Research Article

VISUAL RECALIBRATION OF AUDITORY SPEECH IDENTIFICATION: A McGurk Aftereffect

Paul Bertelson,[1,2] Jean Vroomen,[2] and Béatrice de Gelder[2]

[1]*Université libre de Bruxelles, Brussels, Belgium, and* [2]*Tilburg University, Tilburg, The Netherlands*

Abstract—*The kinds of aftereffects, indicative of cross-modal recalibration, that are observed after exposure to spatially incongruent inputs from different sensory modalities have not been demonstrated so far for identity incongruence. We show that exposure to incongruent audiovisual speech (producing the well-known McGurk effect) can recalibrate auditory speech identification. In Experiment 1, exposure to an ambiguous sound intermediate between /aba/ and /ada/ dubbed onto a video of a face articulating either /aba/ or /ada/ increased the proportion of /aba/ or /ada/ responses, respectively, during subsequent sound identification trials. Experiment 2 demonstrated the same recalibration effect or the opposite one, fewer /aba/ or /ada/ responses, revealing selective speech adaptation, depending on whether the ambiguous sound or a congruent nonambiguous one was used during exposure. In separate forced-choice identification trials, bimodal stimulus pairs producing these contrasting effects were identically categorized, which makes a role of postperceptual factors in the generation of the effects unlikely.*

The question of how sensory modalities cooperate in forming a representation of the environment is the focus of much current work at both the behavioral and the neuroscientific levels. A substantial part of that work is carried out with *conflict* situations, in which incongruent information about potentially the same distal event is presented to different modalities (see review by Bertelson & de Gelder, in press). Exposure to conflicting inputs produces two main effects, *immediate biases* and *aftereffects*. Immediate biases are effects of incongruent inputs in a distracting modality on perception of corresponding inputs in a target modality (see review by Welch & Warren, 1980). For example, in the *ventriloquist effect*, the perceived location of target sounds is displaced toward light flashes delivered simultaneously at some distance, despite instructions to ignore them (Bermant & Welch, 1976; Bertelson & Aschersleben, 1998; Bertelson & Radeau, 1981; Klemm, 1909). Aftereffects are similar shifts observed following exposure to an intermodal conflict, when data in one or in both modalities are later presented alone. For the ventriloquism situation, researchers have reported aftereffects in which unimodal auditory localization, and sometimes also visual localization, is displaced in the direction occupied by the distractors in the other modality during the preceding exposure phase (Radeau & Bertelson, 1974, 1976; see review by Radeau, 1994a). Similarly, exposure to discordant visual and proprioceptive information, as when the apparent location of a hand is displaced through a prism, results in aftereffects in both visual and proprioceptive localization (Welch, 1986). Aftereffects show that exposure to conflicting inputs recalibrates processing in the respective modalities in a way that reduces the conflict.

Both immediate biases and aftereffects have been consistently demonstrated for spatial conflict situations, as in the case of ventriloquism. But the existing evidence is less complete for conflicts regarding event identities; biases have been reported in several cases, but to our knowledge, no aftereffects have been reported so far.

The main example of immediate bias in the case of object identities is the conflict resulting from the acoustic delivery of a particular speech token in synchrony with the visual presentation of a face articulating an incongruent token. This situation can produce strong biases of auditory identification toward the lip-read distractor, generally known as *McGurk effects* (McGurk & MacDonald, 1976). In the most spectacular case, originally reported by McGurk and MacDonald, a spoken /ba/ syllable dubbed onto a visual presentation of /ga/ was reported as /da/ on a majority of trials. However, no identity recalibration subsequent to exposure to McGurk pairs has yet been reported, and moreover, some results, which we discuss later, have been cited as meaning that recalibration does not occur for audiovisual speech (Rosenblum, 1994). Immediate biases but so far no recalibration have also been reported for conflict situations involving musical sounds accompanied by the sight of incongruent performing gestures (Saldaña & Rosenblum, 1993) and a face displaying a particular emotion accompanied by a voice with an incongruent tone (de Gelder & Vroomen, 2000).

