**Acoustic communication and sound degradation:**

**How do the individual signatures of the songs of an Amazonian bird transmit over distance?**

**Problématique :** Chez un animal communiquant par des vocalisations, les individus forment des réseaux de communication parcourus par des échanges d’information. Pour comprendre le fonctionnement de ces réseaux, il faut mesurer comment la dégradation des sons au cours de leur propagation dans l’environnement impacte ces échanges d’information.

Nous prendrons comme modèle un oiseau de la forêt amazonienne, le pihau hurleur *Lipaugus vociferans*. Chez cet oiseau, les mâles chantent depuis des arbres distants de quelques dizaines de mètres. Dans la forêt, ils ne peuvent se voir et se reconnaissent les uns les autres uniquement par leur chant. Le chant du pihau informe en effet sur l’identité du chanteur. Ces signatures vocales sont cependant altérées au cours de la propagation du son à travers la végétation de la forêt (perte d’intensité, filtration, réverbération…). L’information « identité du chanteur » est donc caractérisée par un certain « espace actif », au-delà duquel elle ne pourra plus être perçue par les individus récepteurs. Caractériser l’espace actif de la signature vocale est essentiel pour appréhender le fonctionnement des réseaux de communication acoustiques de l’oiseau.

**Objectif du projet :** Utiliser une approche *machine learning* pour mesurer la fiabilité à différentes distances de l’émetteur de l’information « identité individuelle » portée par le chant.

**Dataset :** Enregistrements réalisées lors d’expériences de propagation du son en forêt.

1. individus \* 10 chants par individu

* 100 chants propagés

5 distances (12.5, 25, 50, 100, 200 mètres)

* **500 chants enregistrés au total**

+ 100 chants originaux non propagés

Meta-données associées à chaque chant : identité de l’individu, distance de propagation, moment de l’enregistrement.

+ prendre une séquence de bruit de fond suivant chaque chant, de même durée.

* **500 séquences de bruit de fond**.

**Grand Total de 1000 enregistrements**.

**Méthodes :**

Ecouter les bandes enregistrées car on donne pas mal d’information dessus.

Par exemple hauteur du micro, HP (5 mètres), distances, date des manips, heure de la manip, meteo, matériel employé. Toute information qu’il faudra détailler dans les méthodes.

On donne aussi les mesures en dB du bruit ambiant + des signaux émis.

Le HP employé : on a la réponse du HP dans une publi de Thèvenet (voir folder Speaker\_specifications).

Cleaner les fichiers :

On peut utiliser PRAAT pour visualiser les spectrogrammes

Filtrer dans la bande passante du Lipaugus (frequence max = 6 000 Hz ; filtrer tout ce qu’il y a au-dessus)

Extraire les chants propagés – automatisation (mesurer au préalable les timing)

Récupérer les chants entiers, les rourous seuls, les pihaus seuls (2 parties du chant)

+ le bruit de fond (juste avant le premier rourou).

Soit 500 chants entiers

Eventuellement : + 500 rourous seuls + 500 pihaus seuls

Calculer les matrices MPS sur ces fichiers : bien choisir la fenêtre du spectro ET du MPS. Pour bien caractériser la modulation de fréquence au cours du chant.

Prendre le script MPS auprès de Quentin (voir aussi publi PloSOne Mouterde et al.)

Pour l’analyse acoustique, je pense qu’il faut employer 2 méthodes différentes : MPS et spectros. De plus, il faut qu’on tente différentes bandes de fréquence.

On peut aussi faire une analyse spectrale comme dans Mouterde et al.

On se retrouve avec 1 matrice MPS par chant et par bruit.

Total de 2000 MPS (500 chants entiers, 500 rourous, 500 pihaus, 500 bruits de fond).

classifs

* Non supervisés : par ex. PCA basées sur MPS (ou sur spectros)

(penser à tester différentes bandes de fréquence)

Prendre signaux pihau seuls ET signaux pihaus + bruits

Clustering

On verra sans doute des clusters correspondants aux distances de propag

Mais on peut penser qu’à courte distance, on aura des clusters par individus

Puis, avec la propag, ces clusters individuels deviendront indifférenciés.

