The Neural Correlates and Mechanisms Mediating the Integration of Auditory and Visual Information in the Human Brain.

Dissertation

zur Erlangung des Grades eines
Doktors der Naturwissenschaften

der Fakultät für Biologie
und
der Medizinischen Fakultät
der Eberhard-Karls-Universität Tübingen

vorgelegt
von

Sebastian Werner
aus Halle/Saale, Deutschland

Februar 2010

Tag der mündlichen Prüfung:	15. September 2010
Dekan der Fakultät für Biologie:	Prof. Dr. H. A. Mallot
Dekan der Medizinischen Fakultät:	Prof. Dr. I. B. Autenrieth
1. Berichterstatter:	Prof. Dr. Heinrich H. Bülthoff (Dr. med. Uta Noppeney, Ph.D.)
2. Berichterstatter:	Prof. Dr. Wolfgang Grodd
3. Berichterstatter:	PD Dr. Toemme Noesselt
Prüfungskommission:	Prof. Dr. Rolf Ulrich Prof. Dr. Wolfgang Grodd Prof. Dr. Dirk Wildgruber PD Dr. Toemme Noesselt Dr. med. Uta Noppeney, Ph.D.

Erklärung

Ich erkläre, dass ich die zur Promotion eingereichte Arbeit mit dem Titel:

„The Neural Correlates and Mechanisms Mediating the Integration of Auditory and Visual Information in the Human Brain.“

selbstständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen als solche gekennzeichnet habe. Ich versichere an Eides statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Mir ist bekannt, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

Tübingen, den 05.02.2010

(Unterschrift)

Table of contents

Abstract

To respond more quickly to events in natural environments the human brain merges information from multiple senses into a more reliable percept. Multisensory integration processes have been demonstrated in a distributed neural system encompassing sensory-specific, higher association and prefrontal cortices. Using fMRI and psychophysical methods this dissertation investigates the functional similarities, differences and constraints that govern the integration of auditory and visual information in different regions of the human cerebral cortex. Characterizing their temporal response codes, effective connectivity patterns and underlying computations for combining multisensory inputs, this work provides evidence for the integration of specific types of information at 3 functionally specialized processing stages. At the first stage, multisensory interactions in sensory-specific regions indicate the same sensory source by integrating spatiotemporally aligned auditory and visual inputs to enhance stimulus detection. At the second stage, multisensory interactions in higher association regions integrate complex environmental features into higher order representations, forming a unified percept and mediating multisensory benefits in object recognition. At the third stage, multisensory interactions in the prefrontal cortex mediate response selection processes based on perceptual information from auditory and visual modalities with multisensory facilitations of reaction time. This dissertation constitutes the first systematic attempt to dissociate the contributions of sensory-specific, higher association and prefrontal areas to audiovisual integration in the human brain.

Synopsis

Background

Most objects and events can be detected by more than one sensory system. Thus, to form a coherent percept and enable effective interactions with the environment the brain needs to combine information from different senses. The integration of inputs from multiple sensory modalities provides ubiquitous behavioral advantages over unisensory situations, allowing more accurate perceptual judgments and faster responses for the detection, discrimination or categorization of stimuli (Calvert et al., 2004). The first neuroscientific advances to understand this phenomenon arose from cellular recordings in the superior colliculcus (SC) that contains neurons responsive to visual, auditory and tactile inputs (Meredith and Stein, 1983). These early electrophysiological investigations demonstrated that multisensory interactions in SC are governed by tight spatial and temporal principles with non-linear (i.e. super-additive) response enhancements signaling the co-localization and co-incidence of multisensory stimuli. These findings demonstrated the intuitive notion that the brain should only integrate inputs that indeed originate from the same object or event while representing independent sensory sources separately (Stein and Stanford, 2008). Another influential principle of integration postulated the inverse effectiveness of sensory inputs. It was based on the observation that non-linear multisensory enhancements were maximal when the responses to individual sensory cues were weakest. This principle was also intuitive, since integration should be especially beneficial when no sense by itself can disambiguate the sensory scene but a combination of senses will (Stanford and Stein, 2007). These principles of integration guided researches for more than two decades in characterizing multisensory neurons not only in the SC but also across the cerebral cortex. Cortical neurons integrating inputs from different sensory modalities were mainly observed at the transitions zones between sensory-specific regions denoting the higher association regions (Wallace et al., 2004). These findings were consistent with prevalent hierarchical models of sensory processing (Felleman and van Essen, 1991;Rauschecker et al., 1997), proposing that information undergo successive processing stages with increasingly larger spatial and/or temporal receptive fields in their unisensory pathways (Hubel, 1988;Rolls, 2000;Griffiths et al., 1998;Hasson et al., 2008), before they are combined in higher order areas such as the superior temporal sulcus (Calvert, 2001). However, with the advent of functional magnetic resonance imaging (fMRI) it became clear that multisensory processes were not only confined to subcortical structures or higher association and prefrontal areas, but were also present in lower visual and auditory cortices (Calvert et al., 1997;Amedi et al., 2005;Driver and Noesselt, 2008). The multisensory influence on sensory-specific areas was traditionally believed to be purely mediated by feedback inputs from association regions that arrive after multisensory signals converged at a late stage of processing (Driver and Spence, 2000;Calvert, 2001). Although clear evidence for those feedback influences has been demonstrated (Schroeder and Foxe, 2002), the exclusiveness of this view was challenged by human electroencephalography (EEG) studies showing surprisingly early audiovisual interactions over visual cortex (Giard and Peronnet, 1999;Molholm et al., 2002). Moreover, electrophysiological recordings in monkeys actually demonstrated multisensory interactions in auditory cortex with clear feedforward activation profiles (Schroeder and Foxe, 2005;Foxe and Schroeder, 2005). Hence, the multisensory influences on low-level sensory areas seem to be not only dependent on feedback processes from association regions but might additionally rely on feedforward (thalamocortical) or lateral connectivity between sensory cortices (Schroeder et al., 2003). The ground breaking discovery of feedforward multisensory interactions in auditory cortex not only started a new era in multisensory science (Foxe, 2008) but also questioned the functional specificity of early sensory regions, leading to the provocative declaration of the entire neocortex as “multisensory” (Ghazanfar and Schroeder, 2006).

