How does musical expertise shape speech perception? Visual evidence from Auditory Classification Images.

Tianyun Wang[1,3], Léo Varnet[1,3], Chloé Peter[1,3], Gustavo Estivalet[2,3], Fanny Meunier[2,3], Michel Hoen[1,3]

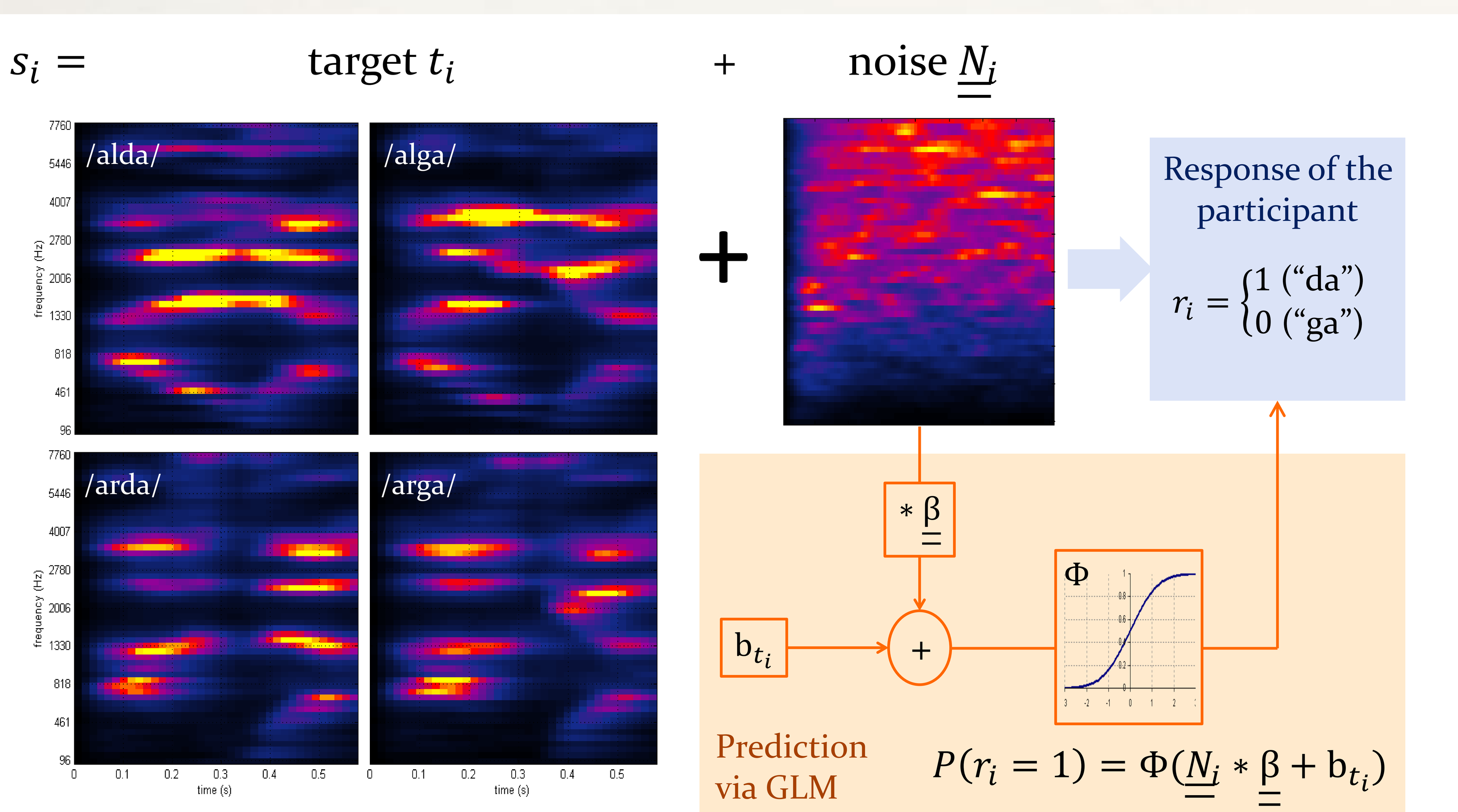
[1]Lyon Neuroscience Research Center, Lyon, France, [2]Laboratoire sur le Langage le Cerveau et la Cognition, Lyon, France, [3]Université de Lyon, Lyon, France.

