

Des airs de parole : du son et du rythme aux représentations symboliques de la musicalité intrinsèque à la langue

Résumés

Cet article présente une synthèse des principales dimensions impliquées par les recherches s'intéressant à la question des informations véhiculées par la musicalité intrinsèque à la langue. De par ses propriétés segmentales et prosodiques, chaque langue possède une empreinte sonore et rythmique qui lui est spécifique. Cette contribution examine, à travers le parallèle à la musique, différents aspects qui contribuent à dépeindre cette musicalité intrinsèque, et tente de cerner les rôles variés que celle-ci peut servir. Les éléments sonores, au-delà même du sens qu'ils dénotent, se posent dans cette perspective tant comme repère incontournable de l'accès au sens (notamment dans l'acquisition de la langue maternelle et l'apprentissage des langues étrangères), que comme manifestation musicale et esthétique. Comme l'illustrent des études récentes rendant compte de situations de contact entre langue et musique, la frontière entre fonctions linguistiques et substance musicale est parfois difficile à délimiter, et s'atténue dans certains contextes au point de voir la musique venir prolonger l'usage de la parole.

This paper surveys the major directions taken by current research investigating the signification conveyed by the intrinsic musicality of language. Through its segmental and prosodic properties, each language has its specific sound profile. Through the parallel to music, this contribution aims to focus on various aspects which contribute to shape the intrinsic musicality of language and identifies the various functions it carries out. Phonic characteristics shall be regarded in this perspective as a key component allowing access to meaning (e.g. in the framework of L1 and L2 acquisition), as well as an aesthetic and musical manifestation of language. As illustrated by some recent studies involving contact between music and language, the boundary between linguistic functions and musical substance is sometimes difficult to define to the point that music happens in some cases to prolong the use of articulated speech.

Mots-clés : perception de la parole, musique, prosodie, phonétique, acoustique

Keywords: speech perception, music, prosody, phonetics, acoustics

Ainsi la cadence et les sons naissent avec les syllabes, la passion fait parler tous les organes et pare la voix de tout leur éclat ; ainsi les vers, les chants, la parole ont une origine commune. Autour des fontaines dont j'ai parlé, les premiers discours furent les premières chansons : les retours périodiques et mesurés du rythme, les inflexions mélodieuses des accents firent naître la poésie et la musique avec la langue [...]¹– Rousseau, J.-J.

Artefacts culturels, langue et musique partagent de nombreuses similarités. Organisant la matière sonore, toutes deux sont en capacité de stimuler des représentations mentales, et de construire de la signification². À l'instar de la musique, la durée, les temps forts et faibles, le phrasé mélodique, les couleurs sonores, le symbolisme intrinsèque des sons, sont autant d'éléments nécessaires à la description de la parole. La voix, tantôt garante de la transmission du sens, tantôt vecteur esthétique, est très certainement l'une des illustrations les plus évidentes à la lisière de ces deux pratiques. Les théoriciens qui se sont intéressés à leur description ont d'ailleurs depuis longtemps établi une étroite connexion entre la voix et l'instrument, entre la parole et le phrasé mélodique. Rousseau déjà adressait dans son *Essai sur l'origine des langues* les particularités chantantes du langage. La parole primitive ayant été selon lui, dans sa forme originelle, musicale. La transmission du savoir a longtemps été principalement orale et le reste encore dans de nombreuses langues. Sous la forme de textes déclamés, de chants traditionnels, ou encore d'incantations rituelles, la rencontre entre langue et musicalité arbore de nombreux visages. Les exemples de langage tambouriné en Afrique, les langues sifflées présentes à travers le monde, le chant dans sa pluralité sont autant de manifestations de la frontière parfois ténue qui peut s'instaurer entre la langue et la musique.

Ces diverses réalisations sonores n'ont pourtant qu'une seule et même origine : le son. Devenu objet de science, phonéticiens comme musicologues participent à toujours mieux comprendre les processus impliqués dans sa production, sa transmission et sa perception, soient-elles langagière ou musicale. Les progrès technologiques des dernières décennies ont permis de faciliter l'étude des phénomènes sonores, par nature volatils. La possibilité d'être en mesure de

* Nous tenons à remercier Jacqueline Vaissière et Claire Pillot-Loiseau pour leurs commentaires sur une version précédente de cet article, ainsi que les deux relecteurs anonymes de la revue *Savoirs en Prisme*. Ce travail a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du programme Investissements d'Avenir portant la référence ANR-10-LABX-0083.

1 Rousseau, 1993 : 102.

2 Si les analogies entre parole et musique restent certes fondées, il convient de garder à l'esprit que les potentiels discursifs du langage et de la musique se démarquent toutefois l'un de l'autre à différents égards (Wolf, 1999 ; Nattiez, 1987). Autant l'on retrouve cette capacité partagée à susciter de la signification sur la base d'unités discrètes, en tirant parti de l'union de signifiants et de signifiés ; autant là où le langage verbal se base principalement sur une relation hautement arbitraire du signe, et fonctionne primordialement dans son processus de signification sur la base de références extrinsèques, la musique possède davantage un caractère autoréférentiel, et a de ce fait essentiellement recours dans son processus de signification à l'utilisation de références intrinsèques. Il résulte de cette différence majeure que la musique n'est pas en mesure d'évoquer chez différents réceptifs des significations tout aussi similaires et objectives que ne le permet le langage (Wolf, 1999).

fixer une structure phonique, de pouvoir la visualiser et la réécouter de manière infinie a permis d'en analyser les composantes et de toujours davantage mettre en lumière la complexité des rapports qui sous-tendent leurs interactions.

