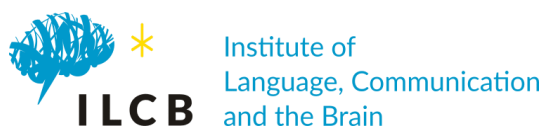


Illustration : Sylvie Pic, « The myth of interiority »



Transdisciplinary Workshop (French/English) Maison des Astronomes, IMéRA, Marseille 9 & 10th of May, 2022

Outcome: a collective publication on methodological and theoretical issues

Abstract

The complexity of animal communication signals (including humans) can be evaluated on the basis of the combinatorics of the units that make up these signals (property also referred to as “double articulation”). The first articulation concerns the arrangement of the semantic units (words) into sentences and the second corresponds to the arrangement of elementary sounds (e.g. human phonemes) into meaningful units (e.g. human words). The possibility to build different combinations at each level allows to produce an exponential number of new statements.

To compare the complexity of the vocal communication signals of different animal species, including humans, we propose to investigate possible combinatorial units in non-human animals.

For this, we intend to address two principal issues:

- 1- The segmentation and clustering methods / principles used in the study of human and non-human vocalizations (mainly non-human primates and cetaceans).
- 2- The association/combinability of units from different sensory channels (acoustic, visual, tactile) and the processing of these multimodal/multicomponent units

This will be done using the tools and concepts from linguistics, psycholinguistics, comparative psychology, philosophy, bioacoustics, neurosciences, computer sciences and ethology.

La complexité¹ des signaux de communication animaux (dont humain) repose notamment sur la **combinatoire** des unités qui composent ces signaux (propriété aussi désignée sous le terme « double articulation » par le linguiste André Martinet, 1960 ; voir aussi Zuidema & de Boer, 2018). La première articulation concerne l'agencement des unités sémantiques (mots) dans la phrase et la seconde l'agencement des sons élémentaires (phonèmes) dans l'unité. La possibilité de produire des combinaisons différentes à chacun de ces deux niveaux correspond à la combinatoire, laquelle permet de produire un nombre exponentiel de nouveaux énoncés.

Pour comparer la complexité des signaux de communication vocaux de différentes espèces animales, dont l'espèce humaine, nous proposons de nous interroger sur l'existence de structures homologues aux unités décrites ci-dessus chez les animaux non-humains.

Pour cela, nous étudierons :

- le problème de la **segmentation** des flux acoustiques non-humains (qu'est-ce qui constitue une unité de son, une unité de sens, comment les isoler ?),
- la diversité, la taille des **unités** qui composent ces signaux acoustiques non-humains
- leur comparaison structurelle et fonctionnelle avec les unités constitutives de la parole humaine
- leur **association/combinabilité** avec des signaux d'autres canaux sensoriels (visuels, tactiles) pour former des unités dites multimodales/multicomposants.

Il s'agira de croiser les **outils** et **concepts** issus de la linguistique, psycholinguistique, psychologie comparée, philosophie, bioacoustique, neuroscience, informatique et éthologie.

Axe 1- Segmentation et unités sonores (chez les espèces principalement vocales)

Nous commencerons par rappeler rapidement quelles sont les méthodes de segmentation du signal acoustique les plus utilisées (quel modèle math/informatique pour quel type de segmentation ?) chez l'espèce humaine, le primate non-humain (marmoset) et plusieurs espèces de cétacés.

La **segmentation** du signal aboutit à la constitution de clusters d'unités soit structurels, soit (plus rarement) fonctionnels. Les clusters structurels sont purement descriptifs, ils servent à étudier la structure des séquences de vocalisation (et rechercher leur organisation interne). Les clusters fonctionnels (« nuages sémantiques ») ont un sens pour l'animal mais sont

¹ La complexité est envisagée ici dans le cadre de la théorie de l'information, et opérationnalisée par le biais de l'algorithme de Shannon (e.g., Cowan, Hanser & Doyle, 2002).

difficilement accessibles à l'entendement humain. Peut-on dépasser cette opposition entre segmentation structurelle et fonctionnelle ?

Les unités considérées dans les approches comparatives interspécifiques sont principalement les unités sonores, mais les silences, la durée des sons et la durée des silences peuvent également être pris en compte (e.g., Ravignani & Norton, 2017 ; De Gregorio et al., 2021 ; Weir et al., 2007). Si le **rythme** est porteur de signification, comment l'intégrer à l'analyse des sons et de leurs variations de fréquences ?

Nous nous interrogerons également sur le **grain** des unités acoustiques. Chez l'espèce humaine, les transitions de formants, les formants, le *voice-onset time* (durée entre la prononciation de la consonne initiale et la prononciation de la voyelle) jouent un rôle important dans la décomposition des mots en phonèmes. Doit-on et peut-on rechercher des analogies chez les autres animaux vocaux ?