Objective

Although multisensory interactions seem to be virtually present on all levels of sensory processing, the functional relevance of these interactions is unclear. This dissertation tries to characterize the functional similarities, differences and constraints that govern the integration of auditory and visual information in different regions of the human cerebral cortex. Using fMRI and psychophysical methods this work specifically attempts to dissociate the contributions of sensory-specific, higher association and prefrontal areas to audiovisual integration while characterizing their temporal response codes and underlying computations for combining multisensory inputs.

Hypothesis

The general framework of this dissertation encompasses a set of hypotheses guiding the experimental approaches and analysis methods. In this simplistic framework, behavioral responses to audiovisual objects and events are hypothesized to require the integration of specific types of information at 3 distinct processing stages, which are functionally specialized to represent and combine that information.

At the first stage, the brain detects the co-incidence and co-localization of auditory and visual signals to recognize that these inputs originate from the same source. Regions mediating this function require narrow spatiotemporal receptive windows to ensure precise timing and spatial processing, but limiting the possibility for integrating complex sensory details.

At the second stage, higher order perceptual information is combined resulting in implicit recognition of audiovisual object and events with enhanced perceptual accuracy. Regions mediating this function need to be able to represent complex visual and auditory features (e.g. visual form and motion; auditory frequency spectrum and envelope), with wider spatiotemporal receptive windows that allow the integration of slower sensory signals at more flexible spatial scales. According to the basic organizational principles of sensory systems, slow and spatially extended signals are understood as more complex since their representation requires the convergence of information from lower regions that process inputs at a more restricted spatial dimension and lower temporal scale.

At the third stage, information is integrated at a decisional level where perceptual information guides the selection of an appropriate action in response to audiovisual object and events with facilitated reaction times. Regions mediating this function must be able to accumulate higher order auditory and visual information until sufficient evidence for response execution is obtained. Given the dependency on perceptual information and the relation to responsiveness, the third stage is best understood in establishing a mapping between perceptual space and response alternatives, i.e. representing multisensory features in motor coordinates.

Overview

To characterize audiovisual integration effects at these distinct processing stages, 4 studies were conducted. These studies manipulated for instance audiovisual stimulus properties, task contexts or subjects' behavioral performances, while reporting the corresponding hemodynamic responses of brain regions, their functional connectivity or their transient and sustained response codes. The first study dissociated the functional contributions of sensory-specific cortices, higher association regions and prefrontal areas to audiovisual object integration, revealing the integration of low-level spatiotemporal, higher order perceptual and decision-related information at 3 distinct processing stages (Werner and Noppeney, 2010a). The second study investigated the neural systems mediating categorical decisions based on multisensory information, showing that prefrontal areas accumulate perceptual evidence from visual and auditory senses to aid appropriate behavioral responses (Noppeney et al., 2010). The third study examined the neural mechanisms underlying the integration of higher order auditory and visual object features, demonstrating that multisensory interaction profiles in higher association areas depend on the informativeness of auditory and visual inputs and are functionally relevant for audiovisual object recognition (Werner and Noppeney, 2010b). Finally, the fourth study investigated the contributions of transient and sustained responses to audiovisual integration dissociating sensory-specific and higher association regions on the basis of their transient response codes and distinct integration profiles at early (i.e. onset) and late (offset) stages of multisensory processing (Werner and Noppeney, 2010c). As all these studies possessed specific objectives, they will be presented here in more detail, including their main results and specific discussions. After these summaries I will turn to the synergetic discussion of all the studies closing with final conclusions.

Study 1

The first experiment was designed to dissociate the 3 stages of audiovisual object integration hypothesized in the general framework of this dissertation (Werner and Noppeney, 2010a). More precisely, multisensory interactions at the first stage indicate the co-stimulation with audiovisual inputs and indicate the co-incidence and co-localization of the sensory sources. Interactions at the second stage signal the integration of higher order object features and the formation of a multisensory object percept. Finally, multisensory interactions at the third stage signal an explicit semantic retrieval enabling the selection of an appropriate action in response to the audiovisual object.

To dissociate the three corresponding neural processing levels, the experiment manipulated the audiovisual input and the task context while exploiting the intrinsic variability in subjects' behavioral performance. In all experimental conditions, subjects were presented with noisy dynamic auditory and/or visual signals emanating from everyday object actions. Subjects either explicitly categorized the objects or passively processed them. Crucially, the experiment included two 2x2 factorial designs that enabled the computation of audiovisual interactions at two levels. (i) The Unimodal Input design [UI] manipulated the absence and presence of the auditory and visual inputs. In the [UI] design, audiovisual interactions emerge due to both, ‘co-stimulation’ per se and the integration of higher order object features. (ii) The Unimodal Object [UO] design provided low level auditory and visual inputs in all conditions but manipulated their informativeness (about objects) by adding different amounts of noise. Hence, the [UO] design controlled for effects of ‘co-stimulation’ and selectively focused on integration of higher order object information. To dissociate the 3 neural processing levels, the following rationale was used. (1) The effect of ‘co-stimulation’ was revealed by comparing the audiovisual interactions of the [UI] relative to those of the [UO] design. (2) Regions associated with the formation of a multisensory object percept were identified by relating subjects’ multisensory benefits in object categorization to their audiovisual interactions (i.e. the superadditive, additive or subadditive multisensory integration profiles). (3) Multisensory facilitations of semantic retrieval and response selection processes were revealed by comparing the audiovisual interactions of the explicit categorization relative to those of the implicit processing task.

The study dissociated 3 patterns of audiovisual interactions at distinct levels of the sensory processing hierarchy.

(1) In primary auditory cortex (PAC), superadditive audiovisual interactions were observed for the integration of low level sensory inputs ([UI] design) but not higher order object information ([UO] design). These superadditive interactions were observed both, when subjects actively categorized and passively perceived the stimuli, indicating that audiovisual response amplifications were of automatic nature. These results suggest that audiovisual co-stimulation (i.e. the co-incidence and co-localization of low level sensory input) within a narrow spatiotemporal window leads to superadditive interactions in primary sensory areas that might enhance stimulus salience and thus enables a coarse initial audiovisual scene segmentation. Further,

effective connectivity analyses demonstrated that these response amplification were mediated via both, direct and indirect (via STS) connectivity to visual cortices confirming both, multisensory feedback influences on sensory cortices from higher order association regions as well as feedforward or lateral connectivity between sensory specific areas.

