



**HAL**  
open science

# Le musilangage : une hypothèse d'origine commune entre la proto-musique et le proto-langage

Caroline Boe

► **To cite this version:**

Caroline Boe. Le musilangage : une hypothèse d'origine commune entre la proto-musique et le proto-langage. 2019. hal-02004782v2

**HAL Id: hal-02004782**

**<https://hal.science/hal-02004782v2>**

Preprint submitted on 1 Nov 2019 (v2), last revised 30 Jan 2020 (v3)

**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.

Copyright

# Le *musilangage* : une hypothèse d'origine commune entre la proto-musique et le proto-langage.

Caroline Boë

caro@zoulous.com

**Abstract.** Selon une hypothèse défendue de façon transdisciplinaire – anthropologique, linguistique, musicologique –, il existerait à l'origine de l'humanité une sorte de *musilangage*, précurseur à la fois de la parole et du chant. Pour étudier cette hypothèse, dont aucune trace archéologique ne témoigne, deux pistes de recherche sont envisagées : une première méthode repose sur l'analyse des sons produits par les animaux proches génétiquement de l'être humain, et une seconde méthode consiste à étudier le *mamanais* ou *motheres*, mode de communication entre la mère et le bébé, puis à élargir à l'étude de la communication entre deux enfants ne maîtrisant pas encore le langage articulé. Ces deux directions permettent d'envisager une sémiotique sonore, qui postulerait un modèle comportemental pré-langagier de l'homme de Neandertal. Protomusique et protolangage auraient une origine commune, déterminée par les capacités phonatoires d'individus physiquement incapables d'articuler.

**Keywords:** protolangage, protomusique, musilangage, *mamanais*, *motheres*, Hummmm, système de communication animale (SCA, S.C.A.), mode de communication des bébés.

## 1 Introduction.

Une étude musicologique sémio-herméneutique des *Récitations* de Georges Aperghis m'a engagée à m'interroger sur les origines du langage. En effet, dans cette partition d'Aperghis pour voix solo, l'association de rythmes, d'intervalles et de phonèmes m'a plongée dans un état primitif, comme si ce vocabulaire et cette musicalité se réfèrent à l'origine même du langage. C'est donc lors de cette étude que j'ai découvert l'hypothèse qui intéresse cet article, hypothèse partagée par une pluralité de chercheurs tels qu'archéologues, anthropologues de la préhistoire, pédopsychiatres, linguistes et musicologues : la thèse du *musilangage*. Ce *musilangage* serait un ancêtre commun entre la musique et le langage. La musique aurait ainsi une origine vocale, et le langage une origine musicale. Il s'agit bien d'une hypothèse, car aucune trace archéologique ne peut la valider.

De récentes analyses effectuées sur des squelettes de néandertaliens tentent de démontrer que ceux-ci possédaient la capacité du langage articulé. D'une part, Nicholas Wade, dans un article de 2007 paru dans le New York Times, annonce qu'un test ADN effectué sur un os de Neandertal révèle que ce néandertalien possédait le même gène FOXP2 que l'homme moderne [19]. Ce gène est nécessaire à la production vocale. Il est présent sous cette forme chez les oiseaux chanteurs, et est absent chez les souris par exemple. D'autre part, l'université du Kansas publie en 2013 un article à propos de l'os hyoïde néandertalien [7]. L'os hyoïde est nécessaire à l'articulation phonatoire. Celui de Neandertal était très proche de l'hyoïde de l'homme moderne. Le protolangage, non articulé aurait donc pu se développer avant leur ancêtre commun, soit avant 600 000 ans.

S'intéresser au protolangage dans ces conditions implique nécessairement d'imaginer ce qu'il aurait pu être. Si ces découvertes scientifiques démontrent une possibilité physique d'élocution, s'interroger sur la forme et sur le sens des émissions vocales antérieures au langage articulé nécessite d'étudier les productions vocales d'individus incapables d'articuler, c'est à dire ceux dont l'os hyoïde n'est pas compatible à l'articulation. L'os hyoïde est important dans l'évolution, puisqu'il serait descendu lorsque les quadrupèdes ont évolué en bipèdes [3].

