VISUAL CALIBRATION OF AUDITORY DISTANCE PERCEPTION

Lubos Hladek[1], Christophe Le Dantec[2], Aaron R Seitz[2], Norbert Kopco[1,3],

[1]Institute of Computer Science, P. J. Safarik University in Kosice, Slovakia [2]Department of Psychology, University of California Riverside, CA, USA, [3]Harvard Med. School – Martinos Center, MGH, Charlestown, MA, USA

MGH/HST Athinoula A. Martinos Center for Biomedical Imaging

Background

Visual (V) signals can influence the perceived location of **auditory** (A) stimuli. This interaction has been extensively studied in horizontal dimension:

Ventriloquism effect (VE) - perceived origin of a sound is shifted towards (or "captured by") the location of concurrently presented V stimulus when the stimuli are at separate locations (Jack and Thurlow, 1973).

Ventriloquism aftereffect (VA) - perceived A location shifts after repeated presentation of horizontally mismatched A-V stimuli, even after V is removed (Recanzone, 1998). This demonstrates rapid short-term recalibration of auditory localization (Shams et al., 2011).

Proximity image effect - in anechoic space, A object is unified with a closer V target (Gardner, 1968).

A-V 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 VA than farther V stimuli (Min, Mershon 2005).

Farther visual adaptors induced stronger VA than closer adaptors (Hladek et al., 2013).

Short-term A-V re-calibration can be linear or logarithmic (in horizontal dimension studied by Shinn-Cunningham et al., 2005).

Motivation and Hypotheses

Previous study (Hladek et al., 2013)

- **VE decreases with distance**
- **VA increases with distance**
- **VE is stronger in V-Closer**
- **VA is stronger in V-Farther**
- **no baseline measurements (taregt distance used instead)**
- **symetrical persistnet VA**

Examine to what extent the previous results were affected by the used reference.

Measure performance in V-Aligned condition and compare it to V-Misaligned (V-Farther and V-Closer) in both immediate and persistent audio-visual effects.

Hypotheses:

H1: The baseline V-Aligned performance will be biased re. actual target locations in both V and AV conditions. This will partially explain the asymmetry between the V-Closer and V-Farther effects in both VE and VA condition.

H2: There will be transference of training. I.e., the effects will be modulated by the order of sessions V-Aligned followed by V-Misaligned or vice versa.

