A Unified Model of the Reference Frame of the Ventriloquism Aftereffect Considering Auditory Saccade Adaptation

Peter Lokša & Norbert Kopčo

Institute of Computer Science, P. J. Šafárik University in Košice
Jesenná 5, 040 01 Košice
peter.loksa@upjs.sk, norbert.kopco@upjs.sk

Abstract

Loksa & Kopco (2023) introduced a model of the reference frame of the ventriloquism aftereffect (RFoVAE) that described many but not all available RFoVAE data by assuming that the auditory spatial map, natively using the head-centered reference frame, is adapted by visual signals in both eye-centered and head-centered reference frames. Here, the model is extended mainly by considering that, when saccade-to-sound responses are used in RFoVAE experiments, the saccades also undergo an adaptation – in the eye-centered reference frame. Also, the model proposes that the saccade-related and ventriloquism-related contributions to adaptation are combined using a normalized sum. The extended model can explain all available data, suggesting that the RFoVAE is largely head-centered, while saccadic adaptation accounts for the mixed reference frame observed experimentally.

1 Introduction

The neural representations of visual and auditory space use different reference frames. Vision is referenced relative to the direction of eye-gaze (eye-centered), while hearing is referenced relative to the head orientation (head-centered). The current study examines how these two representations are aligned at higher level of spatial processing to allow visually guided adaptation of auditory spatial perception.
Existing models of the audio-visual (AV) RF alignment only consider integration when in the auditory and visual stimuli are presented simultaneously (i.e., the ventriloquism effect; VE) (Razavi et al., 2007; Pouget et al., 2002). We proposed a model of the visually guided adaptation of auditory spatial representation in VAE (Lokša & Kopčo, 2023) to describe behavioral data of Kopčo et al. (2009, 2019). Here extensions of the model are introduced to characterize the mixed RF of VAE observed in Kopčo et al. (2009) and to provide a unified account of conflicting results of Kopčo et al. (2009, 2019).

In addition to auditory space representation in HC RF, the current model and the dHEC model (Lokša and Kopčo, 2023) model consider 3 candidate mechanisms underlying these effects: eye-centered signals influencing auditory space representation, fixation-position-dependent attenuation in auditory space adaptation, adaptation in the saccades used for responding in the experiments.

Finally, Kopco et al. (2019) observed a new adaptive phenomenon induced by aligned audiovisual stimuli presented in the periphery that is also considered.

2 Model

The snHC model (**Fig. 1**) focuses on explaining both central and peripheral data for both AV-aligned and AV-misaligned data using two mechanisms.
The first one ('s' component i.e. auditory-saccade adaptation) assumes that auditory space is adapted by visual signals only in HC RF (like in the basic version of dHEC model), while the saccades, used for responding, are also adapted – in EC RF. Specifically, it assumes that (1) during training, the saccades are adapted to be hypometric or hypermetric, depending on the locations of FP, A component, and V component, (2) during testing, the adapted saccades either enhance or reduce the bias due to auditory space representation, depending on the A target location vs FP.
The second one ('n' component i.e. normalization) assumes normalization for limiting the overall output of the neural channel representing the combined space (Dahmen et al., 2010).

Table 1: Performances of various versions of snHC model in comparison to dHEC model.

Model version	Performance		
	AICc	ΔAIC	MAE
nHC	345.6	6.9	1.30
<u>sHC</u>	340.8	2.0	1.19
<u>snHC</u>	338.7	-	1.17
dHEC	352.7	14.0	1.31

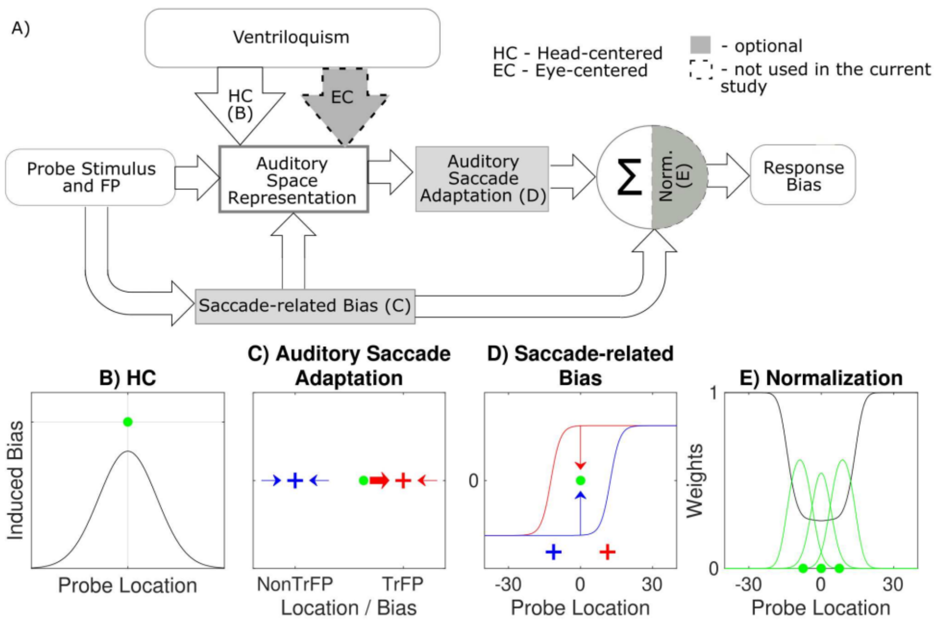

Fig. 1: Diagram of the snHC model.

