A Model of the Reference Frame of the Ventriloquism Aftereffect using a priori bias

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 reference frame (RF) used by audio-visual (AV) spatial representation is likely to be head-centered or eye-centered, aligned with the RFs of either the unimodal auditory (head-centered) or visual (eye-centered) representations. Results of previous RFAV studies are inconsistent, suggesting that the RF is either mostly head-centered, when examined in the periphery, or a mixture of head-centered and eye-centered, when examined in the central field (Kopčo et al., 2009; Lokša & Kopčo, 2016). Here, a model is proposed, assuming a form of a priori bias is combined with the adaptation due to AV stimuli. This model can explain the results in the baseline conditions, but not when ventriloquism aftereffect is induced. Therefore, additional mechanisms are likely to determine the AV RF.

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 a supramodal spatial representation exists , directly used in motion planning etc.

The current study models data from a previous study which examined the RF of the ventriloquism aftereffect (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 RFs is because respectively (1) auditory and (2) visual space is represented in these RFs (Brainard and Knudsen, 1995; Razavi et al., 2007).

In a previous study of the RFVAE, the observed aftereffect was compared between two conditions: eyes not shifted from the fixation point (FP) of ventriloquism aftereffect inducement, eye shifted to a new FP. By eye shift we mean change of fixation position, i.e., direction of the eye gaze when the stimulus is presented. It was hypothesized that if the aftereffect shifted with the eye shift, RFVAE would be probably eye-centered. If it didn't shift, it would be head centered, since head shifts neither. The goal of the previous modelling was to evaluate a possible mechanism causing this inconsistency of results with respect to the above hypotheses.

In the current study we first show the behavioral results, followed by an extension of the model and its evaluation.

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 the experiment the subject was sitting in a dark quiet room with his head fixed. The target speakers and LEDs (visual adaptor) were used to provide stimuli to the 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 FPs of all such trials were same within session (training fixation point (TrFP); Fig. 1A).

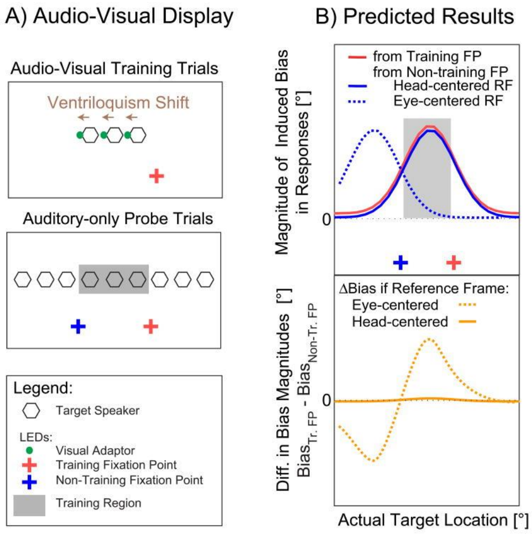
To measure the aftereffect magnitude in the condition of eye not shifted from position of ventriloquism aftereffect inducement, the localization errors were identified according to responses to auditory-only (A-only) trials in TrFP in stimuli range -30° to 30°. Analogically was done for condition of 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 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 FPs azimuth axis can be flipped that TrFP would on 11.8° and NTrFP on -11.8° for each session.

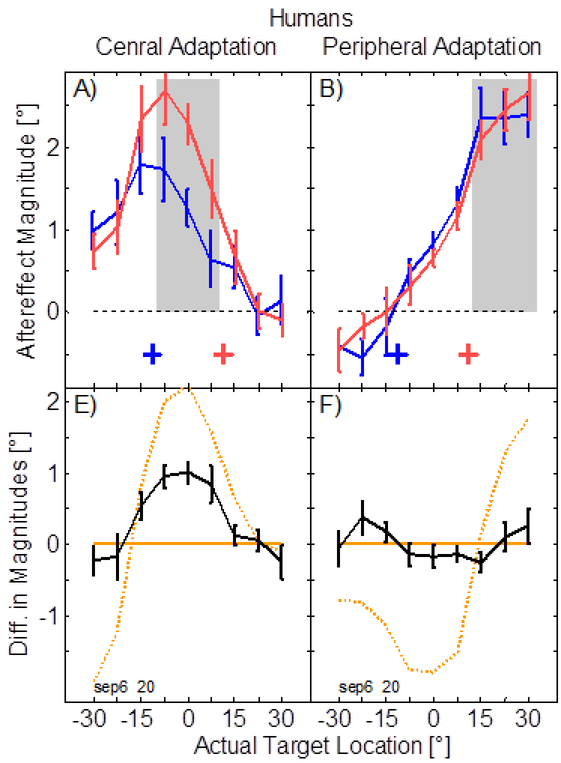
Because discrimination abilities in center vs. periphery are inconsistent (Maier et. al., 2009), two different training regions of aftereffect inducement were used, but the same one within session. These two we call center and periphery. In Obr. 1: central one is shown.