L'os hyoïde du bébé, particulièrement haut pour favoriser sa déglutition lors de l'allaitement, ne permet pas l'articulation vocale à ce stade. Le larynx descend lors de la croissance de l'enfant, lorsque celui-ci commence à s'alimenter avec des ingrédients solides, et qu'en même temps il adopte une posture verticale de bipède et commence à marcher [15].

C'est pourquoi l'étude des sons produits par certains primates génétiquement proches de l'humain ainsi que celle des émissions vocales des nourrissons permet d'envisager une sémiotique du sonore, à la fois musicale et langagière [14].

## 2 Hummm et S.C.A.

A la suite de la découverte de Nicholas Wade à propos du gène FOXP2, Steven Mithen émet l'hypothèse selon laquelle les néandertaliens possédaient un système de langage, qu'il nomme « Hummm », c'est à dire Hollistique, manipulateur, multimodal, musical et mimétique [12].

Ceci confirmerait l'hypothèse de Bruce Richman, musicologue évolutionniste, qui pense que les premiers éléments lexicaux, c'est à dire les mots, sont apparus comme des fragments abrégés de chansons communes [14]. Pour Richman, à mesure que les membres du groupe accumulaient un répertoire de chansons en expansion pour différentes occasions, les schémas interpersonnels d'appel et de réponse, évoluaient le long d'une trajectoire, pour prendre la forme linguistique. Pendant ce temps, le long d'une trajectoire divergente, le chant polyphonique et d'autres types de musique, sont devenus de plus en plus spécialisés et sophistiqués.

Pour étudier cet hypothétique musilangage, le linguiste Derek Bickerton s'intéresse aux systèmes de communication animale (S.C.A.). Il analyse les sons produits par les cercopithèques, des singes africains qui lancent des cris d'alerte [2]. Si le cri d'alerte succède à une vocalise, le danger est au loin, la succession vocalise-cris d'alerte suppose

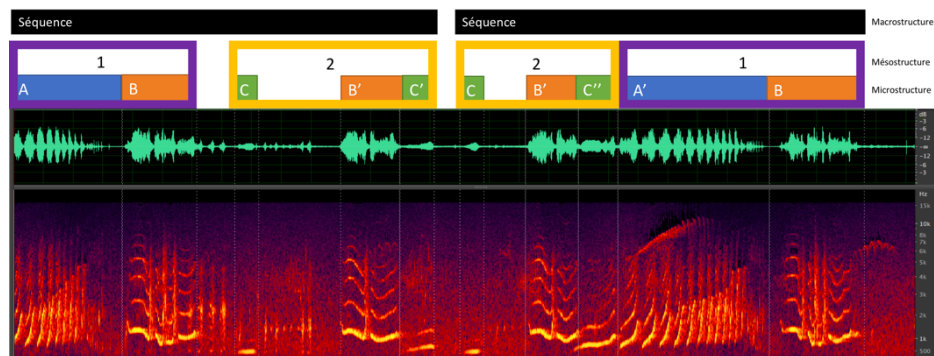
l'emploi du futur. Par contre, lorsque le cri d'alerte est seul, le danger est imminent, il s'agit d'un futur presque présent. La combinaison est dans ce cas signifiante, elle se réfère à une temporalité future plus ou moins proche.

Puis, en étudiant les singes vervets d'Afrique de l'ouest, Bickerton montre que des cris différents sont émis, pour prévenir de l'approche d'un aigle, d'un léopard ou d'un serpent. L'objectif, de la différenciation des cris, est d'obtenir une réponse différente : observer le ciel, ou trouver un arbre proche par exemple.