Methods

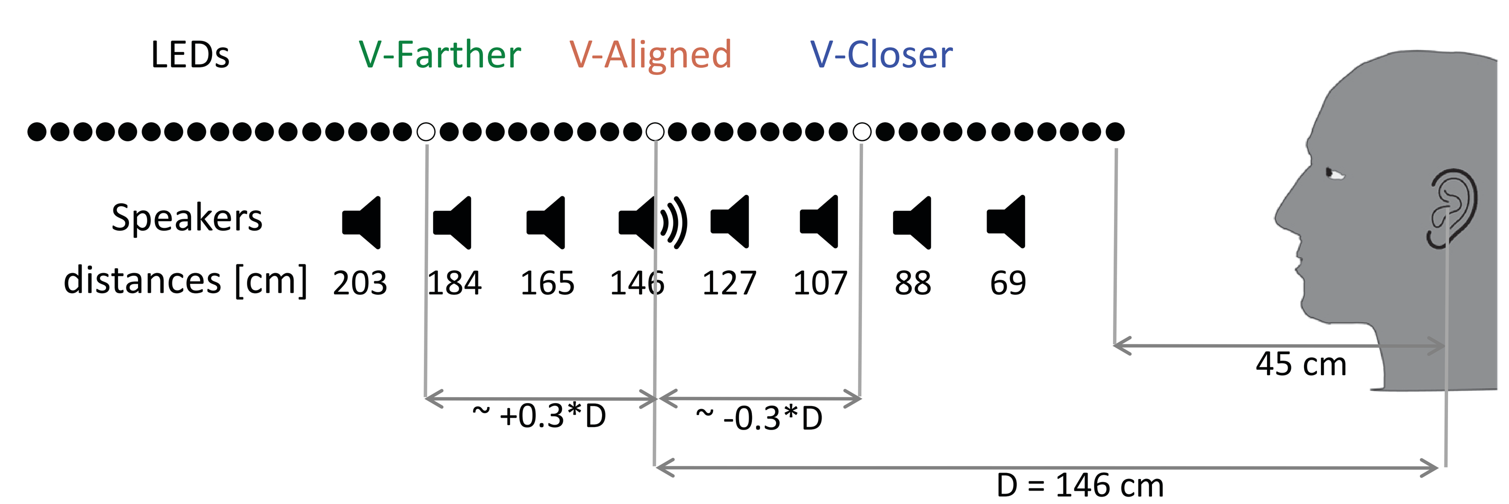

FIGURE 1 Experimental setup and stimuli. Circles represent LEDs (open = LED on, filled = LED off). In the AV presentations, only one LED and one speaker was on at any given time. The LED was aligned with the speaker in AV-Aligned condition. In the V-Closer and V-Farther conditions, the LED was approximately 30% closer or further, respectively, than the active speaker.

Organization of the experiment
Example of group 1. V-Aligned 2. V-Closer

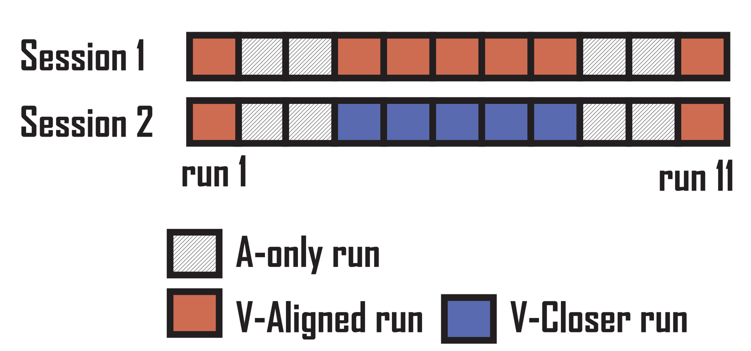

FIGURE 2 Organization of the experiment. Two rows represent two sessions and each block within the row represents one type of the experimental run. Subjects were exposed to simultaneous audio-visual presentation or audio-only presentation.

Conditions (Fig. 1):
V-Aligned - LED at the same distance as sound.
V-Closer - LED 30% closer than sound.
V-Farther - LED 30% farther than sound.

Experiment:
80 subjects (20 per group)
2 one-hour sessions, one of conditions
-V-Aligned vs. V-Farther
-V-Aligned vs. V-Closer
-V-Farther vs. V-Aligned
-V-Closer vs. V-Aligned
fixed within session.
Each session contained 11 runs (Fig. 2).
64 trials per run (self-paced), 500 ms inter-trial pause.

Two types of run:
- AV runs - 75% of AV trials randomly interleaved with A-only (probe) trials (25%);
- A-only runs - 100% of A-only trials

Setup (Fig. 1):
9 speakers covered by sound-transparent cloth in front of the subject at the ear level (closest speaker not used).
Custom made array of LED lights mounted 20 cm above the speaker array.
Stimuli presented via TDT RX8 and Crown CTs 8200 amplifier.

Stimuli:
A-only stimuli - 300 ms broadband noise presented at fixed level; received level range 54-56 dB(A).
AV stimuli - A component identical to A-only; V component (LED light) turned on and off in synchrony with A.

Task:
Subjects indicated perceived sound distance by selecting the closest LED using a trackball.
Subjects instructed to ignore visual stimuli and focus on the perceived sound distance.

