Temporal adaptation to audiovisual asynchrony generalizes across different sound frequencies

Jordi Navarra[1]*, Joel García-Morera[1] and Charles Spence[2]

[1] *Fundació Sant Joan de Déu, Parc Sanitari Sant Joan de Déu – Hospital Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain*
[2] *Department of Experimental Psychology, University of Oxford, Oxford, UK*

The human brain exhibits a highly adaptive ability to reduce natural asynchronies between visual and auditory signals. Even though this mechanism robustly modulates the subsequent perception of sounds and visual stimuli, it is still unclear how such a temporal realignment is attained. In the present study, we investigated whether or not temporal adaptation generalizes across different auditory frequencies. In a first exposure phase, participants adapted to a fixed 220-ms audiovisual asynchrony or else to synchrony for 3 min. In a second phase, the participants performed simultaneity judgments (SJs) regarding pairs of audiovisual stimuli that were presented at different stimulus onset asynchronies (SOAs) and included either the same tone as in the exposure phase (a 250 Hz beep), another low-pitched beep (300 Hz), or a high-pitched beep (2500 Hz). Temporal realignment was always observed (when comparing SJ performance after exposure to asynchrony vs. synchrony), regardless of the frequency of the sound tested. This suggests that temporal recalibration influences the audiovisual perception of sounds in a frequency non-specific manner and may imply the participation of non-primary perceptual areas of the brain that are not constrained by certain physical features such as sound frequency.

Keywords: audiovisual asynchrony, adaptation, temporal recalibration, sound frequency, primary and non-primary sensory areas

INTRODUCTION

Audiovisual signals referring to the same external event often arrive asynchronously to their corresponding perceptual brain areas. This is both because different kinds of energy (e.g., light and sound waves) do not travel at the same velocity through the air (300,000,000 vs. 340 m/s, respectively), and also because the speed of neural transmission is different for vision and audition (see Spence and Squire, 2003; Schroeder and Foxe, 2004; King, 2005). As approximate synchrony is crucial for multisensory binding at the neuronal level (e.g., King and Palmer, 1985; Meredith et al., 1987), realigning asynchronous signals is typically required if one is to perceive the outside world coherently, and hence react appropriately in potentially dangerous situations. Multisensory neurons (e.g., in the superior colliculus of the cat) tend to increase their firing rate when the visual and the auditory signals are detected in the same spatial location at approximately the same time (see Stein and Stanford, 2008, for a review). There is, however, a temporal window (from approximately 50–80 ms when audition leads to 150–250 ms when vision leads, depending on the complexity of the stimulus) within which multisensory perception can be still clearly observed both behaviorally (Munhall et al., 1996; van Wassenhove et al., 2007; Vatakis and Spence, 2010; Vroomen and Keetels, 2010) and neurally (Stein and Meredith, 1993). The very existence of such a temporal window, and the fact that more asynchrony is tolerated when visual inputs lead auditory inputs (as when we perceive distant events) could easily be seen as an adaptative consequence of perceiving an intrinsically asynchronous external world.

Along similar lines, researchers have demonstrated that brief exposure (lasting no more that 3 min) to an asynchronous pairing of signals presented in different sensory modalities can have a profound effect upon our perception of subsequent incoming multisensory signals (Fujisaki et al., 2004; Vroomen et al., 2004). Thanks to this flexible process of temporal adaptation, the asynchrony experienced while seeing and hearing a tennis racket hitting the ball becomes, for someone sitting in the upper part of the stadium, less and less evident as the match goes on. Despite the enormous impact that this phenomenon has on the perception of both concurrent (Navarra et al., 2005) and subsequently presented sensory signals (Fujisaki et al., 2004; Vroomen et al., 2004; Heron et al., 2007; Vatakis et al., 2007; Hanson and Heron, 2008; Harrar and Harris, 2008), there is not much consensus regarding its underlying cognitive and neural mechanisms or its generalizability as a function of changes in stimulus attributes.

According to Roach et al. (2011), temporal adaptation is attained by populations of multisensory neurons that are tuned to specific asynchronies. Recalibration does not influence the processing of visual and auditory stimuli at a unimodal level. Despite of the fact that this theory represents a simple and parsimonious means of describing the mechanisms underlying temporal adaptation, it still needs to accommodate previous evidence indicating that adaptation to asynchrony modulates the speed of our behavioral responses to both unimodally presented sounds (Navarra et al., 2009) and visual stimuli (Di Luca et al., 2009). In line with these two studies, research using event-related potentials (ERPs)

has also demonstrated that the latency of early (unimodal) neural responses to sounds can be modulated by preceding visual information (e.g., we see the lips close more than 100 ms before hearing the sound/p/; van Wassenhove et al., 2005; Besle et al., 2008; see also Stekelenburg and Vroomen, 2007; Navarra et al., 2010). Taken together, this evidence suggests that the representation of the temporal relation between auditory and visual signals can, in fact, be modulated by multisensory experience.

