Adaptation in distance perception induced by audio-visual stimuli with spatial disparity

Ľuboš Hládek[1], Aaron R Seitz[2], Norbert Kopčo[1]

[1] **Institute of Computer Science**, P. J. Šafárik University in Košice, Slovakia
[2] **Department of Psychology**, University of California, Riverside, USA

Introduction

Audio-visual (AV) integration can be studied as ventriloquism effect (VE) and aftereffect (VAE), which refer to localization biases due to audio-visual disparity and subsequent plasticity. (Recanzone, 1998)

In distance
- Proximity image effect was reported in anechoic space, auditory (A) object is unified with a closer V target (Gardner, 1968).
- AV unification in VE is more effective for closer V stimuli (Mershon, 1980; Zahorik, 2003), but experiments were performed only with a single fixed V stimulus.
- Closer V stimuli tend to induce stronger VAE than farther V stimuli (Min, Mershon 2005).
- None of the studies used a range of stimuli with fixed AV distance ratio.
- Very few of them in real reverberation.

Current study
Reanalysis of Hládek et al. (2013) with new subjects.
Systematically study VE and VAE in distance dimension in a real room for a range of target distances (directly ahead of listener).
Induce VE and VAE using multiple speaker (A) + LED (V) pairs with a fixed A-V distance ratio, by placing V 30% further or closer than A.

Questions
Is the strength of induced VE & VAE constant across the examined range?
Is it equal in V-Further and V-Closer conditions?
Is there a direct relationship between observed VA and VE patterns?
Does the mechanism of A-V alignment operate on linear or log scale?

Methods

- 136 young NH subjects

Environment
- Small semi-reverb room (T_{60}=408ms)
- 8 targets (Fig.1)
- in medial plane in front of the subject
- LED array above the loudspeakers served to present visual stimuli and collect responses

Target stimuli
- A: 300 ms white noise at 49 (spkr. 1) - 53 (spkr. 8) dB(SPL)

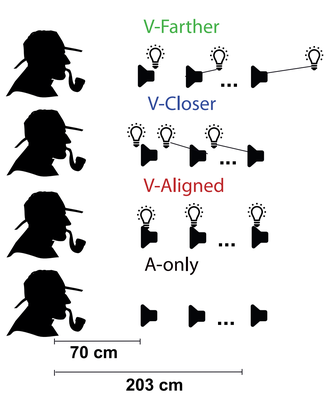


Fig.1 - Experimental setup and conditions. Subjects sat in front of 8 target loudspeakers. The visual stimuli in the AV trials (300 ms LED flas simultaneously presented with sound) were either Aligned, Closer, Farther , or no present at all.The condition was fixed within run.

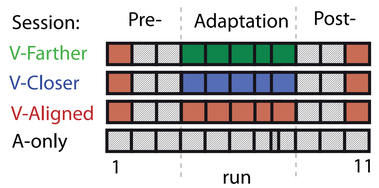


Fig.2 - Types of experimental sessions. Each line represents one type of session, each square represents one run. Color represents condition in each run.

-AV: A + concurrent LED flash, which was either aligned (V-Aligned) or 30% closer (V-Closer), or 30% farther (V-Farther)
- A and AV were interleaved with ratio 1:3

Procedures (Fig.2)
- 2 sessions, e.g., V-Closer and V-Farter was CF group (other groups: FA,CA,AA), each session consited of 11 runs of 64 trials (8 targets x 8 repetitions)
-pre-,post-,adaptation runs
- condition was fixed within adaptation period
- each subject performed 2 different conditions

Task
- indicate the perceived distance using a trackball and the LED array

Analysis
- in log space
- data were grouped by sessions

Acknowledgements
This work was supported by the EU H2020-MSCA-RISE-2015 Grant No. 69122 and VEGA 1/1011/16.

References:
Gardner MB (1968) Proximity image effect in sound localization. J Acoust Soc Am 43:163.
Hládek Ľ, Le Dantec C, Kopčo N, Seitz A (2013). Ventriloquism effect and aftereffect in the distance dimension, In: POMA,19, 050042, 2-7 June 2013;
Mershon DH, Desaulniers DH, Amerson J (1980) Visual capture in auditory distance perception: Proximity image effect reconsidered. J Aud Res 20:129–136.
Min Y, Mershon D (2005) An Adjacency effect in auditory distance perception. Acta Acust united with Acust 91:480–489.
Recanzone GH (1998) Rapidly induced auditory plasticity: the ventriloquism aftereffect. Proc Natl Acad Sci U S A 95:869–75.
Zahorik P (2003) Auditory and visual distance perception: The proximity image effect revisited. J Acoust Soc Am 113:2270–2270.

Results

Baseline

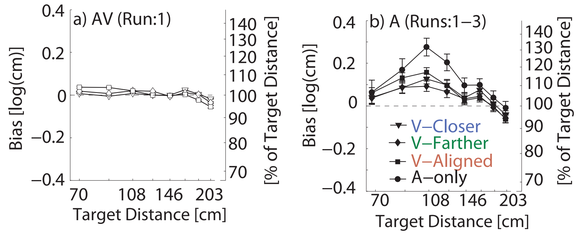


Fig.3 - Baseline perception of auditory distance for each session type. (A) AV trilas in run 1 (i.e., V-Aligned runs) (B) Responses of A trials in runs 1-3.

AV - localization is overall accurate,
- data are aligned across groups.

