Modeling of the Reference Frame of the Ventriloquism Aftereffect

Ing. Peter Lokša, doc. Ing. Norbert Kopčo PhD.

Institute of Computer Science, Faculty of Science, Šafárik University in Kosice
peter.loksa@gmail.com, norbert.kopco@upjs.sk

Abstract

The human brain extracts information from various senses in order to represent the physical space. To integrate spatial information from the visual and auditory modalities, the modalities need to be aligned as each of them represents spatial information in a different reference frame (RF). The visual reference frame is aligned with the direction of eye gaze while the auditory one is aligned with the orientation of the head. The aligned audiovisual spatial representation is most likely using one of the two RFs as well. Previous experimental data attempting to identify the aligned RF are inconsistent. This article presents modeling attempts aiming at resolving this inconsistency and identifying the reference frame of the ventriloquism aftereffect.

1 Introduction

Vision plays an important role in how the brain processes auditory information (Alais, Burr, 2004). In the spatial domain, vision provides guiding signals for calibration of spatial auditory processing. This can be illustrated by the ventriloquism aftereffect illusion in which repeated pairings of spatially mismatched visual and auditory stimuli produce shifts in the perceived locations of sound sources that persist even when the sounds are presented by themselves (Alais, Burr, 2004; Knudsen, Knudsen, 1985; Knudsen, Knudsen, 1989). It might be that the supramodal spatial representation could be the ultimate one in sense of being directly used in motion planning etc.

The current study models data from a previous study which examined the RF of the ventriloquism aftereffect (abbr.: RFVAE) (Kopčo et al., 2009). RFVAE might by identical or connected to RF of general supramodal spatial adaptation.

There were two basic hypotheses considering properties of RFVAE so would be: (1) head- and (2) eye-centered, in case of holding of which the RF is spatially fixated to specific body part (1) head itself (2) eyeball. The reason for choosing such ones as possible RF-s is because respectively (1) auditory and (2) visual localization go on in these RF-s (Brainard, Knudsen, 1995; Razavi et. al., 2007).

In a previous study attempting to identify RFVAE the so called aftereffect magnitude was compared between following two conditions: eye not shifted from position of ventriloquism aftereffect inducement, eye shifted so. By eye shift we mean change of so called initial fixation point, in which the eye is right before providing the stimulus. And if such aftereffect magnitude shifted with eye, RFVAE would be probably eye-centered. If it didn't shift, it would be head centered, since head shifts neither.

Modelling was performed because of inconsistency of results according to basic hypotheses. In current article we will also show such basic results first.

2 Experimental data

The experimental data used here are taken from a previous study that investigated the reference frame of ventriloquism aftereffect (Kopčo et al., 2009).

2.1 Materials and methods

Obr. 1: illustrates the experimental setup and the hypothesized results.

In experiment given subject was sitting in dark quiet room with his head fixed. The target speakers and LEDs (visual adaptor) were placed in order to provide stimuli to subject. The saccadic responses to stimuli were recorded.

To induce ventriloquism aftereffect the AV training trials with constant shift of light from sound were induced in specific azimuth region, while FP-s of all such trials are same within session (training fixation point (TrFP)).

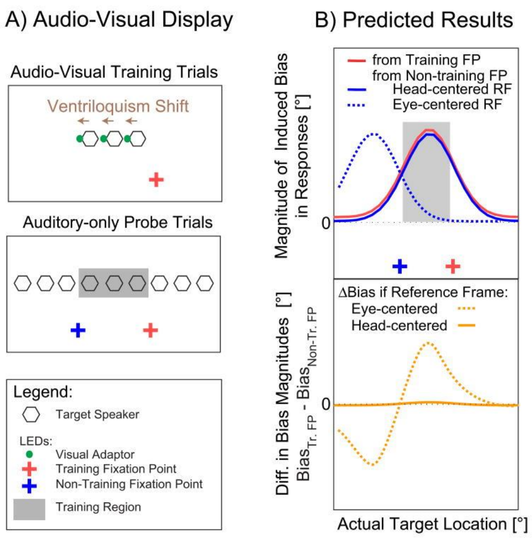
In order to measure aftereffect magnitude in condition of eye not shifted from position of ventriloquism aftereffect inducement, the localization errors were identified according to responses to auditory-only (A-only) trials in TrFP in stimuli range -30° to 30°. Analogically was done for condition of eye shifted in so called Non-training fixation point (NTrFP). So within session AV trials were in TrFP and there were A-only trials in TrFP and A-only trials in NTrFP. These three kinds of trials were interleaved.