The occurrence of recalibration occupies an important place in the current picture of cross-modal interactions. On the one hand, it reveals the existence of a genuine perceptual contribution to the causation of such interactions. This argument has been used (Bertelson, 1999; Bertelson & Radeau, 1976; Radeau, 1994a; Vroomen, 1999), as have others (Bertelson & de Gelder, in press), against reductionist interpretations of cross-modal biases in terms of postperceptual processes (Choe, Welch, Gilford, & Juola, 1975; Welch, 1999; Welch & Warren, 1980). On the other hand, recalibration probably plays a central role in developing and later maintaining cross-modal coordination (Held, 1965). Recalibration thus links cross-modality research with evolutionary accounts of perceptual function. Such ecological arguments would lose much of their strength if recalibration coexisted with immediate biases only for some cognitive domains. For that reason, in the current study, we examined whether exposure to a cross-modal conflict at the level of event identity can recalibrate stimulus identification. We took as our starting point the conflict between heard and lip-read speech that forms the basis of the McGurk effect.

An obvious way of running this investigation might be to transpose to audiovisual speech the tasks typically used to study recalibration in cases of spatial conflicts, such as ventriloquism (Radeau & Bertelson, 1974, 1976; Recanzone, 1998). In the basic experimental condition, participants would perform forced-choice identifications of an auditory test token before and after repeated exposure to the same token

Address correspondence to Paul Bertelson, Laboratoire de Psychologie Expérimentale, Université libre de Bruxelles, CP 191, 50 Av. F. D. Roosevelt, B-1050 Bruxelles, Belgium; e-mail: pbrtlsn@ulb.ac.be.

synchronized with an incongruent visual distractor. If interpretations consistent with the visual distractor were more frequent on posttests than on pretests, this would be proof of visual recalibration.

One simple, and at first thought adequate, procedure would be to expose subjects to, for example, auditory /ba/ or /da/, each combined with an incongruent visual correspondent (/da/ or /ba/, respectively), and see if items from a /ba/-/da/ auditory continuum are reported as /ba/ more often after exposure to auditory /da/ combined with visual /ba/ than after exposure to the opposite pair. However, the interpretation of this result would be complicated by the phenomenon known as *selective adaptation to speech*, first described by Eimas and Corbit (1973): Repeated presentation of a particular speech utterance by itself, and in the absence of any distractor, causes a reduction in the frequency with which that token is reported in subsequent identification trials. It is thus an adaptation phenomenon that, like recalibration, manifests itself by aftereffects, but, unlike recalibration, does not depend on the occurrence of conflicting inputs. Like analogous visual cases of sensory adaptation, such as motion aftereffects (Anstis, 1986), selective adaptation to speech probably reveals fatigue of some of the relevant processes, although criterion-setting operations, resulting in range or contrast effects, can also be involved under particular conditions[1] (see reviews by Jusczyk, 1986; Samuel, 1986).

In the recalibration experiment under consideration, with repeated exposures to auditory /ba/ combined with visual /da/, selective adaptation to /ba/ would result in more test auditory tokens being interpreted as /da/. This is exactly the result predicted for recalibration by a /da/ visual distractor. Thus, if this recalibration paradigm was run with end-point auditory tokens, the observed effect could be attributed equally to visual recalibration and to selective speech adaptation.

In the current study, we tried to isolate the effect of recalibration by using an ambiguous synthetic token, intermediate between the two end-points, as the auditory component of the adapting stimulus pairs. Items of this kind have been shown to cause no selective adaptation (Sawusch & Pisoni, 1976) and to be susceptible to strong McGurk biases (Bertelson, Vroomen, Wiegeraad, & de Gelder, 1994). We predicted that exposure to bimodal stimulus pairs combining an ambiguous auditory token with either of the two corresponding extreme visual tokens would increase the number of postexposure judgments consistent with the visual distractor. This would provide a nonambivalent demonstration of visual recalibration.

EXPERIMENT 1: VISUAL RECALIBRATION WITH AMBIGUOUS AUDITORY SPEECH

The purpose of the first experiment was to determine whether our proposed bimodal-exposure condition effectively produces visual recalibration. We used a place-of-articulation auditory continuum ranging from /aba/ to /ada/. Identification tests carried out at the start of the session provided for each participant the continuum token that came closest to producing the two judgments (/aba/ vs. /ada/) in equal proportions. This item became the participant's ambiguous token. It was paired with one of the two visual distractors, /aba/ or /ada/, on the exposure trials. The effects of these exposures were assessed through forced-choice identification of similarly ambiguous auditory tokens. Visual recalibration was expected to result in a higher proportion of /aba/ judgments after exposure to /aba/ than after exposure to /ada/.