La structure de la parole n'est d'ailleurs pas sans rappeler les éléments impliqués dans la production musicale. On y retrouve deux facettes complémentaires : d'une part, les segments et la façon dont ceux-ci se combinent, d'autre part la prosodie, qui réunit des éléments tels que l'accentuation, le rythme ou encore l'intonation, qui viennent se superposer à ce niveau segmental³. Tous ces éléments, s'ils peuvent, comme nous le verrons plus loin, contribuer à la création d'un certain esthétisme sonore, constituent dans la parole de précieux indices, que les locuteurs savent utiliser pour extraire du sens de la variabilité intrinsèque au flux sonore. Les nouveau-nés s'appuient d'ailleurs sur ces différents éléments, pour identifier et reconnaître leur langue maternelle à travers sa musique. Les éléments prosodiques, souvent assimilés à la musique de la langue, apparaissent en effet dans un premier temps primordiaux ; ils posent les bases de l'accès au sens et les fondements des attentes perceptives spécifiques à la langue (Millotte & Christophe, 2009 ; Hirsh-Pasek *et al.*, 1987). Si les éléments segmentaux et suprasegmentaux ont avant tout un rôle fonctionnel, le caractère sensible que véhicule toute expression musicale, et la force conative qui s'y rattache, ne sont toutefois pas exclusifs à la musique. Les productions langagières écrites ou orales ont aussi souvent joué plus ou moins consciemment avec la forme, et ont fait le jeu des écrivains, des poètes et des orateurs depuis la nuit des temps. Des complaintes des troubadours au Moyen-Âge aux productions imagées des poètes surréalistes du début du xx^e siècle, en passant par moult campagnes publicitaires, les exemples de jeux savants reposant sur la musique intrinsèque à la langue, et sa portée esthétique, restent innombrables. Car si cet aspect sensible de la parole influence indiscutablement l'auditeur, il sollicite de surcroît l'appel à son imaginaire. L'impression laissée par la musicalité de la langue joue aussi d'ailleurs dans ce prolongement un rôle décisif sur l'impression globale que nous nous faisons de la langue elle-même.

Nous essaierons à travers cette contribution de préciser ce que désigne la musique de la langue – expression souvent entendue mais en elle-même assez vague-, et ce à travers les rôles variés que celle-ci peut servir. Dans cette perspective, nous présenterons, à travers le prisme des recherches actuellement engagées dans différents domaines de spécialité, les principales dimensions qu'englobe la réflexion menée autour des informations véhiculées par la musique de la langue. L'accent portera dans un premier temps à poser des mots sur ce qu'est l'élément sonore dans sa forme la plus simple, et plus particulièrement sa réalisation vocale. Nous verrons qu'à cette identité acoustique et articulatoire vient

3 Dans le cadre d'une comparaison adressant spécifiquement la structure et l'organisation de la langue et de la musique, cette analogie porte cependant davantage sur les aspects prosodiques en jeu que sur les segments eux-mêmes. En effet, dès lors que la mise en parallèle concerne la segmentation d'unités à un niveau moins global (la mise en parallèle de notes individuelles avec des phonèmes, ou encore les correspondances entre des notes individuelles, des extraits de mélodies ou des accords avec des mots ou des morphèmes), les correspondances sont alors beaucoup moins transparentes et évidentes à établir (Wolf, 1999).

se superposer une dimension symbolique, évocatrice, porteuse de sens, parfois plus difficile à définir, mais pourtant largement partagée au sein d'un groupe linguistique donné. La primauté de la musicalité intrinsèque aux éléments composant la langue sera ensuite discutée au travers de son implication primordiale, tant dans les processus d'acquisition de la langue, que dans le contexte de son apprentissage par des locuteurs non-natifs. Nous évoquerons enfin quelques exemples de situations de communication donnant tout leur sens aux parallèles entre musique et langue, et dans lesquels musique et parole viennent se relayer pour garantir la bonne transmission d'informations et contourner les contraintes dues à des environnements géographiques particuliers.

Au commencement était le son

La production du son : L'homme ne possède pas d'appareil phonatoire spécifiquement dédié au langage. La réalisation des sons de la parole peut être représentée par un modèle source-filtre, assimilant la production du son à l'effet conjoint d'une source sonore et d'un banc de filtres déterminé par la configuration du conduit vocal (Fant, 1960). En d'autres mots, l'émission d'un son nécessite avant tout une impulsion rendant possible l'émission d'une onde sonore, par la suite modelée par la forme des cavités du conduit vocal. Ce sont les poumons qui génèrent dans un premier temps cette impulsion en agissant telle une soufflerie, propulsant un fin filet d'air vers le larynx. La pression sous-glottique augmentant provoque alors la mise en vibration des plis vocaux, si ces derniers sont mollement resserrés. Ce phénomène est en principe assimilable à la mise en vibration de l'air, qui s'échapperait d'un ballon de baudruche dont on pincerait les extrémités, ou encore au bourdonnement qu'émet le trompettiste à l'embouchure de son instrument. Le conduit vocal joue ensuite le rôle de caisse de résonance et altère la forme spectrale de ce bourdonnement à travers la forme adoptée par les résonateurs supraglottiques (le pharynx, la cavité buccale et la cavité nasale). Le passage de l'air alors libre se traduit par la réalisation de sons vocaux, caractérisés par la présence de formants⁴. Le rétrécissement ou la fermeture du conduit vocal entravent le passage de l'air, et donnent lieu en revanche à des sons de nature consonantique, caractérisés alors pour beaucoup par la présence de bruit résultant de la constriction du conduit vocal. En ce sens, la configuration des résonateurs supraglottiques agit donc comme un filtre et favorise la mise en valeur de certaines résonances, permettant à l'oreille de caractériser chaque son de la parole.

La composition du son : Si elle peut sembler propice aux appréciations subjectives, la description des sons musicaux ou de la parole répond de fait à un ensemble de propriétés physiques bien délimitées. Chaque son peut ainsi principalement être défini par sa hauteur, son intensité et son timbre (Heffner, 1969).

4 Nous entendons par formant le rassemblement d'un groupe d'harmoniques présentant une intensité renforcée, en comparaison à ceux environnants. Les formants sont notamment en grande partie responsables du timbre des voyelles.