A - near distances are overestimated, far distances are more accurate, - aligned across groups.
- A-only are even less accurate.

Adaptaion

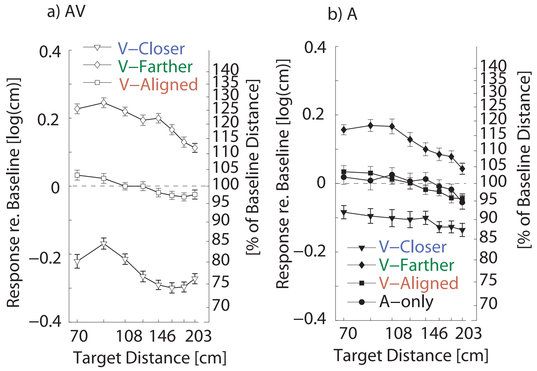
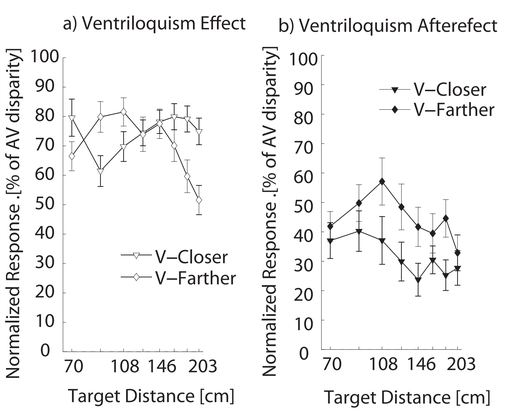


Fig.4 - Auditory distance perception during adaptation period with respect to Baseline in different sessions (symbols). (A) AV data (B) A data

Localization is shifted in the expected direction. AV is shifted more than A. Bias towads middle of targets range was observed in all conditions.
AV - strong compression even in the V-Aligned data. V-Closer produces stronger shift than V-Farther.

A - V-Farther produces stronger shift than V-Closer. A-only is similar to V-Aligned.

Ventriloquism effect (VE) and aftereffect (VAE)



To compute VE and VAE the V-Closer and V-Farther were normalized by V-Aligned and exact value of the AV disparity.

VE - bacame almost independent on target distance and condition. V-Farther is slightly lower at far distances.
VAE - approximately constant with distance, stronger in V-Farther than V-Closer.

Fig.5 -VE and VAE effects normalized by V-Aligned responses and by the physical AV disparity..

Build-up and decay of VE and VAE

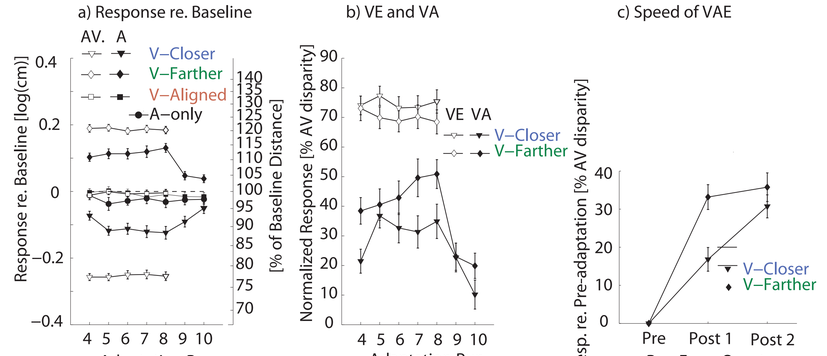


Fig.6 - (A) Temporal profile of mean response bias . Data were pooled across target distances and expressed for each run within a session. (B) VE and VAE as function of run. (C) Rate of VAE buildup / decay averaged across post-onset and post-offset runs and referenced to the pre-onset/offset run.

Fig.6A - AV - constant over runs in all three conditions. A - buildup and decay of VAE visible over several runs.
Fig.6B - VE - equal for the two directions of shift,
VAE - V-Farther increases over runs and it has steeper onset and offset than the V-Closer -
Fig. 6C - difference in buildup/decay is particularly visible when post-onset and post-offset data averaged.

Summary and discussion

VE reached 72% of the AV disparity, VAE reached in the V-Farter 44% V-Closer 31% of the AV disparity.
The constant AV disparity lead to approximately constant VE but the V-Farther was slightly diminished towards the end of the response range. This decrease can relate either to the edge effect of the response range or the perceptual properties of the V adaptors.
The VAE magnitude was approximately constant over the range of target distances, consistently stronger for V-Farther than V-Closer.
The temporal profile of the VE was constant, which points to immediate low level mechanism. The temporal profile of the VAE was much slower which points to the adaptation of the auditory map, as opposed to integration of immediate A and V signals. In the V-Farther condition it built-up quickly and continued to increase over the adaptation period, then decayed quickly. In the V-Closer condition the temporal progress was slower, and the magnitude was constant over the adaptation period.
The observed differences between VE and VAE reflect that VE is an immediate integration of inconsistent V and A signals while VAE is a result of a visually guided adaptation in the spatial auditory map.
The current VE data are inconsistent with previous data which showed VE asymmetry with V-Closer effect stronger than V-Farther. A possible explanation is that the previous studies did not use the correct normalization by V-Aligned responses and physical disparity.
This study is, to our knowledge, the first to observe that V-Farther VAE is stronger and faster than V-Closer VAE. A possible explanation of this asymmetry is, e.g., that the representation of distance in the adapted map is not logarithmic, as assumed here.