Enfin, nous étudierons le rôle de la **motricité** et des **boucles perception/action** dans la détermination du grain de ces unités. Chez l'humain, on sait que la perception auditive des phonèmes et des syllabes se rapporte aux rythmes moteurs des mouvements phonatoires (phénomènes qui s'expliquent dans le cadre de la théorie motrice de la perception, voir Barnaud, Schwartz, Bessière & Diard, 2019). Peut-on établir un parallèle similaire entre motricité et unités perceptives chez les animaux à dominante vocale non-humains ? Ces considérations seront à rapporter aux **contraintes biomécaniques** (morphologie de l'appareil phonatoire) exercées sur la mise en place des processus de production vocale (e.g., Boë et al., 2017 ; Dediu et al., 2021).

Axe 2 - Unités multimodales/multicomposants

Les systèmes de communication animaux sont rarement unimodaux (uniquement vocaux par exemple) mais très majoritairement multimodaux. Les signaux en provenance de différents canaux sensoriels sont émis simultanément ou séquentiellement. Selon Bohn et collaborateurs, la suppression d'un seul de ces signaux rend le message inopérant (Bohn et al., 2022).

La perception multimodale d'objets chez l'humain est bien étudiée. Elle répond à un ensemble de lois et processus bien décrits (superadditivité, binding, simultaneity, hebbian learning ; voir Calvert, Spence & Stein, 2004). Ces lois et processus s'appliquent-ils également à la communication multimodale ? Peut-on enrichir notre compréhension des systèmes de communication multimodaux sous l'éclairage des données issues des neurosciences cognitives en matière de perception multimodale ?

L'éthologie distingue les signaux multimodaux (plusieurs modalités sensorielles simultanément) des signaux multicomposants (plusieurs signaux provenant d'un même

canal comme par exemple un geste et une posture ; Aychet & Lemasson, 2021). Dans leur article, Partan et Marler (2005) proposent une classification des signaux multimodaux/multicomposants et de leurs propriétés qui ne sont pas sans rappeler les propriétés de la perception multimodale telle qu'abordées par les neurosciences (voir la figure ci-dessous à titre d'illustration).

En parallèle à la question des convergences entre neurosciences et éthologie, nous chercherons à définir ce qu'est une unité multimodale/multicomposants et comment considérer et analyser l'association d'informations provenant de différentes sources.

Tout au long de cet atelier, nous insisterons sur les points de vue croisés des diverses disciplines concernées.

Nous nous répartirons le travail en fonction des domaines d'expertise de chacun.e. Nous discuterons sur la base d'articles choisis, méthodologiques et théoriques, et de présentations orales synthétiques. Il s'agira de croiser et trier les idées les plus pertinentes.

SEPARATE COMPONENTS		MULTIMODAL COMPOSITE SIGNAL	
signal	response	signal	response
Redundancy	a → □	a + b → □	Equivalence (intensity unchanged)
	b → □	a + b → □	Enhancement (intensity increased)
Nonredundancy	a → □	a + b → □ and ○	Independence
	b → ○	a + b → □	Dominance
		a + b → □ (or □)	Modulation
		a + b → △	Emergence

Figure extraite de Partan et Marler (2005) proposant une classification des signaux animaux multimodaux/multicomposants et de leurs effets sur la réponse du récepteur. L'*enhancement* ici présenté pourrait correspondre à la superadditivité décrite en neurosciences de la perception multimodale.