Method

The material was based on digital audio and video recordings of a male native speaker of Dutch, pronouncing /aba/ and /ada/. The video recordings showed the speaker facing the camera with the frame extending from neck to forehead. The audio recordings were synthesized using the Praat program (Boersma & Weenink, 1999) to produce two 640-ms-long synthetic stimuli with 240-ms stop closure. Varying the frequency of the F2 formant by equal steps of 39 Mel provided a nine-token place-of-articulation continuum. Bimodal stimulus pairs for use in the adaptation phase of the experiment were prepared by dubbing particular continuum tokens (described later) onto each of the two videos.

Participants were 11 Tilburg University students, all naive concerning the aims of the study. Tested individually in a soundproof booth, they sat 75 cm from a 17-in. monitor on which the videos were displayed. Sounds were delivered over a loudspeaker placed just below the monitor.

The experimental session started with a series of 98 *auditory pretests* that consisted of oral forced-choice /aba/-/ada/ judgments on all tokens from the continuum (but with presentation frequency biased in favor of the five central tokens; see the caption to Fig. 1). Each participant's phoneme boundary was estimated using probit analysis, and the token closest to that boundary became that participant's *ambiguous auditory token* (denoted as A?) for the rest of the session.

Testing for recalibration was organized by adaptation blocks of 8 *bimodal-exposure trials*, each block followed immediately by six au-

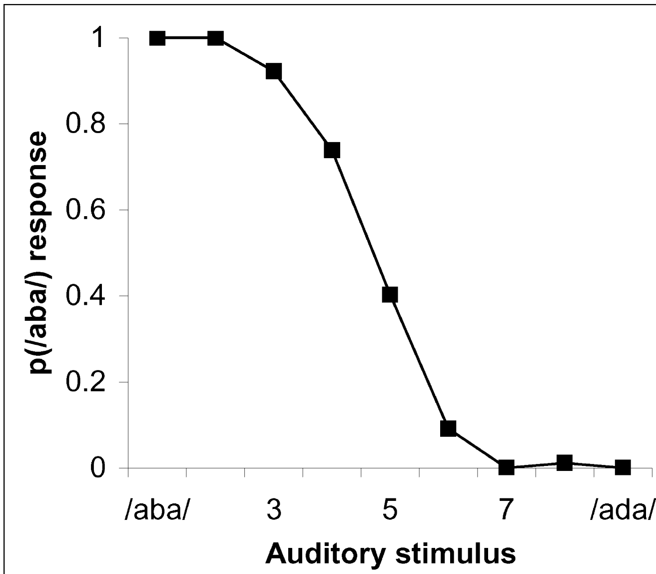

Fig. 1. Results of the auditory pretests in Experiment 1. The graph shows the mean proportion of /aba/ judgments across the 11 participants for each item in the auditory continuum. Items 3 through 9 were each presented on 14 trials, Items 2 and 10 were presented on 8 trials, and Items 1 and 11 were presented on 6 trials.

1. The nature of the underlying mechanisms has no implication for the present investigation, in which selective speech adaptation is considered only as a possible confounding factor in the study of visual recalibration. We use the term selective speech adaptation in a purely descriptive way, and, for example, by no means as an opposite to contrast or range effect.

ditory posttests. On every bimodal-exposure trial, the ambiguous token A? was dubbed onto one of the two videos, giving pairs A?Vb (ambiguous auditory token, visual /aba/) and A?Vd (ambiguous auditory token, visual /ada/). The same pair was used throughout the 8 exposure trials of each block. During these exposure trials, attention was focused on the face by having the participant monitor it for the occasional appearance of a small (12 pixels) white spot on the upper lip (the spot appeared 10 times over the 128 exposure trials). Eight adaptation blocks were run with each bimodal pair, in randomized order. Each auditory posttest, just like the auditory pretests, involved a forced-choice /aba/-/ada/ judgment. The three tokens closest to the phoneme boundary, A? and its two neighbors on the continuum (A? − 1 and A? + 1), were each presented on two of the six posttests. Both exposure trials and posttests followed each other at 4-s intervals.

At the end of the session, each of the two bimodal stimulus pairs of the exposure trials was presented in 10 forced-choice audiovisual identification trials, to check on the effectiveness of the material in producing the McGurk effect. Items were presented in randomized order, and the participants were instructed to say after each presentation whether they had heard /aba/ or /ada/.

Results

Mean judgments in the auditory pretests are shown in Figure 1. Individual ambiguous tokens (A?) chosen on the basis of these tests were Item 4 for 3 participants and Item 5 for the other 8. The mean percentage of /aba/ judgments on these tokens was 48% (*SD* = 18).