(2) In higher order association regions of the superior temporal (STS) and intraparietal (IPS) sulci, the pattern of interaction depended on subjects' multisensory benefit in categorization accuracy implicating this network of regions in the integration of higher order object features and the perceptual formation of audiovisual object representations. Hence, in contrast to the stimulus driven superadditive interactions in PAC, audiovisual integration in the STS and IPS depended on the subjects' percept with subadditive interactions (AV<V+A) observed when perception did not benefit from audiovisual integration and additive (AV=V+A) to superadditive (AV>V+A) integration profiles when perceptual performance improved during audiovisual stimulation. The finding demonstrates the essential contribution of higher order association areas to audiovisual integration in mediating multisensory benefits at the level of object recognition.

(3) In the ventrolateral prefrontal cortex (vlPFC), strong subadditive (AV<V+A) audiovisual interactions were observed for the explicit categorization task but not during implicit processing. These interactions were present when subjects integrated audiovisual input [UI] and audiovisual object information [UO] with the suppressed audiovisual BOLD-response relative to both unisensory responses reflecting the reaction time pattern under those conditions. Thus vlPFC may mediate multisensory facilitations of semantic retrieval and response selection, suggesting that vlPFC is involved in audiovisual integration processes and/or receives already integrated information from other regions, e.g. the STS or IPS. Given the need to respond, these integration processes in vlPFC might take place while object information from multiple sensory modalities is accumulated until sufficient evidence is obtained for an

appropriate response. Interestingly, the task-dependent subadditive interactions were located in the transition zones between auditory- and visual-dominant regions in the vlPFC. This specific location suggests that these areas may indeed play an essential role in audiovisual integration per se rather than just processing already integrated information.

In conclusion, the first study demonstrated that multisensory integration emerges at multiple processing levels within the cortical hierarchy with distinct functions. Primary sensory regions detect and integrate spatiotemporally aligned auditory and visual inputs into more salient units through automatic superadditive interactions. In higher order association areas the audiovisual integration profiles predicted subjects' multisensory perceptual benefits suggesting a role in integrating higher order object features into perceptual representations. The suppressive interactions in vlPFC during explicit categorization reflect multisensory facilitation of semantic retrieval and selection of an appropriate action.

Study 2

The second study was designed to identify the neural systems mediating explicit categorical decisions based on multisensory information (Noppeney et al., 2010). Decision making processes are formalized as a stochastic accumulation of auditory and visual evidence over time, until the decisional threshold is reached at which an appropriate response is elicited. Hence, sensory reliability and congruency are possible factors that influence these processes, thus allowing detailed predictions about accumulation rates that translate into reaction times (as markers for the time to decisional threshold) and a corresponding response profile of a putative audiovisual accumulator region.

To identify candidate regions that accumulate information from multiple senses, subjects were presented with audiovisual movies of tools and musical instruments. In a visual selective attention paradigm they categorized the visual information while ignoring the concurrent sound tracks that were either semantically congruent or incongruent with the visual stimulus. Both, visual and auditory information could be intact or degraded. Hence, the 2x2x2 factorial design manipulated the visual reliability (intact, degraded), auditory reliability (intact, degraded) and semantic congruency. A stochastic model of decision making predicted the accumulation rates of all conditions and their corresponding reaction times. For instance, with increases in visual reliability, accumulation rates increase and reaction times decrease. Similarly, reaction times are longer for incongruent than congruent trials (i.e. incongruency effect). Hence, with subjects responding to the visual information, an interaction between incongruency and sensory reliability is predicted with incongruency effects that decrease with higher visual reliability and increase with higher auditory reliability. Convolving the predicted accumulation process, i.e. ramps of neuronal activity with the hemodynamic response function, qualitatively predicts the BOLD-responses of a putative audiovisual accumulator region. Similar to the reaction times, BOLD-response magnitudes were expected to show the incongruency-by-reliability interaction, with incongruency effects decreasing with visual reliability and increasing with auditory reliability. The following 3 objectives were addressed to test the predictions of the stochastic model of multisensory decision making: (1) Evaluation of reaction times for incongruency-by-reliability interactions. (2) Identification of regions with incongruency-by-reliability BOLD-response interactions. (3) Assessment of inter-trial variability in subjects' response times to identify regions of multisensory decision making.

The study identified regions within the lateral prefrontal cortex as the audiovisual accumulator region consistent with the reaction times and the predictions of the stochastic model of multisensory decision making.

(1) For reaction times, significant interactions between (i) incongruency and visual reliability and (ii) incongruency and auditory reliability were observed. More specifically, degraded vision increased the incongruency effect, while degraded

auditory stimuli decreased the incongruency effect. Hence, visual and auditory reliability exerted opposite effects on the incongruency effects in reaction times as predicted by the stochastic model of multisensory decision making.

(2) The BOLD-responses of left inferior frontal sulcus (IFS) and inferior precentral sulcus (iPrCS) showed an interaction between incongruency and visual reliability, where the incongruency effect emerged primarily when the visual input was unreliable. Further, the left IFS also exhibited an interaction between incongruency and auditory reliability, where the incongruency effects were greater for reliable compared to unreliable auditory input. Hence, the visual and auditory reliabilities modulate the incongruency effects in opposite directions as predicted by the model: Increasing the reliability of the visual input that needs to be categorized reduces the incongruency effect, while increasing the reliability of the interfering auditory input amplifies the incongruency effect. Further, effective connectivity analyses demonstrated that these response interactions in IFS were mediated by dynamically weighting the connectivity to the auditory and visual cortices according to the sensory reliability and behavioral relevance of the audiovisual stimulation.

(3) The activations of the left IFS and iPrCS were positively predicted by reaction times, showing greater activations on trials when subjects took longer time to respond.

In summary, the second study identified regions along the IFS that process and accumulate audiovisual information over time to form decisions guiding appropriate behavioral responses. Consistent with the stochastic model of multisensory evidence accumulation that incorporated sensory reliability and congruency factors, the left IFS/iPrCS was the only region that showed the predicted response interactions between incongruency and sensory reliability. Likewise, reaction times as markers for the time to decisional threshold showed the same incongruency-by-reliability interaction and positively predicted activations in the left

IFS/iPrCS on a trial-by-trial basis. Hence, the IFS/iPrCS seems to accumulate audiovisual object evidence to form decisions that guide behavioral responses.

Study 3

The third study was designed to characterize the neural mechanisms underlying the integration of visual and auditory higher order object information (Werner and Noppeney, 2010b). More specifically, the study intended to investigate the factors that determine whether audiovisual inputs are computationally combined with superadditive, additive or subadditive integration profiles, while showing the functional relevance of these integration modes.