REFERENCES BIBLIOGRAPHIQUES

- Aychet, J., Blois-Heulin, C., & Lemasson, A. (2022). Sequential and network analyses to describe multiple signal use in captive mangabeys, *Animal Behavior*.
- Barnaud M-L, Schwartz J-L, Bessière P, Diard J (2019) Computer simulations of coupled idiosyncrasies in speech perception and speech production with COSMO, a perceptuo-motor Bayesian model of speech communication. *PLoS ONE* 14(1): e0210302. <https://doi.org/10.1371/journal.pone.0210302>
- Boe L-J, Berthommier F, Legou T, Captier G, Kemp C, Sawallis TR, et al. (2017) Evidence of a Vocalic Proto-System in the Baboon (*Papio papio*) Suggests Pre-Hominin Speech Precursors. *PLoS ONE* 12(1): e0169321. doi:10.1371/journal.pone.0169321
- Bohn, Liebal, K., Tessler (2022). Great ape communication as contextual social inference: a computational modeling. <https://psyarxiv.com>
- Calvert, G., Spence, C., & Stein, B. (2004). *The handbook of multisensory processes*. MIT Press.
- Dediu D, Moisik SR, Baetsen WA, Bosman AM, Waters-Rist AL. (2021). The vocal tract as a time machine: inferences about past speech and language from the anatomy of the speech organs. *Phil. Trans. R. Soc. B.*, 376: 20200192. <https://doi.org/10.1098/rstb.2020.0192>
- De Gregorio, C., Valente, D., Raimondi, T., Torti, V., Miaretsoa, L., Friard, O., Giacoma, C., Ravignani, A. & Gamba, M. (2021). Categorical rhythms in a singing primate. *Current Biology* 31, R1363–R1380.
- McCowan, B., Hanser, S., & Doyle, L. (2002). Using Information Theory to Assess the Diversity, Complexity, and Development of Communicative Repertoires. *Journal of Comparative Psychology*, Vol. 116, No. 2, 166–172.
- Martinet, A. (1960). *Eléments de linguistique générale*. Ed. Colin.
- Partan, S. & Marler, P. (2005). Issues in the Classification of Multimodal Communication Signals. *The American Naturalist*, 166, 231-245
- Ravignani, A., & Norton, P. (2017). Measuring rhythmic complexity: A primer to quantify and compare temporal structure in speech, movement, and animal vocalizations. *Journal of Language Evolution*, 2017, 4–19 doi: 10.1093/jole/lzx002.
- Weir, C., Frantzis, A., Alexiadou, P., & Goold, J. (2007) The burst-pulse nature of ‘squeal’ sounds emitted by sperm whales (*Physeter macrocephalus*) *J. Mar. Biol. Ass. U.K.*, 87, 39–46.
- Zuidema, W. & de Boer, B. (2018). The evolution of combinatorial structure in language. *Current Opinion in Behavioral Sciences*, 21:138–144

Programme

Lundi 9 mai - Axe 1 – Segmentation et Unités sonores

8h45 - Welcome café

9h-9h30 - Situer de la question du jour (segmentation comparée) - Marie Montant

9h30-10h - Segmentation/clustering de la parole humaine - Noel Nguyen, Christine Meunier (Thierry Legou)

10h-10h30 – Unité de son, unité de sens, rôle de la prosodie dans la segmentation de la parole humaine - Cristel Portes

Pause café

11h- 11h30 - Segmentation/clustering vocalisations monkeys –Manon Obliger, Charly Lamothe (Pascal Belin)

11h30-12h - Segmentation/clustering vocalisations monkeys –Alban Lemasson

12h- 12h30 – Unité structurelle/unité fonctionnelle monkeys –Alban Lemasson

12h30 – 13h – Discussion

Déjeuner

14h30 – 15h00 – Segmentation/clustering vocalisation des baleines –Julie Patris & Franck Malige

15h – 15h30 - Vocalisations et répertoire des dauphins –Fabienne Delfour

Pause café

16h – 17h – Marine sound-landscapes and repertoires (cetaceans) –Gianni Pavan (introduced by Fran Cagnacci)

17h– 18h30 – Discussion, synthèse de la journée, pistes et idées à retenir, à développer

Apéro

Mardi 10 mai - Axe 2 - Unités Multimodales/Multicomposants

8h45 - Welcome café

9h-9h30 – Situer la question du jour (multimodalité) - Marie Montant

9h30 – 10h - Studying multimodal communication in the wild –Francesca Cagnacci

10h – 10h30 - Multimodalité des signaux de communication monkeys –Juliette Aychet

10h30 – 11h - Comment traiter le chevauchement/l'onset des signaux multimodaux ? – Lise Habib Dassetto

11h – 12h – Discussion

Déjeuner

13h30h-14h - Boucle perception-action et rôle des rythmes moteurs dans la perception de la parole humaine – Benjamin Morillon

14h-14h30 – Anatomie du tractus vocal et production de la parole humaine : variations inter-individuelles et variations inter-linguistiques – Dan Dediu

14h30 – 15h15 - Preliminaries to non-human primate vocalizations analysis (in relation to the anatomy of the vocal tract) - Louis Jean Boë

Pause café

15h45- 16h15 –Reconnaissance cross-modale de formes proto-symboliques (audio-visuelles) – Konstantina Margiotoudi

16h15 – 16h45 - La reconnaissance multimodale d'objets – des principes et des lois exportables aux signaux de communication ? - Anne Kavounoudias

16h45-17h15 – Multimodalité, intermodalité et binding. Clarification de ces notions – Salomé Fazi

17h30- 18h30 – Discussion finale

Qu'a-t-on appris ? Que peut-on transposer d'un champ à un autre ? Qu'est-ce qu'une unité (grain)/répertoire ? Rôle du corps ? Rôle des silences, du rythme ? Umwelten et pragmatique ...

What's next ?

Recommandations aux speakers :

Soyez synthétique, mettez l'accent sur les principes théoriques et la méthode qui permettent de répondre aux questions de l'atelier (il ne s'agit pas de présenter un panorama de vos propres recherches).

N'oubliez pas que les autres ne sont pas forcément de votre domaine !

Le must : si vous pouvez voyez à l'avance un article de travail représentatif de la question que vous abordez !