In the final audiovisual identification tests, A?Vb was judged as /aba/ on 98% of trials, and A?Vd was judged as /ada/ on 97% of trials, demonstrating almost total visual biases in identifying the ambiguous auditory token A?. A substantial bias was necessary for obtaining an adequate test of the recalibration hypothesis.

The visual catch stimuli of the exposure trials were detected on 100% of their presentations. The data from the auditory posttests appear in Figure 2. The proportion of /aba/ judgments was, for every test token, higher after exposure to the bimodal pair with visual /aba/ than after exposure to the bimodal pair with visual /ada/. In a 2 (visual adapter) × 3 (auditory test stimulus) analysis of variance (ANOVA) on that proportion, the main effects of visual adapter, $F(1, 10) = 7.06$, $p < .024$, and of test stimulus, $F(2, 20) = 88.6$, $p < .001$, and their interaction, $F(2, 20) = 11.0$, $p < .001$, were all significant. When the data were averaged over test tokens, the larger proportion of /aba/ judgments after visual /aba/ was observed in 10 of the 11 participants (the 11th showed no difference), $p < .005$ by sign test.

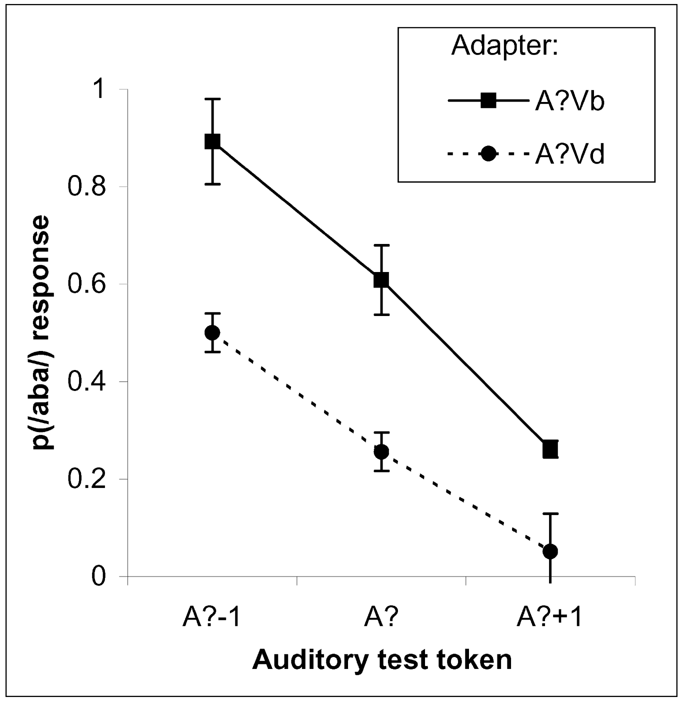

Fig. 2. Results of the auditory posttests in Experiment 1. The graph shows the proportion of /aba/ judgments after exposure to an adapter consisting of the participant's ambiguous auditory token (A?) combined with either visual /aba/ (A?Vb adapter) or visual /ada/ (A?Vd adapter). On each posttest, the presented auditory token was either the ambiguous token (A?) or one of its two neighbors on the auditory continuum (A? − 1 or A? + 1).

Discussion

As predicted, exposing participants to a particular visual speech token combined with an ambiguous auditory token moved the distribution of postexposure auditory judgments in the direction of that visual token. Thus, incongruent lip-read information can recalibrate auditory speech identification.

In the introduction, we mentioned that some earlier results have been considered as supporting the conclusion that recalibration does not occur for audiovisual speech. These results, however, came from two studies (Roberts & Summerfield, 1981; Saldaña & Rosenblum, 1994) that were primarily addressing a different question. They asked whether selective speech adaptation takes place at a phonetic level of processing, as proposed originally by Eimas and Corbit (1973), or at a more peripheral, earlier-reached level. The classical speech-adaptation paradigm was run with either a unimodal auditory token (in Saldaña & Rosenblum's study, /ba/) or an audiovisual token (auditory /ba/ paired with visual /va/, nearly always perceived as /va/) as the adapter. The effect was identical in the two conditions (an increased number of /va/ judgments), contrary to the prediction from the phonetic hypothesis, according to which there would be adaptations in opposite directions in the two conditions (more /va/ after unimodal /ba/, more /ba/ after pairs of auditory /ba/ with visual /va/). Roberts and Summerfield obtained basically similar results, although their bimodal adapters (auditory /bɛ/ with visual /gɛ/) produced more variable percepts. In both studies, the authors' only conclusion was that speech adaptation is a purely auditory phenomenon; they did not mention the issue of visual recalibration.