On peut traiter les subsets suivants de manière indépendante :

* 500 chants entiers + 500 bruits de fond
* 500 chants entiers
* 500 rourous
* 500 pihaus
* 5 subsets se focalisant sur chaque distance (et reprenant soit les chants entiers, soit les rourous, soit les pihaus, soit les bruits de fond).
* …

On teste les bruits de fond car l’ambiance sonore (bruit de fond) varie certainement au cours de l’expérience de propagation et donc entre les enregistrements. On teste que ce n’est pas le bruit de fond qui pilote les résultats obtenus avec le classifieur (ou en tout cas on cherche à caractériser cet effet).

* Classifications supervisées

On entraîne le classifieur avec des chants de chaque individu enregistrés à la distance la plus courte, puis on teste avec les chants enregistrés aux différentes distances.

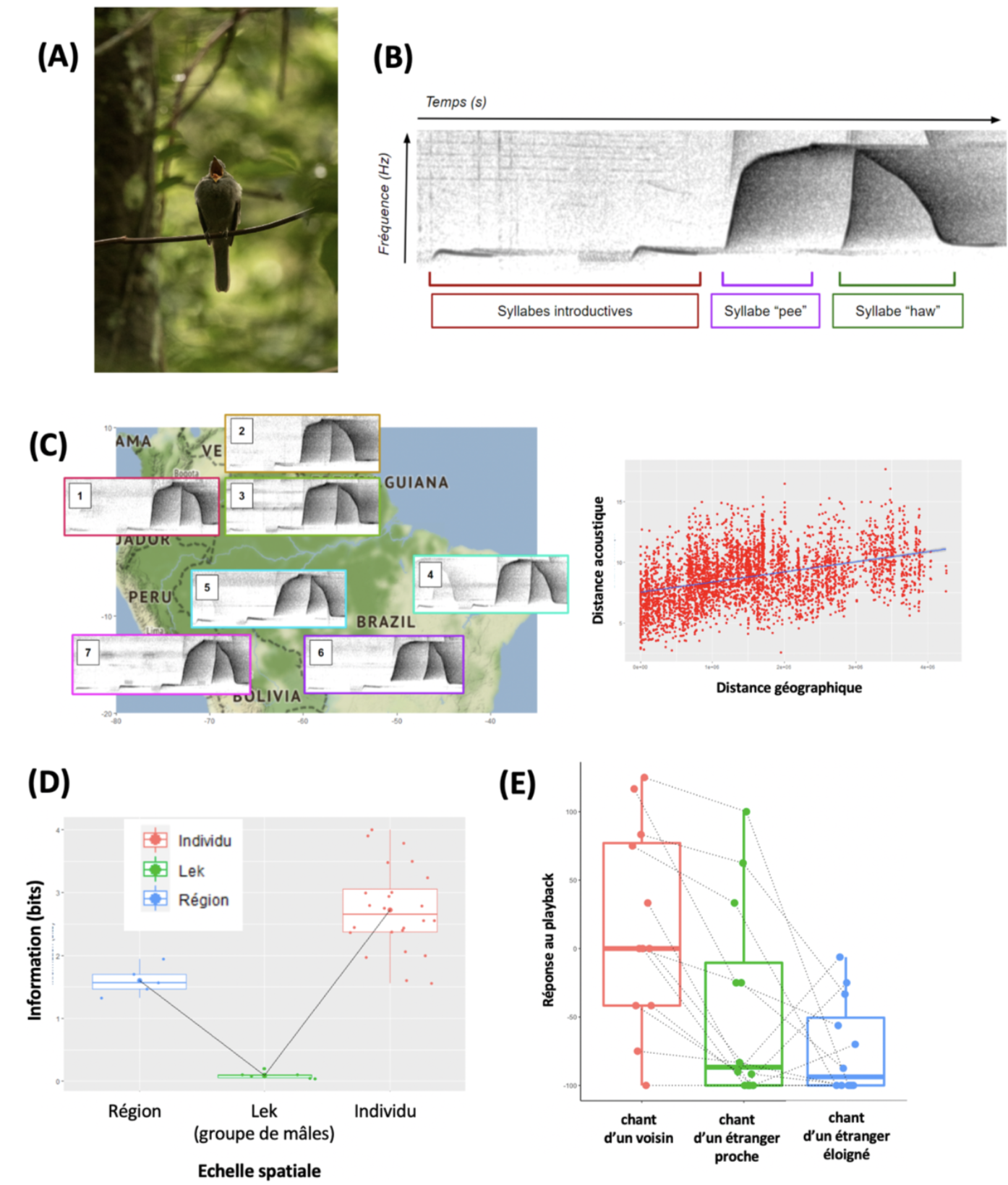
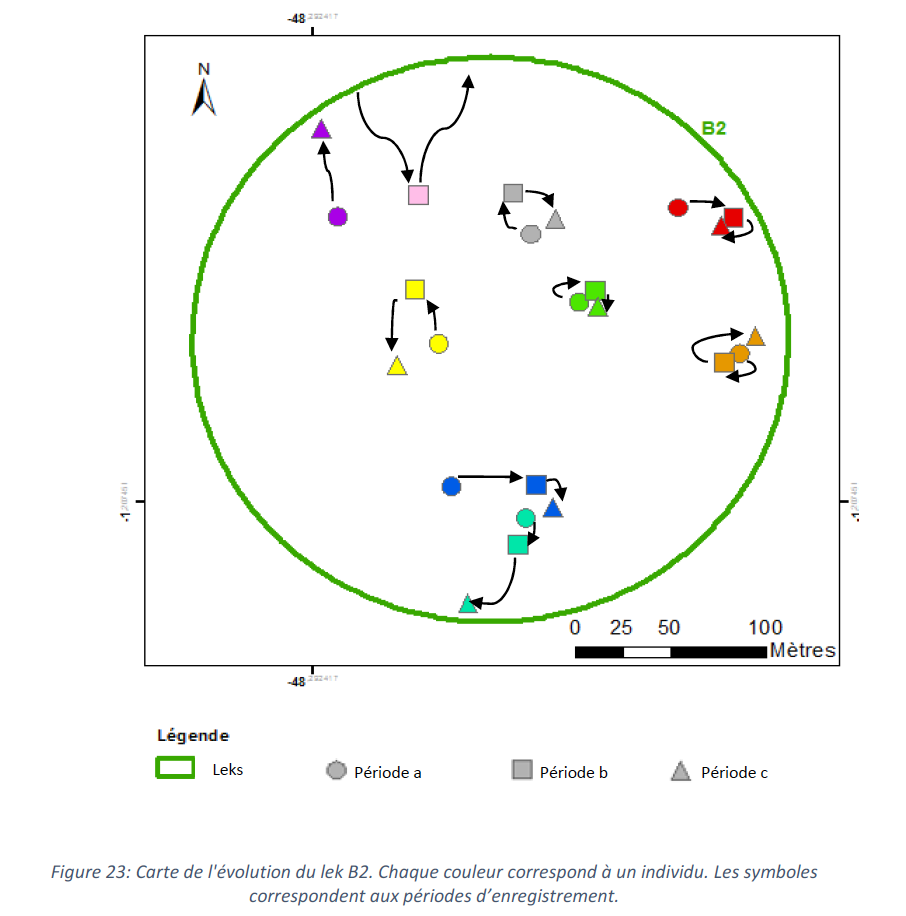
On peut aussi tester en entraînant à n’importe quelle distance et en testant avec les autres distances (peut-on apprendre à identifier un individu en l’entendant à distance, ou faut-il être le plus près possible ?). Autrement dit : même subsets que plus haut.

Voir la publi Mouterde + wapiti pour ces différents stratégies de training datasets

* Calcul de quantité d’information à partir des matrices de confusion (voir publi pic).

On pourra représenter une heat map de propagation de l’information identité individuelle en situation réelle, en utilisant les positions connues d’oiseaux sur un lek (voir données Fred Sèbe).

\*\*



**Biblio**

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0102842>

Cet article est exactement dans ton topic. Les signaux sont décrits avec les MPS, des spectros et des spectres. Cependant, les MPS ne sont pas vraiment utilisés pour mesurer l’impact de la propag sur la classification des individus.