To see whether ventriloquism aftereffect is symmetrical or not, the session differed in (1) in shift of visual component of AV trials from its auditory component, and (2) in training fixation point. There were three kinds of shifts of visual component: no shift (sound and light have same azimuth), positive shift (visual component is shifted by 5° to the side, toward which the TrFP is from 0°). Regarding FP-s azimuth axis can be flipped that TrFP would on 11.8° and NTrFP on -11.8° for each session.

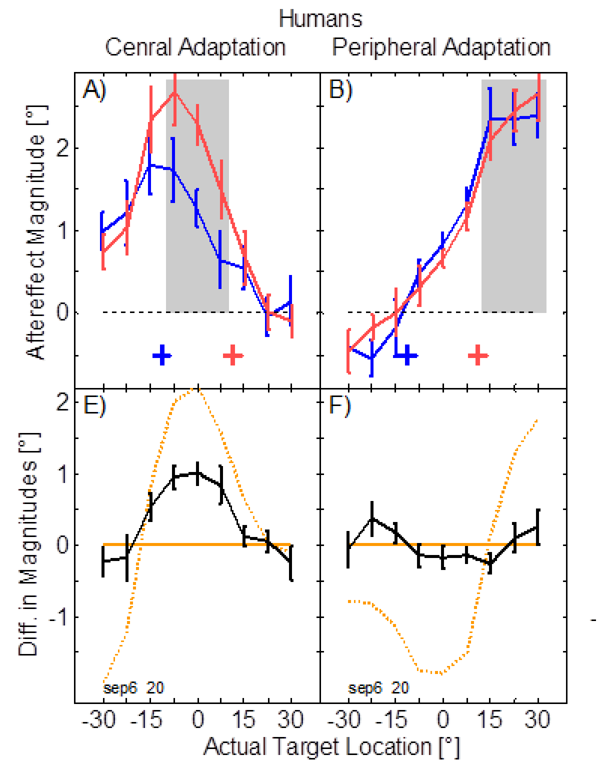
Because discriminability in center vs. periphery is inconsistent (Maier et. al., 2009), two different so called training regions of aftereffect inducement were used, but the same one within each session. These two we call center and periphery. In Obr. 1: central one is marked.

The 9 speakers were displaced within same horizontal plane, while holding: distance of each speaker from center of the listener's head is equal; angle difference of the speaker from adjacent one is equal (7.5°) (see Obr. 1:).

According to Diff in bias magnitude (bias of NTrFP A-only trials subtracted from bias of TrFP A-only trials; Obr. 1:) the RFVAE had to be identified.



Obr. 1: A) Symbols in "Audio-Visual Training Trials" panel mark the azimuths of stimuli provided to subjects in audiovisual training trials, in the way that the azimuthal relative shift between physical location of stimuli, that are synchronous, are constant within given experimental session for each session. The symbols in "Auditory-only Probe Trials" panel mark azimuths of auditory-only trials, which were interleaved with the already mentioned training ones. B) This panel visualize hypothetical experimental data for cases of questioned reference frame being head- vs. eye centered. "Magnitude of Induced Bias in Responses" here means localization error of them toward the shift in given session for each session.



Obr. 2: Magnitude of ventriloquism aftereffect and reference frame determination according to difference between training vs. non-training trials. Red/Blue line - separation of probe auditory-only trials according to the pre-trial eye gaze azimuths (marked by '+' of given color). But eye gazes of all audio-visual training trials are preceded by the red one so this is called training fixation point (FP), and the blue one non-training one. The black line can be arithmetically described as the subtraction of blue line from the red line and we call it aftereffect FP dependence. The orange lines reflect hypothetical Aftereffect FP dependences: the solid one for the case of eye-centered head centered and the dotted one.