In this study, subjects were presented with visual images and auditory sounds emanating from everyday objects. The experiment manipulated the informativeness (intact, degraded, noise) of the concurrently presented auditory and visual stimuli in a 3x3 factorial design. The different levels of informativeness (i.e. degradation) were created by applying different levels of Fourier phase scrambling. Thus, all the experimental conditions were equated with respect to low-level stimulus characteristics (i.e. spatial/temporal frequency contents and low-level image/sound statistics) but differed in the availability of higher order object information. Similar to the first study, subjects either actively categorized or passively processed the stimuli. The experimental design served three purposes. First, equating all conditions with respect to low-level stimulus characteristics, the design allowed a selective focus on the integration of higher order object information, because low-level audiovisual interaction effects due to ‘co-stimulation’ were controlled. Second, independently manipulating intact and degraded levels of auditory and visual informativeness (as a surrogate of stimulus efficacy) enabled the formal investigation of the principle of inverse effectiveness (i.e. non-linear multisensory response enhancements are larger when the responses to individual sensory cues are weakest). Third, using both an active and a passive task enabled the distinction between perceptual (passive) and task-induced (active) integration effects. The following three objectives were addressed. (1) The neural systems that integrate higher order object information were identified. (2) The inverse effectiveness principle was investigated (i.e. $\text{superadditivity}_{\text{degraded}} > \text{superadditivity}_{\text{intact}}$). (3) The functional relevance of the audiovisual integration profiles was evaluated by relating them to subject's multisensory benefits in categorization accuracy.

The study dissociated distinct audiovisual interaction profiles in the superior temporal sulcus (STS) underlying the integration of higher order object information.

(1) Prominent subadditive audiovisual interactions for intact stimuli were observed along the STS of both hemispheres. These interactions were suppressive with the audiovisual response being consistently smaller than the most effective (i.e. auditory) response.

(2) Brain regions obeying the inverse effectiveness principle were observed along the STS of both hemispheres and partially overlapped with regions in (1). These STS areas exhibited strong subadditive audiovisual interactions for intact stimuli and additive to superadditive response profiles for degraded stimuli. The different integration profiles for intact and degraded conditions mapped onto distinct performance patterns during categorization. While no audiovisual benefit in categorization accuracy was found for intact stimuli, a clear multisensory performance enhancement was observed for the degraded conditions. These findings show that the multisensory integration profiles in STS are dictated by the informativeness of auditory and visual sensory inputs, i.e. the perceptibility of higher order audiovisual features necessary for crossmodal object recognition.

(3) To investigate the functional relevance of superadditive, additive and subadditive integration profiles, they were regressed onto subjects' multisensory behavioral benefit in categorization accuracy. The subjects' perceptual benefits (for degraded stimuli) selectively predicted their multisensory integration profiles in the posterior STS bilaterally, regions partially overlapping with activations observed in (1)

and (2). In the posterior STS, subjects without multisensory benefits exhibited subadditive audiovisual interactions, whereas those with high multisensory benefits showed superadditive interactions. These findings demonstrate that superadditive and subadditive integration profiles are functionally relevant and related to behavioral indices of multisensory integration with superadditive interactions mediating audiovisual perceptual benefits. All (1, 2, 3) effects reported here were observed under both, passive and active task conditions and were actually stronger during passive processing of the stimuli. This finding suggests that the observed integration effects were of perceptual rather than task-induced nature. Hence, multisensory interaction profiles in the STS mediate the integration of higher order object features and the perceptual formation of audiovisual object representations.

In conclusion, the third study characterized the computational mechanisms underlying the integration of higher order audiovisual object information in STS. Consistent with the inverse effectiveness principle, multisensory integration profiles were dictated by stimulus efficacy and depended on the informativeness of sensory-specific inputs. Further, the distinct integration profiles were functionally relevant and predicted subjects' perceptual benefits in audiovisual object categorization. The additive and superadditive multisensory profiles might mediate efficient integration of near-threshold inputs from multiple senses. In contrast, subadditive interactions might reflect more efficient processing when the stimulus can already be recognized in at least one sensory modality.

Study 4

The fourth study was designed to characterize the temporal response codes of brain regions mediating the integration of auditory and visual information at different levels of the cortical hierarchy. More precisely, the study examined whether sensory-specific and higher association areas are dissociable on the basis of their transient or sustained response components and their specific audiovisual interaction profiles (Werner and Noppeney, 2010c).

To evaluate the contributions of transient and sustained responses to audiovisual integration, subjects were presented with random dot kinematograms of radial motion, auditory noise or the combination of both in blocks of variable durations. Both, the velocity of the motion and the amplitude of the sound were sinusoidally modulated at 0.1Hz to provide continuous synchrony cues for audiovisual integration. The fMRI design was optimized to model stimulus responses as linear combination of canonical hemodynamic components, parameterizing transient onset and offset, sustained and modulatory responses to visual, auditory and audiovisual stimulations. Multisensory interaction effects for the individual response components were evaluated, i.e. for the (1) onset, (2) sustained and (3) offset responses.

The study dissociated distinct audiovisual interaction profiles at different levels of the cortical hierarchy according to transient response components at early (i.e. onset) and late (offset) stages of stimulation.

(1) Superadditive audiovisual interactions for onset responses were primarily observed in low and higher visual and auditory areas, including the calcarine sulcus (CaS), the fusiform gyrus (FFG), the Heschl’s gyrus (HG) and superior temporal gyrus (STG). The superadditive onset transients were mediated by two components. First, the onsets of auditory information lead to deactivations in visual cortex while visual stimulus onsets lead to deactivations in auditory areas. Second, the responses to concurrently presented audiovisual inputs were amplified in both sensory-specific cortices. Hence, superadditive interactions in low and higher visual and auditory regions were driven by complementary mechanisms: a mutual unisensory inhibition and a multisensory co-excitation at stimulus onsets. These response characteristics might be mediated by feedforward or lateral connectivity between sensory-specific areas where neuronal populations with small temporal integration windows allow a

precise detection of co-incident multisensory events by superadditive responses that might enhance stimulus salience and enable a coarse initial multisensory scene-segmentation.

(2) No significant audiovisual interactions were detected for sustained brain activations.