Dans un contexte musical, la hauteur traduit le concept de note. Plus globalement, la hauteur d'un son correspond à la perception du caractère grave ou aigu qui caractérise l'élément sonore. La hauteur est liée pour la parole à la fréquence de vibration des plis vocaux. Cette fréquence de vibration, appelée fréquence fondamentale, est déterminée par des paramètres tels que la longueur, la tension ou encore l'épaisseur des dits plis vocaux. Ces caractéristiques physiologiques ont une influence directe sur le nombre de cycles d'ouverture et de fermeture des cordes vocales et explique en grande partie qu'une voix féminine ou celle d'un enfant soit plus aiguë qu'une voix masculine, pour laquelle la vitesse de vibrations des plis vocaux, plus longs et plus épais, est moins élevée. Une autre propriété importante de l'élément sonore est l'intensité avec laquelle il est émis. Celle-ci est physiologiquement déterminée par la pression d'air sous-glottique, et se manifeste acoustiquement à travers l'amplitude de l'onde sonore et la répartition de l'énergie dans les différentes plages de fréquences. Souvent associée chez les non-spécialistes à une voix puissante, l'intensité n'est toutefois pas l'unique garante de l'efficacité vocale.

Enfin, le timbre de la voix ou celui d'un instrument, parfois désigné par le terme de couleur sonore ou de grain de voix, est déterminé par la richesse en harmoniques rentrant dans la composition du son, ainsi que par les transitoires d'attaque et d'extinction ; il permet de différencier deux sons aux fréquences et amplitudes similaires. C'est grâce à lui que nous sommes par exemple en mesure de distinguer une même note de même intensité jouée respectivement par une trompette et une clarinette, ou le jeu de deux artistes jouant le même instrument (Barthet, 2008), c'est encore entre autres grâce à lui qu'il nous est possible de distinguer une voix d'une autre. Les qualités de timbre de voix contribuent en effet à la singularité et à l'identification de chaque voix, et participent aussi en partie à l'expression des émotions (Scherer *et al.*, 2003 ; Fónagy, 1983). L'amplitude des tout premiers harmoniques est un corrélat important à la caractérisation de la qualité de voix (Klatt & Klatt, 1990) : une qualité de voix soufflée, souvent associée à la féminité ou encore à l'intimité, se caractérisera ainsi par un renforcement de l'amplitude du premier harmonique aux dépens du second, tandis qu'à l'inverse une voix plus serrée, vocalement plus efficace, s'illustrera par un deuxième harmonique plus intense que le premier.

Ces particularités physiologiques et acoustiques permettent d'offrir une nomenclature de base pour la description des éléments sonores, soient-ils de nature musicale ou linguistique. Ces mêmes caractéristiques sur lesquelles reposent la production et le décodage des énoncés linguistiques, ainsi que les interprétations qui en sont faites, s'accompagnent également de qualités esthétiques donnant lieu à des interprétations de nature plus subjective et faisant référence à une dimension symbolique parfois peu aisée à formaliser.

La musicalité de la langue, une symbolique universelle ?

La musicalité de la langue apparaît dans les textes dits ou chantés, aussi bien au travers de la prosodie, que dans la combinaison adroite des segments la composant. Les sons, le jeu des sonorités et les symbolismes qui s'y rattachent tiennent un rôle central en poésie, dans l'opéra et dans tout exercice oratoire requérant une mobilisation de l'imaginaire et du ressenti de l'auditeur. Fónagy (1979, 1983), dans une démarche d'approche psychoacoustique de la parole, a longuement exploré les caractéristiques attribuées intuitivement à différents types de phonèmes. Ses recherches ont mis en évidence les convergences observées, en termes de symbolisme et de champs d'associations attribués aux éléments sonores composant la langue. Les expériences conduites ont mis en avant que la perception de ces qualités n'est pas aussi aléatoire qu'on ne pourrait le penser, mais qu'au contraire, il existe une certaine constance dans les qualités que chaque son de la parole peut véhiculer. Le choix de désignations métaphoriques (voyelles féminines ou masculines, minces ou larges, etc.) suit par ailleurs une certaine concordance avec la nature des traits acoustiques propres à chacun des dits phonèmes. Les attributs *profond*, *grave*, *sombre* tendront ainsi à être plus facilement associés au phonème [u], et à l'inverse des qualités telles que *étroit*, *aigu* ou *clair* auront davantage de chance d'être attribuées au phonème [i]. Cet aspect purement psychologique de la perception des caractéristiques sonores, s'il est naturellement à la base de nombreuses figures rhétoriques visant à magnifier le sens, en faisant notamment usage au travers de procédés d'allitérations et d'assonances, a aussi été exploré dans le cadre de l'enseignement de la prononciation, pour mettre en valeur certaines caractéristiques des phonèmes de la langue. La perception de qualités musicales et esthétiques est alors transposée et utilisée pour rendre plus mémorables certains aspects articulatoires et perceptifs, ayant trait aux phonèmes et aux éléments suprasegmentaux.

L'iconicité ou le symbolisme dégagé par certains phénomènes sonores a aussi été mis en parallèle avec une origine éthologique (Vaissière, 2003 ; Ohala, 1984). Ce postulat expliquerait ainsi qu'une fréquence fondamentale basse tende davantage à être catégorisée comme un attribut de dominance, ou de menace (Ohala, 1994), tandis qu'à l'inverse, une fréquence fondamentale plus élevée ou montante puisse traduire une expression de politesse ou encore de soumission. La façon dont la mère s'adresse à son bébé, en plus de sa fonction communicative évidente (Fernald, 1985 ; Fernald & Kuhl, 1987), illustre aussi d'une certaine manière ce phénomène, avec une voix souvent volontairement rehaussée, s'accompagnant d'une exagération des caractéristiques prosodiques. Le symbolisme intrinsèque de la voix trouve donc ses racines dans les manifestations les plus primitives de la parole jusque dans l'élaboration volontaire de jeux avec la substance phonique. La voix du poète, comme celle du chanteur, exploite à bon escient cette dimension évocatrice de la voix. Le chant lyrique, qui s'illustre dans les œuvres d'opéra, attribue par exemple à différents types de personnages différents types de voix. La couleur, la tessiture, le registre participent à façonner la crédibilité du personnage et dépeignent les principaux traits de caractère

des intervenants. Une voix de basse, la plus grave des voix d'homme, peindra un personnage mûr et influent, tandis qu'une voix de ténor plus aiguë et agile dépeindra par exemple avec plus de justesse la bravoure et la fougue de jeunes héros.