Room:
Sound-attenuated small (2.3 m x 3.3 m) reverberant room.
Background noise 35 dB(A).
T_{60} - 406ms

Immediate effects

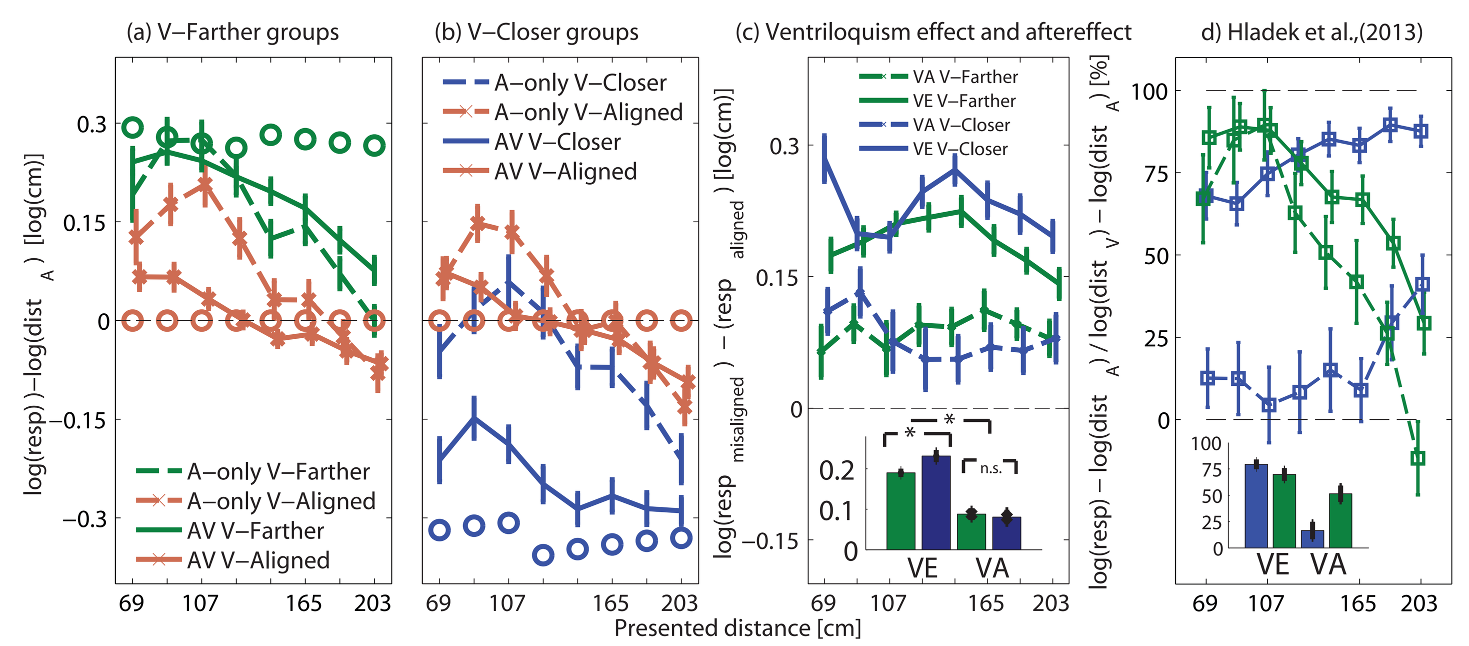

FIGURE 3 Accuracy of localization as a function of target distance averaged across session ordering. Y-axis denotes localization bias expressed in logarithimc units computed as a difference of mean subject's responses throghout runs 4-9 (adaptation) and actual target distance. Mean subject responses in AV trials are plotted using solid lines, in A-only trials using dashed lines. Different colors represent different sessions: V-Aligned, V-Farther, and V-Closer. Small circles represent the location of the V component in AV trials. Each panel (a) and (b) combine data from two groups of subjects with the same conditions (but see Fig. 5). Panel (c) combines data from all 4 groups of subjects and plots VE and VA as a difference of V-Aligned and V-Farther/V-Closer conditions from panels (a) and (b). In-panel graphic at the bottom (c) shows across target means and statistical significance. Error bars show across-subject SEM. Panel (d) shows VE and VA from Hladek et al.,(2013) - thin lines with squares. Actual position of speaker was used as a reference. Data in (d) were adjusted to small discrepancies between LED lights.