(3) Subadditive audiovisual interactions for offset responses were observed in higher visual and association areas of the brain, i.e. in the human middle temporal area (hMT+/V5+), the anterior intraparietal sulcus (aIPS) and the posterior superior temporal sulcus (pSTS). Subadditive interaction profiles in pSTS were characterized by comparable (i.e. amodal) visual, auditory and audiovisual offset responses, while the aIPS showed suppressed audiovisual responses with respect to auditory offset transients. The three regions differed in sensory selectivity as expressed in the onset and sustained responses. Surprisingly, even within a single region sensory selectivity differed for onset and sustained responses. For instance, the aIPS showed auditory onset but visual sustained responses. In contrast, pSTS showed auditory selectivity for onset responses but a rather amodal response profile for sustained and offset responses. Critically, the sinusoidal modulation of the sustained response was significantly stronger for audiovisual than unisensory inputs suggesting that slow continuous signals in both sensory modalities influence the response dynamics in pSTS. This complex response profile for onset, offset and sustained components suggests that higher visual and association regions may contain multiple neuronal populations differing in terms of temporal selectivity and sensory preference. From the perspective of predictive coding, ‘offset’ neuronal populations might form higher order ‘amodal’ stimulus representations via perceptual learning during the course of the stimulation block with an ‘amodal’ rebound of activity reflecting a prediction error signal induced by the abrupt change in stimulus structure at stimulus offset. The enhanced sensitivity of the pSTS to the slow dynamics of the combined auditory and visual intensity profiles suggest the presence of neuronal populations with wider

temporal integration windows mediating the integration of continuous audiovisual signals on lower time scales.

In summary, the fourth study characterized the transient and sustained response codes of brain regions mediating the integration of auditory and visual information. First, audiovisual interactions emerged primarily for transient responses at stimulus onset and offset highlighting the importance of rapid stimulus transitions for multisensory integration. Second, the audiovisual integration profiles differed across response components. They were superadditive for onset responses, yet subadditive and for offset responses. Third, the audiovisual onset and offset interactions were anatomically segregated. Superadditive interactions at stimulus onset were observed primarily in low-level visual and auditory areas possibly mediating an increase in stimulus salience. In contrast, subadditive interactions at stimulus offset were located in higher visual and association areas and may reflect the formation of higher order representations as a result of perceptual learning.

General Discussion

The objective of this dissertation was to characterize the functional similarities, differences and constraints that govern the integration of auditory and visual information in different regions of the human cerebral cortex. The project was guided by a general framework hypothesizing the integration of specific types of information at 3 distinct processing stages, i.e. (1) low-level spatiotemporal information indicating the same sensory source, (2) higher order perceptual information mediating recognition and (3) decisional information driving response selection processes. This dissertation dissociated the 3 processing levels and demonstrated the functional contributions of sensory-specific cortices, higher association regions and prefrontal areas to audiovisual integration. In the following, the collected experimental evidence for each individual processing level is conjointly discussed with respect to the predictions of the general framework and the present state of multisensory research.

Integration at a low 'stimulus-driven' level, signaling the same sensory source. Multisensory interactions at the first stage were hypothesized to indicate the co-incidence and co-localization of sensory sources with narrow spatiotemporal receptive windows that ensure a precise timing and spatial processing. This functional specialization should prevent from integrating more complex sensory details since their feature space covers slower and spatially extended signals.

The experimental evidence of two studies suggests that this first processing stage resides in low and higher sensory-specific regions of the brain (Werner and Noppeney, 2010a;Werner and Noppeney, 2010c). In the first study the primary auditory cortex integrated low-level sensory inputs but not higher order object information, showing that co-incidence and co-localization rather than complex features drive superadditive interactions in sensory-specific regions (Werner and Noppeney, 2010a). Extending this finding to the visual cortex, the fourth study showed that these audiovisual interactions selectively occur at stimulation onset when rapid stimulus changes co-occur, suggesting that receptive windows are functionally specialized for precise timing and spatial processing (Werner and Noppeney, 2010c). This notion is supported by neurophysiological recordings in non-human primates demonstrating that multisensory processes in sensory-specific cortices are governed by tight temporal (Kayser et al., 2008) and also spatial constraints (Lakatos et al., 2007). Furthermore, the current studies revealed that audiovisual interactions in visual and auditory regions are driven by two complementary mechanisms with unisensory inhibitory processes between sensory modalities turning into facilitatory mechanisms for concurrent multisensory stimulations (Werner and Noppeney, 2010a;Werner and Noppeney, 2010c). This interdependency between auditory and visual cortices might be mediated via direct feedforward (thalamocortical) or lateral connectivity between sensory-specific areas or via indirect feedback modulations from higher association regions (Schroeder et al., 2003). Previous fMRI studies have only provided evidence for feedback influences using Granger Causality (van Atteveldt et al., 2009) or Directed Information Transfer (Noesselt et al., 2007). However, in the first study audiovisual interactions in primary auditory cortex were observed independent of task contexts, suggesting automatic integration mechanisms and direct interactions between sensory-specific regions (Werner and Noppeney, 2010a). Further, the fourth study showed that rapid stimulus changes at onset are integrated in low-level sensory regions in the absence of interactions in association areas, suggesting the independence of higher order processing stages (Werner and Noppeney, 2010c). These claims were indeed confirmed by the analyses of effective connectivity between sensory-specific and association regions (STS) in the first study, demonstrating that multisensory integration in sensory-specific regions rely not only on 'indirect' feedback processes, but also depend on the presence of 'direct' feedforward or lateral mechanisms (Werner and Noppeney, 2010a). These findings are consistent with anatomical studies showing direct connections between visual and auditory areas as well as feedback projections from STS to sensory-specific areas (Falchier et al., 2002;Rockland and Ojima, 2003). Further, EEG and intracranial recordings have shown both, early audiovisual interactions that are most likely mediated via direct connectivity (Molholm et al., 2002;Besle et al., 2008) and late audiovisual interactions most probably the result of indirect mechanisms (Busse et al., 2005;Bonath et al., 2007). Collectively, these studies suggest that audiovisual integration in sensory-specific cortices can occur at both, early and late stages of sensory processing. On the one hand, at the early stage direct interactions between sensory-specific regions might detect the co-incidence and co-localization of audiovisual signals with narrow spatiotemporal receptive windows, signaling a unified sensory source to subsequent processing stages. On the other hand, at a later stage feedback influences from higher association regions might adaptively change the width of the receptive windows in sensory-specific regions, allowing a unified percept even when audiovisual inputs are in slight spatiotemporal conflict. An example for such conflicts and its resolution is the ventriloquist illusion. A ventriloquist can create the illusion that his voice emerges from the visibly moving mouth of a puppet. Such mislocalizations of sounds towards a co-occurring visual stimulus have been demonstrated to depend on late audiovisual interactions in auditory cortex that create a shift of auditory spatial perception towards visual locations (Bonath et al., 2007). Hence, in a “late phenomenon” like the ventriloquist illusion, sensory-specific regions at the early stage of direct interaction might actually detect the spatial discrepancy between visual and auditory information and signal independent sensory sources to subsequent processing stages. However as the spatial conflict is quite small, the same information might be nevertheless integrated in higher association regions, since their spatiotemporal receptive windows are larger and thus operate with less acuity. With the arrival of these integrative feedback signals in sensory-specific regions, late audiovisual interactions might align the visual and auditory streams into a unified percept, ensuring that the sounds are adaptively localized to the position of the visual stimulus. These feedback-dependent integration processes in sensory-specific regions might represent mechanisms postulated by maximum likelihood models of integration, where unisensory perceptual estimates combine as a linear sum after weighting them by their sensory reliabilities (Ernst and Banks, 2002). In the case of the ventriloquist illusion (Alais and Burr, 2004), feedback-dependent interactions in auditory cortex align visual and auditory percepts by “pulling” the (spatially) less reliable sounds towards the more reliable position provided by the visual modality (Bonath et al., 2007). Conversely, when audition would be hypothetically more reliable than vision, integration would be mediated by feedback-dependent interactions in visual cortex aligning a less reliable visual percept towards the more reliable information provided in the auditory modality.