La musicalité de la langue ne repose pas uniquement sur les « éléments sonores » à proprement parler. Il va sans dire que le rythme, et la prosodie dans sa globalité, font partie intégrante de l'empreinte sonore de la langue, et constituent eux-aussi une dimension tout aussi capable d'évoquer des représentations symboliques. S'il n'est pas rare que le rythme constitue en soi les fondations d'un langage poétique (Zumthor, 1982), il traduit dans la parole de manière générale un aspect évocateur tout aussi important. C'est ainsi qu'une vitesse d'élocution rapide évoquera par exemple davantage un état de joie qu'un état de tristesse, tandis qu'un débit d'élocution plus lent aura tendance au contraire à être plus facilement associé à l'expression d'un état de tristesse (Scherer, 2003 ; Murray & Arnott, 1993). De la même manière, l'organisation temporelle de la parole, l'emploi de pauses dans le discours, peuvent refléter et symboliser un positionnement au sein d'une hiérarchie sociale. Duez (1999) dans son analyse du discours politique relève par exemple qu'une articulation lente et des pauses longues caractérisent le « style de discours dominants ». La symbolique du statut social politique/auditeur, gagnant/perdant transparaît alors entre autres à travers les éléments liés à l'organisation temporelle de la langue. Les études récentes s'intéressant aux parallèles entre musique et langage tendent enfin, elles aussi, à toujours mettre davantage en lumière l'importance de cette dimension rythmique et structurelle, dans les représentations symboliques que peut transmettre la musique (Jandausch, 2012 ; Zbikowski, 2009).

Parole et musique ont donc ceci en commun d'être en capacité d'activer à travers l'utilisation des caractéristiques sonores, des représentations conceptuelles et des impressions perceptives, aussi bien universellement partagées que construites socialement.

Des enfants, musiciens ?

L'enfant est exposé très rapidement aux sonorités et à la mélodie de sa langue. À travers la paroi abdominale, le fœtus perçoit dès la période intra-utérine la voix de sa mère. Ce premier contact, quoique encore très approximatif, sensibilise semble-t-il le futur bébé à la redondance de certains aspects sonores. À la naissance, les nouveau-nés sont sensibles à tous les contrastes phonétiques, et sont alors en théorie potentiellement capables d'acquiescer n'importe quelle langue. Cependant, sous l'influence de l'input linguistique à leur disposition, ceux-ci développent très vite une préférence pour leur langue maternelle (Jusczyk, 1997 ; Boisson-Bardies, 2005). Cette préférence est d'abord en grande partie guidée par la musique de la langue, et plus particulièrement sa prosodie (DeCasper & Fifer, 1980 ; Hallé, 2006). En seulement quelques mois, le bébé se spécialise, et perd progressivement toute sensibilité pour les contrastes qui ne

portent pas de rôle fonctionnel dans son système phonologique. Le rôle structurant de la prosodie apparaît ainsi primordial dès les premiers mois de vie, et détermine fortement l'accès au sens du jeune enfant, en lui permettant de segmenter toujours davantage le flux sonore (Hirsh-Pasek *et al.* ; 1987 ; Jusczyk *et al.*, 1992). Avant même d'avoir conscience de l'existence de propositions, de syntagmes, ou encore de mots, les nouveau-nés accordent beaucoup d'importance aux contours intonatifs de la langue et à son rythme, les segments jouant dans un premier temps un rôle plus négligeable. L'importance des caractéristiques sonores de la langue est très vite vérifiable : les nouveau-nés montrent ainsi un intérêt très limité à l'écoute d'un enregistrement en langue maternelle joué à l'envers, car celui-ci dénature en fin de compte la prosodie propre à la voix maternelle (Mehler *et al.*, 1988). Un autre fait intéressant concernant cette sensibilité aux caractéristiques prosodiques est que, sans devoir y être forcément exposés, les bébés sont capables de distinguer deux langues l'une de l'autre, si tant est que celles-ci appartiennent à des classes rythmiques différentes (Nazzi, Bertoncini & Mehler, 1998 ; Moon, Cooper & Fifer, 1993). Des recherches ont toutefois dans certains cas aussi démontré l'habileté des nouveau-nés à discriminer deux langues de classe rythmique identique, à condition que l'une d'entre elles soit la langue maternelle de l'enfant (Bosch & Sebastian-Gallés, 1997 ; Nazzi, Jusczyk & Johnson, 2000).

L'acquisition des propriétés sonores de la langue constitue un enjeu majeur des premiers mois de vie. La cristallisation des habitudes articulatoires et perceptives propres à la langue se met en place dans un espace-temps réduit. Les jeunes enfants s'approprient très vite la musique de la langue et apprennent en s'appuyant sur l'environnement linguistique à leur disposition à être en adéquation avec les exigences spécifiques à leur langue maternelle. Hallé, Boysson-Bardies et Vihman (1991) ont par exemple observé à travers la production de vocalisations disyllabiques, chez de jeunes enfants français et japonais âgés de 18 mois, la tendance de chaque groupe linguistique à réaliser des productions en accord avec les attentes perceptives de leur langue respective. Les enfants français ont ainsi montré une tendance à produire des contours montants sur la syllabe finale ainsi qu'un allongement terminal, tendance absente chez les enfants japonais, qui en revanche commençaient à ébaucher des tons lexicaux.

Cette prégnance musicale de la langue et la capacité du jeune enfant à y porter déjà en bas âge une attention toute particulière se prêtent naturellement à un regard croisé sur les parallèles existant entre la pratique de la musique et celle de la langue. C'est d'ailleurs non sans intérêt que les recherches continuent toujours à essayer de mieux cerner ces parallèles (Patel, 2008) et à peser notamment l'influence potentielle de la pratique musicale et du chant sur l'appropriation et la perception des caractéristiques sonores de la langue. Une oreille entraînée à l'écoute musicale semble ainsi avoir davantage de facilité à percevoir certaines particularités segmentales et suprasegmentales d'une langue étrangère (Dodane, 2003). Une batterie d'expériences conduites sur des enfants musiciens et non-musiciens a montré une tendance des enfants ayant une pratique musicale à mieux discriminer certains contrastes vocaliques de l'anglais, ainsi

qu'une aptitude à restituer avec plus d'aisance les caractéristiques prosodiques de cette langue. Une hypothèse intéressante réside par ailleurs dans le fait que l'entraînement musical tendrait entre autres à stimuler une mémorisation à long terme, favorisant à son tour une meilleure restitution de l'empreinte sonore de la langue cible. De tels constats invitent tout naturellement à vouloir exploiter ces transferts de compétence entre écoute musicale et perception de la langue, et à en tirer parti dans le cadre de l'enseignement de la prononciation (Cornaz, Henrich & Vallée, 2010), voire de la rééducation vocale (Pillot-Loiseau & Vaissière, 2009).

