

OBJECT IDENTITY DECISIONS: AT WHAT PROCESSING LEVELS? OR: WHY THE CANTALOUPE MIGHT WORK

Commentary on Felice BEDFORD: "Toward a general law of numerical/object identity"

Paul Bertelson¹², Jean Vroomen², Gisa Aschersleben³ and Béatrice de Gelder²¹

¹Université libre de Bruxelles (Belgium)

²Tilburg University (The Netherlands)

³Max-Planck Institute for Psychological Research, Munich (Germany).

Correspondence:

Paul Bertelson

Laboratoire de Psychologie Expérimentale, CP 191

Université libre de Bruxelles

50 Av. F.D. Roosevelt

B-1050 Bruxelles

BELGIUM

E-mail: pbrtln@ulb.ac.be

Felice Bedford (FB)'s new enterprise - grafting considerations from modern geometry on the traditional issue of perceptual unity - is certainly an important new development. The potential ability of the approach to deal with the vexing problem of stimulus similarity is impressive. In the present comment, however, we shall limit ourselves to the factual database and in particular to the parts that belong to our own domain of expertise, multisensory perception.

In her initial description of phenomena implying object identity decisions, FB does rarely ask at what level of processing these decisions occur. All she offers in that line is a one-paragraph caveat (p. **) that the decisions need not be conscious, while whether conscious penetration can help is an open question. As a matter of fact, the issue takes different forms for the different phenomena she considers. For apparent movement and for stereopsis as well, conscious post-perceptual deliberation is not a possible factor, because the pre-unity elements are simply not accessible to consciousness. The same healthy situation exists for priming, when the prime is effectively masked, and, if one moves outside the visual domain, for stereophony. The cases of prism adaptation, of ventriloquism, of Gestalt grouping, are different. There, the elements are sometimes available to conscious inspection, or, in other words, the experimental situations present variable degrees of transparency, so that responses suggestive of identity formation can be generated both automatically within perceptual processing proper, or in more controlled ways at post-perceptual levels. The question for FB would then be whether the same geometries apply in these two cases.

For the particular example of ventriloquism, many of the often-quoted studies only showed that participants' responses to the bimodal input were influenced by both the auditory and the visual component, but left the locus of the effect much in doubt. Some of our own recent work on the subject has been focused on isolating automatic from eventual strategic components. We took the manifestation known as the visual bias of auditory location, i.e. the fact that the apparent location of sounds shifts in the direction of visual events occurring simultaneously in a discordant place (Bermant & Welch, 1975;

Bertelson & Radeau, 1981; Pick, Warren & Hay, 1969). It has usually been studied by asking subjects to report the apparent location of the sounds, while ignoring the visual distracters. The fact that bias occurs in spite of these selective attention instructions has sometimes been taken as proving its automatic nature, and has encouraged comparisons with Stroop phenomena. The argument is a weak one, however, for once the discordance is detected, it is still up to the participant to decide what to do about the experimenter's exhortation. As noted earlier, the problem belongs to the social psychology of subject-experimenter interaction rather than to the study of perception (Bertelson, 1994).

We have now tried to address the problem by a psychophysical staircase procedure. On each trial, the subjects has to decide whether target sounds, whose apparent azimuthal location is controlled stereophonically, came left or right of the median plane. Following the staircase principle, after a "left" response, the next target on the same staircase is moved one step to the right, and vice versa after a "right" response. Two staircases, starting respectively far to the left and far to the right, are run in random alternation, and converge progressively towards the center, until response reversals (responses different from the preceding one on the same staircase) begin to occur. From that moment, the subject can no longer discriminate left from right deviations, and no voluntary strategy can affect his performance in any systematic way. In our experiments, reversals began to occur at significantly larger distances from the center when flashes were produced in a central lamp in exact synchrony with the target sounds than in the absence of visual distracters (Bertelson & Aschersleben, 1998). This result demonstrates that ventriloquism originates partly at least in an automatic phenomenon.

A converging argument for automaticity has been obtained from brain-lesioned patients with severe left unilateral neglect. When presented with bright flashes of light left or right of the midpoint, these patients detected none of those on the left, but their localization of a target sound delivered frontally in synchrony with the flashes was nevertheless shifted significantly in the direction of the undetected left ones (Bertelson, Pavani, Ladavas, Vroomen & de Gelder, 2000). That visual bias can take place without awareness of the

occurrence of the visual distracter provides a demonstration of its automaticity that is still stronger than the one from the staircase studies, in which the subjects were aware of the distracter's presence and only ignored its location relative to the target.

Further insight regarding the cognitive locus of ventriloquism has resulted from work showing that visual bias does not depend on whether attention is focused on the visual distracter or somewhere else. This was found both for voluntary (or endogenous) orientation, imposed by a secondary target-monitoring task (Bertelson, Vroomen, de Gelder, & Driver, 2000) and for automatic (exogenous) capture by a singleton distracter (Vroomen, Bertelson, & de Gelder, 2001). These findings imply that crossmodal interaction reorganizes the multi-sensory spatial scene before the operation of selective attention.

Thus, data from ventriloquism studies can effectively provide good examples of automatic identity decision occurring at a pre-attentional stage and independently of any strategic influence, but only when adequate experimental controls were applied.