1. Background

- An essential step in understanding the processes underlying speech comprehension is to identify which phonetic primitives are extracted from the acoustic signal and used to categorize a speech stimulus as specific phonemes.
- Musical training triggers neurophysiological plasticity in the auditory system, sharpening the tuning of cochlear filter-responses and reinforcing the pathway between brainstem and primary auditory cortex, thus improving neurophysiological encoding of speech sounds[1].
- These advantages enhance auditory discrimination, resulting in a better perception even in a higher noise background[2]. Nevertheless, the processes underlying musicians' advantage in speech-in-noise are not completely understood and further research must be carried out.
- In the present study, we used a **psychoacoustic imaging method** which isolates acoustic cues from natural stimuli in a speech-in-noise situation[3]. We applied the **Auditory Classification Image (ACI)** technique to identify the listening strategies of a musician group with intense and prolonged musical training and compare their results with those of a non-musician group.

2. Materials and Methods

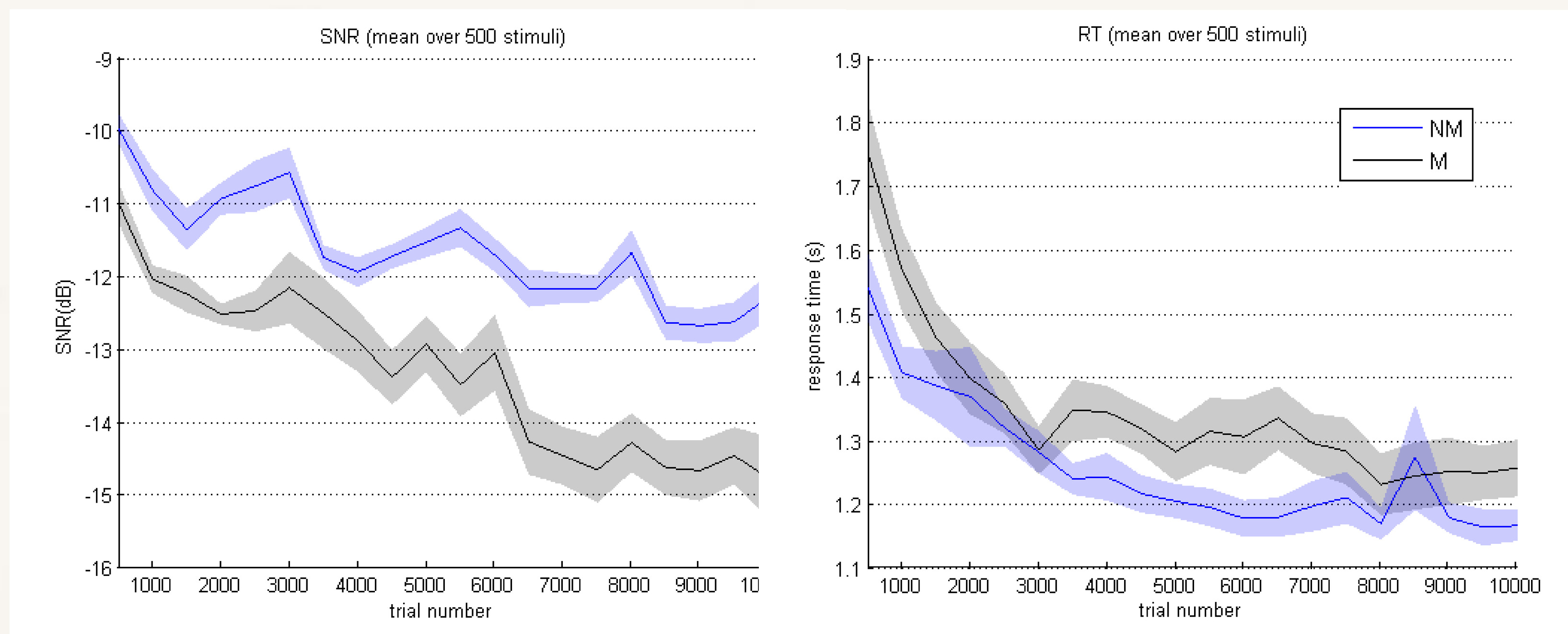
- **Stimuli**: 4 natural male speech productions of **/alda/**, **/alga/**, **/arda/**, and **/arga/**, equated in duration and RMS power, presented in random Gaussian noise.
- **Task**: Participants performed **10 000 phoneme categorizations** (20 sessions of 500 trials over 4 days), indicating whether the last syllable of the stimuli was **/da/** or **/ga/**.
- Signal-to-noise ratio (SNR) continuously adapted to ensure a correct response rate of 79%.
- **Data analysis:** The probability of **/da/** answer is linked to the time-frequency distribution of noise in each trial and the target actually presented through a **Generalized Linear Model**. The ACI (β) shows how the presence of noise at each point impacts the decision (i.e. which parts of the stimulus serve as cues for categorization).