L'étude de ces systèmes de communication animale permet de constater que l'homme n'est pas le seul animal sémiotique. Le protolangage serait déjà manipulateur, c'est-à-dire qu'un cri attend une réponse, et ceci met en jeu la fonction conative, déterminée par Jakobson, fonction utilisée pour influencer le destinataire [9].

Bickerton s'interroge ensuite sur l'origine du chant et du langage en observant les gibbons qui chantent des duos. L'origine se situerait dans le chant amoureux. Les gibbons sont totalement monogames. Les échanges de duos longs renforceraient la solidarité entre les deux partenaires. « Vient ensuite, la défense du territoire, avec l'éventuelle intrusion des prédateurs », et, enfin « la nécessité de rester en contact avec les autres membres du groupe » [2].

Ces trois éléments – renforcement de la solidarité entre les partenaires, défense du territoire et nécessité de rester en contact avec les autres membres du groupe – ont attiré mon attention sémiotique. J'ai cherché, à partir d'un enregistrement de sons produits par des gibbons à favoris blancs, à définir une structure à partir de la combinaison d'éléments sonores récurrents, de rythmes, de fréquences et d'intensités différentes. L'enregistrement étudié est extrait d'une vidéo captée dans le parc animalier de la Vallée des Singes de Romagne, France.



**Fig. 1.** Visualisation de l'extrait du chant de gibbons

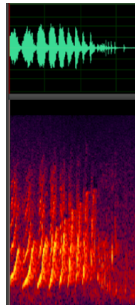
Sur cette visualisation du spectre sonore apparaissent trois éléments, que j'ai appelés cellules A, B et C, avec des variantes comme A', B', C' et C''. Avec la méthodologie d'analyse herméneutique musicale de Bernard Vecchione [18], et à partir de ces trois éléments classés dans la microstructure, nous pouvons découvrir une mésostructure composée de deux assemblages : 1 = AB et 2 = CBC.

La macrostructure, selon cet assemblage de combinaisons apparaît sous la forme : 1-2 / 2-1. Cet agrégat singulier relève d'une forme en miroir. Et cette organisation en miroir se retrouve dans la combinaison CBC. Ceci nous permet de remarquer que la microstructure reproduit le schéma de la macrostructure, dans un agencement de type fractal. Cette interprétation de l'arrangement sonore construit en fractale nous oriente vers la biologie en rappelant par exemple la structure géométrique du chou Romanesco [11]. Ceci pourrait tisser des liens avec la musique dodécaphonique [5], qui utilise volontiers des procédés de miroir et de fractales dans la composition musicale.

Mais si nous cherchons une interprétation linguistique de ces signes sonores, il est nécessaire d'envisager un tout autre point de vue. En se référant aux théories de phonologie de Jakobson, nous pouvons nous intéresser aux limites, aux intensités, aux fréquences dominantes et aux rythmes de ces cellules A, B et C.

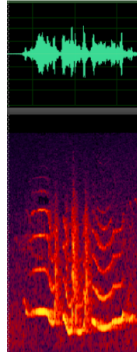
[...] la cadence, l'intonation tombante en fin de phrase, signale la fin de l'unité de sens, présentée dans la phrase. Dans sa fonction subordonative, l'accent signale directement l'importance du mot accentué dans l'énoncé. Nous pouvons ne pas comprendre les mots d'une phrase, mais nous savons que la cadence annonce sa fin, nous savons que le nombre des accents égale le nombre des termes de la phrase, nous savons que l'accent le plus fort signale le terme le plus important, celui dont la signification sert de point de départ à la phrase [10].

La cellule A, est composée d'une séquence rythmique qui démarre de façon franche, s'accélère tout en produisant un crescendo puis un décroscendo, tandis que l'intonation monte, puis brutalement se termine par une sorte de gazouillis sur une intonation descendante.