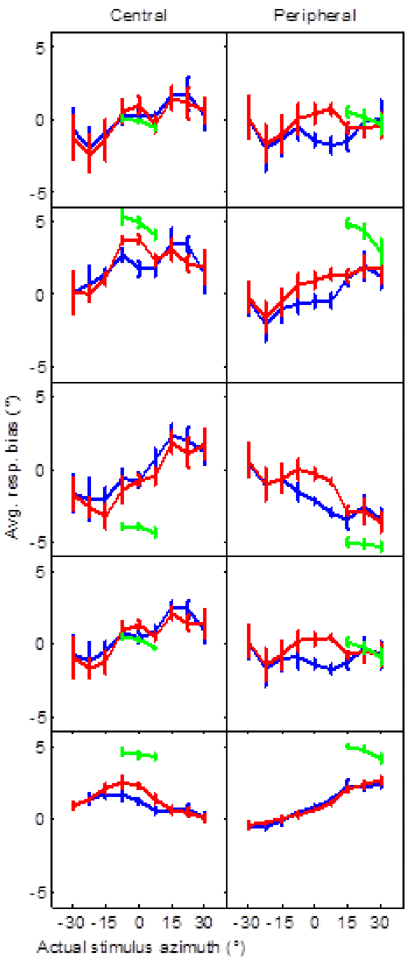
2.2 Early analysis

According to Obr. 2: the result of RFVAE is inconsistent: according to central adaptation this RF seems to be mixed of head- and eye centered, while for peripheral adaptation it seems to be purely head-centered.

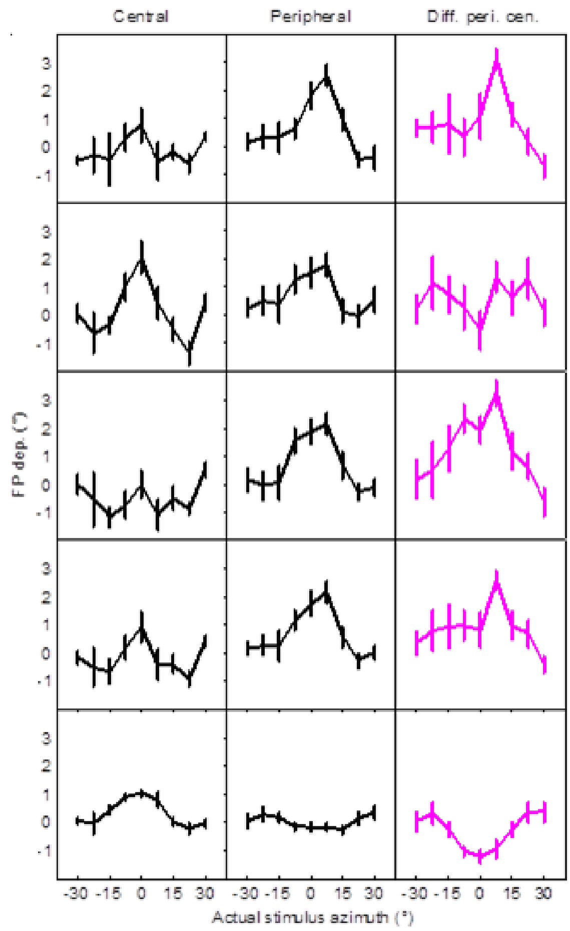
In order to resolve this inconsistency, we attempted to model experimental data. But in order to model it we first displayed results for different conditions Obr. 3:.

It was unlikely in the brain that two different forms of reference would be utilized for same representation. On the other side, there are multiple other explanations for difference we observed, related to other forms of adaptation that might have occurred in this experiment: the undershooting, the expansion of auditory space, saccade adaptation. The Kopčo et al. (2009) data showed another form of plasticity for which the study

has not been designed, and in modelling we will explore first two above-mentioned explanations.
Undershooting means shortening saccades in comparison of them in case when saccade endpoints ended in location where the stimulus is perceived.



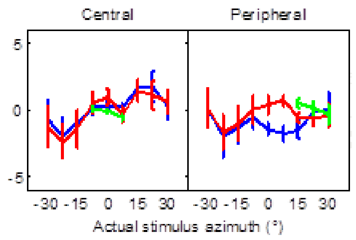
Obr. 3: Mean localization error of human subject experimental data and SEM across 7 subjects. Red line – A-only trials - training fixation point, blue line A-only data – non-training fixation point, green line – AV (training) trials, black line – difference between training vs. non-training A-only trial mean (FP dependence), magenta line – difference between peripheral vs. central adaptation FP dependence. Conditions according to rows respectively: 1. no shift, 2. positive shift, 3. negative shift, 4. mean across shifts, 5. aftereffect magnitude. The graphs in the 5th except of magenta lines row are little different with Obr. 2: A, B, E, F, except of yellow lines only because of technical errors and outliers removal.



Obr. 4: Continuation of Obr. 3:.