Auditory Classification Image technique.

3. Behavioral results

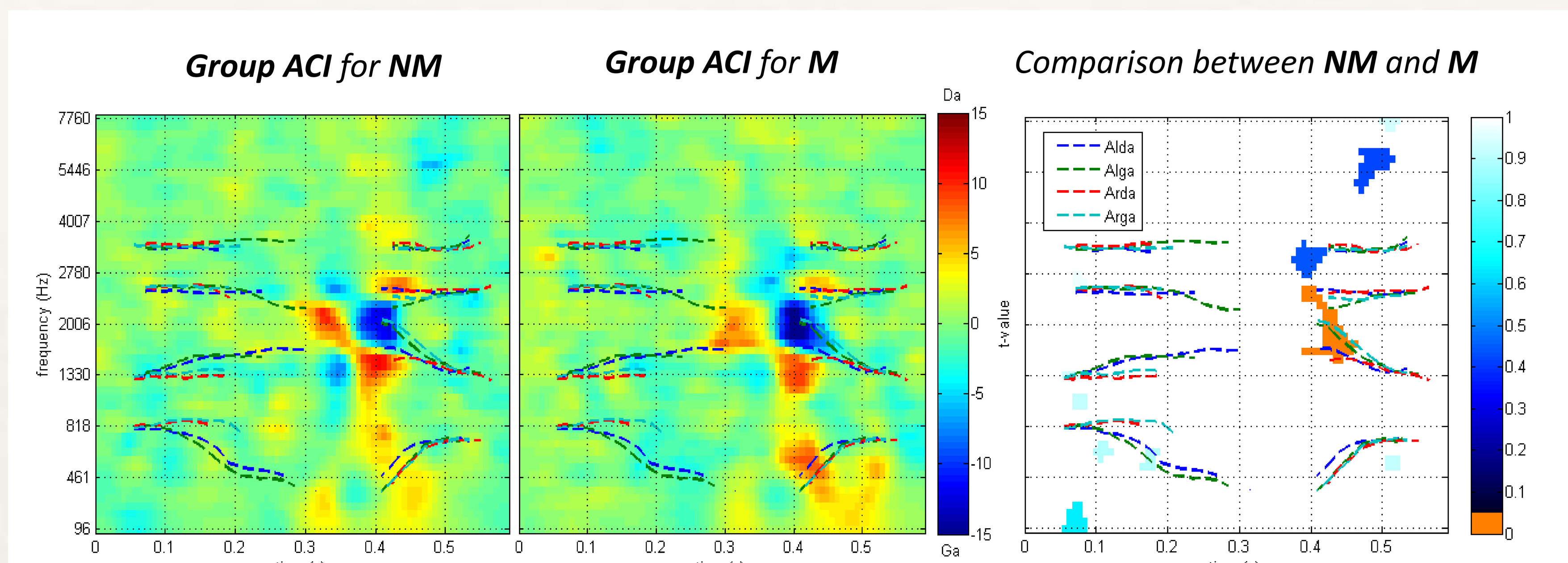
- **30 participants** (native French speakers). **Musician (M) group**: **N=15** (9 women, mean age 22.93), music practice average 16.27 years, perfect pitch N=8. **Non-musician (NM) group**: **N=15** (11 women, mean age 22.67).
- Mean correct response rate ≈ **79%** throughout the experiment, for each participant.
- A 2-way ANOVA showed significant effects of session number and group on SNR (both p<.0005), with a significant interaction effect. Response Times were only affected by session number (p<.0005) ⇒ M work with significantly lower SNR than NM while both groups exhibit SNR and RT learning effects over the course of the experiment.