Since visual information usually carries more accurate information regarding the time of occurrence of distant events (due to the higher velocity of transmission of light), it makes sense that the processing of sounds will, in certain circumstances, be accommodated to the visual timing of multisensory (that is, asynchronous) events appearing from afar (see Navarra et al., 2009). Even so, it is currently unclear whether the consequences of this temporal recalibration apply only to the adapted sounds (e.g., a 250-Hz tone) or whether instead they generalize to other non-adapted auditory stimuli (a 2500 or 300-Hz tone) that may appear following exposure to asynchrony[1]. Determining whether temporal adaptation to audiovisual asynchrony generalizes across different sound frequencies is by no means a trivial issue. Indeed, it could provide clues as to the possible neural mechanisms underlying this multisensory phenomenon. Generalization across different frequencies could be taken to implicate the involvement of brain areas (e.g., multisensory portions of the superior temporal sulcus, STS; see Beauchamp, 2005) that are, unlike other early and tonotopically organized sensory areas (Formisano et al., 2003; Humphries et al., 2010) able to compute information over different frequencies and not only over those frequencies that happen to be perceived at a given moment.

The transfer of adaptation effects to novel stimuli has been documented from one audiovisual display (e.g., a light flash and a beep) to another (e.g., two visual stimuli moving toward each other and then colliding with a beep; see Fujisaki et al., 2004). However, it is possible that such effects may simply reflect the use of exactly the same sound during both exposure to asynchrony and during the test phase (where the effects of adaptation were measured). Therefore, in the present study, we investigated whether or not adaptation to audiovisual asynchrony occurs exclusively for the adapted sound. More specifically, we investigated whether or not temporal adaptation to audiovisual combinations of stimuli including a specific 250 Hz tone not only modulates the perception of audiovisual stimulus pairs that include the same tone (250 Hz), but also the multisensory perception of other low (300 Hz) or high (2500 Hz) frequency tones[2].

---

[1] Note that by investigating the possible generalization of temporal recalibration to different frequency ranges we are not denying (1) that temporal recalibration may also influence the perception of visual stimuli (as in Di Luca et al., 2009); and (2) that the mechanisms of temporal adaptation could be achieved in multisensory (rather than unisensory) brain areas, as suggested by Roach et al. (2011).