**Fig. 2.** Cellule A

Selon ce critère, la cellule A, avec sa cadence finale d'intonation tombante serait la fin d'une phrase et non le début, comme ce que j'ai suggéré dans l'interprétation fractale. Pour continuer le dépliage herméneutique avec une visée linguistique, observons la cellule B :



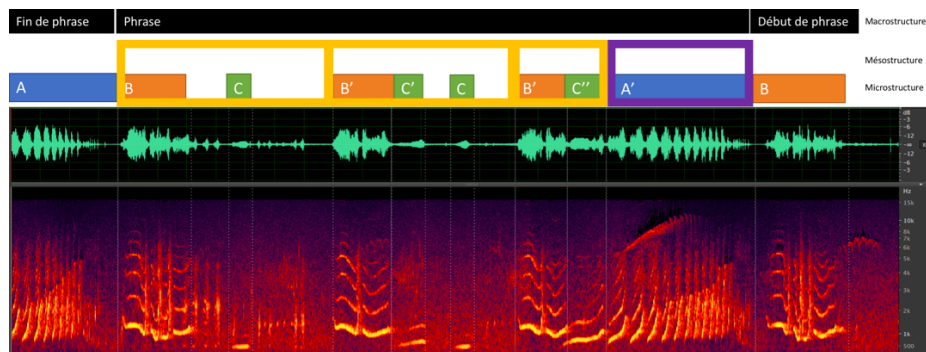
**Fig. 3.** Cellule B

La cellule B est délimitée par des silences, son intonation est montante, ce qui selon Jakobson, annonce l'énonciation d'une autre unité sémantique. L'intonation montante, suivie d'un silence aurait l'effet d'une conjonction de coordination. Avec sa grande tessiture vocale, cette cellule semble insister sur un sujet particulier, et marquer une signification qui sert de point de départ à une phrase.

En ce cas, et toujours dans l'hypothèse de Richman qui pense que la structure linguistique est préexistante dans le chant, la répétition de A-B montrerait la répétition d'une phrase, A étant la fin de phrase et B le début – pour mémoire, l'agencement de l'extrait est le suivant : A-B-C-B'-C'-C-B'-C''-A-B.

La cellule C semble être émise par le second individu, au vu de la différence de niveau sonore. Les interventions en C, C' et C'' ressemblent à une réponse à B, et ceci confirmerait à nouveau les hypothèses de Richman sur le chant responsorial, originel du langage articulé.

Nous pouvons désormais envisager une nouvelle macrostructure :



**Fig. 4.** Structure Jakobsonienne du chant des gibbons

L'orientation linguistique de l'observation du chant des gibbons pourrait bien être une piste pour spéculer sur l'origine commune du langage et du chant. Steven Brown, musicologue, en est convaincu :

Le "musilangage" [. . .était] précurseur à la fois de la parole et du chant, plus proche de la musique que du langage parlé moderne, et comportait des caractéristiques telles que le "ton lexical" (c'est-à-dire l'utilisation de la hauteur afin de véhiculer un sens sémantique), ou encore la formation combinatoire de phrases mélodiques et rythmiques. [...] Notre opinion est qu'il y a eu passage direct, à partir d'un stade "musilinguistique" [...] ou d'un stade de "protolangage musical" [...] à une différenciation consciente entre les deux manifestations du langage : le parler et le chant [4].

Ces deux interprétations – biologique et linguistique – fondamentalement différentes montrent l'importance du point de vue dans la sémiotique du sonore, selon que l'on se place, dans cet exemple, dans un objectif biologique ou dans un objectif linguistique. J'ai choisi, par cette lecture à deux faces, de mettre l'accent sur l'aspect subjectif de l'interprétation sémiotique.

Après ces remarques à propos du chant des gibbons, je vous invite à explorer le chant des bébés, animaux sémiotiques incapables d'articuler, mais qui révèlent une volonté de communication, et donc construisent au cours de leur croissance leur aptitude à s'exprimer de façon musicale, par la voix.