3 Unexpected form of plasticity

In Obr. 5: we observed inconsistency. In this figure we can see different azimuth and different condition that there are two types of cases for localization error being (1) depending (2) not depending on initial eye fixation point visualized as (1) similar or (2) dissimilar value of red vs. blue line: 1. all central azimuths, azimuths -30° to -15° in periphery and azimuths 15° to 30° in periphery. 2. azimuths -7.5° to 7.5° in periphery. This unexpected plasticity could be possible reason for inconsistency of central vs. peripheral RF-s of ventriloquism aftereffect appearance.
In order to explain this we attempted to model data present in this visualization (Typical property of this visualization is consistency of audiovisual training trials that affect localization errors (so called no-shift) as the selection key for data included.



Obr. 5: Localization error for no-shift condition for different training regions.

4 Modelling

In this section there is the attempt to model newly observed phenomenon, and also test of relevant qualities of its result using a model that assumes that two adaptive processes occur, unrelated to the ventriloquism aftereffect, and that their effect combines additively.

4.1 Description

Basic idea of this modelling is to consider following two factors in additive relation: saccade hypometry, expansion outside training region.

Saccade hypometry is in other words undershooting of saccades (Harris, 1994). Saccade hypometry shortens the saccade in comparison with the case when saccade would end in location where the response is perceived. We considered it because according to Obr. 5: peripheral data, azimuths -7.5° to 7.5° the localization errors appear to be shifted from each other toward related fixation point.

The effect which we call outside training region expansion has zero value inside training region and its absolute value increases with distance from this region. The reason for this is that the data appear to reflect this phenomenon.

Established variables and functions:

t ...target azimuth

$b(t, FP, trreg)$ ...bias in response to auditory target at azimuth t from eyes initially fixated at FP when the training region was $trreg$ (predicted variable).

$eotr(t, trreg)$ ...expansion of response on target with azimuth t outside training region $trreg$

$h(t, FP)$ ...saccade hypometry on target t with eyes initially fixated on FP

$dtr(t, trreg)$ ...relative distance of target t from training region $trreg$

Free parameters: ak_1, ak_2, ek_1, ek_2

Established equations:

$$b(t, FP, trreg) = eotr(t, trreg) + h(t, FP)$$

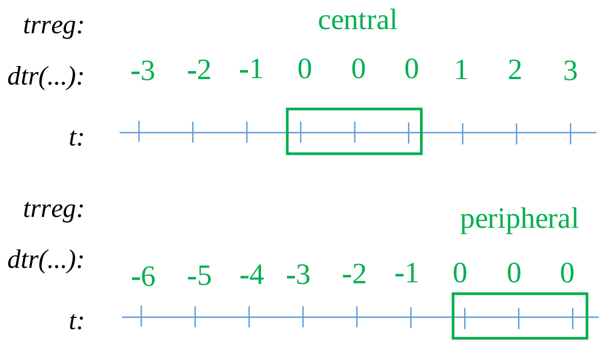
...additive relation of given concepts.

$$eotr(t, trreg) =$$

$$= ek_1 \cdot \left(\frac{1}{1 + e^{ek_2 \cdot (dtr(t, trreg))}} - \frac{1}{2}\right);$$

$$h(t, FP) =$$

$$= ak_1 \cdot \left(\frac{1}{1 + e^{ak_2 \cdot (t - FP)}} - \frac{1}{2}\right);$$



Obr. 6: Illustration of the relative distance of t from training region $trreg$. $trreg$ is training region, t is target azimuth and dtr is relative distance of t from training region $trreg$. (There is arithmetical distinction between this and the addend of the model, but that depend on this, according to equation in current subchapter)

4.2 Performance

The free parameters for model were fitted in MATLAB function nlinfit was used which is using iterative least squares estimation. The data in Obr. 5: are used as the base for observational data for this model.

In first step the hypometry was fitted. This was done in domain of FP dependence. This means that observational data and fitted model as the input to this tool were converted to this format. Results are displayed in 0 In this step parameters ak_1 and ak_2 were fitted.

In second step was done on residual of data after subtracting results of the first step, expansion outside

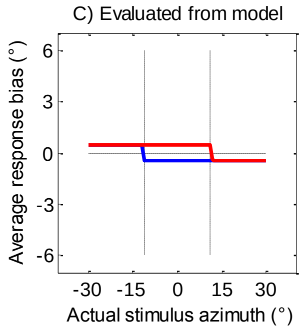
training region was fitted. This was done in domain of bias (localization error). Results are displayed in Obr. 7: In this step parameters ek_1 and ek_2 were fitted.
Resulting coefficients are following:

$ak_1 = 0.94,$

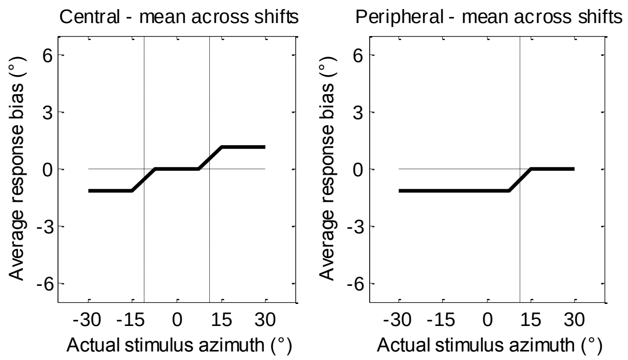
$ak_2 = 151.13$

$ek_1 = -2.26$

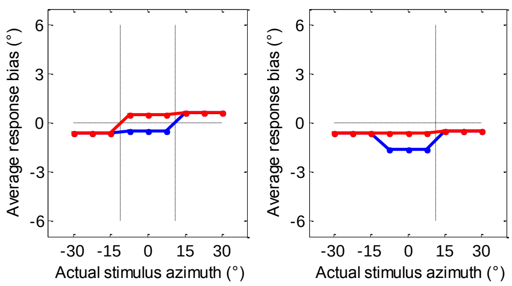
$ek_2 = 130.62$



Obr. 7: The visualization of the saccade hypometry according (h) to the resultant model fitted on experimental data. Colors have meaning analogical to whole article.



Obr. 8: The visualization of the expansion outside training region according ($eotr$) to the resultant model fitted on experimental data.



Obr. 9: Approximation of behavior of model of bias (b) in no-shift condition for central (left graph) and peripheral (right graph) adaptation according to given model. Colors have meaning analogical to whole article.

Figure of current modelling results (Obr. 9:) show no difference between FP dependences of central vs. peripheral adaptation (FP dependence is difference between training vs. non-training fixation biases. (Difference in red vs. blue)). Absence of such difference is inconsistent with experimental data (Obr. 5:) and with explanation of unexpected form of plasticity.
In Obr. 9: there were some FP independent conditions (-30° to -15° and 15° to 30° for both training regions) and some FP dependent conditions (-7.5° to 7.5° for both training regions). There is the difference with experimental data because for -7.5° to 7.5° for model data they are independent.

4.3 Proof of current model inappropriateness

We can even prove mathematically that current model cannot explain current unexpected form of plasticity.
To show this, we define:
$FPdep(t, trreg)$ ...FP dependence of biases on training region $trreg$
$TrFP$ ...azimuth of training fixation point.
$NtrFP$ ...azimuth of non-training fixation point

$$FPdep(t, trreg) =$$
$$= b(t, TrFP, trreg) - b(t, NtrFP, trreg);$$

We infer:

$$FPdep(t, trreg) =$$
$$= b(t, TrFP, trreg) - b(t, NtrFP, trreg) =$$
$$= eotr(t, trreg) + h(t, TrFP) -$$
$$(eotr(t, trreg) + h(t, NtrFP)) =$$
$$h(t, TrFP) - h(t, NtrFP);$$

We see that according to current model FP dependence ($FPdep$) does not depend on training region ($trreg$), so the model cannot describe such a dependence.
In experimental data FP dependence depend on training region in azimuth range -7.5° to 7.5°. But according to this proof model cannot do so.

5 Conclusion

We have described a previous study examining the reference frame of the ventriloquism aftereffect and its main results, which contain some ambiguity. We examined a part of the experimental data from that study, and we described a new adaptive phenomenon. We made the attempt to model these data and we have proven that the proposed additive model combining hypometry and auditory space expansion is inappropriate for the explanation of the newly observed phenomenon in manner.

One of the alternatives for current modelling is instead of additivity of factors the function composite would be used, so these factors would be related hierarchically such that one operates on the output of the other one. Alternatively, completely other factors might play role. Additional modeling is currently underway to examine these alternatives.

Acknowledgement

This contribution was supported by Slovak Research and Development Agency under grant project APVV-0452-12 Spatial Attention and Listening in Complex Environments.

References

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

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

Maier, J. K., McAlpine, D., Klump, G. M., Pressnitzer, D. (2009). Context Effect in the Discriminability of Spatial Cues. *JARO* 11(2): 319-328

Brainard, M. S., Knudsen, E. I. (1995) Dynamics of visually guided auditory plasticity in the optic tectum of the barn owl. *J Neurophysiol* 73:595–614

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

Harris, C. M. (1994), Does saccadic undershoot minimize saccadic flight-time? A Monte-Carlo study. *Vision Res.* 35(5): 691-701