One aspect of ventriloquism that FB presents as particularly demonstrative of the occurrence of an identity decision is the alleged role of familiarity of the bimodal scene in bringing about the effects. She reasons, for example, that the illusion created by performing ventriloquists that their speech comes from the dummy they hold must depend on the facts that “the dummy’s appearance is human-like, and the knowledge that humans talk” and would not work “if the ventriloquist held a plain cantaloupe in his lap”(p.**). In the literature on multimodality there is indeed a long tradition of stressing the role of top-down factors of semantic nature (see Welch, 1999; Welch & Warren, 1980). The empirical support for that view is nevertheless somewhat short of overwhelming (see Bertelson, 1999, for a discussion).

There are three points to consider here.

First, all the classical manifestations of ventriloquism (impression of perceptual fusion of the discordant inputs, immediate bias, aftereffects of exposure to discordant inputs) have been obtained not only with realistic

simulations of familiar scenes, but also with purified situations, consisting of completely arbitrary pairings of sound bursts and light flashes, with near temporal coincidence as the only essential condition. The use of familiar audiovisual pairs is thus certainly not necessary.

The second point is whether realism can increase ventriloquism beyond the level obtained with purified situations. FB quotes well-known results (by Jackson and by Jack & Thurlow) suggesting that it does, but other studies have failed to support the notion. Radeau & Bertelson (1977) found no difference between the aftereffects resulting, in one experiment, from exposure to percussion sounds paired with either the sight of the hands playing the instruments, or with diffuse light flashes synchronized with the beats, or, in another experiment, auditory speech paired either with the sight of the talker's moving lips or again with flashes synchronized with the amplitude peaks of the speech. On the other hand, one classical result suggesting a pure top-down effect of knowledge was that sounds were attracted towards the displaced location in which a dummy loudspeaker was seen (Pick et al. 1969). Trying to replicate the result, Radeau (1992) obtained no trace of immediate bias, nor of aftereffect, with the loudspeaker, while the two effects were obtained once more in the usual sound-flash situation. The evidence on that point is thus contradictory.

The third point concerns the processing level at which semantic effects, when obtained, would have originated. In another field, a well-published semantic phenomenon, the effect of object size familiarity in distance estimation, has classically been observed with explicit direct judgments, but disappeared when distance perception was measured implicitly with a non-transparent method (Predebon & Wooley, 1994). For ventriloquism, the effects of context familiarity reported in the literature were all obtained in clearly transparent situations, and should now be replicated with staircases.

Thus, going back now to FB's mind experiment with the performing ventriloquist, the actual evidence suggests that any conspicuous object, cantaloupe or not, held at the right distance and agitated with the right timing, would probably capture the ventriloquist's speech. Whether a dummy with

movable lips would outperform the cantaloupe is, on the other hand, still an open question.

REFERENCES

- Bermant, R.I. & Welch, R. B. (1976) The effect of the degree of visual-auditory stimulus separation upon the spatial interaction of vision and audition. Perceptual and Motor Skills, 43, 487-493.
- Bertelson, P. (1994) The cognitive architecture behind auditory-visual interaction in scene analysis and speech identification. Current Psychology of Cognition, 13, 69-95.
- Bertelson, P. (1999) Ventriloquism: A case of crossmodal perceptual grouping. In Cognitive Contributions to the Perception of Spatial and Temporal Events. G. Aschersleben, T. Bachmann, J. Müsseler (Eds.), Elsevier, 347-362.
- Bertelson, P. & Aschersleben, G. (1998) Automatic visual bias of perceived auditory location. Psychonomic Bulletin & Review, 5, 482-489.
- Bertelson P., Pavani F., Ladavas E., Vroomen J, & de Gelder B. (2000) Ventriloquism in patients with unilateral visual neglect. Neuropsychologia, 38, 1634-1642.
- Bertelson, P. & Radeau, M. (1981) Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. Perception & Psychophysics, 29, 578-587.
- Bertelson, P., Vroomen, J., de Gelder, B. & Driver, J. (2000) The ventriloquist effect does not depend on the direction of deliberate visual attention. Perception & Psychophysics, 62, 321-332.
- Pick, H.L., Warren D.H. & Hay, J.C. (1969) Sensory conflict in judgments of spatial direction. Perception & Psychophysics, 6, 203-205.
- Predebon, J. & Wooley, J. S. (1994) The familiar-size cue to depth under reduced-cue viewing conditions. Perception, 23, 1301-1312.
- Radeau, M. (1992) Cognitive impenetrability in auditory-visual interaction. In J. Alegria, D. Holender, J. Morais & M. Radeau (Eds.) Analytic approaches to human cognition. Amsterdam: Elsevier, 41-55.
- Radeau, M. & Bertelson, P. (1977) Adaptation to auditory-visual discordance and ventriloquism in semi-realistic situations. Perception & Psychophysics, 22, 137-146.
- Welch, R. B. Meaning, attention and the “unity assumption” in the intersensory bias of spatial and temporal perceptions. In Cognitive Contributions to the Perception of Spatial and Temporal Events. G. Aschersleben, T. Bachmann, J. Müsseler (Eds.), Elsevier, 371-388
- Welch, R.B. & Warren, D.H. (1980) Immediate perceptual response to intersensory discrepancy. Psychological Bulletin, 88, 638-667.