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. Results are displayed in Obr. 2:.



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.

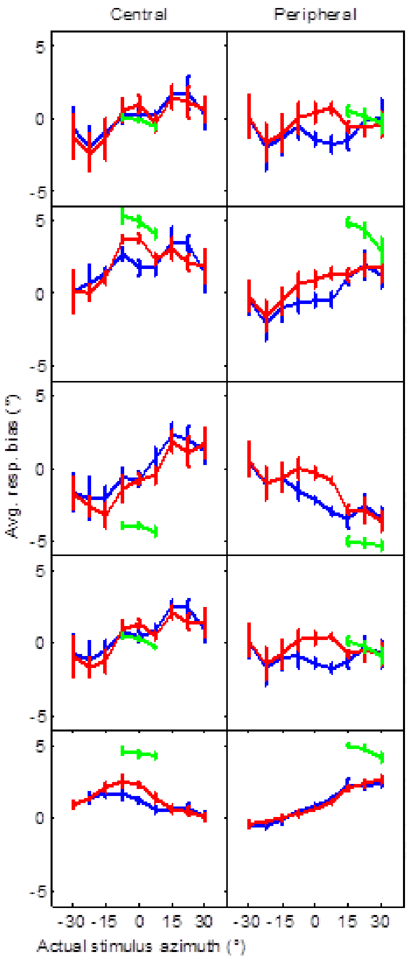
2.2 Data analysis

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

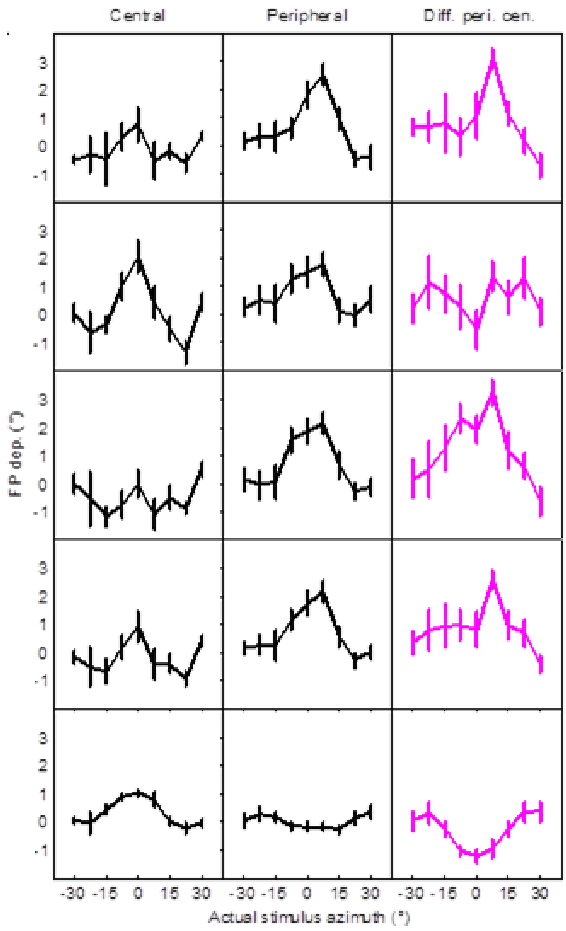
To resolve this inconsistency, we attempted to model the experimental data. To better understand the causes of the inconsistency, Obr. 3:, Obr. 4: shows the detailed behavioral results for different conditions.

It is unlikely in the brain that two different forms of reference would be utilized for the 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 saccadic hypometria (undershooting), the expansion of auditory space, saccade adaptation.

The Kopčo et al. (2009) data showed another form of plasticity, described in the following section,for which the study has not been designed, and in modelling first two above-mentioned explanations were explored in previous article Lokša & Kopčo (2016).



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.

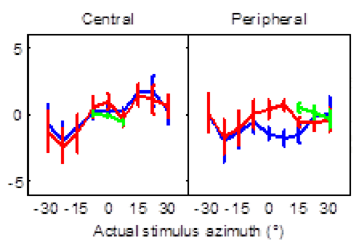


Obr. 4: Continuation of Obr. 3:.