La musicalité comme ancrage mnémotechnique des caractéristiques sonores de la langue

Des témoignages comme celui rapporté par Akira Mizubayashi (2011), écrivain japonais et professeur de littérature, montrent à travers le récit de sa découverte tardive du français, que la musique de la langue est un levier indéniable à l'éveil linguistique et à la motivation nécessaire pour s'initier à une langue étrangère. Ainsi narre-t-il son premier contact avec cette « langue venue d'ailleurs », survenu au travers de la diffusion d'émissions radiophoniques consacrées à l'apprentissage du français :

[...] le contenu de chaque leçon se réduisait à des sons à la fois clairs et veloutés, produits par les deux invités. C'était pour moi comme un récital à deux voix, un concert retransmis en différé où la voix de l'homme et celle de la femme se cherchaient, se répondaient, se confondaient, s'entrelaçaient dans leur mouvement phonique délicat et soigneusement défini.⁵

La voix, la musicalité de la langue constitue souvent le premier contact avec une langue étrangère. Un document sonore, un extrait télévisé, une chanson, une comptine, la première impression n'est bien souvent qu'éphémère, difficile à fixer. La parole s'efface et laisse alors place à une impression globale. À cette empreinte vocale, opaque, précédant toute compréhension, viennent alors se substituer des attributs divers pour en décrire ses consonances, son profil rythmique, sa ligne mélodique. Les dimensions subjectives et émotionnelles qui interviennent dans la musique sont tout autant mobilisées dans la perception initiale des langues. Comme en témoigne la découverte heuristique du jeune Akira Mizubayashi, la langue est avant tout musique :

Je n'ai pas fait de musique. Dans mon enfance, le piano fut un compagnon, mais forcé. Je ne m'y attachai pas. Bientôt, je voulus me débarrasser de cet importun pour adhérer aux jeux d'enfant. De

5 Mizubayashi, 2011 : 23.

guerre lasse, mes parents l'acceptèrent. Pourtant, la musique ne me quitta pas : le français en prit la place. C'était pour moi un instrument qui faisait chanter une musique particulière. [...] j'eus une musique à moi, à moi seul, c'était le français.⁶

26 | Une langue *abrupte, douce, joyeuse, colorée*, au rythme *syncopé* ou *legato* sont autant de termes révélateurs, relayant des impressions émises spontanément sur la composition segmentale et les phénomènes prosodiques d'une langue encore inconnue. Un travail d'écoute régulier, la pratique de l'imitation, la compréhension des phénomènes linguistiques sous-jacents à ces intuitions souvent sibyllines sont la garantie de l'accès à une bonne maîtrise de la langue, de sa diction, et de sa prononciation. Les caractéristiques musicales de chaque langue, à l'instar de la musique, sont enclines à motiver la création d'associations visuelles et à éveiller des sensations perceptives spécifiques. Ce dernier point peut donner lieu à des prolongations pédagogiques intéressantes. Par exemple, l'évocation de certaines particularités rythmiques peut être réactivée grâce à l'utilisation de métaphores (Ibarrondo, 2013) ou de gestes motivés (Llorca, 2001), facilitant l'appropriation des tendances rythmiques d'une langue donnée.

La musique offre en outre, de par sa notation, des outils descriptifs permettant de mettre à plat la structure sonore d'une phrase, de la décrypter. Les équivalences de durée à la base de la notation musicale, la pulsation, les temps fort et faible, le phrasé, sont autant de notions pouvant être réinvesties dans la description des phénomènes prosodiques, et permettent de mettre en évidence des régularités exploitables dans le cadre de l'enseignement de la prononciation (Wioland, 1991). Les parallèles entre structures linguistiques et structures musicales offrent des perspectives d'applications fructueuses. En ce sens, la conscientisation des régularités phonétiques structurant l'apprentissage de la prononciation d'une langue, doit amener l'apprenant à adopter la posture, non plus d'un auditeur passif, mais celle du musicien amateur. L'étude phonétique de la langue, si elle arbore une approche scientifique rigoureuse, peut par ailleurs tirer efficacement parti dans sa vulgarisation à des fins pédagogiques, de l'exploitation de parallèles à la musique, qui se prêtent à une explicitation simple et concise des phénomènes phonétiques régissant l'oralité de la langue. Par l'écoute active, quasi-musicale, et le recul théorique s'y superposant, la structure de la langue apparaît sous un nouveau jour. Les principes de l'approche verbo-tonale (originellement développée dans les années 1970, sous l'impulsion de Petar Guberina), ont dans cette continuité aussi accordé une importance prépondérante à l'écoute ; la correction phonétique a consacré dans ce contexte une place de choix à la dimension perceptive, ainsi qu'à la prise en compte de la prosodie (Renard, 2002 ; Renard, 1971), l'apprenant devant avant tout être sensibilisé aux spécificités sonores de la langue. L'exploitation de contextes facilitants (mettant en évidence les traits caractéristiques aux phonèmes travaillés, et favorisant leur

6 Mizubayashi, 2011 : 57.

reproduction) reste d'ailleurs une influence indéniable de cette approche, et se retrouve toujours dans de nombreuses méthodes ayant trait à la prononciation du français (Callamand, 1981 ; Kaneman-Pougatch & Pedoya-Guimbretière, 1991 ; Abry & Chalaron, 2011).