### **3 Le mamanais et autres langages des bébés du XXI<sup>ème</sup> siècle**

Le terme de *mamanais* concerne le langage établi entre la mère et le nourrisson, lorsque l'appareil phonatoire de l'enfant ne lui permet pas encore d'articuler : son larynx est encore haut pour permettre une meilleure déglutition. À ce stade, lorsque les adultes s'adressent au bébé, ils utilisent souvent un langage non standard, « ce style se caractérise par une hauteur élevée et une exagération des intonations » [16].

Cette exagération des intonations transforme le langage parlé en une sorte de chant. La volonté de communication pendant cette période semble employer la *musicalisation* de la voix.



**Photographie 1.** Conversation entre une mère et son bébé. Anonyme.

Le *mamanais*, apparaît curieusement comme un mode de communication à la fois sophistiqué et artistique, et à la fois naturel, spontané, naïf. L'exagération des intonations et la *musicalisation* de la parole est involontaire et en même temps, elle est le moyen d'exprimer la volonté de communication. Le son émis en est le médium, et la qualité d'écoute du destinataire, issue de ce contexte particulier et intime – qualité d'écoute partagée par l'adulte et l'enfant – va déterminer la façon d'être au monde, à travers un geste musical : « L'expérience de l'écoute dévoile notre rapport au milieu, notre rapport aux autres » [13].

L'expérience de l'écoute est fondamentale dans ces recherches sur la sémiotique du sonore, et c'est elle qui va déterminer l'invention musicale dans la communication, et approuver l'hypothèse d'origine commune du langage et de la musique.

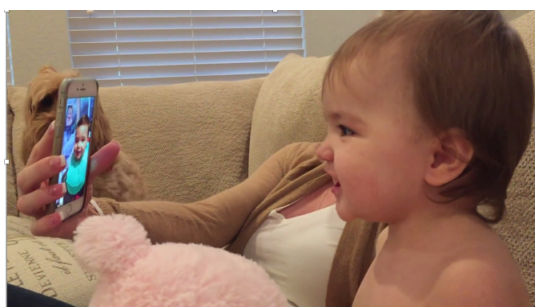
C'est ainsi que je me suis intéressée à la communication entre deux bébés, spontanée et créative, isolée du contexte éducatif lié à la relation maman-bébé. À partir d'une vidéo postée sur le site Dailymotion [1] qui montre deux bébés en conversation vidéo via un téléphone portable, j'ai cherché à comprendre le processus de construction de la communication, à l'aide d'outils d'analyse musicale classique.

Certes, l'outil, avec ses approximations, oriente le résultat de la recherche, mais le choix de celui-ci m'a semblé adapté au contexte socio-culturel dont sont issus ces bébés états-unien du XXI<sup>ème</sup> siècle. Cet exemple singulier ne tente donc pas d'établir des invariants, mais de montrer une interprétation subjective de la mécanique de construction du langage, à partir de la musique, et plus précisément ici à partir de la musique occidentale tonale dont sont très probablement imprégnés ces bébés. L'étude est ici réduite à un seul exemple, extrait d'un corpus provenant d'une minuscule population de culture occidentale, et il serait intéressant de l'étendre à des contextes de civilisation différentes, à l'aide de recherches extrême-orientales, arabes, océaniques ou africaines par exemple.

Il est néanmoins amusant de constater que l'on pourrait penser musicologiquement l'origine de la musique dans le langage, tandis que d'un autre côté, certains linguistes ou psychologues cherchent l'origine du langage dans la musique : « la musique est une forme prélinguistique de communication, qui prépare l'émergence du langage et facilite le développement de l'empathie » [8]. L'hypothèse du musilangage, en tant qu'origine commune de la musique et du langage pourrait raisonnablement réconcilier ces deux perspectives.



Dans l'exemple qui suit, je me suis concentrée sur le processus de co-création – observé d'un point de vue musical – inventé par deux bébés animés par un désir de communication.