Taken together these studies suggest that early and late mechanisms of audiovisual integration in sensory-specific cortices serve different functions. Early processes that are mediated by direct interactions between sensory-specific regions integrate spatiotemporally aligned auditory and visual inputs with narrow receptive windows into more salient representations for stimulus (or novelty) detection. However, late integration processes in sensory-specific cortices rely on feedback influences from higher association regions that potentially adjust receptive windows and stimulus processing to overcome small spatiotemporal discrepancies between audiovisual inputs and thus mediate a unified percept even in the face of slight sensory conflict.

Integration at a higher ‘perceptual’ level, mediating the formation of multi-sensory object representations.

Multisensory interactions at the second stage were hypothesized to combine higher order perceptual features with wider spatiotemporal receptive windows that allow the integration of slower sensory signals at more flexible spatial scales. The functional specialization of this processing stage should promote integration of complex sensory details and formation of a multisensory object percept, i.e. object recognition with enhanced accuracy.

The cumulative experimental evidence of three studies suggests that this second processing stage resides in higher association regions of the brain, most prominently in posterior portions of the superior temporal sulcus (pSTS) (Werner and Noppeney, 2010c;Werner and Noppeney, 2010a;Werner and Noppeney, 2010b). In the third study, where all conditions were equated with respect to low-level stimulus characteristics but differed in the informativeness of sensory inputs, pSTS integrated higher order information demonstrating that complex object features rather than low-level sensory cues drive audiovisual interactions in higher association regions. Further, the integrative mechanism obeyed the principle of inverse effectiveness, as multisensory integration in pSTS was dictated by stimulus efficacy and depended on the degradation of object features. Furthermore, these effects were stronger during passive stimulus processing but were also present during active categorization, suggesting that these integration effects were observed on a perceptual processing stage (Werner and Noppeney, 2010b). A functional specialization for the integration at the perceptual level was further strongly supported by tight correlations between subjects' audiovisual benefits in object perception and their multisensory interaction profiles in pSTS – a predictive relationship that was demonstrated in the first and replicated in the third study using different subjects and stimulus degradation methods (Werner and Noppeney, 2010a;Werner and Noppeney, 2010b). Hence, in contrast to the stimuli-driven 'automatic' audiovisual interactions in sensory-specific cortices, integration processes in higher association regions seem to be critically dependent on the perceptual interpretation of such stimuli. The formation of an audiovisual percept at a higher stage of multisensory convergence is consistent with hierarchical models of sensory processing that propose a progressive transformation from simple to more complex representations along unisensory pathways, before higher perceptual representations are "mapped" onto each other and combined in multisensory cortices (Calvert, 2001). Multisensory association regions such as the STS are best suitable to perform such mappings, as their anatomical connections (Falchier et al., 2002;Rockland and Ojima, 2003) allow complex interactions with visual and auditory areas (Ghazanfar et al., 2008;Kayser and Logothetis, 2009) and thus can mediate the formation of a unified percept and benefits in object recognition (Werner and Noppeney, 2010a;Werner and Noppeney, 2010b). In line with these findings, the fourth study additionally revealed comparable offset responses for visual, auditory or audiovisual stimuli, suggesting that pSTS forms higher order 'amodal' stimulus representations via perceptual learning during the course of the stimulation with 'amodal' responses possibly reflecting a prediction error signal induced by the abrupt change in stimulus structure at stimulation offset (Werner and Noppeney, 2010c).

With a functional specialization promoting the integration of complex sensory features and perceptual information, higher processing stages should reveal characteristic spatiotemporal receptive windows that allow the integration of slower sensory signals at larger spatial scales. To test this prediction, the fourth study presented visual, auditory and audiovisual stimuli with slow (0.1Hz) sinusoidal intensity modulations. Interestingly, only the sustained response in pSTS exhibited a stronger modulation at 0.1Hz for audiovisual than unisensory stimulations suggesting that slow continuous signals from both sensory modalities drive the temporal dynamics of STS responses (Werner and Noppeney, 2010c). Previous studies have shown that along higher hierarchical levels of visual or auditory processing, unisensory information is accumulated over progressively longer time scales with larger receptive windows in higher association regions (Griffiths et al., 1998;Boemio et al., 2005;Hasson et al., 2008;Overath et al., 2008). The fourth study extended these findings, demonstrating that the association areas of the pSTS actually sustain the integration of slower audiovisual dynamics (Werner and Noppeney, 2010c).

Collectively these studies suggest that higher association regions such as pSTS integrate complex audiovisual object features into perceptual object representations. Due to their extensive connectivity to sensory-specific regions and the presence of wider receptive windows tuned to higher order sensory detail, higher association regions mediate the formation of unified percepts and audiovisual benefits in object recognition.

Integration at a 'decisional' level, mediating response selection processes based on perceptual information.

Multisensory interactions at the third stage were hypothesized to integrate audiovisual information to form decisions that guide appropriate behavioral actions with faster response times. With the functional specialization to establish a behavioral response mapping based on perceptual information, this processing stage should allow an accumulation of higher order auditory and visual information until sufficient evidence for a response is obtained.