Localization accuracy during adaptation (Fig3a, Fig3b)

- AV responses in V-Aligned are slightly biased towards the middle of the response rage,
- A-only responses in V-Aligned condition overshot AV trials for nearby targets by as much as 20%,
- V-Misaligned (V-Farther/V-Closer) produced shifts in expected directions in both A-only and AV trials (re. respective V-Aligned responses)
- VE and VA as defined in Hladek et al., (2013) is similar to localization bias in current data, except the sign in V-Closer (compare Fig. 3a Fig. 3d)

Ventriloquism effect and immediate aftereffect (Fig3c, 3d)

- reference is V-Aligned condition
- dependence on target distnace mostly removed in current experiment re. Hladek (2013)
- VE change with distance - except the very first target, the strongest is in the mid of the response range
- VE is stronger in V-Closer than V-Farther
- VA has equal magnitudes accross distances and orientations
- interaction of (VE, VA) vs. (V-Farther,V-Closer) is significant $F(1,78) = 5.05$ ($p=0.0275$)

Persistent effects

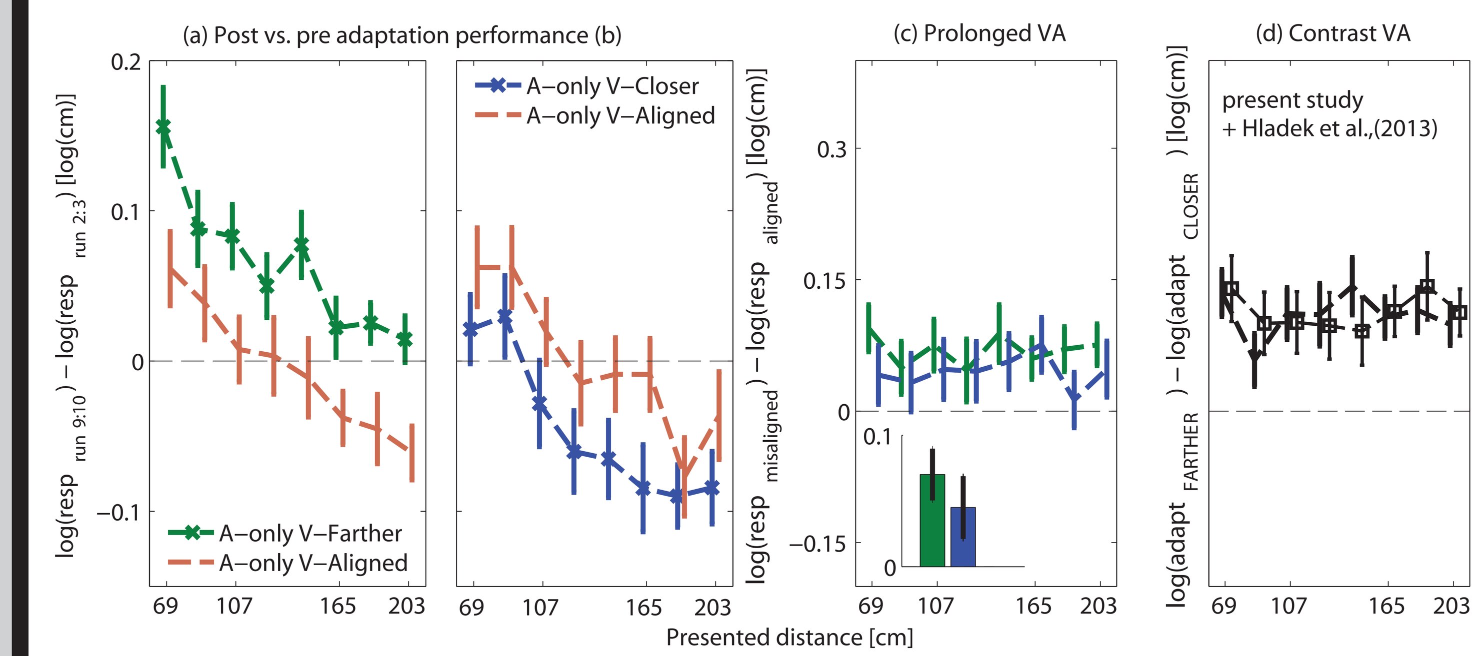

FIGURE 4 Effect of AV training on post-training A-only performance averaged across the session-ordering. Panels (a)(b) express on y-axis the change in a response bias on log scale due AV training as a function of target distance. Post-adaptaiton runs 9-10 are contrasted with pre-adaptation runs 2-3. Each line represents data from