**Photographie 2.** Bébés en visio-conférence. Anonyme.

En partant du principe que le sonore est musical, j'ai transcrit les enregistrements des productions vocales des bébés sur partition, en utilisant le système de notation musicale occidentale classique. Cette transcription est la simplification d'un signal sonore complexe. J'ai extrait uniquement les voix des bébés, et j'ai ignoré celles des mamans – qui certes ont un rôle affectif très encourageant. Enfin, pour avoir une vue synthétique, je n'ai pas respecté les silences entre les interventions. A propos du tempo, 120 noires à la minute, je l'ai choisi en fonction du rythme cardiaque moyen du bébé entre 6 mois et un an, qui varie entre 80 et 160 battements par minutes. Ce choix est déterminé par la théorie selon laquelle le rythme cardiaque et le plaisir musical sont liés [17].

Sur la partition ainsi élaborée apparaissent deux portées par système, une pour la voix du bébé A et l'autre pour celle du bébé B – celui dont la voix est diffusée via le téléphone. Les signes visuels de communication sont aussi ignorés dans cette analyse, pour se focaliser uniquement sur le sonore, objet de la journée d'études intitulée « Vers une sémiotique du sonore » qui donne lieu à cette publication.

Avec cette notation, il est possible d'analyser la construction de motifs, qui se réalise à deux, en co-création. Nous allons voir que cette improvisation est très structurée :

**Exemple musical 1.** Transcription d'une conversation entre deux bébés. Analyse musicale.

Au départ, encadré en bleu, un ré est posé avec un glissando jusqu'à l'octave du dessous. C'est le départ d'une phrase. Ce ré est répété, avec une inflexion au  $\frac{1}{2}$  ton inférieur, sur un ré bémol, encadré en orange. Le ré bémol est utilisé pour débiter la construction du motif. L'intervalle de  $\frac{1}{2}$  ton, encadré en violet sert tantôt d'antécédent, tantôt de conséquent dans le motif, avec différentes transpositions.

Le motif est encadré en vert, mesures 2-3-4. Le mécanisme de variation est celui employé en musique classique : rétrogradation (c'est à dire que le thème est joué à l'envers), variations rythmiques et modification des notes de début ou fin, en gardant le dessin général.

Mesure 4 (à la fin du premier système), le bébé A chante uniquement la fin du motif rétrogradé, ce que remarque le bébé B, qui enchaîne avec le début du motif transposé mesure 5, deuxième système.

Puis mesures 5-6-7, l'encadré vert, montre l'évolution du motif : Les intervalles se resserrent puis s'élargissent pour aboutir à notre octave de ré initial. Cette longue phrase construite à deux se termine.

Nous avons ensuite la naissance d'un second motif, encadré en jaune, qui varie autour de la note do. Le système se complexifie, chaque bébé ne se contente pas de faire des mots ascendants ou descendants, comme dans le motif vert, mais des motifs au dessin ascendant et descendant (visualisés avec des paraboles jaunes sur le troisième système).

Nous voyons encore une fois que les bébés se complètent dans l'invention sonore, par exemple mesure 9, le bébé A rajoute un conséquent au motif tout juste inventé par le bébé B.

Enfin, sur le 4<sup>ème</sup> système, nous avons un retour dans la tonalité initiale, de ré, avec la construction progressive d'un accord parfait majeur. À partir de ré, sont ajoutés successivement la quinte, l'octave, et la tierce. Il semble indiscutable que c'est très construit, et que ça se construit à deux.

Enfin, dans la dernière mesure, une sorte de drapeau final termine la séquence.

Cet exemple est à considérer dans son contexte : une communication singulière entre bébés du XXI<sup>ème</sup> siècle, sans aucun doute imprégnés d'une musique tonale dominante dans les médias. Dans un autre contexte, le résultat serait différent.

La seule conclusion possible semble être que ces deux bébés états-uniens du XXI<sup>ème</sup> siècle s'approprient le langage au cours de leurs évolutions phonatoires, en passant par la musique. Ils cherchent à communiquer, et façonnent leurs appareils phonatoires ainsi que leurs capacités socio-cognitives par la pratique vocale et la co-création.