Evolution of mean SNR and mean RT by session (± s.e.m) over the course of the experiment

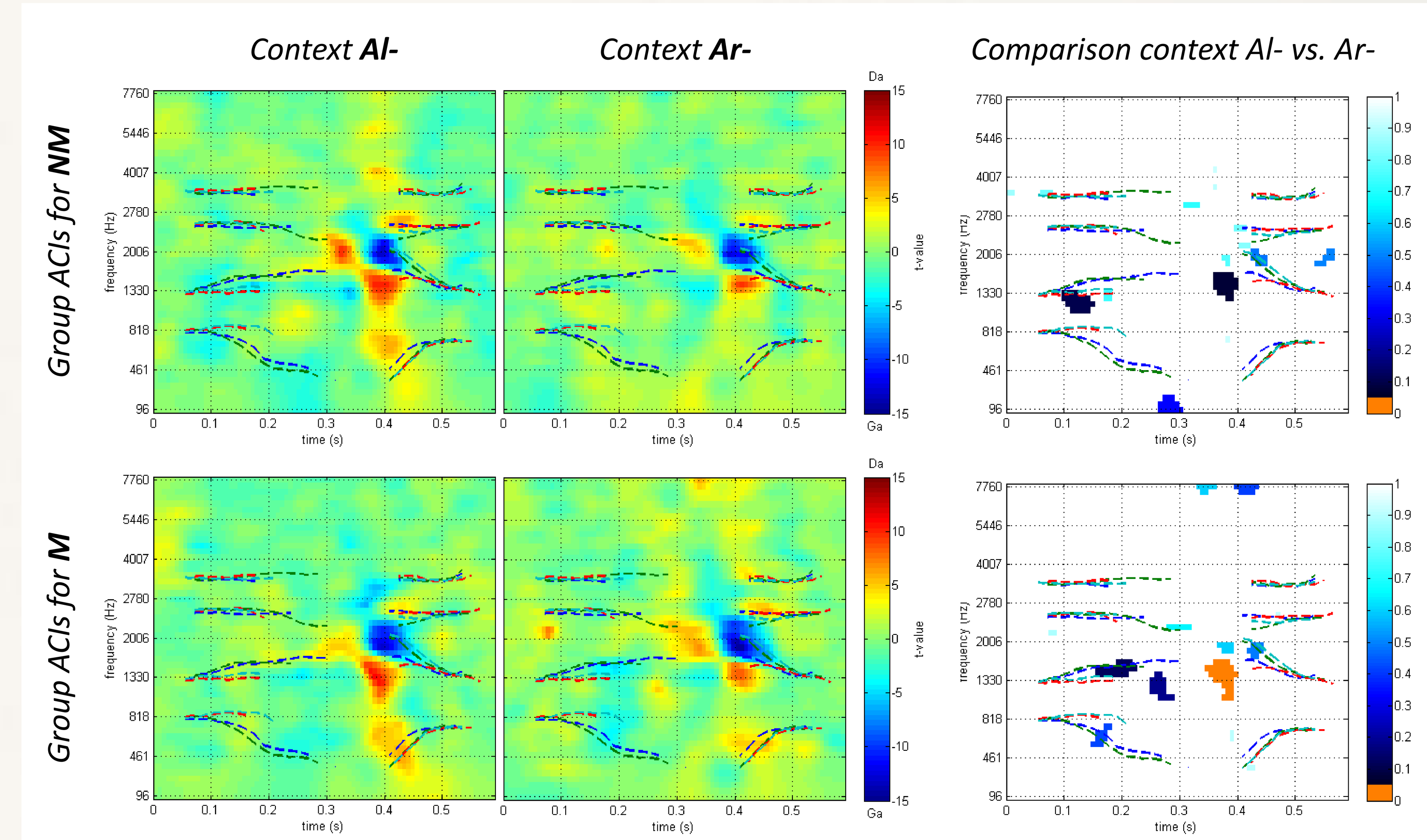
4. ACIs for non-musician participants

- Individual ACIs are derived and group averaged. High positive (red) and negative (blue) weights are time-frequency space where noise biases the response of the participant towards **/da/** or **/ga/**, respectively.
- Averaged ACIs of both groups show significant clusters of weights on **F2 and F3 formant onsets** (corrected t-test, FDR<.01) ⇒ these cues are critical for correct categorization of /da/ or /ga/.
- A cluster-based nonparametric test indicated that M listeners placed greater weights on the central negative cue (p=.04), suggesting **better selectivity for task-relevant cue** which more closely traces the second formantic transition in M.



Weights differing significantly are shown in orange, p(cluster)<.05

- Weights corresponding to high frequencies (>4000 Hz) are significantly lower in M's ACIs, demonstrating that M participants are **less influenced by noise in non-task-relevant frequency bands**.
- When comparing ACIs derived separately in context Al- and Ar-, M exhibited stronger differences in listening strategy based on preceding context than NM (M: p(cluster)=.04; NM: p(cluster)=.08).



Weights differing significantly are shown in orange, p(cluster)<.05

5. Conclusions

- As expected, M demonstrated better speech perception in background noise than NM; **musical expertise enhances hearing resistance to noise**.