The experimental evidence of two studies suggests that this third processing stage resides in regions of the ventrolateral prefrontal cortex (vlPFC), most prominently in the inferior prefrontal sulcus (IFS) and the inferior precentral sulcus (iPrCS) (Werner and Noppeney, 2010a;Noppeney et al., 2010). In the first study, the vlPFC integrated higher order information showing that complex object features rather than low-level sensory cues drive audiovisual interactions in prefrontal regions. However, these interactions occurred only, when subjects explicitly categorized and responded to the object stimuli and not during passive processing (Werner and Noppeney, 2010a). This finding is consistent with the proposed functional specialization of the processing stage to establish a mapping between perceptual features and response alternatives represented in a decisional process that is absent during passive exposure. Additionally, these studies demonstrated ‘decision-related’ activations only in the left hemisphere of the vlPFC – contralateral to the subjects’ response hand – a finding that is also consistent with a proposed function of linking percept with behavioral planning (Werner and Noppeney, 2010a;Noppeney et al., 2010). Further, the behavioral response mapping based on combined audiovisual information was performed faster than based on unisensory information, with the responses in the vlPFC resembling the reaction time measures under those conditions (Werner and Noppeney, 2010a). These behavioral and neuroimaging data are consistent with models of perceptual decision making formalizing decisions as stochastic diffusion or accumulator processes that integrate sensory information over time until sufficient evidence for a response is obtained (Smith and Ratcliff, 2004). For instance, in putative ‘decision’ regions of non-human primates neuronal firing rates have been shown to build up until a decisional threshold is reached and a response is selected (Roitman and Shadlen, 2002;Huk and Shadlen, 2005). This ramp-like neuronal activity has been suggested to reflect the accumulation of noisy sensory evidence provided by lower sensory areas (Gold and Shadlen, 2007). Hence, in the case of a combined audiovisual stimulation, accumulation rates will be higher than for unisensory stimuli potentially based on a linear superposition of activations elicited by two inputs (Schwarz, 1994), which consequently resulted in an earlier reach of the decisional threshold and facilitated reaction times (Schwarz, 2006). These putative processes of multisensory decision making are reflected in the BOLD-response profile of the vlPFC. While for unisensory information higher activations were observed due to longer accumulation processes, the combined audiovisual stimulation allowed the obtainment of the threshold value in shorter time, i.e. with smaller BOLD activation (Werner and Noppeney, 2010a).

The accumulation of visual and auditory information in vlPFC was even more explicitly demonstrated in the second study that manipulated the reliability and congruency of audiovisual information (Noppeney et al., 2010). With subjects usually responding slower to incongruent relative to congruent information, the stochastic model of multisensory decision making predicted distinct changes of these incongruency effects depending on the reliability of the sensory inputs. For both, the reaction times as well as BOLD-responses in the vlPFC the incongruency effects decreased with the reliability of visual inputs that were categorized but increased with the reliability of the auditory input that needed to be ignored (Noppeney et al., 2010). This finding demonstrates that reaction times and BOLD responses are based on the same stochastic accumulation processes that integrate visual and auditory information over time, with accumulation rates determining the reaction times and the corresponding response profile of the region that mediated the formation of the perceptual decision.

Collectively these studies suggest that audiovisual information is integrated in the vlPFC at a decisional level where perceptual information drives the selection of an appropriate action with facilitated reaction times. The vlPFC thereby performs a mapping between perceptual features and behavioral options. The mapping is reflected in a decisional process that is characterized by a stochastic accumulation of visual and auditory information over time until a decisional threshold is reached and a response is selected. Hence, vlPFC activations necessarily depend on perceptual features, as they depend on the number of alternative behavioral responses.

Conclusion

This dissertation demonstrated that audiovisual integration emerges at multiple processing stages and levels within the cortical hierarchy. At the first stage, early multisensory interactions in sensory-specific regions indicate the same sensory source by integrating spatiotemporally aligned auditory and visual inputs with narrow receptive windows to enhance stimulus detection. At a second stage, multisensory interactions in higher association regions integrate complex audiovisual object features into perceptual object representations. Due to the presence of wider receptive windows tuned to higher order sensory detail in addition to pronounced feedback connectivity to sensory-specific regions, higher association regions mediate the formation of a unified percept and audiovisual benefits in object recognition even in the face of slight sensory conflict. At the third stage, multisensory interactions in the prefrontal cortex mediate response selection processes based on perceptual information from different sensory modalities. This decisional process can be formalized as stochastic accumulation of visual and auditory information until a threshold is reached at which the appropriate response is elicited. The integration processes at the 3 distinct processing stages are most likely not only occurring in a serial but also in parallel fashion with a high degree of interaction between processing stages.

Reference List

Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. Current Biology 14:257-262.

Amedi A, von Kriegstein K, van Atteveldt NM, Beauchamp MS, Naumer MJ (2005) Functional imaging of human crossmodal identification and object recognition. Experimental Brain Research 166:559-571.

Besle J, Fischer C, Bidet-Caulet A, Lecaignard F, Bertrand O, Giard MH (2008) Visual Activation and Audiovisual Interactions in the Auditory Cortex during Speech Perception: Intracranial Recordings in Humans. Journal of Neuroscience 28:14301-14310.

Boemio A, Fromm S, Braun A, Poeppel D (2005) Hierarchical and asymmetric temporal sensitivity in human auditory cortices. Nat Neurosci 8:389-395.

Bonath B, Noesselt T, Martinez A, Mishra J, Schwiecker K, Heinze HJ, Hillyard SA (2007) Neural basis of the ventriloquist illusion. Current Biology 17:1697-1703.

Busse L, Roberts KC, Crist RE, Weissman DH, Woldorff MG (2005) The spread of attention across modalities and space in a multisensory object. Proceedings of the National Academy of Sciences of the United States of America 102:18751-18756.

Calvert GA (2001) Crossmodal processing in the human brain: Insights from functional neuroimaging studies. Cerebral Cortex 11:1110-1123.

Calvert GA, Bullmore ET, Brammer MJ, Campbell R, Williams SC, McGuire PK, Woodruff PW, Iversen SD, David AS (1997) Activation of auditory cortex during silent lipreading. Science 276:593-596.

Calvert GA, Spence C, Stein BE (2004) The handbook of multisensory processes. Cambridge, Massachusetts: MIT Press.

