



**HAL**  
open science

# Identifying acoustic cues for English voiced stop consonants with reverse correlation: a pilot study

Simon Major, Léo Varnet, Maria Giavazzi

## ► To cite this version:

Simon Major, Léo Varnet, Maria Giavazzi. Identifying acoustic cues for English voiced stop consonants with reverse correlation: a pilot study. SPIN2026 – 17th Speech in Noise Workshop, Jan 2026, Paris, France. 2026, <10.5281/zenodo.18186587>. <hal-05467966>

**HAL Id: hal-05467966**

**<https://hal.science/hal-05467966v1>**

Submitted on 20 Jan 2026

**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.

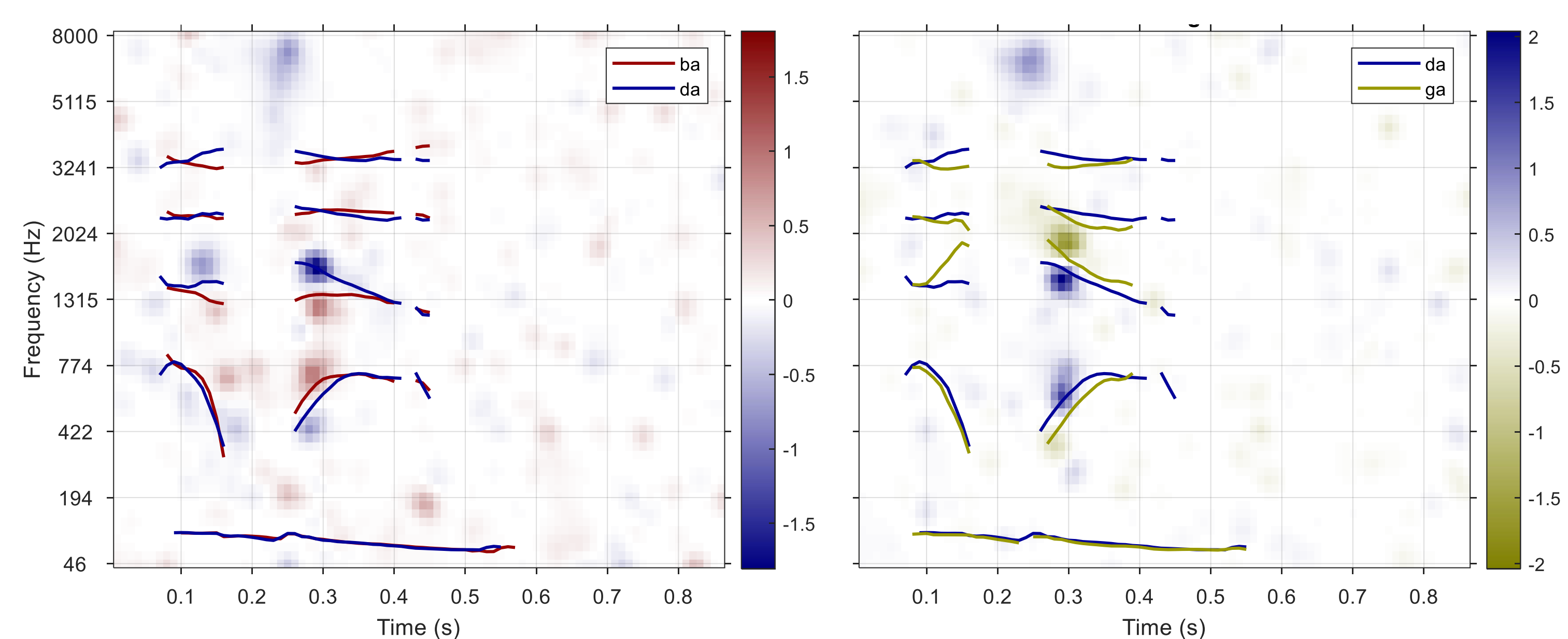
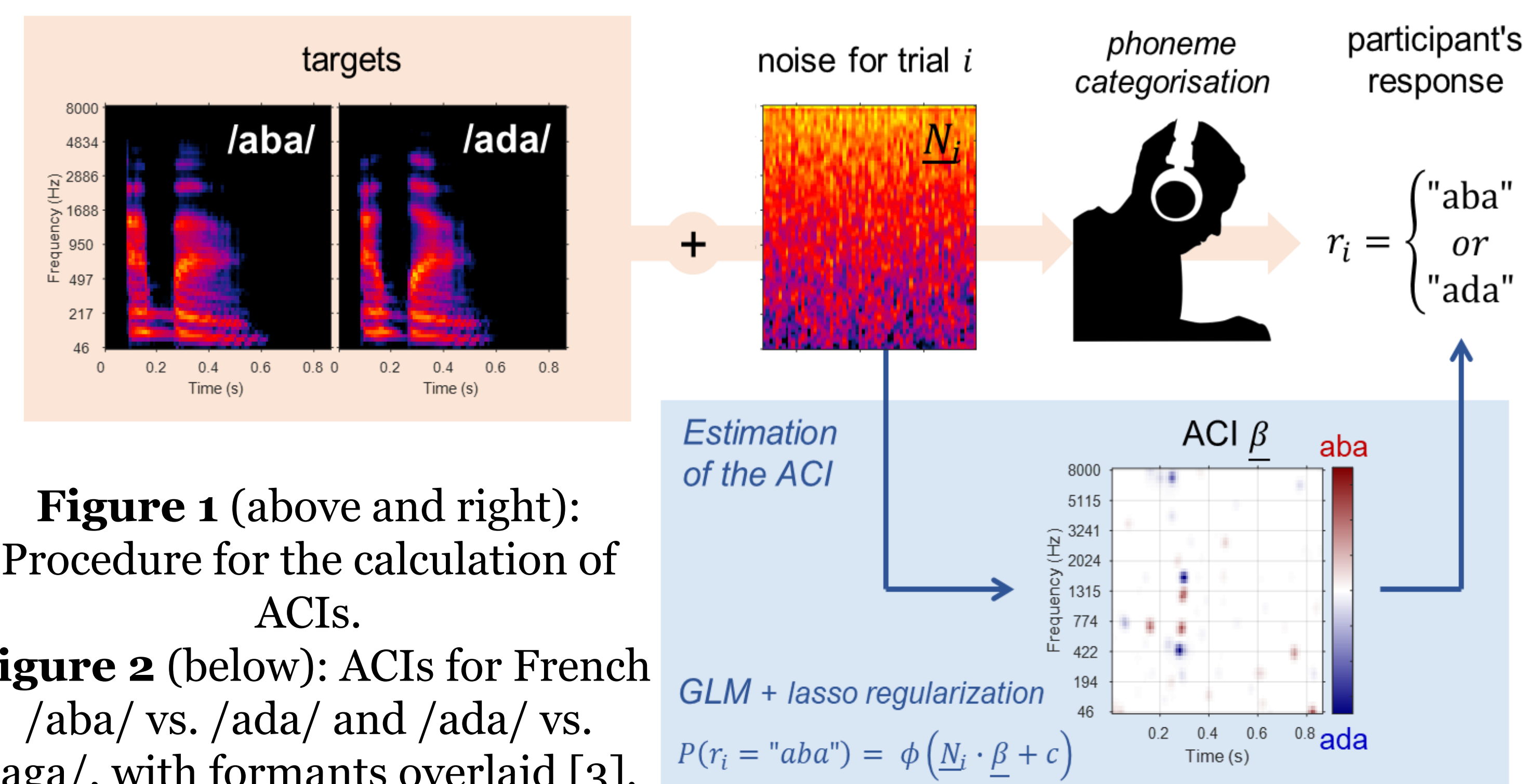
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