La voix chantée

Il serait difficile d'évoquer l'ancrage musical de la langue et la musicalité intrinsèque aux éléments qui la composent sans évoquer bien évidemment le chant. Des hymnes proférés à la gloire des héros grecs à la naissance des premières polyphonies, des chants diphoniques mongols aux plus grands opéras italiens, la voix se révèle sous de multiples facettes. Tantôt soliste, portée au-devant de la scène, tantôt davantage fondue au sein d'un ensemble, l'utilisation de la voix dans le chant obéit à de nombreuses conventions, dictées à la fois par les cultures musicales locales et par l'époque. Si la compréhension de ces paradigmes dépasse le cadre de cette contribution, il est peut-être intéressant de noter que le rapprochement ne serait-ce que descriptif de la musique et par là du chant à la parole ne doit pas amener à penser que ces pratiques soient en tous points similaires, en matière de gestion de voix. Les exigences de la voix chantée ne sont en effet pas nécessairement identiques à celles requises dans le contexte de l'interaction verbale, et impliquent des compromis de nature diverse. Les contraintes liées à l'intelligibilité du message doivent bien souvent cohabiter avec des exigences musicales divergentes. Les actions articulatoires servent avant tout dans le chant une intention musicale. Roubeau (2014) souligne ainsi le fait que « *l'esthétique règle le choix des priorités dans les qualités acoustiques de la voix* », mais qu'« *elle règle aussi le fait qu'une qualité dans un contexte esthétique donné peut être un défaut dans un autre* ». De la même manière, différents aspects (productifs, perceptifs, communicationnels) concourent à définir la notion d'efficacité vocale, la prépondérance de chacun de ces critères étant étroitement dépendante de l'objectif visé. L'efficacité vocale recherchée dans le chant accorde ainsi davantage d'importance à des critères tels que le timbre et la portée de la voix (Pillot, 2004), là où la transmission du sens, ou tout du moins son intelligibilité, reste centrale dans le cadre de la parole. Cette différenciation entre voix chantée et voix parlée se retrouve en fait à différents niveaux. Articulatoirement, la distinction entre phonèmes vocaliques est par exemple plus difficile à maintenir dans les aigus, et pousse ainsi le chanteur à accepter certains compromis entre précision articulatoire et rendu sonore. L'étude acoustique de la voix chantée lyrique a par ailleurs par exemple mis en évidence un renforcement des harmoniques autour de 3000Hz, zone de grande sensibilité de l'oreille. Ce phénomène, désigné comme *formant du chanteur* (Sundberg, 1969), garantit aux voix entraînées de passer plus facilement au-dessus de l'orchestre, et ainsi de porter tout en ménageant l'effort vocal. Pillot (2006) souligne ainsi que « *selon un point de vue physiologique, donner de l'énergie et porter la voix ne sont pas si connectés. C'est dire que plus un chanteur porte sa voix, moins il a*

besoin de donner de l'énergie ». Comme le rappellent ces quelques constats, le rapprochement entre voix chantée et parole doit se faire en tenant compte des particularités propres à chacun de ces modes de réalisation, soient-elles d'ordre articulatoire, acoustique ou perceptif.

Quand la langue se passe de mots

28

L'esthétisme de la parole sollicite des dimensions expressive et conative primordiales à l'œuvre musicale. Les œuvres de musiciens contemporains enchevêtrent, parfois au point d'en effacer la maigre délimitation, la voix vecteur de communication, et sa réalisation artistique et musicale à la recherche constante de l'esthétisme sonore. Wim Mertens, compositeur belge, illustre par exemple dans sa *Stratégie de la rupture* l'utilisation d'une langue issue du fruit de son imagination, et parvient malgré tout à retransmettre sans recours à des mots intelligibles une palette d'états émotionnels divers. À défaut de rendre compte du sens précis des logatomes articulés, la musique relaie par ses codes la transmission des émotions.

Mais, le sens disparaît-il forcément lorsque la langue se passe de mots ? Pas toujours. Le *silbo* est une transposition de la langue espagnole, répertoriée sur l'île de la Gomera, dans les Canaries. Elle a de particulier que ses locuteurs ne la parlent pas. Ils la sifflent. Le *silbo* espagnol est un des langages sifflés répertoriés dans le monde. Pensées pour prendre le relais du langage articulé afin de transmettre des messages sur des distances éloignées, par exemple dans des régions montagneuses, les « langues sifflées » ont cela de particulier qu'elles semblent venir se substituer aux mots. Sont-elles assimilables à de la musique ou encore à un langage ? Des études ont montré que, chez les locuteurs du *silbo*, les zones cérébrales correspondant aux fonctions langagières étaient stimulées (Carreiras *et al.*, 2005), tandis que pour les locuteurs ignorant ce langage, ces mêmes aires n'étaient pas sollicitées. Les sons pouvant pour certains s'apparenter à une mélodie semblent donc bien parfois pouvoir prendre la place des mots. Les études acoustiques du *silbo*, et d'autres systèmes de communication semblables, ont montré que les siffleurs transmettent en fait les indices acoustiques les plus importants, calqués sur le langage articulé (structure formantique, éléments prosodiques), permettant à la personne recevant leur message d'être en mesure de décoder les signaux envoyés (Meyer & Gautheron, 2006 ; Rialland, 2005). D'autres langages, tambourinés, illustrent également l'implication primordiale des qualités musicales de la langue dans le processus de transmission du sens. De nombreuses langues africaines tirent en effet parti de l'utilisation de schèmes tonaux pour construire le sens. Ces langues à tons confèrent une valeur distinctive aux variations de hauteur, permettant d'établir des distinctions tant lexicales que grammaticales. La reproduction canonique des hauteurs et du rythme par le biais de l'instrument permet au tambourinaire de transmettre sous forme musicale des énoncés linguistiques, et aux destinataires d'avoir accès au sens en

réinterprétant le parallélisme existant entre les séquences entendues et le langage articulé⁷.

Les informations segmentales et suprasegmentales façonnent l'identité sonore de la langue. Elles sont constamment mobilisées par les locuteurs pour hiérarchiser et ordonner le flux d'informations contenu dans la parole. Qu'il s'agisse du patron formantique, des tons, ou d'autres éléments prosodiques, chaque élément participe à donner au locuteur natif des repères identifiables, lui permettant de cerner des invariants et de passer au-delà de la variabilité présente dans le signal sonore. Les exemples ci-dessus illustrent des cas de figure singuliers, où la maîtrise de la langue arbore une nouvelle dimension, et introduit des contextes particuliers, où l'environnement local a amené les populations à s'approprier des pratiques nouvelles -autres que celle du langage articulé-, également approchées dans le cadre de la communication animale (le sifflement), voire débordant sur le terrain de l'expression musicale (les percussions). La musique intrinsèque à la langue devient alors le prolongement de la parole elle-même et par là même son propre instrument.