[2] The auditory stimuli used in the present study were chosen on the basis of previous research (Formisano et al., 2003; Humphries et al., 2010). Note that all of the sounds used in the present study were equated in terms of their physical intensity, but not in terms of their 'perceived' intensity (see Methods section). The fact that sounds were probably perceived differently not only in terms of their frequency, but also in terms of their perceived intensity, gave us the opportunity to test the generalization in a more 'extreme' situation than might otherwise have been the case.
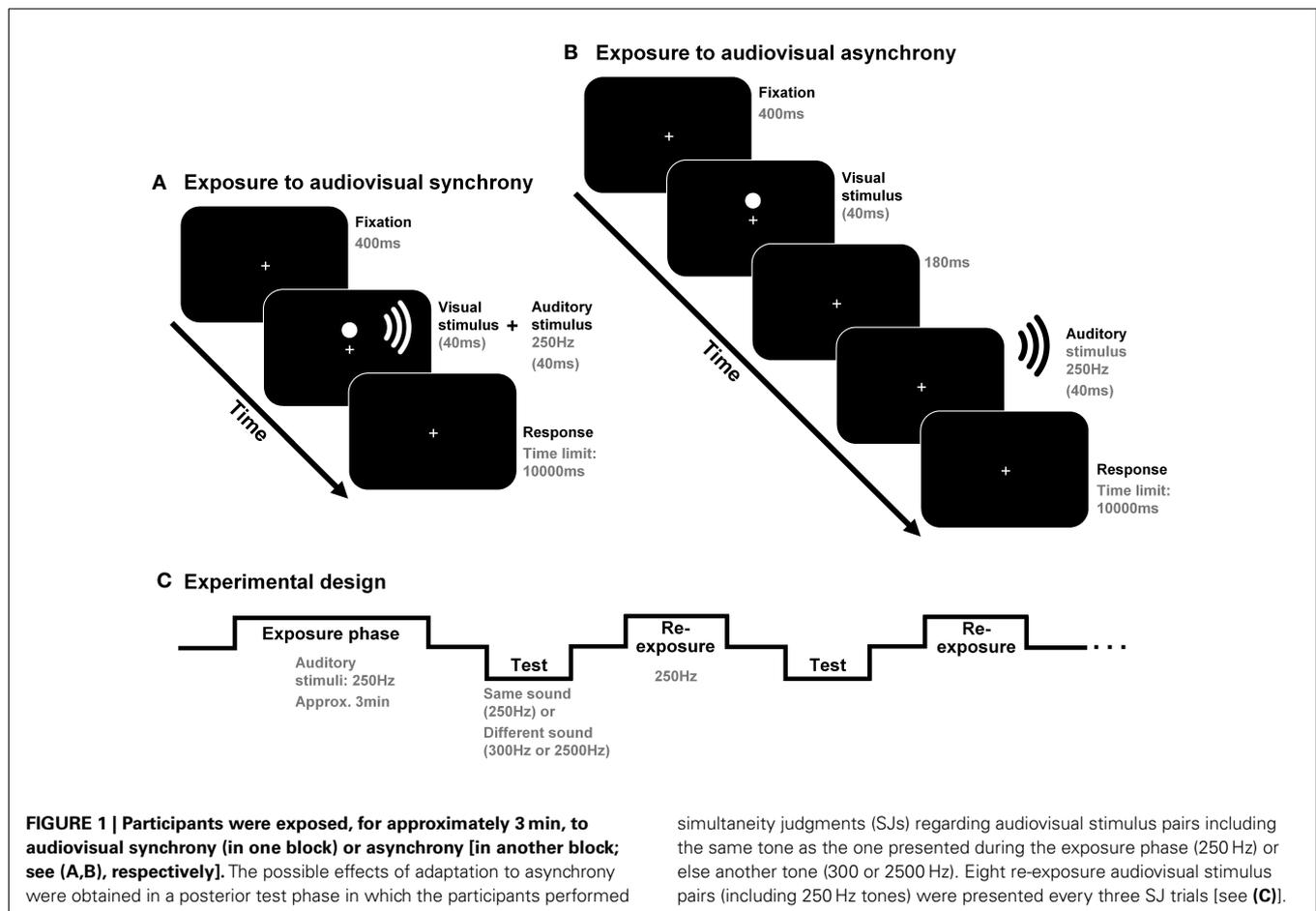
MATERIALS AND METHODS
PARTICIPANTS
Seven participants (four female; mean age: 23 years) took part in the study. All were naive as to the purpose of the study, reported normal hearing and normal or corrected-to-normal vision, and received 15 Euros in return for their participation. The experiment was non-invasive, was conducted in accordance with the Declaration of Helsinki, and had ethical approval from Hospital Sant Joan de Déu (Barcelona, Spain).

STIMULI AND PROCEDURE
The study was conducted in two sessions (each lasting about 45 min); one designed to test for generalization effects from 250 to 2500 Hz (that is, from one frequency range to another) and the other to test for generalization effects from 250 to 300 Hz (both falling in the same frequency range). Each session contained two different blocks, both starting with a 3-min exposure phase where pairs of visual and auditory stimuli were presented simultaneously (in the "synchrony" block) or with the tone lagging by a fixed time interval of 220 ms (in the "asynchrony" block; see **Figure 1**). In a subsequent test phase, the participants performed a simultaneity judgment (SJ) task regarding auditory and visual stimuli presented at nine different stimulus onset asynchronies (SOAs) using the method of constant stimuli (± 330, ± 190, ± 105, ± 45, or 0 ms; negative values indicate that the sound was presented first). Each SOA appeared 12 times, for each sound and condition (i.e., synchrony and asynchrony), during the experimental session. Eight re-exposure/top-up audiovisual pairs of stimuli (identical to those presented in the previous adaptation phase) were presented following every three SJ trials (see **Figure 1**). SJs after exposure to synchrony and asynchrony were compared in order to analyze possible recalibration effects.