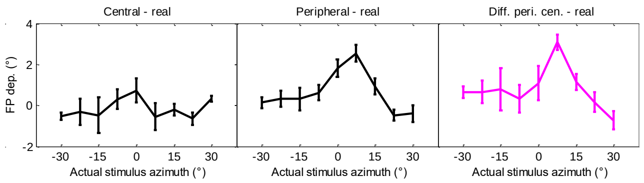
3 Unexpected form of plasticity

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 RFs of ventriloquism aftereffect appearance.

In order to explain this inconsistency 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.



Obr. 6: Continuation of Obr. 5: FP dep. means difference between biases of TrFP vs. NtrFP. (The magenta graph for model of (Lokša & Kopčo, 2016) would be zero constant - black ones would be equal)

4 Modelling

This section presents a model of the newly observed adaptation, and also tests of relevant qualities of the model. The model assumes that a priori bias is modified by so called AV effect.

Previous modelling assumed that unexpected form of plasticity is caused by saccadic hypometry and expansion outside training region, specifically their additive composite, and this was proven to be insufficient to describe the data (Lokša & Kopčo, 2016).

4.1 Description of current model

Basic idea of this modelling is a priori bias affected by biases of responses to AV stimuli (vertical position of green line). A priori bias is attracted by vertical position of green dots. Attraction is represented by weighted mean of A priori bias and vertical positon of green points. Weight of a priori bias is one of the free parameters ($w_{a-priori}$). Weight of given vertical position of green point is product of (1) implicit free parameter ($1-w_{a-priori}$) and of (2) Gaussian function of: horizotal position of green point as the center of Gaussian function, azimuth of auditory-only stimulus as the main input of it and another free parameter as the width of the Gaussian function (wdt).

A priori bias is sigmoid function modified to be odd (inflection point at vertical 0 instead of 0.5), where its horizontal center ($-FP\cdot c$), its height (hg) and its slope (sp) are adjustable by free parameters.

Established variables and functions:

Free parameters:
c ... coefficient of horizontal position of inflection point according to FP for a priori bias,
hg, sp ... height and slope of a priori bias,
$w_{a-priori}$ ...weight for a priori bias
wdt ...width of AV effect.

Input variables:
azi_{stim} ... actual azimuth of stimulus,
FP ... fixation point (red vs. blue line)
$azi_{AV-stim}$ ... azimuths of A component of AV stimuli (green - abscissa (horizontal))
$bias_{AV-resp}$ ... bias of response to AV stimuli (green - ordinate (vertical))

Established equations:

$$bias_{a-priori}(x,FP)=$$

$$=2\cdot hg\cdot\frac{1}{1+e^{-sp\cdot(x-(-FPc))}}-1;$$ ...A priori bias function;

$c, hg, sp > 0;$

$Gauss(x,\mu,\sigma);$ ...Gaussian function,

$$f_{wdt}(x,\mu,\sigma)=\frac{Gauss(x,\mu,\sigma)}{\sum_{i=-8}^{8}Gauss(7.5\cdot i,0,\sigma)};$$ ... width function,

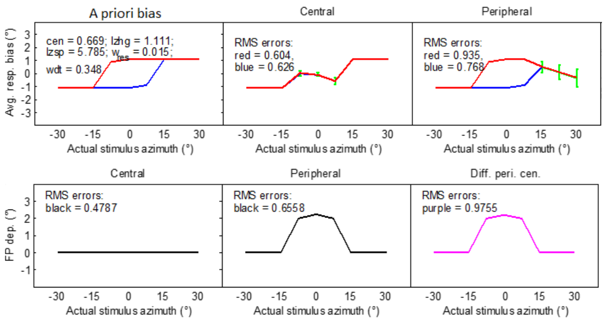
$$bias_{A-only-resp}(azi_{stim},FP,azi_{AV-stim},bias_{AV-resp})=$$

$$=\frac{bias_{a-priori}(azi_{stim},FP)\cdot w_{a-priori}+(1-w_{a-priori})\cdot\sum_{i=1}^{count_{AV}}bias_{AV-resp}(i)\cdot f_{wdt}(azi_{stim},azi_{AV-stim}(i),wdt)}{w_{a-priori}+(1-w_{a-priori})\cdot\sum_{i=1}^{count_{AV}}f_{wdt}(azi_{stim},azi_{AV-stim}(i),wdt)}$$

...core function (This is weighted mean of $bias_{AV-resp}(i)$ with $bias_{a-priori}(azi_{stim},FP)$ and this core function describes effect of bias of response to AV stimulus on a priori bias).