## 4 Conclusion

Concernant l'hypothèse du *musilangage* comme origine commune du langage et de la musique, ces deux cas singuliers étudiés dans une perspective musicologique et linguistique pourraient apporter une contribution. Cependant, il ne me semble ni opportun, ni possible de déterminer qui, de la musique ou du langage seraient à l'origine du langage ou de la musique. Le concept de *musilangage*, en tant qu'origine commune reste toutefois efficient pour se questionner d'un point de vue archéologique sur les liens de la musique et du langage. Cette question me semble fondamentale lorsque l'on s'interroge sur la sémiotique du sonore.

## Références

1. Anonyme : Deux bébés discutent sur FaceTime (2016)  
<https://www.dailymotion.com/video/x5c761z>
2. Bickerton, D.: La langue d'Adam. Dunod La Recherche, Paris (2010)
3. Boë L.J., Granat J., Autesserre D., Perrier P., Peyre É.: Variation et prédiction de la position de l'os hyoïde de l'Homme moderne à Neandertal. (2007)  
<https://www.researchgate.net/publication/281473539>
4. Brown, S.: The 'musilanguage' model of music evolution. In : The Origins of Music, B. Merker et S. Brown, (dir.) N. Lennart Wallin, p. 269. Cambridge: MA, MIT Press, (2000)
5. Castanet, P.A.: Art-Science et Musique Contemporaine: Les Métaphores de la Fractalité 1. New Sound: International Journal of Music, 01 01 2017, pp. 91-112 (2017)
6. D'Anastasio, R., et al.: Micro-Biomechanics of the Kebara 2 Hyoid and Its Implications for Speech in Neanderthals. Polos One, pp. 1-7 (2013)

7. Dodane, C., Martel, K.: La prosodie chez le bébé au stade pré-linguistique : premières formes stables. Actes des XXVIIème Journées d'Etude sur la Parole, 9-13 juin 2008. Avignon (2008)
8. Jakobson, R.: Linguistique et Poétique. In: Essais de linguistique générale, pp. 209-248, Minuit, Paris (1963)
9. Jakobson, R.: Six leçons sur le son et le sens. Éditions de minuit, Paris (1976)
10. Lehning, H., Francis C., et Benoit R. : Les Fractales. Pole, Paris (2004)
11. Mithen, S.: The singing neanderthals: The origins of music, language, mind, and body. Harvard University Press, Cambridge, Massachusetts (2006)
12. Pardo, C.: Pratique des Écoutes, Écoutes des Pratiques. IMERA, Marseille (2018)
13. Richman, B.: How music fixed "nonsense" into significant formulas: on rhythm, repetition, and meaning. In: The Origins of Music: An introduction to Evolutionary Musicology, de N.L. Wallin, B. Merker, & S. Brown, pp. 301-314. MIT Press, Cambridge, Massachusetts (2000)
14. Ruaux, C.: Production de la Parole. (2018) <https://clemedicine.com/1-production-de-la-parole/>
15. Siegler, R.S.: *Enfant et raisonnement : Le Développement Cognitif de l'Enfant.* De Boeck, Bruxelles (2000)
16. Tomatis, A.: *Pourquoi Mozart ?* Fixot, Paris (1994)
17. Vecchione, B.: Une Poétique du Motet Médiéval : Textes, Hypotextes et Niveaux de Discours dans l'Ave Regina Celorum / Tenor [Joseph] / Mater Innocencie de Marchetto da Padova. In: Beyond 50 years of Ars Nova Studies at Certaldo, 1959–2009., de Ziino Agostino , Zimei Francesco Gozzi Marco, pp. 69-133. LIM, Lucca (2014)
18. Wade, N.: Neanderthals Had Important Speech Gene, DNA Evidence Shows. *The New York Times*, 19 October 2007.