The idea that these results disproved the possibility of visual recalibration in the case of audiovisual speech was first discussed later, in peer commentaries on a target article by Radeau (1994a) concerning the relations between audiovisual speech and ventriloquism (de Gelder & Vroomen, 1994; Radeau, 1994b; Rosenblum, 1994). Closer examination shows, however, that the results do not in fact have this implication. In the two cited studies, bimodal adaptation was run with nonambiguous auditory adapters that, as the results from the unimodal auditory condition clearly showed, produced speech adaptation. The procedure was in fact exactly the one we considered in the introduction, and decided not to use, because the possible effect of visual recalibration would be confounded with a speech adaptation effect in the

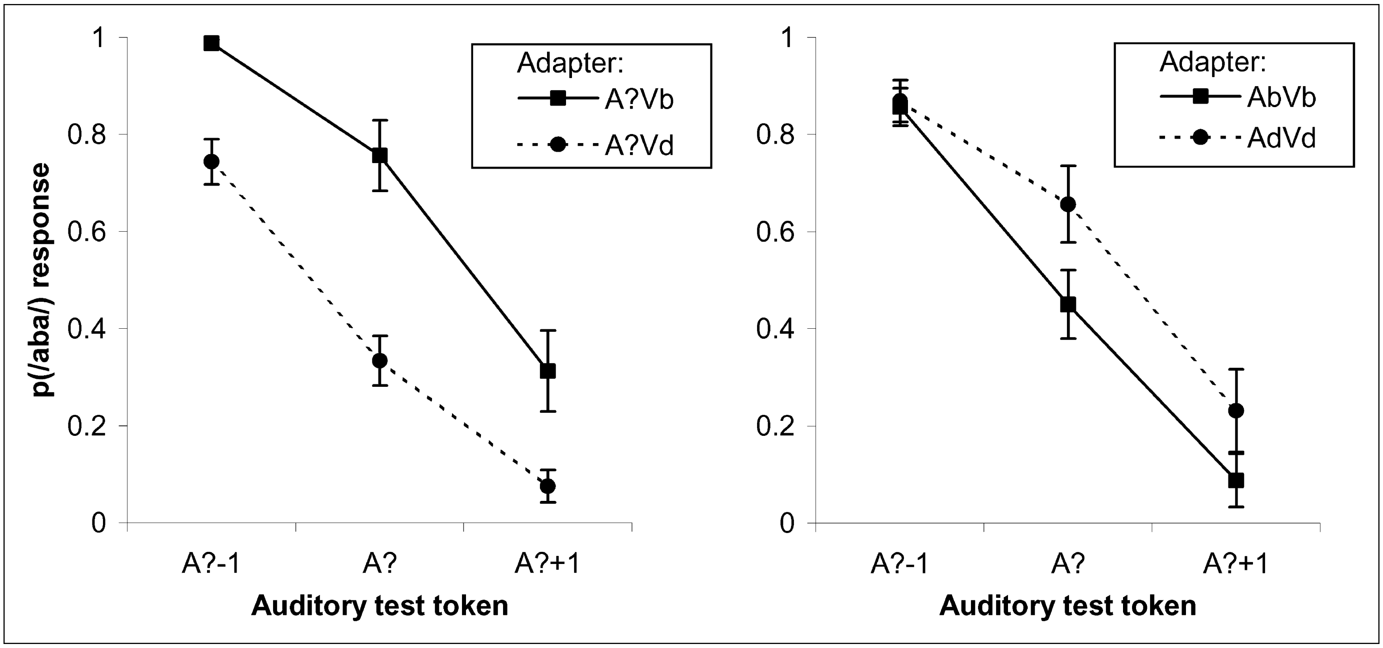

Fig. 3. Results of the auditory posttests in Experiment 2. The graphs show the proportion of /aba/ judgments after exposure to an adapter consisting of (left panel) the participant's ambiguous auditory token (A?) combined with either visual /aba/ (A?Vb adapter) or visual /ada/ (A?Vd adapter) or (right panel) a nonambiguous auditory token (Ab or Ad) combined with the congruent visual stimulus, /aba/ (AbVb adapter) or /ada/ (AdVd adapter). On each posttest, the presented auditory token was either the ambiguous token (A?) or one of its two neighbors on the auditory continuum (A? − 1 or A? + 1).

same direction. Under such conditions, the only available proof of recalibration would be a larger shift with bimodal than with unimodal auditory adapters. Now that the present study has demonstrated recalibration with a more powerful method, the remaining question is why, when combined with speech adaptation, recalibration did not result in a larger final shift.[2] There is no completely convincing answer to this question, but one possibility (to be tested in later research) is that same-direction effects produced by separate adaptation processes do not sum.

