical surrogacy. In : Daniela Sammler (ed.). *The Oxford Handbook of Language and Music*. Oxford : Oxford University Press (OUP)

Meyer Julien, Dentel Laure, Seifart Frank (2012). A methodology for the study of rhythm in drummed forms of languages : bora of Amazon. *Proceedings of Interspeech 2012 Annual Conference of the International Speech Communication Association* , 686–690. doi: 10.21437/Interspeech. 2012-210

Meyer Julien, Dentel Laure, Meunier Fanny (2013) Speech Recognition in Natural Background Noise. *PLoS ONE* 8(11) : e79279

- Meyer Julien., Meunier Fanny, Dentel Laure (2015) Speech Recognition Experiment in 'Natural Quiet' Background Noise. *Proceedings of the International Congress of Phonetic Sciences 2015*. Glasgow.
- Meyer Julien, Meunier Fanny, Dentel Laure, Do Carmo Blanco Natalia, Sèbe Frédéric (2018) Loud and Shouted Speech Perception at Variable Distances in a Forest. *Proceedings of Interspeech 2018*, Sep 2018, Hyderabad, India. Interspeech 2018, pp.2285-2289
- Moles Abraham (1970). Etude sociolinguistique de la langue sifflée de Kusköy. *Revue de Phonétique Appliquée*, 14/15 : 78-118.
- Mondes Sonores Amazoniens, imitations animales et parole sifflée des Amérindiens Wayãpi* (2023). Réalisé par Julien Meyer et Damien Davy [Court Métrage], Lyon : Beto Acosta et Tiffanie Deschamps. ([https://youtu.be/F\\_BZ4GYVCiM](https://youtu.be/F_BZ4GYVCiM) et <https://vimeo.com/810960718>)
- Niles Don (2010). Editor's Introduction to Zemp and Kaufmann. *Kulele Occasion. Pap. Pacif. Music Dance* 4, 1–50.
- Nketia Joseph H. K.(1963). *Drumming in Akan Communities of Ghana*. Edinburgh : Thomas Nelson and sons.
- Nketia Joseph H. K. (1971). "Surrogate languages of Africa" in *Speech Surrogates: drum and Whistle Systems*, eds T. A. Sebeok and D. J. Umiker-Sebeok (1976, Mouton: La Hague and Paris), 825–864.
- Novák Alexej, Vokřál Jan (2000) The speech intelligibility at the opera singing. *Sbornik Lekarsky* 101(2), 153-64.
- Padgham Mark (2004). Reverberation and frequency attenuation in forests - implications for acoustic communication in animals. *J. Acoust. Soc. Am.* 115, 402–410. (doi:10.1121/1.1629304)
- Pepper Herbert (1950) *Musique centre-africaine*. Paris (FR) : Encyclopédie Coloniale et Maritime, 1950
- Pepper Herbert (1954) *Considérations sur le langage tambouriné et autres langages musicaux d'Afrique centrale, sur la pensée musicale Africaine*. Ivo C.I.A.O., Fernando Pô Guinée Espagnole
- Persoon Gerard A., Schefold Reimar (2018) Frightened by the eagle Recording songs and music from the Island of Siberut, Mentawai Island, *Wacana Journal of the Humanities of Indonesia* 18(3), 581-613.
- Rialland Annie (2005). Phonological and phonetic aspects of whistled languages. *Phonology* 22, 237–271. doi: 10.1017/S0952675705000552
- Ridouane Rachid, Turco Giuseppina, Meyer Julien (2018). Length Contrast and Covarying Features: Whistled Speech as a Case Study. *Proceedings of Interspeech 2018*, Sep 2018, Hyderabad, India, 1843-1847
- Risset, Jean-Claude (1968). Sur certains aspects fonctionnels de l'audition. *Ann. Télécommun.* 23, 91–120. doi: 10.1007/BF02996236
- Ros, Sofiya (2021). Rhythm-speech correlations in a corpus of Senegalese drum language. *Frontiers in Communication* 6, article 643683.
- Rossing Thomas D. (2000). *Science of percussion instruments*. Singapore; River Edge, N.J.: World Scientific
- Rouget, Gilbert (1964). Tons de la langue en gùn (Dahomey) et tons du tambour. *Revue de musicologie* 50 (128), 3-29.

- Schefold, Reimar (1973) Schlitztrommeln und Trommelsprache in Mentawai, *Zeitschrift für Ethnologie* 98(1), 36-73.
- Schneider Marius (1952). Zur trommelsprache der duala. In Sebeok T. & Umiker-Sebeok D.J. (Eds). *Speech surrogates: drum and whistle systems*. 1976 La Haye and Paris, Mouton, 669-678
- Sebeok Thomas A., and Umiker-Sebeok Donna J. (1976). *Speech Surrogates: drum and Whistle Systems*. Mouton: La Hague and Paris. doi: 10.1515/9783110804416
- Seifart, Frank, Meyer, Julien, Grawunder, Sven, and Dentel, Laure (2018). Reducing language to rhythm: amazonian Bora drummed language exploits speech rhythm for long-distance communication. *R. Soc. Open Sci.* 5:170354. doi: 10.1098/rsos.170354
- Sicoli Mark (2016). Repair organization in Chinantec whistled speech. *Language* 92(2):411–32
- Snyders J., Haudricourt A. (1968). Le langage par tambours à San Cristoval, British Solomon Islands. *Journal de la Société des océanistes* 24, 133-138
- Stern Theodore (1957). Drum and whistle languages: an analysis of speech surrogates. *Am. Anthropol.* 59, 487–506. doi: 10.1525/aa.1957.59.3.02a0 0070
- Stevens Kenneth. N. (1998). *Acoustic phonetics*. Cambridge, MA: MIT Press.
- Thiesen Wesley (1969). The Bora signal drums. *Lore* 19, 101–103.
- Varnet Léo, Meyer Julien, Hoen Michel, Meunier Fanny (2012). Phoneme resistance during speech-in-speech comprehension. *Proceedings of Interspeech 2012* Portland, USA, 598-602.
- Wiley R. Haven, Richards Douglas G. (1978). Physical constraints on acoustic communication and the atmosphere: Implications for the evolution of animal vocalizations. *Behavioral Ecology and Sociobiology* 3, 69–94.
- Wiley R. Haven, Richards Douglas G. (1982). Adaptations for acoustic communication in birds: Sound transmission and signal detection. In E.H. Kroodsma, D.E. Miller (Eds.), *Acoustic Communication in Birds*, Vol. 1 (pp. 131–181) New York: Academic.
- Winter Yoad (2014). On the Grammar of a Senegalese Drum Language. *Language* 90, 644–668. doi: 10.1353/lan.2014.0061
- Zemp Hugo, Kaufman Christian (1969). Pour une transcription automatique des « langages tambourinés » mélanésiens (Un exemple kwoma, Nouvelle-Guinée), *L'Homme* 9(2), 38-88