Attention, Ventriloquism and Audio-Visual Speech

David Cottrell
James Cook University, Australia

Background

Typically, performance in almost any audio-visual task is enhanced when the auditory and visual stimuli are close together in time and space (c.f. Calvert, Spence, & Stein, 2004)

Thought to underly this enhanced performance are the overlapping receptive fields of multisensory cells in the superior colliculus (c.f. Stein & Stanford, 2008).

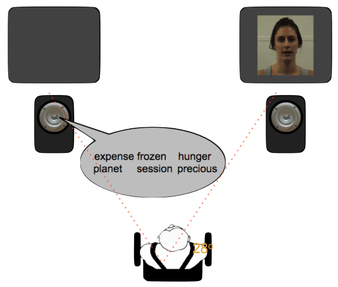
Driver (1996) reported data, unique in the literature, where performance in a difficult audio-visual shadowing task was improved when the sound source was displaced from the talker (highlighted row in Table).

Despite being widely cited (>228 citations) there are almost no published replications of the effect.

Research Question

Are we really better at recognising audio-visual speech when the sound only appears to come from the the person talking BUT doesn't really?

Method: Driver (1996) Expt. 1



Driver presented 12 participants with 112 word triplets, all in the same female voice which were carefully synchronised to the lip movements in a video of the talker.

Half the triplets matched the talker's lip movements, half did not.

The participant's task was to shadow the words that matched the lip movements, and ignore the other words.

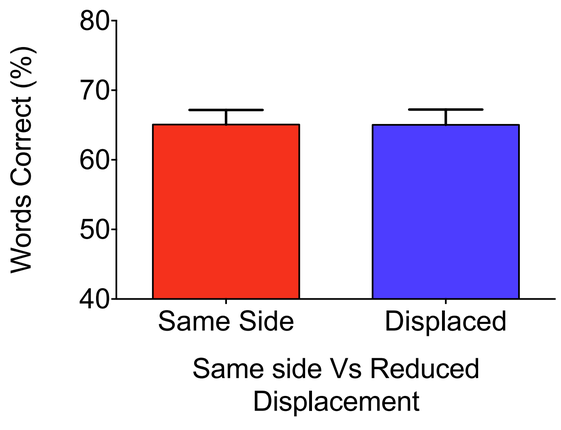
On half the trials the all the spoken words were presented from the speaker under the active video monitor, on the other half the sound and visual stimuli were displaced (28º).

Results

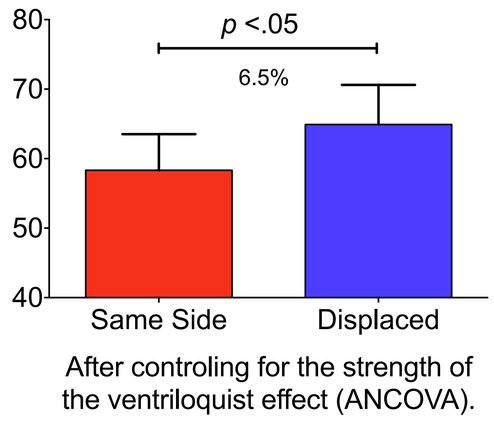
A Difficult Effect to Replicate...

Experiment	n	Words correct (%)		Displacement Advantage	Ventriloquist Illusion Reported
		Same Side	Displaced		
Driver (1996) Experiment 1	12	59	76	17%*	✔
Experiment 1 (Replication)	43	51 (11)	59 (13)	8%*	✔
Experiment 2 (Replication)	16	55 (8)	56 (9)	1%	✗
Experiment 3 (Data Projector)	35	48 (11)	49 (10)	1%	✗
Experiment 4 (HRTF)	38	69 (9)	68 (9)	-1%	✗
Experiment 5 (Single Words)	46	89 (13)	89 (12)	0%	?
Experiment 6 (Multiple Choice)	27	85 (8)	83 (8)	-2%	?

The importance of the Ventriliguist effect: Experiment 7



Same side Vs Reduced Displacement



After controling for the strength of the ventriloquist effect (ANCOVA).

Method: Current Experiments

The stimuli were digital recordings of 480 words.

Replications

Experiments 1 & 2 replicated Driver's original study but with slight (unplanned) changes to the room configuration between experiments.

Experiment 3 was conducted in a larger room with the visual stimuli presented via a projector.

Experiment 4 attempted to remove specific room effects and used headphones and generic Head Related Transfer Functions (HRTF) to manipulate the apparent sound source.

Simplified responding and memory requirements

Experiment 5 simplified the presentation to word pairs (target + distractor) to decrease memory load.

Experiment 6 simplified the response to a multiple choice but presented the words in triplets as in all experiments other than Experiment 5.

Enhancing the ventriloquist illusion

Experiment 7 was designed to optimise the ventriloquist effect. Driver asserted the advantage in shadowing words in the displaced condition was a consequence of the ventriloquist effect producing an illusory spatial separation between the target and distractor words.

Conclusions

Displacement-aided speech perception phenomena is very difficult to demonstrate.

In 7 Experiments the effect was demonstrated only twice.

In 15 unpublished experiments from other laboratories, 2 successful replications.

In all successful replications, participants reported a strong ventriloquist effect.

It is possible that displacement-aided speech perception depends on the ventriloquist effect as Driver suggested but the 17% advantage he reported is at the upper end of the range of the effect.

References

Driver, J. (1996). Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature*, 381(6577), 66–68. doi: 10.1038/381066a0

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, 9(4), 255–266. doi:10.1038/nrn2331

Calvert, G. A., Spence, C., & Stein, B. E. (Eds.). (2004). *The handbook of multisensory processes*. MIT press.