Les différents éléments sonores évoqués jusqu'ici, des phonèmes jusqu'aux tons, en passant par le rythme, définissent la musicalité d'une langue. Si les locuteurs, mêmes natifs, ne sont pas toujours en mesure de la décrire, il n'en reste pas moins que leurs attentes perceptives en sont en grande partie dépendantes. La musicalité intrinsèque aux éléments sonores est par définition multiple. Tantôt porteuse de sèmes perçus inconsciemment, elle est alors à l'origine de métaphores, jouant avec les sonorités et avec la forme. Tantôt structurant et véhiculant le sens, elle relaie des informations sans lesquelles la compréhension du message serait mise à mal. S'il ne fait pas de doute que la musique et la langue sont deux systèmes sémiologiques régis par des codes leur étant propres, le recoupement des propriétés indissociables à leur description et leur réalisation vocale peuvent toutefois offrir des points de comparaison utiles, dont les perspectives d'application suscitent toujours un intérêt pluridisciplinaire important, comme en témoignent les recherches actuelles regroupant des spécialistes de domaines aussi variés que l'ethnomusicologie, la phonétique ou encore la phoniatry.

Ludovic Ibarrondo

Université Sorbonne-Nouvelle – Paris 3, UMR 7018 LPP
ludovic.ibarrondo@outlook.fr

7 Pour une illustration de ce type de communication, le lecteur pourra consulter l'intervention de S. Arom (2008) *Entre parole et musique : les langages tambourinés d'Afrique Centrale*, au Collège de France.

Œuvres citées

- ABRY, Dominique, Chalaron, Marie-Laure (2011) : *Les 500 exercices de phonétique B1/B2*. Paris : Hachette.
- AROM, Simha (consulté le 28.01.2015) : « Entre parole et musique : les langages tambourinés d'Afrique centrale ». *Aux origines du dialogue humain : parole et musique* (Paris, France, 16-17 octobre 2008). <http://www.college-de-france.fr/site/colloque-2008/symposium-2008-10-16-14h40.htm>
- BARTHET, Mathieu (2008) : *De l'interprète à l'auditeur : une analyse acoustique et perceptive du timbre musical*. Thèse de doctorat. Université Aix-Marseille II.
- BOSCH, Laura, Sebastián-Gallés, Núria (2003) : « Language experience and the perception of a voicing contrast in fricatives: Infant and adult data ». *15th ICPhS* (Barcelone, Espagne, 3-9 août 2003).
- BOYSSON-BARDIES, Bénédicte de (2005) : *Comment la parole vient aux enfants*. Paris : Odile Jacob.
- CALLAMAND, Monique (1981) : *Méthodologie de l'enseignement de la prononciation. Organisation de la matière phonique du français et correction phonétique*. Paris : Clé International.
- CARREIRAS, Manuel, LOPEZ, Jorge, RIVERO, Francisco, CORINA, David (2005) : « Linguistic perception: Neural processing of a whistled language ». *Nature*. 433 : 31-32.
- CORNAZ, Sandra, HENRICH, Nathalie, VALLÉE, Nathalie (2010) : « L'apport d'exercices en voix chantée pour la correction phonétique en langue étrangère : le cas du français langue étrangère appliqué à des apprenants italiens d'âge adulte ». *Cahiers de l'APLIUT*. 29.2 : 103-119.
- DODANE, Christelle (2003) : *La langue en harmonie : influences de la formation musicale sur l'apprentissage précoce d'une langue étrangère*. Thèse de doctorat. Université de Besançon.
- DUEZ, Danièle (1999) : « La fonction symbolique des pauses dans la parole de l'homme politique ». *Faits de langues*. 13 : 91-97.
- FANT, Gunnar (1960) : *Acoustic Theory of Speech Production*. La Haye : Mouton.
- FERNALD, Anne (1985) : « Four-month-old infants prefer to listen to motherese ». *Infant Behavior & Development*. 8 : 181-195.
- FERNALD, Anne, KUHL, Patricia (1987) : « Acoustic determinants of infant preference for motherese speech ». *Infant Behavior & Development*. 10 : 279-293.
- FÓNAGY, Iván (1979) : *La métaphore en phonétique*. Ottawa : Didier.
- (1983) : *La vive voix : essai de psychophonétique*. Paris : Payot.
- HALLÉ, Pierre, BOYSSON-BARDIES, Bénédicte de, VIHMAN, Marilyn (1991) : « Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants ». *Language and Speech*. 34.4 : 299-318.
- HALLÉ, Pierre (2006) : « La musique des premiers mots ». *Cerveau et psycho*. 17 : 40-43.
- HEFFNER, Roe-Merrill (1969) : *General Phonetics*. Madison : University of Wisconsin Press.
- IBARRONDO, Ludovic (2013) : *Étude des voyelles antérieures non-arrondies en allemand, français, et finnois, et applications en vue d'une didactique de la prononciation*. Thèse de doctorat. Université Paris III.
- JANDAUSCH, Alexandra (2012) : « Conceptual metaphor theory and the conceptualization of music ». *Proceedings of the 5th International Conference of Students of Systematic Musicology* (Montréal, Canada, 24-26 mai 2012).
- JUSZYK, Peter (1997) : *The discovery of spoken language*. Cambridge, MA : MIT Press.