Pour les stats, c’est la DFA qui est utilisée. Avec une méthode bien détaillée. Mais voir la publi dans PloS Comput Biol pour compléter.

Il y a plusieurs scénarios, où l’on utilise seulement les enregistrements à 2 m pour entrainer, ou la totalité des enregistrements. Ces scénarios sont bien justifiés du point de vue biologique. Ce sera aussi à essayer + à reprendre dans ton papier.

La figure 9 est aussi très intéressante. Surtout si on la fait sur les spectros.

<https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1010325>

Papier très intéressant. Avec tous les scripts R.

3 classifieurs sont comparés : DFA, SVM et xgboost

Tu pourrais (éventuellement) les essayer tous les 3.

Ce papier explique dans le détail les méthodes (voir la partie VI du papier)

Il y a aussi une step-by-step demo avec SVM :

<http://github.com/keruiduo/SupplMatBonobos>

Les recommandations indiquées dans cet article pour les SUNG datasets sont :

1. comparing several acoustic parameterizations as they may be complementary;
2. visualizing the dataset with Supervised UMAP to examine the species acoustic space;
3. adopting SVM (Support Vector Machines) rather than discriminant functional analysis as the baseline classification approach;
4. explicitly evaluating data leakage and possibly implementing a mitigation strategy.

<https://www.nature.com/articles/s41467-020-18772-3>

Ce papier est intéressant en particulier pour le calcul de la quantité d’information à partir des matrices de confusion. Calculer la quantité d’info (en bits) permet d’estimer le nombre d’individus discriminables, et de répondre à une question comme « Combien d’individus peut-on discriminer si on les entend tous après 12m de propag ? après 50m ? etc.

On pourrait faire aussi une cartographie (heatmap) de l’information sur le lek des Lipaugus, ce qui permettra d’estimer à partir de quand on ne peut plus du tout distinguer entre des individus + à combien de signatures individuelles au total un individu est-il exposé (pour discuter de la mémoire des signatures -voir le papier Elie & Theunissen sur ce thème, mais avec le zebra finch comme modèle).

<file:///Users/mathevon/Downloads/wapiti(1)/03.supervised_classification01.html>

Ce document web donne la méthode employée avec des cris de wapiti, dans le même genre de manip de propag que pour le Lipaugus.

Pour l’analyse acoustique, Vincent Arnaud a utilisé les MFCC. Nous ne prendrons plutôt les MPS.