The auditory stimuli (presented in a dark sound-attenuated booth during both the exposure and test phases) lasted for 40 ms and were presented at 80 dB(A), as measured from the participant's head position (located approximately 50 cm in front of the computer monitor). The onsets and offsets of all auditory stimuli had a 5-ms ramp (from 0 to 100%, and from 100 to 0% sound intensity, respectively). The sounds were presented via two loudspeaker cones (Altec Lansing V52420, China), one placed on either side of the computer monitor. Low (250 Hz) tones were presented during the exposure phase in all of the experimental sessions. Either the same 250 Hz tone or a different tone appeared randomly in the SJ trials (along with the visual stimulus) in the test phase. The non-exposed sounds (i.e., tones that were different from the 250-Hz tone presented during the adaptation phase) consisted of a 2500-Hz tone, and a 300-Hz tone. The order of presentation of the sessions (each one including just one of the two different non-exposed sounds in the test) was counterbalanced across participants. The visual stimulus (kept constant during the experiment) consisted of a red disk (1.5 cm in diameter), always appearing, for 40 ms, 1 cm above a continuously and centrally presented fixation point on a CRT monitor screen (Hyundai Q770, South Korea; 17″; refresh rate = 75 Hz). The presentation of all the stimuli was controlled by a program running on DMDX (see Forster and Forster, 2003).

FIGURE 1 | Participants were exposed, for approximately 3 min, to audiovisual synchrony (in one block) or asynchrony [in another block; see (A,B), respectively]. The possible effects of adaptation to asynchrony were obtained in a posterior test phase in which the participants performed simultaneity judgments (SJs) regarding audiovisual stimulus pairs including the same tone as the one presented during the exposure phase (250 Hz) or else another tone (300 or 2500 Hz). Eight re-exposure audiovisual stimulus pairs (including 250 Hz tones) were presented every three SJ trials [see **(C)**].

In order to ensure that participants attended to the visual stimuli during the exposure phase, they were instructed to press a key on the computer keyboard whenever they detected a smaller red disk that appeared with a probability of 10%. Before the start of the main experiment, the participants received a brief practice session containing a 1-min version of the exposure-to-synchrony phase and 60 SJ trials regarding audiovisual stimuli that were identical to those presented in the exposure phase and which were presented at SOAs of ±330, ±190, and 0 ms. As in the main experiment, eight top-up stimulus pairs were presented after every three SJs during this training phase.

RESULTS

Participants' SJ data were fitted, after passing a Shapiro–Wilk normality test for both synchrony and asynchrony conditions (Shapiro and Wilk, 1965), into a three-parameter Gaussian function (mean correlation coefficient between observed and fitted data = 0.96). This provided two different indicators of participants' temporal perception: the point of subjective simultaneity (PSS), indicating the SOA at which the visual and auditory stimuli were most plausibly perceived as being simultaneous, and the standard deviation (SD), which provides a measure of the participants' sensitivity to objective synchrony.

A temporal shift effect was observed in the PSS in the direction of the pre-exposed asynchrony both for audiovisual combinations including the adapted (33 ms on average)[3] and the non-adapted tones (15 ms for the 2500-Hz tone; and 31 ms for the 300-Hz tone; see **Figure 2**). A non-parametric Wilcoxon signed-rank test, comparing the PSS after exposure to synchrony and asynchrony, showed that this effect was significant for all the tested sound frequencies ($Z = -2.37$, $p = 0.018$; $Z = -2.37$, $p = 0.018$; and $Z = -2.2$, $p = 0.028$, respectively). Further analyses revealed that while the difference in PSS after exposure to synchrony and asynchrony was statistically equivalent for the adapted test tone and the 300 Hz test tone (Wilcoxon signed-rank test: $Z = -0.17$, $p = 0.87$), there was a difference between the adapted (250 Hz) and the 2500 Hz-tone conditions (Wilcoxon signed-rank test: $Z = -2.4$, $p = 0.018$). In other words, even when the temporal shift was consistently observed in all participants for audiovisual pairs including a 2500 Hz test tone, this effect was significantly smaller in this condition than when the same tone was used in both the exposure and the test phases.

The SD widened significantly after exposure to asynchrony for the same (250 Hz; Wilcoxon signed-rank tests: $Z = -2.03$,