We suggest that the reason that recalibration was demonstrated in our study and not in earlier ones may be the degree of ambiguity of the auditory adapters that we used. Experiment 2 provides a direct test of this hypothesis.

EXPERIMENT 2: ADAPTATION WITH AMBIGUOUS AND NONAMBIGUOUS AUDITORY SPEECH

The purpose of Experiment 2 was to explore more directly the relations between visual recalibration of auditory speech and selective speech adaptation. We tested the assumption that the probability of observing either of these two phenomena depends on the degree of ambiguity of the auditory speech tokens used during exposure.

Method

Ten new participants from the same pool were tested in a single session. The procedure was the same as in Experiment 1, except for the introduction of two new exposure conditions, in which the same /aba/ and /ada/ visual components were combined with the corresponding end-point auditory tokens (condition AbVb: auditory /aba/ and visual /aba/; condition AdVd: auditory /ada/ and visual /ada/). Because of the nonambiguous speech components, these conditions were expected to produce speech adaptation, and because of the absence of cross-modal incongruence, to produce no recalibration. Eight adaptation blocks of each of these nonambiguous-auditory-token conditions were run, mixed in randomized order with eight blocks of each of the ambiguous-auditory-token conditions (A?Vb and A?Vd) used in Experiment 1. For these latter conditions, we made the opposite predictions—no selective speech adaptation and recalibration producing judgments consistent with the visual adapter. In all four conditions, postexposure auditory tests were run with the same three tokens centered on the participant's ambiguous token, as determined at the start of the session. Finally, the procedure for the audiovisual identification trials was the same as in Experiment 1, except that all four bimodal stimulus pairs used in the exposure trials were tested.

Results and Discussion

During exposure, the catch stimuli were detected on a large majority of presentations (97%). Posttest results appear in Figure 3. In the

2. Recently, Shigeno (2002) replicated these results when similarly using numerous exposure trials followed by test trials. But he obtained larger aftereffects with bimodal audiovisual than with unimodal auditory exposure when using a "paired-contrast" paradigm, in which each identification trial is preceded by a single presentation of an end-point stimulus. He called this result an "anchoring effect," considering it completely different from adaptation. As noted by Samuel (1986), this argument assumes that single presentations produce no fatigue.

ambiguous-auditory-token conditions (left panel), the proportion of /aba/ judgments was higher after exposure to visual /aba/ than after exposure to visual /ada/, just as in Experiment 1. In the nonambiguous conditions (right panel), the opposite pattern, fewer /aba/ judgments after exposure to visual /aba/ than after exposure to visual /ada/, was obtained. In a 2 (visual adapter) × 3 (auditory test token) × 2 (auditory-adapter ambiguity) ANOVA, the main effect of visual adapter, $F(1, 18) = 8.09$, $p < .02$, and the critical interaction between visual adapter and auditory-adapter ambiguity, $F(1, 18) = 51.1$, $p < .001$, were highly significant. Also significant were the main effects of visual adapter and of auditory test token and the three-way interaction of visual adapter, auditory test, and auditory ambiguity.

In the audiovisual identification trials, the two incongruent bimodal pairs with the ambiguous auditory token (A?Vb and A?Vd) were judged in accordance with the visual component on 98% of the trials; the congruent bimodal pairs with nonambiguous auditory tokens, AbVb and AdVd, were, unsurprisingly, judged respectively as /aba/ and /ada/ on 100% of the trials. Thus, exposure pairs with the same visual component but different, ambiguous versus nonambiguous, auditory components apparently gave rise to the same percepts.

As predicted, exposure to a particular visual speech token combined with the corresponding nonambiguous auditory token resulted in a reduced tendency to perceive that token (i.e., the typical selective speech adaptation effect). The same visual token combined instead with the ambiguous auditory token resulted, as in Experiment 1, in the opposite shift, more frequent perception of that token, indicative of cross-modal recalibration. Thus, dissociation between these two adaptation effects was obtained under otherwise identical conditions, just by manipulating the ambiguity of the auditory speech presented during adaptation. The fact that adapting stimulus pairs causing these opposite effects were identically categorized in the separate audiovisual identification trials makes it very unlikely that these effects reflected any postperceptual strategies.