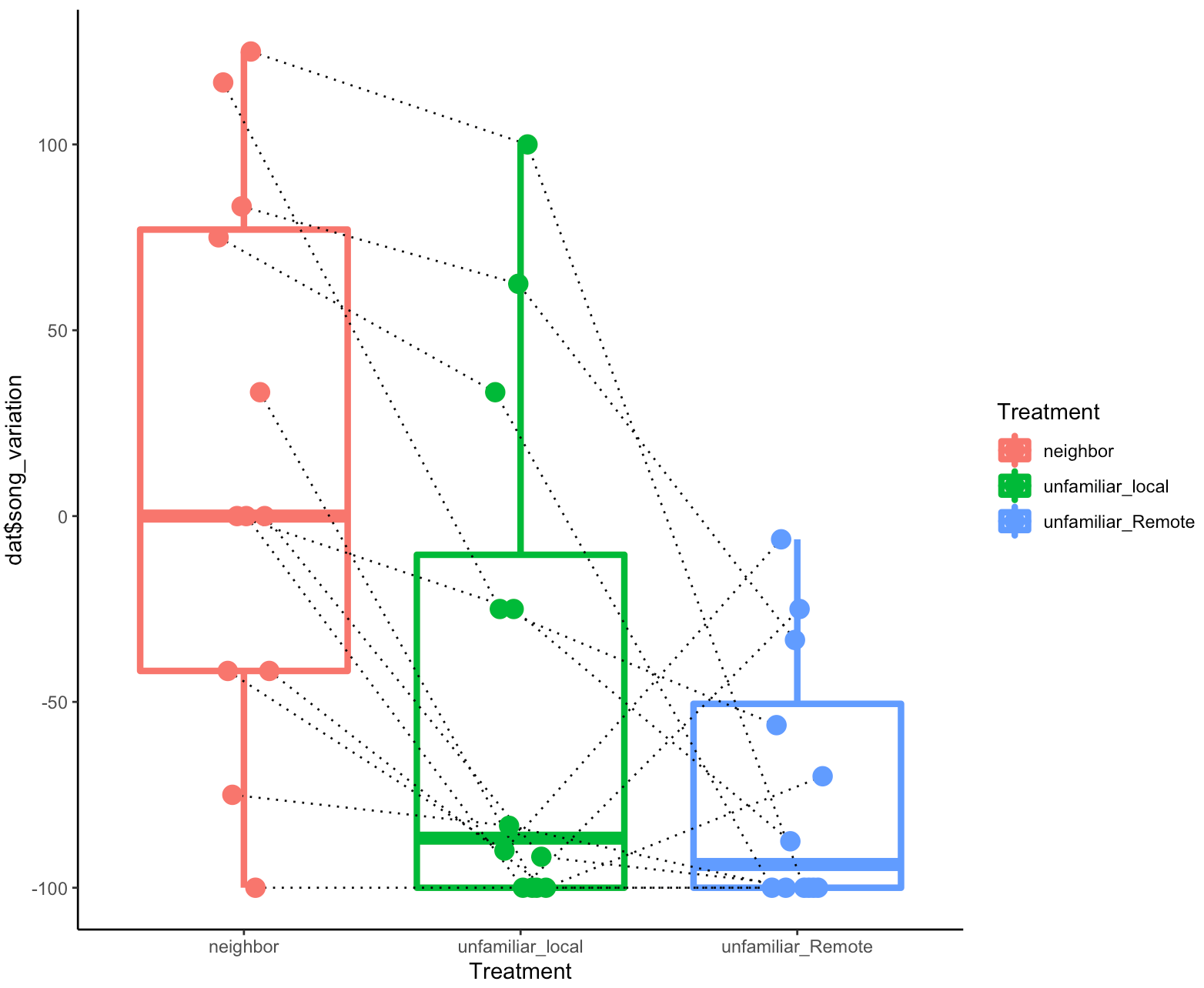
Pour l’analyse stat, il a utilisé des SVM (voir la publi bonobo dans Plos Comput Biol).

\*\*\*

**Playback experiments**

Il faudrait reprendre les data + refaire (éventuellement) les tests stats (en bayésien – ici c’est du non paramétrique) + refaire la figure.

Sur la figure, tu peux éventuellement mélanger les 2 catégories d’étrangers (mais il faut tenir compte de ces catégories dans les tests stats).



**Figure. Neighbor-stranger vocal discrimination in the screaming piha.** Each bird (n = 12) was successively challenged with 3 types of stimuli : a neighbor’s song (“neighbor”), a song from an unfamiliar individual from another lek in French Guyana (“unfamiliar\_local”), and a song from an unfamiliar individual recorded in the south-western part of the Amazonian basin (“unfamiliar\_remote”). Each of the three stimuli was played back three times, with a 10 seconds interval between each rendition. There was a minimal interval of 10 minutes between each stimuli series. The order of the stimuli series was balanced between the tested subject. Each stimulus was used in only one playback trial, thus avoiding pseudo-replication. The behavioral reaction of the birds was assessed for each playback trial by measuring the change in song occurrence between the first minute preceeding the start of the playback and the two minutes following the start of the playback, according to the formula : ((number of songs during the two minutes following playback/2) – number of songs during the minute preceeding playback) / number of songs during the minute preceeding playback. (Friedman rank sum test : Friedman chi-squared = 14.372, df = 2, p-value = 0.0007571; Pairwise comparisons using Wilcoxon signed rank test, “neighbor” vs “unfamiliar\_local”: p = 0.012, “neighbor” vs “unfamiliar\_remote” : p = 0.017, “unfamiliar\_local” vs “unfamiliar\_remote” : p = 0.262).

\*\*\*

A voir si on teste aussi la signature de lek

Pas encore préparé ces signaux

\*\*\*