---

[3]In order to equate the SJ tests across the two experimental sessions, they both included 250-Hz beeps (as well as exemplars of just one of the non-adapted tones). As the SJ test was performed twice for the 250-Hz beep condition, all of the analyses including data from this condition were performed over the average (mean) between the two PSS (or SD) values obtained in each of the two sessions separately.
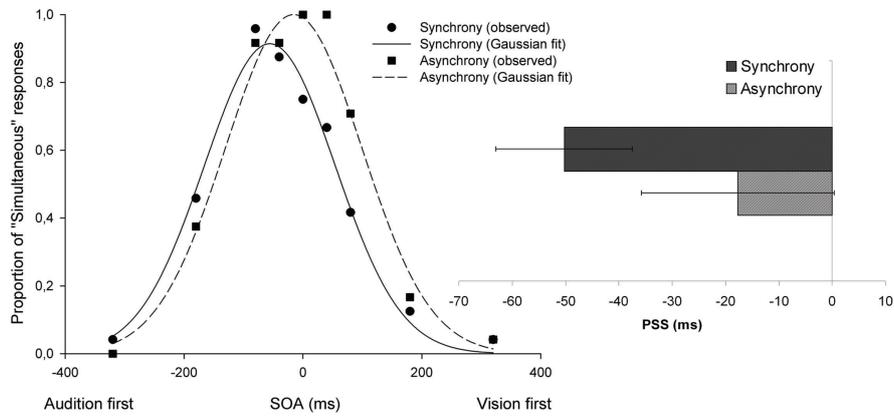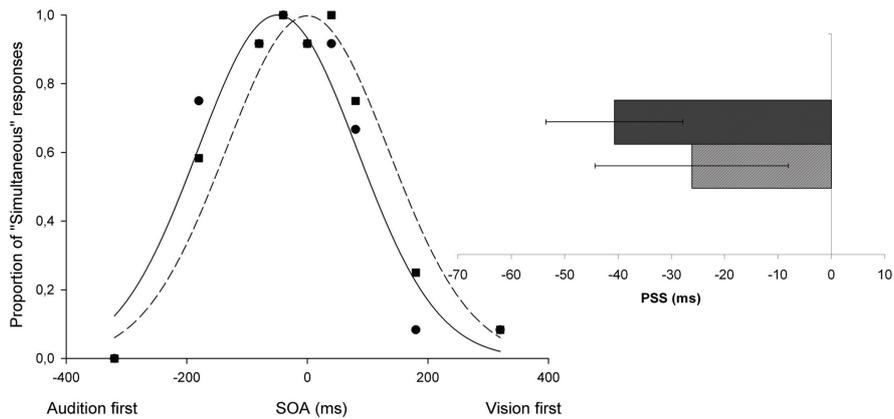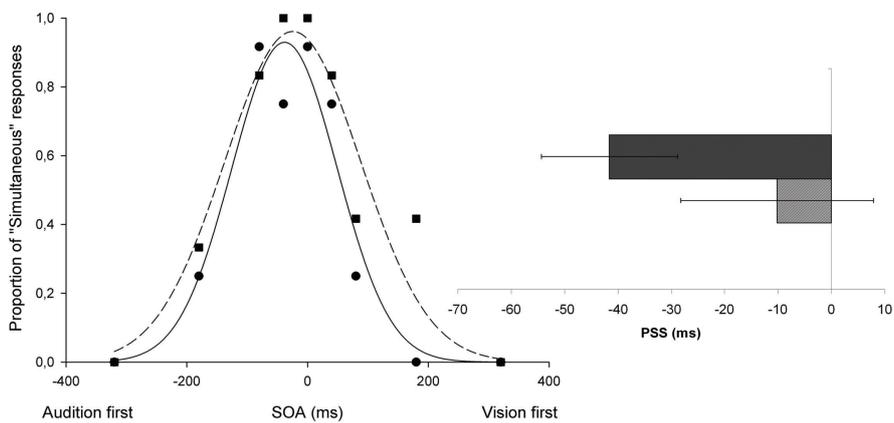
FIGURE 2 | Temporal recalibration effects seen in audiovisual SJs including beeps of 250, 2500, and 300 Hz [shown in (A–C), respectively]. The proportion of "simultaneous" responses across the different SOAs was fitted with a Gaussian function. The graph shows the observed and fitted data from one of the participants in the experiment. The bar graph shows the point of subjective simultaneity (PSS) group average for synchrony and asynchrony conditions, including data (mean and standard error) from all of the participants. When audiovisual stimuli were presented asynchronously during the exposure phase, a temporal shift was observed (see dashed lines), in the PSS, toward the direction of the asynchrony (vision-first), and regardless of the sound being tested. The sensitivity to physical synchrony decreased (i.e., the SD increased) as a result of exposure to asynchrony.

$p = 0.043$) and the 300 Hz tones ($Z = -2.37$, $p = 0.018$). This effect was marginally significant for the 2500 Hz tone ($Z = -1.69$, $p = 0.091$).