4.2 Performance

For spatially congruent AV stimuli the model looks like Obr. 7:. It was fitted by 'nlinfit' Matlab nonlinear regression fitting function.



Obr. 7: Modelling results

Resulting coefficients for fitting on no-shift data:

$$c = 0.669,$$

$$hg = 1.111,$$

$$sp = 5.785,$$

$$w_{a\text{-}priori} = 0.015,$$

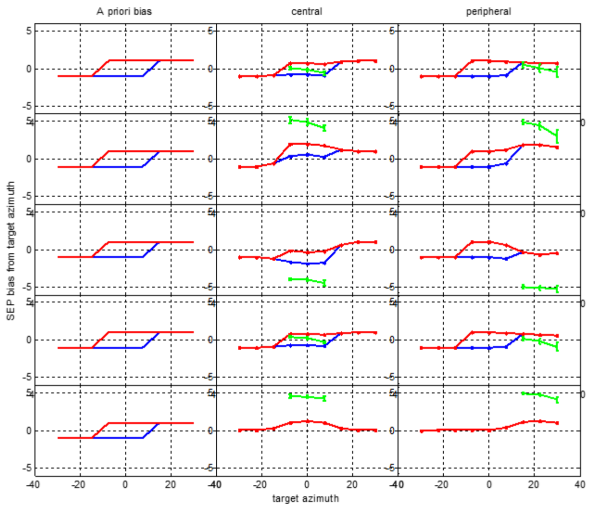
$$wdt = 0.348;$$

Experimental data in Obr. 5: are well-explained by current model. You can see that for central adaptation the red and blue lines are almost equal and this is also the case for prediction according to current model (Obr. 7:). You can also see that difference in red vs. Blue line for peripheral adaptation is present in central 3 azimuths for both experimental data and prediction. Magenta line is also similar for experimental data and prediction according to model.

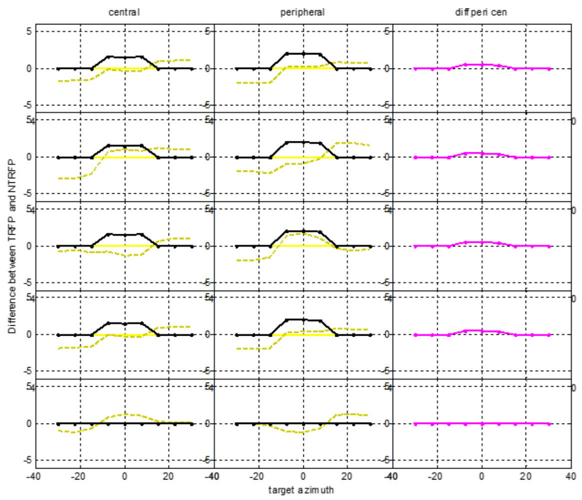
But the prediction of magenta line according to previous model (Lokša & Kopčo, 2016) would be zero constant for no-shift condition.

The model is less successful when we use fitting of also experimental data other than no-shift ones (Obr. 3: and Obr. 4:). We did it and here are the results (Obr. 8: and Obr. 9:).

The problem is that 5th row of these graphs, which displays difference of rows 2 and 3 divided by two, gives no difference between red and blue line, and that is inconsistent with experimental data (Obr. 3:, Obr. 4:). This result shows the model's limitation in that the FP-specific shift (i.e., the difference between the red and blue lines in Fig. 8) is independent of the direction or magnitude of the visually-induced adaptation. Thus, the model cannot describe any adaptation that is eye-centered. This can be proven analyticaly, as shown in the following section.



Obr. 8: Prediction of no-shift, positive shift, negative shift, mean across shifts, mean across shifts oriented as positive ones, respectively for given rows. Prediction was done on experimental data from Obr. 3:.



Obr. 9: Continuation of Obr. 8:.

Resulting coefficients for fitting on data on all 3 shift conditions:

$$c = 8.980,$$

$$hg = 0.989,$$

$$sp = 8.930,$$

$$w_{a\text{-}priori} = 0.260,$$

$$wdt = 4.939;$$

Mean square error of model fitted on all 3 shift conditions:

$$\mathrm{MSE} = 1.375;$$

4.3 Proof of current model inappropriateness

This section shows the weak aspect of current model in the manner that proves its weakness algebraically.