## 1. Introduction

- Auditory reverse correlation (revcorr) is a psychophysical paradigm whereby participants in a listening experiment identify stimuli as belonging to category *a* or *b*. By introducing random fluctuations of interfering noise into the stimuli over thousands of trials, this allows to determine the acoustic cues participants use in their identification of the categories.
- This methodology has been deployed in a number of phoneme identification tasks, such as Varnet et al. (2013) [1], Varnet et al. (2016) [2] and Carranante et al. (2024) [3]. They employ the fastACI toolbox [4], which produces Auditory Correlation Images (ACIs) in the form of time-frequency maps (see **Figure 1** for an example overview of the process).
- In this study, we use the revcorr method to investigate the acoustic cues used in English voiced stop (/b, d, g/) identification. It builds upon previous research on stops [5, 6], which find that post-consonantal F2 is a significant cue, as well as the temporal and frequency placement of the post-consonantal burst.
- Carranante et al. is of particular note, having conducted their experiment, using the revcorr method, on the French equivalents. Their findings (**Figure 2**) also report pre-consonantal F2 cues and post-consonantal F1 cues. We build upon this research in two ways:
  - By using the same methodology, we can directly compare our results with those of Carranante et al..
  - Previous revcorr studies have involved a two-way categorisation task. However, in the case of English and French, voiced stops have a three-way distinction. We explore whether the method is viable as a three-way categorisation task.



**Figure 3** (left): Mean ACIs of all seven participants overlaid with formants. From top to bottom: /aba/ vs. not-/aba/; /ada/ vs. not-/ada/; /aga/ vs. not-/aga/. Positive cues are the red/blue/yellow pixels, negatives the green.

## 2. Experiment

- The stimuli, of the form /aCa/, were recorded by a male native English speaker from Hertfordshire, England.
- Seven participants were recruited, including the speaker, from the same town. The gender split is 3M/3F/1NB, with an average age of 21.14 years and a standard deviation of 1.07. None report any hearing problems.
- They listened to 6,000 trials (2,000 per condition), and identified what they heard as 'aba', 'ada', or 'aga'.
- While the total volume was unchanging, the signal-to-noise ratio was varied based on performance, in order to achieve a 70% proportion of correct responses.

## 3. Results

- Figure 3 confirms the key roles of both pre-consonantal F2 and post-consonantal F1 and F2 in phoneme identification.
- There are mix of key positive and negative cues for each ACI, and greater weight is assigned to post-consonantal F1 than expected.
- As expected, F0 is not used, as there is no voicing contrast.
- Burst cues indicating /d/ are present but very faint – only reliably used for one participant – in contrast to in **Figure 2** (although note that there was also interparticipant variation here too).
- There are unexpected pixels early on which are hypothesised to represent acoustic features cueing for the onset of the speech target within the noise
- Overall, there is significant inter-participant variation, to the extent that some speakers share no cues at all.

## 4. Conclusions

- The revcorr approach to phonemic cue identification works with a three-way categorisation task, although it remains to be seen if the /x/ vs. not-/x/ ACIs are the best way to visualise the results.
- English speakers largely use expected cues to identify /b, d, g/, but they seem to apply more weight to F1 and less to the post-consonantal burst on average compared to the results in **Figure 2**.
- Speakers are not a homogenous group, and are sensitive to a variety of temporal and frequency-based cues.

**References:** [1] Varnet et al. (2013). Using auditory classification images for the identification of fine acoustic cues used in speech perception. *Frontiers in Human Neuroscience*, 7, 865. [2] Varnet et al. (2016). Speech reductions cause a de-weighting of secondary acoustic cues. *Interspeech 2016*. [3] Carranante et al. (2024). Mapping the spectrotemporal regions influencing perception of French stop consonants in noise. *Scientific Reports*, 14 (1), 27183. [4] Osses et al. (2025). FastACI: a toolbox for investigating auditory perception using reverse correlation. *Frontiers in Psychology*, 16, 1668690. [5] Liberman et al. (1954). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychol. Monogr. Gen. Appl.* 68, 1–13. [6] Li et al. (2011). Manipulation of consonants in natural speech. *IEEE Trans. Acoust. Speech Signal Process.* 19, 496–504.

**Acknowledgements:** MG is supported by the ANR grant DIA-SYN-PHON (ANR-21-CE28-0008).