DISCUSSION

The results of the present study help to characterize the mechanisms underlying audiovisual temporal realignment – a basic but, as yet, not well characterized neural function. In line with the results of previous research (Fujisaki et al., 2004; Vroomen et al., 2004), we demonstrate that exposure to visual-leading asynchronies (i.e., the kind of asynchrony that perceivers typically experience in everyday life) leads to temporal realignment between visual and auditory signals. Our results go further by demonstrating that this temporal readjustment between sensory signals appears to act not just upon audiovisual stimuli including sounds of the same frequency as the adapted sound, but also upon audiovisual stimulus pairs that include other sounds, regardless of how different in frequency they are with respect to the adapted sound.

When taken together with previous studies, our results suggest that temporal recalibration is a malleable mechanism that is highly dependent on the perceptual characteristics of each situation: in conditions where attention is driven toward a limited number of visual and auditory stimuli (as in the present study), temporal recalibration appears to spread over different frequency ranges. However, the fact that recalibration can be simultaneously achieved in both vision-first and audition-first directions for different audiovisual stimuli (Hanson et al., 2008; Roseboom and Arnold, 2011) may suggest that this mechanism can modulate the perception of different audiovisual events (including different sounds) occurring at different locations. Further research will be needed in order to investigate whether the same frequencies can be realigned in different locations simultaneously or not.

With visual and auditory stimuli appearing from the front[4], we have previously demonstrated that recalibration can modulate the speeded detection of sounds (but not the detection of visual stimuli; see Navarra et al., 2009). As suggested in the Introduction, it makes sense for the processing of sounds to be modulated during adaptation to asynchrony, because vision gives us more precise information regarding the exact instant at which an audiovisually perceived distal event occurred (note, once again, that light travels at 300,000,000 m/s and sound at 340 m/s). Consequently, there is no reason to believe that the effects found in the present study were exclusively based on a purely "visual" temporal shift that influenced all possible audiovisual combinations (see Di Luca et al., 2009). Previous studies suggesting that temporal recalibration takes place at a multisensory (as opposed to unisensory) level (Roach et al., 2011) that is influenced by other processes (e.g., spatial grouping; Yarrow et al., 2011) would also contradict this idea. It is, however, necessary to further investigate whether the temporal recalibration

really implies an adjustment of the speed of neural transmission of auditory (or visual) signals or not.

In line with Roach et al.'s (2011) study, our results provide evidence consistent with the idea that non-primary sensory areas of the brain (e.g., the STS) that are less dependent of the physical attributes (e.g., frequency, intensity) of the perceived stimulus would be implicated in reducing the perceived asynchrony between visual and auditory signals. The degree of involvement of this area/s also needs to be clarified. It could, for example, serve as an "asynchrony detector" influencing earlier, and perhaps more "unisensory," sensory processes by means of a feedback system. This hypothesis fits well with previously reported effects of temporal recalibration on subsequent RTs to stimuli presented unimodally (Di Luca et al., 2009; Navarra et al., 2009). There is, in fact, some evidence to suggest that many multisensory processes rely on feedback circuitry (e.g., Driver and Noesselt, 2008; Noesselt et al., 2010). Another hypothesis, however, may imply that recalibration itself takes place in this/these area/s, without percolating through the visual and auditory systems (Roach et al., 2011). Mounting evidence suggesting the existence of early multisensory interaction (Molholm et al., 2002; Cappe et al., 2010; Raij et al., 2010; Senkowski et al., 2011; Van der Burg et al., 2011) and feedback connections between visual, auditory, and integrative areas (see Driver and Noesselt, 2008) may, however, contradict this "recalibration-in-one-site" view of temporal recalibration (and perhaps multisensory integration).

A tentative "biological" explanation for the non-specific recalibration effect found in the present study may be that the perceptual system takes into account the possibility that one distant visual object is often linked to more than a single sound or sound frequency. One could perhaps think of a gazelle gazing at a distant lion: the lion can, in fact, produce different sounds (everything from branches snapping underfoot to the lion roaring, etc.). Therefore, it makes sense that all the possible sounds that are linked to the sight of the lion would be influenced by asynchrony adaptation, rather than necessarily just the sound of its roar.