- M and NM exhibited **similar listening strategies** for da/ga categorization task (acoustic cues on F2 and F3 onsets) in spite of **fine acoustic cue differences**.
- M's performances could be explained by an **enhanced selectivity** for the most behaviorally relevant aspect of the sound: M selectively focus on a small time-frequency region critical for correct da/ga categorization and are less disturbed by the presence of noise in unessential high-frequency region.
- Additionally, a more context-dependent weighting in M could account for finer representations of co-articulated features in the 4 stimuli.
- We show that ACI is a suitable method for studying group differences in auditory plasticity.

References

[1] Kraus, N., Chandrasekaran, B. (2010). Musical training for the development of auditory skills. *Nat Rev Neurosci 11*: 599-605.

[2] Strait, D., Kraus, N. (2011). Playing Music for a Smarter Ear: Cognitive, Perceptual and Neurobiological Evidence. *Music Percept 29*(2):133-146.

[3] Varnet, L., Knoblauch, K., Meunier, F., Hoen, M. (2013). Using auditory classification images for the identification of fine acoustic cues used in speech perception. *Front Hum Neurosci 7*: 865.

Contact: leo.varnet@isc.cnrs.fr. This research was supported by an European Research Council grant to the SpiN project (n°209234) and an Ecole Doctorale Neuroscience et Cognition (ED NSCo) PhD grant awarded to L.V.