Model formula:

$$bias_{A-only-resp}(azi_{stim}, FP, azi_{AV-stim}, bias_{AV-resp}) = \frac{bias_{a-priori}(azi_{stim}, FP)\cdot w_{a-priori} + (1-w_{a-priori})\cdot \sum_{i=1}^{count_{AV}} bias_{AV-resp}(i)\cdot f_{wdt}(azi_{stim}, azi_{AV-stim}(i), wdt)}{w_{a-priori} + (1-w_{a-priori})\cdot \sum_{i=1}^{count_{AV}} f_{wdt}(azi_{stim}, azi_{AV-stim}(i), wdt)}$$

Main substitution I.:

$$o = \{azi_{stim}, azi_{AV-stim}\}$$

Main substitution II.

$$d_{FP}(FP1, FP2, bias_{AV-resp}, o) = bias_{A-only-resp}(o.azi_{stim}, FP1, o.azi_{AV-stim}, bias_{AV-resp}) - bias_{A-only-resp}(o.azi_{stim}, FP2, o.azi_{AV-stim}, bias_{AV-resp})$$

Main hypothesis:

$$\forall(a,b,FP1,FP2,o): (d_{FP}(FP1,FP2,azi_{AV-bias-a},o) = d_{FP}(FP1,FP2,azi_{AV-bias-b},o))$$

4.3.1 Inference

Substitution 1:

$$B_{denom}(o) = w_{a-priori} + (1-w_{a-priori})\cdot \sum_{i=1}^{count_{AV}} f_{wdt}(o.azi_{stim}, o.azi_{AV-stim}(i), wdt)$$

Substitution 2:

$$B_{A-priori}(FP,o) = bias_{a-priori}(o.azi_{stim}, FP)\cdot w_{a-priori}$$

Substitution 3:

$$B_{bias}(bias_{AV-resp}, o) = (1-w_{a-priori})\cdot \sum_{i=1}^{count_{AV}} bias_{AV-resp}(i)\cdot f_{wdt}(o.azi_{stim}, o.azi_{AV-stim}(i), wdt)$$

Inference 1: (Equivalent notation of model formula. To confirm, substitute members of current formula and compare)

$$bias_{A-only-resp}(azi_{stim}, FP, azi_{AV-stim}, bias_{AV-resp}) = \frac{B_{A-priori}(FP,o) + B_{bias}(bias_{AV-resp},o)}{B_{denom}(o)}$$

Inference 2: (Derived from main substitution II. and result of inference 1. We see that d_{FP} is independent from $bias_{AV-resp}$ variable.)

$$d_{FP}(FP1,FP2,bias_{AV-resp},o) = \left(\frac{B_{A-priori}(FP1,o) + B_{bias}(bias_{AV-resp},o)}{B_{denom}(o)} - \frac{B_{A-priori}(FP2,o) + B_{bias}(bias_{AV-resp},o)}{B_{denom}(o)}\right) = \frac{B_{A-priori}(FP1,o) - B_{A-priori}(FP2,o)}{B_{denom}(o)};$$

Main hypothesis contradiction:

$$\exists(a,b,FP1,FP2,azi_{stim}): [d_{FP}(FP1,FP2,bias_{AV-resp-a},o) \neq d_{FP}(FP1,FP2,bias_{AV-resp-b},o)]$$

Contradiction inference: (by substituting main hypothesis contradiction by result of inference 2)

$$\exists(a,b,FP1,FP2,azi_{stim}) \left(\frac{B_{A-priori}(FP1,o) - B_{A-priori}(FP2,o)}{B_{denom}(o)} \neq \frac{B_{A-priori}(FP1,o) - B_{A-priori}(FP2,o)}{B_{denom}(o)}\right)$$

Contradiction is disproved, thus the main hypothesis, that so called FP-dependence is independent of shift direction, is proved. This fact can be visually seen on Obr. 9:: 2nd, 3rd and 5th row.

5 Conclusion

We have described previous studies 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 model of a priori bias affected by AV responses seems appropriate to explain newly observed phenomenon when looking to the no-shift conditions, but inappropriate for the explanation of the difference of reference frames for shifted conditions (Obr. 2:, Obr. 4:).

One of the alternatives for current modelling is use different weights for TrFP vs. NtrFP. Other one is to make NtrFP biases depending on TrFP biases instead directly of AV biases. Alternatively, completely other factors might play role. Additional modeling is currently required 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.

Bibliography

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

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

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

Razavi B, O'Neill WE, Paige GD. (2007) Auditory Spatial Perception Dynamically Realigns with Changing Eye Position. *Journal of Neuroscience*. 2007; 27:10249–10258.