one condition (see in-panel legend). Panel (c) shows persistent VA as a difference of the conditions (black vs. color line). In-panel (c) graphic compares the overall magnitudes of persistent VA. Panel (d) shows a contras VA as a difference of V-Closer and V-Farther.

Within-session visual training

- perceptual shift induced in adaptation persists 5-10 minutes after training
- compression in V-Aligned condition observed during adaptation also persists
- persistent VA could be attributed to the short-term plasticity of auditory map in distance

Persistent ventriloquism aftereffect

- Persistent VE is defined as difference of post and pre adaptation
- V-Aligned reference used to show Persistent VE in panel (c)
- equal in magnitude across locations and direction of induced shift
- somewhat smaller in magnitude compared with aftereffect induced immediately (compare dashed lines in Fig. 4c and Fig. 3c), but still well above zero
- Contrast VA is comparable across studies, in current study equal to sum of prolonged VAs in both V-Misaligned conditions

Carry-over effects

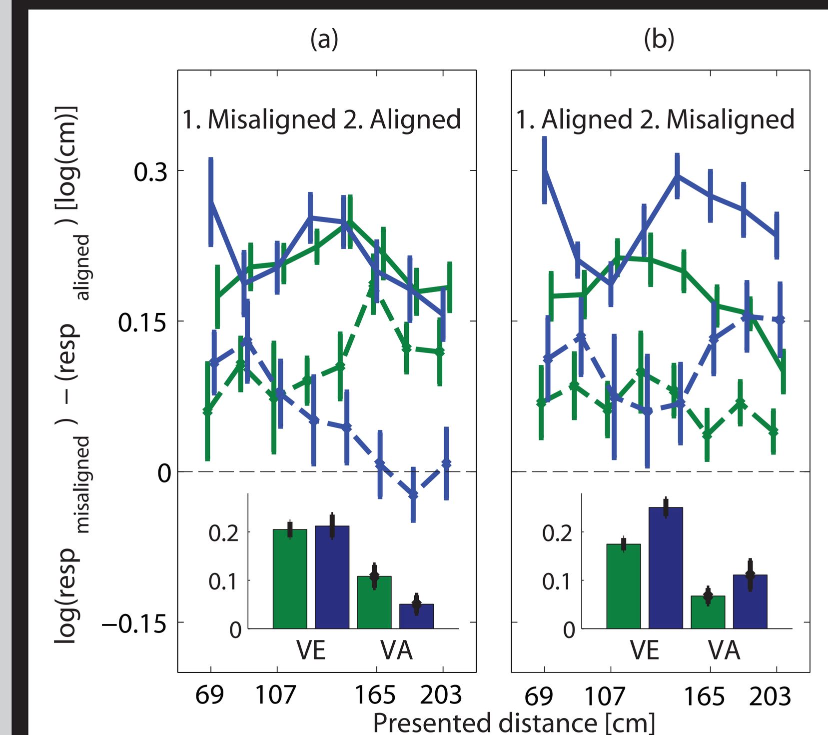

FIGURE 5 Ventriloquism effect and aftereffect for subjects who stared in the (a) V-Misaligned first session (V-Closer / V-Farther) and (b) those who started the first session with V-Aligned. The lines in panels (a) and (b) are equivalent to Fig. 3c (which shows the average of the Figs. 5a and 5b).