Here, it is worth noting that the sounds that one encounters in real life tend to be much more complex than the auditory stimuli typically used in laboratory studies, normally containing different frequencies at the same time (e.g., speech, noise . . .). It is plausible, therefore, that adaptation to asynchronous audiovisual signals that include a specific tone would have some impact on the subsequent multisensory perception of other tones with which it ordinarily co-occurs. The results reported here (i.e., revealing diffuse/general, rather than stimulus-specific, perceptual mechanisms) are also in line with previous research showing stimulus-non-specific effects in the spatial ventriloquism effect (Frissen et al., 2003, 2005; but see Recanzone, 1998). Although this converging evidence may be purely coincidental, it could also support the idea that some spatial and temporal mechanisms underlying multisensory integration are governed by similar basic principles.

The results of the present study also demonstrate that observers need more time between the visual and the auditory stimuli (especially in the same frequency range) in order to "segregate" them after temporal resynchronization. Similar effects have already been reported in several previous studies (Navarra et al., 2005, 2007),

---

[4]Since temporal adaptation mainly results from (frontally perceived) physical distance between the audiovisual event and the perceiver, we do not use headphones in temporal adaptation experiments (see Zampini et al., 2003, for arguments against the use of headphones in experiments on audiovisual integration).

and could result from a mechanism that, acting in parallel with temporal realignment, stretches the temporal window of audio-visual integration. A possible (and, at this point, speculative) explanation for the appearance of this phenomenon may be that this temporal window broadens whenever asynchrony is perceived between vision and audition, as a general/non-specific (and perhaps provisional) way to predict and facilitate integration in certain circumstances (e.g., when multiple multisensory events are perceived distantly).

REFERENCES

Beauchamp, M. S. (2005). See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. *Curr. Opin. Neurobiol.* 15, 145–153.

Besle, J., Fischer, C., Bidet-Caulet, A., Lecaignard, F., Bertrand, O., and Giard, M. H. (2008). Visual activation and audiovisual interactions in the auditory cortex during speech perception: intracranial recordings in humans. *J. Neurosci.* 28, 14301–14310.

Cappe, C., Thut, G., Romei, V., and Murray, M. M. (2010). Auditory-visual multisensory interactions in humans: timing, topography, directionality, and sources. *J. Neurosci.* 30, 12572–12580.

Di Luca, M., Machulla, T.-K., and Ernst, M. O. (2009). Recalibration of multisensory simultaneity: Cross-modal transfer coincides with a change in perceptual latency. *J. Vis.* 9, 1–16.

Driver, J., and Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on "sensory-specific" brain regions, neural responses, and judgments. *Neuron* 57, 11–23.

Formisano, E., Kim, D. S., Di Salle, F., van de Moortele, P. F., Ugurbil, K., and Goebel, R. (2003). Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron* 40, 859–869.

Forster, K. I., and Forster, J. C. (2003). DMDX: a windows display program with millisecond accuracy. *Behav. Res. Methods Instrum. Comput.* 35, 116–124.

Frissen, I., Vroomen, J., de Gelder, B., and Bertelson, P. (2003). The after-effects of ventriloquism: are they sound-frequency specific? *Acta Psychol. (Amst)* 113, 315–327.

Frissen, I., Vroomen, J., de Gelder, B., and Bertelson, P. (2005). The after-effects of ventriloquism: generalization across sound-frequencies. *Acta Psychol. (Amst)* 118, 93–100.

Fujisaki, W., Shimojo, S., Kashino, M., and Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nat. Neurosci.* 7, 773–778.

Hanson, V. M., and Heron, J. (2008). Recalibration of perceived time across sensory modalities. *Exp. Brain Res.* 185, 347–352.

Hanson, V. M., Roach, N. W., Heron, J., and McGraw, P. V. (2008). Spatially specific distortions of perceived simultaneity following adaptation to audiovisual asynchrony. *Poster presented at ECVP*, Utrecht, The Netherlands.

Harrar, V., and Harris, L. R. (2008). The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Exp. Brain Res.* 186, 517–524.

Heron, J., Whitaker, D., McGraw, P. V., and Horoshenkov, K. V. (2007). Adaptation minimizes distance-related audiovisual delays. *J. Vis.* 7, 1–8.

Humphries, C., Liebenthal, E., and Binder, J. R. (2010). Tonotopic organization of human auditory cortex. *Neuroimage* 50, 1202–1211.

King, A. J. (2005). Multisensory integration: strategies for synchronization. *Curr. Biol.* 15, R339–R341.

King, A. J., and Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp. Brain Res.* 60, 492–500.

Meredith, M. A., Nemitz, J. W., and Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J. Neurosci.* 7, 3215–3229.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., and Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res. Cogn. Brain Res.* 14, 115–128.

Munhall, K. G., Gribble, P., Sacco, L., and Ward, M. (1996). Temporal constraints on the McGurk effect. *Percept. Psychophys.* 58, 351–362.