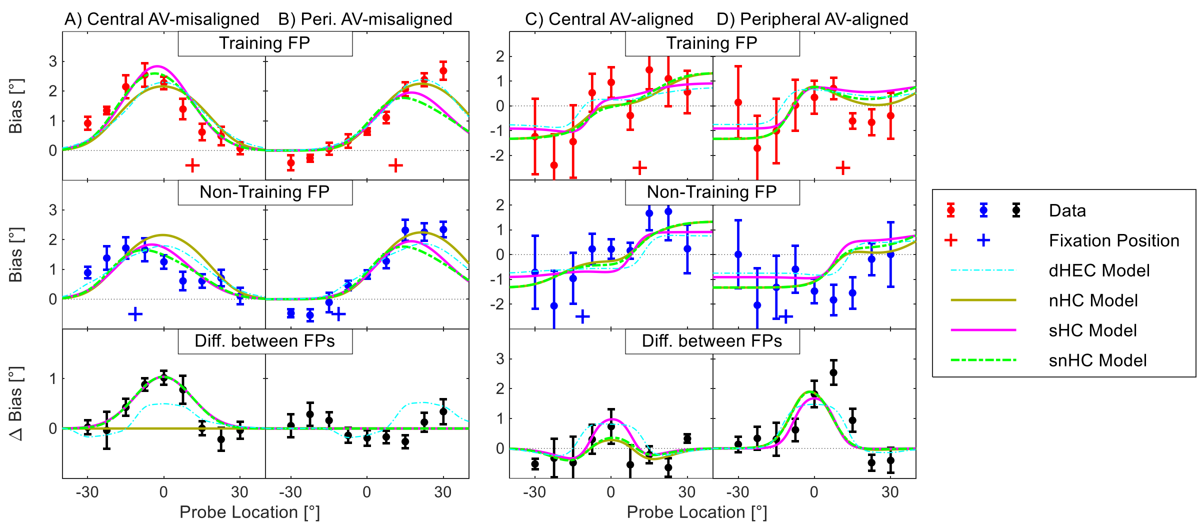

Fig. 2: snHC model evaluation on central & peripheral data. Model predictions (solid lines) and experimental data (symbols).

3 Results

From Table 1 and Fig. 2, we can see the following results:
The best performance was for the full version (snHC – green line) of the model fits the central (A, C) and peripheral (B, D) data for both AV-misaligned (A, B) and AV-aligned condition simultaneously.
The second-best performance was for the version (sHC) with the saccadic adaptation, but without normalization, and despite its performance being weaker, the difference in performance was not significant ($\Delta AIC < 2.5$) according to principle stated by Burnham & Anderson (2004).

Since the model only uses HC RF to induce VAE, this result supports the conclusion that reference frame of VAE is purely head-centered, and the previously observed mixed RF was due to saccade adaptation.

4 Conclusion

We introduced two extensions of a previous model to describe the reference frame of ventriloquism aftereffect data of Kopčo et al. (2009, 2019).

A previous model by Lokša & Kopčo (2023) was able to explain central adaptation and peripheral adaptation results of Kopčo et al. (2009, 2019) when fitted to the data separately, i.e., with different values of model parameters. Thus, the inconsistency between the behavioral results were not reconciled.
The current model, incorporating auditory saccade adaptation, can explain the central and the peripheral AV-misaligned data simultaneously which confirms that RF of VAE is most likely not mixed. And moreover, it can explain AV-aligned condition together with the AV-misaligned condition using the normalization mechanism, which supports limiting the overall output of the neural channel theory (Dahmen et al., 2010).

Next step is to experimentally test the model predictions about saccade-related EC bias and saccade representation adaptation, as well as the prediction that the reference frame of the VAE is purely head-centered.

Acknowledgements

The research is supported by the Slovak Research and Development Agency APVV-23-0054; SK-AT-23-0002.

Literature

Alais, D. and Burr, D., (2004). The ventriloquist effect results from near-optimal bimodal integration. Curr Biol. 14:257-262

Burnham K. P., Anderson D. R. (2004). Multimodel inference understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33(2), 261–304. https://doi.org/10.1177/0049124104268644

Dahmen, J. C., Keating, P., Nodal, F,R,, Schulz, A.L,, King, A.J. (2010) Adaptation to stimulus statistics in the perception and neural representation of auditory space. Neuron. 2010 Jun 24;66(6):937-48. doi: 10.1016/j.neuron.2010.05.018. PMID: 20620878; PMCID: PMC2938477.

Knudsen, E. I. and Knudsen, P. F. (1985). Vision guides the adjustment of auditory localization in young barn owls. Science. 230:545-548

Knudsen, E. I. and Knudsen, P. F. (1989) Vision calibrates sound localization in developing barn owls. Journal of Neuroscience. 9:3306-3313

Kopčo, N., Lin, I., Shinn-Cunningham, B. G. and Groh, J. M., (2009). Reference Frame of the Ventriloquism Aftereffect. Journal of Neuroscience. 29(44): 13809-13814

Kopčo, N., Lokša, P., Lin, I., Groh, J. M., Shinn-Cunningham, B. G. (2019). Hemisphere-specific properties of the ventriloquism aftereffect. JASA 146, EL177

Lokša P, Kopčo N. (2023). Toward a Unified Theory of the Reference Frame of the Ventriloquism Aftereffect. Trends in Hearing. 2023;27.

Pouget, A., Deneve, S., Duhamel, J. R. (2002). A computational perspective on the neural basis of multisensory spatial representations. Nat Rev Neurosci. 2002 Sep;3(9):741-7. doi: 10.1038/nrn914

Razavi, B., O'Neill, W.E., Paige, G.D. (2007) Auditory spatial perception dynamically realigns with changing eye position. J Neurosci 27:10249–10258

Lokša, P. and Kopčo, N., (2016). Modelling of the Reference Frame of the Ventriloquism Aftereffect. Sborník z 16.ročníku konference Kognice a Umělý život. 101-106