GENERAL DISCUSSION

The results with the ambiguous auditory adapter, first obtained in Experiment 1, and then replicated in Experiment 2, show that auditory speech identification can be recalibrated on the basis of discordant lip-read visual speech. Thus, the occurrence of cross-modal recalibration is not, as proposed in earlier assessments of the evidence, an exclusive characteristic of space perception. The present results instead open the possibility that recalibration occurs in all situations that produce cross-modal biases. The generality of this conclusion is a matter for future research.

The results obtained in Experiment 2 with the congruent nonambiguous auditory adapters confirm that they cause speech adaptation, and to that extent support our assumption that earlier studies (Roberts & Summerfield, 1981; Saldaña & Rosenblum, 1994) did not yield proof of visual recalibration because of a confound with selective speech adaptation. We noted before that these studies should not be criticized for failing to produce that proof, because they were designed to address a different question. However, the reason why their bimodal conditions, which must have caused both recalibration and speech adaptation in the same direction, did not result in larger aftereffects than their control unimodal auditory conditions should be the focus of new investigations.

In Experiment 2, either visual recalibration or speech adaptation was obtained, depending uniquely on the degree of ambiguity of the auditory component of the adapting stimulus pairs. However, ambiguous and nonambiguous bimodal stimuli that contained the same visual stimulus were identically categorized in the separate forced-choice identification trials. The latter finding implies that the contrasting adaptation effects cannot have resulted from different response strategies, and must be perceptual.

Finally, the present investigation allowed the separation of two different adaptation phenomena that presumably respond to different underlying mechanisms. Identity recalibration, here demonstrated for the first time, is like the well-known cases of spatial recalibration; it is contingent on exposure to conflicting information from different sources and on some ensuing perceptual learning (Bedford, 1995). On the contrary, selective speech adaptation occurs in the absence of such conflict, and could, to some extent at least, reflect the fatigue of some of the processing mechanisms caused by repeated extreme stimulation. These two forms of adaptation co-occur also in other domains of perception. Conflict-based recalibrations have also been demonstrated, beyond the cases of cross-modal spatial conflict mentioned in the introduction, for analogous intramodal conflicts, such as between different cues to visual depth (see reviews by Epstein, 1975, and Wallach, 1968). Conflict-free adaptation is manifested in visual cases of sensory adaptation, such as color, curvature (Gibson, 1933), and size adaptation (Blakemore & Sutton, 1969), and also in motion aftereffects (Anstis, 1986), already mentioned in the introduction. The two types of phenomena have generally been the objects of separate lines of research, and their relations have rarely been investigated. In an early discussion, Bertelson and Radeau (1975) remarked that optical transformations used to create spatial conflicts between vision and other modalities often result in spatially biased inputs in at least one of the modalities, which may cause sensory adaptation. In the same vein, Welch (1986, pp. 24–26) has described promising ways of avoiding that sort of contamination, but his proposals have so far not been applied in concrete cases. The present study might be the first in which the two forms of adaptation have been dissociated within the same situation.

Acknowledgments—The present work was supported by the Belgian National Fund for Collective Fundamental Research (Contract 10759/2001 to Régine Kolinsky and Paul Bertelson). Thanks are due to Mirjam Keetels for conducting the testing. Constructive criticisms by Arthur Samuel, two anonymous reviewers, and the Editor, Sam Glucksberg, are gratefully acknowledged.

REFERENCES

Anstis, S. (1986). Motion perception in the frontal plane: Sensory aspects. In K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. I, chap. 16). New York: Wiley.

Bedford, F. (1995). Constraints on perceptual learning: Objects and dimensions. *Cognition*, *54*, 253–297.

Bermant, R.I., & Welch, R.B. (1976). The effect of degree of visual-auditory stimulus separation and eye position upon the spatial interaction of vision and audition. *Perceptual and Motor Skills*, *43*, 487–493.

Bertelson, P. (1999). Ventriloquism: A case of cross-modal perceptual grouping. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 347–362). Amsterdam: Elsevier.

Bertelson, P., & Aschersleben, G. (1998). Automatic visual bias of perceived auditory location. *Psychonomic Bulletin & Review*, *5*, 482–489.

Bertelson, P., & de Gelder, B. (in press). The psychology of multimodal perception. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention*. London: Oxford University Press.

Bertelson, P., & Radeau, M. (1975). Cognitive and perceptual factors in adaptation to auditory-visual conflict. In F. Vital-Durant & M. Jeannerod (Eds.), *Aspects of neural plasticity* (pp. 153–160). Paris: Editions INSERM.