Navarra, J., Alsius, A., Velasco, I., Soto-Faraco, S., and Spence, C. (2010). Perception of audiovisual speech synchrony for native and non-native language. *Brain Res.* 1323, 84–93.

Navarra, J., Hartcher-O'Brien, J., Piazza, E., and Spence, C. (2009). Adaptation to audiovisual asynchrony modulates the speeded detection of sound. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9169–9173.

Navarra, J., Soto-Faraco, S., and Spence, C. (2007). Adaptation to audiotactile asynchrony. *Neurosci. Lett.* 413, 72–76.

Navarra, J., Vatakis, A., Zampini, M., Soto-Faraco, S., Humphreys, W., and Spence, C. (2005). Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Brain Res. Cogn. Brain Res.* 25, 499–507.

Noesselt, T., Tyll, S., Boehler, C. N., Budinger, E., Hinze, H.-J., and Driver, J. (2010). Sound-induced enhancement of low-intensity vision: Multisensory influences on human sensory-specific cortices and thalamic bodies relate to perceptual enhancement of visual detection sensitivity. *J. Neurosci.* 30, 13609–13623.

Raij, T., Ahveninen, J., Lin, F. H., Witzel, T., Jääskeläinen, I. P., Letham, B., Israeli, E., Sahyoun, C., Vasios, C., Stufflebeam, S., Hämäläinen, M., and Belliveau, J. W. (2010). Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *Eur. J. Neurosci.* 31, 1772–1782.

Recanzone, G. (1998). Rapidly induced auditory plasticity: the ventriloquism aftereffect. *Proc. Natl. Acad. Sci. U.S.A.* 95, 869–875.

Roach, N. W., Heron, J., Whitaker, D., and McGraw, P. V. (2011). Asynchrony adaptation reveals neural population code for audiovisual timing. *Proc. Biol. Sci.* 278, 1314–1322.

Roseboom, W., and Arnold, D. H. (2011). Twice upon a time: multiple concurrent temporal recalibrations of audiovisual speech. *Psychol. Sci.* 22, 872–877.

Schroeder, C. E., and Foxe, J. J. (2004). "Multisensory convergence in early cortical processing," in *The Handbook of Multisensory Processes*, eds G. A. Calvert, C. Spence, and B. E. Stein (Cambridge, MA: MIT Press), 295–309.

Senkowski, D., Saint-Amour, D., Höfle, M., and Foxe, J. J. (2011). Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *Neuroimage* 56, 2200–2208.

Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* 52, 591–611.

Spence, C., and Squire, S. (2003). Multisensory integration: Maintaining the perception of synchrony. *Curr. Biol.* 13, 519–521.

Stein, B. E., and Meredith, M. A. (1993). *The Merging of the Senses.* Cambridge, MA: MIT Press.

Stein, B. E., and Stanford, T. R. (2008). Multisensory integration: Current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* 9, 255–267.

Stekelenburg, J. J., and Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *J. Cogn. Neurosci.* 19, 1964–1973.

Van der Burg, E., Talsma, D., Olivers, C. N., Hickey, C., and Theeuwes, J. (2011). Early multisensory interactions affect the competition among multiple visual objects. *Neuroimage* 55, 1208–1218.

van Wassenhove, V., Grant, K. W., and Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proc. Natl. Acad. Sci. U.S.A.* 102, 1181–1186.

van Wassenhove, V., Grant, K. W., and Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia* 45, 598–607.

Vatakis, A., Navarra, J., Soto-Faraco, S., and Spence, C. (2007). Temporal recalibration during audiovisual speech perception. *Exp. Brain Res.* 181, 173–181.

Vatakis, A., and Spence, C. (2010). "Audiovisual temporal integration for complex speech, object-action, animal call, and musical stimuli," in *Multisensory Object Perception in the Primate Brain*, eds. M. J. Naumer, and J. Kaiser (New York, NY: Springer), 95–120.

Vroomen, J., and Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Atten. Percept. Psychophys.* 72, 871–884.

Vroomen, J., Keetels, M., de Gelder, B., and Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Brain Res. Cogn. Brain Res.* 22, 32–35.

Yarrow, K., Roseboom, W., and Arnold, D. H. (2011). Spatial grouping resolves ambiguity to drive temporal recalibration. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1657–1661.

Zampini, M., Shore, D. I., and Spence, C. (2003). Audiovisual temporal order judgments. *Exp. Brain Res.* 152, 198–210.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.