- JUSCZYK, Peter, *et alii* (1992) : « Perception of acoustic correlates of major phrasal units by young infants ». *Cognitive Psychology*. 24 : 252-293.
- KANEMAN-POUGATCH, Massia, PEDOYA-GUIMBRETIERE, Élisabeth (1991) : *Plaisir des sons : enseignement des sons du français*. Paris : Hatier-Didier.
- KLATT, Dennis, KLATT, Laura (1990) : « Analysis, synthesis and perception of voice quality variations among female and male talkers ». *Journal of the Acoustical Society of America*. 87 : 820-857.
- LORCA, Régine (2001) : « Jeux de groupe avec la voix et le geste sur les rythmes du français parlé ». *L'enseignement des langues aux adultes aujourd'hui*. Ruth, Johnson, Geneviève-Dominique de, Salins (dir.). Saint-Étienne : Publications de l'Université de Saint-Étienne : 141-150.
- MEHLER, Jacques, JUSCZYK, Peter, LAMBERTZ, Ghislaine, HALSTED, Nilofar, BERTONCINI, Josiane, AMIEL-TISON, Claudine (1988) : « A precursor of language acquisition in young infants ». *Cognition*. 29 : 144-178.
- MEYER, Julien, GAUTHERON, Bernard (2006) : « Whistled Speech and Whistled Languages ». *Encyclopedia of Language & Linguistics, Second Edition, volume 13*. Keith, Brown (Ed.). Oxford : Elsevier : 573-576.
- MEYER, Julien (2008) : « Acoustic Strategy and Typology of Whistled Languages; Phonetic Comparison and Perceptual Cues of Whistled Vowels ». *Journal of the International Phonetic Association*. 38. 1 : 64-90.
- MILLOTTE, Séverine, Christophe, Anne (2009) : « À la découverte des mots : le rôle de la prosodie dans l'acquisition du lexique et de la syntaxe ». *Enfance*. 3 : 283-292.
- MIZUBAYASHI, Akira (2011) : *Une langue venue d'ailleurs*. Paris : Gallimard.
- MOON, Christine, COOPER, Robin, FIFER, William (1993) : « Two-day-olds prefer their native language ». *Infant Behavior and Development*. 16 : 495-500.
- NATIEZ, Jean-Jacques (1987) : *Musicologie générale et sémiologie*. Paris : Christian Bourgeois.
- MURRAY, Iain, ARNOTT, John (1993) : « Toward the simulation of emotion in synthetic speech ». *Journal of the Acoustical Society of America*. 93 : 1097-1108.
- NAZZI, Thierry, BERTONCINI, Josiane, MEHLER, Jacques (1998) : « Language discrimination by newborns: Toward an understanding of the role of rhythm ». *Journal of Experimental Psychology: Human Perception and Performance*. 24.3 : 756-766.
- NAZZI, Thierry, JUSCZYK, Peter, JOHNSON, Elizabeth (2000) : « Language discrimination by English learning 5-month-olds: effects of rhythm and familiarity ». *Journal of Memory and Language*. 43 : 1-19.
- OHALA, John (1984) : « An ethological perspective on common cross-language utilization of Fo of voice ». *Phonetica*. 41 : 1-16.
- (1994) : « The frequency codes underlies the sound symbolic use of voice pitch ». *Sound symbolism*. Leanne Hinton, Johanna Nichols, John Ohala (Ed.). Cambridge : Cambridge University Press : 325-347.
- PATEL, Aniruddh (2008) : *Language, Music and the Brain*. Oxford : Oxford University Press.
- PILLOT, Claire (2004) : *Sur l'efficacité vocale dans le chant lyrique. Aspects physiologiques, cognitifs, acoustique et perceptif*. Thèse de doctorat. Université Paris III.
- (2006) : « Représentation cognitive de l'efficacité vocale dans la parole et le chant : conséquences en rééducation ». *Voix parlée et chantée (Bilans, rééducations, recherche, travail vocal comédiens et chanteurs)*. Carine, Klein-Dallant (dir.). Avray : Klein Dallant : 429-451.

- PILLOT-LOISEAU, Claire, VAISSIÈRE, Jacqueline (2009) : « La portée de la voix parlée et chantée : aspects scientifiques et rééducatifs ». *La voix dans tous ses maux*. Peggy Gatignol (dir.). Isbergues : Ortho-Edition : 243-249.
- RENARD, Raymond (1971) : *Introduction à la méthode verbo-tonale de correction phonétique*. Paris : Didier.
- , (dir.) (2002) : *Apprentissage d'une langue étrangère/seconde 2. La phonétique verbo-tonale*. Bruxelles : De Boeck Université.
- RIALLAND, Annie (2005) : « Phonological and phonetic aspects of whistled languages ». *Phonology*. 22.2 : 237-271.
- ROUBEAU, Bernard (2014) : « Aspects de la motricité laryngée : de la cellule aux synergies musculaires ». *La voix chantée, entre sciences et pratiques*. Nathalie Henrich-Bernadoni (dir.). Bruxelles : De Boeck Université : 21-33.
- ROUSSEAU, Jean-Jacques (1993, 1781) : *Essai sur l'origine des langues*. Catherine Kintzler (dir.). Paris : GF Flammarion.
- SCHERER, Klaus, JOHNSTONE, Tom, KLASMEYER, Gudrun (2003) : « Vocal Expression of Emotion ». *Handbook of Affective Sciences*. Richard Davidson, Klaus Scherer, Hill Goldsmith (Ed.). New-York : Oxford University Press : 433-456.
- SCHERER, Klaus (2003) : « Vocal communication of emotion: A review of research paradigms ». *Speech Communication*. 40 : 227-256.
- SUNDBERG, Johan (1969) : «Articulatory differences between spoken and sung vowels in singers ». *STL QPSR*. 1 : 33-46.
- VAISSIÈRE, Jacqueline (2003) : « Sur les universaux de substance prosodiques ». *Les universaux sonores*. Sophie Wauquier (dir.). Rennes : Presses universitaires de Rennes.
- WIOLAND, François (1991) : *Prononcer les mots du français : des sons et des rythmes*. Paris : Hachette FLE.
- WOLF, Werner (1999) : *The Musicalization of Fiction. A Study in the Theory and History of Intermediality*. Amsterdam/Atlanta : Rodopi.
- ZBIKOWSKI, Lawrence (2009) : « Music, language and multimodal metaphor ». *Multimodal Metaphor*. Charles Forceville, Eduardo Urios-Aparisi (Ed.). Berlin : Walter de Gruyter : 359-381.
- ZUMTHOR, Paul (1982) : « Le rythme dans la poésie orale ». *Langue française*. 56.1 : 114-127.