Bertelson, P., & Radeau, M. (1976). Ventriloquism, sensory interaction and response bias: Remarks on the paper by Choe, Welch, Gilford and Juola. *Perception & Psychophysics*, *19*, 531–535.

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., Wiegeraad, G., & de Gelder, B. (1994). Exploring the relation between McGurk interference and ventriloquism. In *ICSLP 94* (Vol. 2, pp. 559–562). Yokohama: Acoustical Society of Japan.

Blakemore, C., & Sutton, P. (1969). Size adaptation: A new after-effect. *Science*, *166*, 245–257.

Boersma, P., & Weenink, D. (1999). Praat, a system for doing phonetics by computer. Retrieved from http://www.fon.hum.uva.nl/praat/

Choe, C.S., Welch, R.B., Gilford, R.M., & Juola, J.F. (1975). The “ventriloquist effect”: Visual dominance or response bias? *Perception & Psychophysics*, *18*, 55–60.

de Gelder, B., & Vroomen, J. (1994). A new place for modality in a modular mind. *Current Psychology of Cognition*, *13*, 84–91.

de Gelder, B., & Vroomen, J. (2000). Perceiving emotions by ear and by eye. *Cognition & Emotion*, *14*, 289–311.

Eimas, P.D., & Corbit, J.D. (1973). Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, *4*, 99–109.

Epstein, W. (1975). Recalibration by pairing: A process of perceptual learning. *Perception*, *4*, 59–72.

Gibson, J.J. (1933). Adaptation, after-effects and contrast in the perception of curved lines. *Journal of Experimental Psychology*, *18*, 1–31.

Held, R. (1965). Plasticity in sensory-motor systems. *Scientific American*, *213*, 84–94.

Jusczyk, P. (1986). Speech perception. In K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. II, chap. 27). New York: Wiley.

Klemm, O. (1909). Localisation von Sinneneindrücken bei disparaten Nebenreizen [Localization of sensory impressions with disparate distractors]. *Psychologische Studien (Wundt)*, *5*, 73–161.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.

Radeau, M. (1994a). Auditory-visual interaction and modularity. *Current Psychology of Cognition*, *13*, 3–51.

Radeau, M. (1994b). Ventriloquism against audio-visual speech: Or where Japanese-speaking barn owls might help? *Current Psychology of Cognition*, *13*, 124–140.

Radeau, M., & Bertelson, P. (1974). The after-effects of ventriloquism. *The Quarterly Journal of Experimental Psychology*, *26*, 63–71.

Radeau, M., & Bertelson, P. (1976). The effect of a textured visual field on modality dominance in a ventriloquism situation. *Perception & Psychophysics*, *20*, 227–235.

Recanzone, G. (1998). Rapidly induced auditory plasticity: The ventriloquism aftereffect. *Proceedings of the National Academy of Sciences, USA*, *95*, 869–875.

Roberts, M., & Summerfield, Q. (1981). Audiovisual presentation demonstrates that selective adaptation in speech perception is purely auditory. *Perception & Psychophysics*, *30*, 309–314.

Rosenblum, L D. (1994). How special is audiovisual speech integration? *Current Psychology of Cognition*, *13*, 110–116.

Saldaña, H.M., & Rosenblum, L.D. (1994). Visual influences on auditory pluck and bow judgements. *Journal of the Acoustical Society of America*, *95*, 3658–3661.

Samuel, A.G. (1986). Red herring detectors and speech perception: In defense of selective adaptation. *Cognitive Psychology*, *18*, 452–499.

Sawusch, J.R., & Pisoni, D.B. (1976). Response organization in selective adaptation to speech sounds. *Perception & Psychophysics*, *20*, 413–418.

Shigeno, S. (2002). Anchoring effects in audiovisual speech perception. *Journal of the Acoustical Society of America*, *111*, 2853–2861.

Vroomen, J. (1999). Ventriloquism and the nature of the unity decision (commentary on Welch). In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 389–394). Amsterdam: Elsevier.

Wallach, H. (1968). Informational discrepancy as a basis of perceptual adaptation. In S.J. Freeman (Ed.), *The neuropsychology of spatially oriented behaviour* (pp. 209–230). Homewood, IL: Dorsey.

Welch, R.B. (1986). Adaptation of space perception. In K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. I, chap. 24). New York: Wiley.

Welch, R.B. (1999). Meaning, attention, and the unity assumption in the intersensory bias of spatial and temporal perceptions. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 371–387). Amsterdam: Elsevier.

Welch, R.B., & Warren, D.H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*, 638–667.

(Received 8/22/02; Revision accepted 12/17/02)