AV training persists over a day (Fig. 5)

Respective lines in panels (a) vs. (b) show:

- for targets up to 1m, effects are similar for all groups regardless of the order of sessions,
- for distances above 1m, VA (and V-Closer VE) depends on the session ordering .
  - in Fig. 5b VE is stronger in V-Closer while in Fig. 5a is equal
  - in Fig5a, VA is stronger in V-Farther while in Fig 5b in V-Closer
- AV training may persist over a day. Since we compare V-Aligned and V-Misaligned, and we manipulate the its order during the training, both V-Aligned and V-Misaligned might have contributed to observed differece - further analysis is needed
- if subjects start the first session with V-Aligned (Fig. 5b), VA is proportional to VE, but when they start with V-Misaligned (Fig. 5a), VA is asymmetrical re. VE

Conclusions

Results of Hladek (2013) were replicated.

H1 confirmed: V-Aligned baseline responses provided better reference than actual locations, removing most of the distance dependence. Also, the asymmetry between V-Closer and V-Farther VA and VE was considerably reduced when the baseline performance is taken into account.

H2 partially confirmed: Session ordering resulted in different pattern of VA and VE effects, but only for distances larger than 1m. This suggests that distance-ventriloquism effects persist over hours/days.

Effects on different time scales and magnitudes were observed. The perception is shifted in simultaneous AV presentation, A-only presentation after seconds and minutes of discrepant and also aligned AV presentation. These results might point to different neural circuits involved in audio-visual learning.

Room learning processes might have affected the results.

Further research needed to understand the mechanisms.

References

Gardner, M. B. (1968). "Proximity image effect in sound localization," Journal of the Acoustical Society of America 43, 163.

Hladek, L., Le Dantec, Ch., Kopco N., Seitz AR. (2013). "Ventriloquism effect and aftereffect in the distance dimension". ICA Montreal, POMA Volume 19, pp. 050042

Jack, C. E., and Thurlow, W. R. (1973). "Effects of degree of visual association and angle of displacement on the' ventriloquism' effect," Perceptual and Motor Skills

Kopco, N, I-F Lin, BG Shinn-Cunningham, and JM Groh (2009). Reference frame of the ventriloquism aftereffect. Journal of Neuroscience, 29(44):13809-13814

Mershon, D. H., Desaulniers, D. H., and Amerson, J. (1980). "Visual capture in auditory distance perception: Proximity image effect reconsidered," Journal of Auditory Research 20, 129–136.

Min, Y., and Mershon, D. (2005). "An Adjacency effect in auditory distance perception," Acta acustica united with acustica 91, 480–489.

Recanzone, G. H. (1998). "Rapidly induced auditory plasticity: the ventriloquism aftereffect," Proceedings of the National Academy of Sciences of the United States of America 95, 869–75.

Shams, L., Wozny, D. R., Kim, R., and Seitz, A. (2011). "Influences of multisensory experience on subsequent unisensory processing," Frontiers in psychology 2, 264.

Shinn-Cunningham, B. G., Streeter, T., and Gyss, J.-F. (2005)."

Zahorik, P. (2003). "Auditory and visual distance perception: The proximity image effect revisited," J Acoust Soc Am 113, 2270–2270.

Acknowledgement:

This work was supported by EU Marie Curie Grant (FP7-247543), Scientific grant agency of the Ministry of Education of Slovak Republic and the Slovak Academy of Sciences (VEGA-1/0492/12), National Science Foundation (BCS-1057625), Safarik University (PF-VVGS 2013-82), Visegrad Fund (51300869), Slovak Agency for Science and Research (APVV-0452-12), NIH (P41RR14075)