Driver J, Noesselt T (2008) Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. Neuron 57:11-23.

Driver J, Spence C (2000) Multisensory perception: Beyond modularity and convergence. Current Biology 10:R731-R735.

Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415:429-433.

Falchier A, Clavagnier S, Barone P, Kennedy H (2002) Anatomical evidence of Multimodal integration in primate striate cortex. Journal of Neuroscience 22:5749-5759.

Felleman DJ, van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. pp 1-47.

Foxe JJ (2008) Toward the end of a "principled" era in multisensory science. Brain Research 1242:1-3.

Foxe JJ, Schroeder CE (2005) The case for feedforward multisensory convergence during early cortical processing. Neuroreport 16:419-423.

Ghazanfar AA, Chandrasekaran C, Logothetis NK (2008) Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. Journal of Neuroscience 28:4457-4469.

Ghazanfar AA, Schroeder CE (2006) Is neocortex essentially multisensory? Trends in Cognitive Sciences 10:278-285.

Giard MH, Peronnet F (1999) Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. Journal of Cognitive Neuroscience 11:473-490.

Gold JI, Shadlen MN (2007) The neural basis of decision making. Annual Review of Neuroscience 30:535-574.

Griffiths TD, Buchel C, Frackowiak RS, Patterson RD (1998) Analysis of temporal structure in sound by the human brain. Nat Neurosci 1:422-427.

Hasson U, Yang E, Vallines I, Heeger DJ, Rubin N (2008) A hierarchy of temporal receptive windows in human cortex. J Neurosci 28:2539-2550.

Hubel DH (1988) Eye, Brain, and Vision. New York: Scientific American Library.

Huk AC, Shadlen MN (2005) Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. J Neurosci 25:10420-10436.

Kayser C, Logothetis NK (2009) Directed Interactions Between Auditory and Superior Temporal Cortices and their Role in Sensory Integration. Frontiers in Integrative Neuroscience 3.

Kayser C, Petkov CI, Logothetis NK (2008) Visual modulation of neurons in auditory cortex. Cerebral Cortex 18:1560-1574.

Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE (2007) Neuronal oscillations and multisensory interaction in primary auditory cortex. Neuron 53:279-292.

Meredith MA, Stein BE (1983) Interactions Among Converging Sensory Inputs in the Superior Colliculus. Science 221:389-391.

Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE, Foxe JJ (2002) Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. Cognitive Brain Research 14:115-128.

Noesselt T, Rieger JW, Schoenfeld MA, Kanowski M, Hinrichs H, Heinze HJ, Driver J (2007) Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. Journal of Neuroscience 27:11431-11441.

Noppeney U, Ostwald D, Werner S (2010) Perceptual decisions formed by accumulation of audiovisual evidence in prefrontal cortex. Journal of Neuroscience 30:7434-7446.

Overath T, Kumar S, von KK, Griffiths TD (2008) Encoding of spectral correlation over time in auditory cortex. J Neurosci 28:13268-13273.

Rauschecker JP, Tian B, Pons T, Mishkin M (1997) Serial and parallel processing in rhesus monkey auditory cortex. J Comp Neurol 382:89-103.

Rockland KS, Ojima H (2003) Multisensory convergence in calcarine visual areas in macaque monkey. International Journal of Psychophysiology 50:19-26.

Roitman JD, Shadlen MN (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. Journal of Neuroscience 22:9475-9489.

Rolls ET (2000) Functions of the primate temporal lobe cortical visual areas in invariant visual object and face recognition. Neuron 27:205-218.

Schroeder CE, Foxe J (2005) Multisensory contributions to low-level, 'unisensory' processing. Current Opinion in Neurobiology 15:454-458.

Schroeder CE, Foxe JJ (2002) The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. Cognitive Brain Research 14:187-198.

Schroeder CE, Smiley J, Fu KG, McGinnis T, O'Connell MN, Hackett TA (2003) Anatomical mechanisms and functional implications of multisensory convergence in early cortical processing. Int J Psychophysiol 50:5-17.

Schwarz W (1994) Diffusion, Superposition, and the Redundant-Targets Effect. Journal of Mathematical Psychology 38:504-520.

Schwarz W (2006) On the relationship between the redundant signals effect and temporal order judgments: Parametric data and a new model. Journal of Experimental Psychology-Human Perception and Performance 32:558-573.

Smith PL, Ratcliff R (2004) Psychology and neurobiology of simple decisions. Trends in Neurosciences 27:161-168.

Stanford TR, Stein BE (2007) Superadditivity in multisensory integration: putting the computation in context. Neuroreport 18:787-792.

Stein BE, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. Nat Rev Neurosci 9:255-266.

van Atteveldt N, Roebroeck A, Goebel R (2009) Interaction of speech and script in human auditory cortex: insights from neuro-imaging and effective connectivity. Hear Res 258:152-164.

Wallace MT, Ramachandran R, Stein BE (2004) A revised view of sensory cortical parcellation. Proceedings of the National Academy of Sciences of the United States of America 101:2167-2172.

Werner S, Noppeney U (2010a) Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. Journal of Neuroscience 30:2662-2675.

Werner S, Noppeney U (2010b) Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization. Cerebral Cortex 20:1829-1842.

Werner S, Noppeney U (2010c) The contributions of transient and sustained response codes to audiovisual integration. Cerebral Cortex Advanced Access: published September 1, doi: 10.1093/cercor/bhq161.

Personal contributions

Personal contributions to papers and manuscripts

(Darstellung des Eigenanteils bei Gemeinschaftsarbeiten nach §9 (2))

I. **Werner S**, Noppeney U (2010) Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. **Journal of Neuroscience** 30(7): 2662-2675

   I designed the experiment together with Uta Noppeney, recorded the stimuli, created the stimulus manipulations, programmed and carried out the experiment, performed the data analyses and wrote the first draft of the manuscript.

II. Noppeney U, Ostwald D, **Werner S** (2010) Perceptual decisions formed by accumulation of audiovisual evidence in prefrontal cortex. **Journal of Neuroscience** 30(21): 7434-7446

   I recorded the stimuli, and assisted Dirk Ostwald with creating the stimulus manipulations, carrying out the experiment and performing the data analyses.

III. **Werner S**, Noppeney U (2010) Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization. **Cerebral Cortex** 20(8): 1829-1842 (Epub 2009 Nov 18)

   I designed the experiment together with Uta Noppeney, recorded the stimuli, created the stimulus manipulations, programmed and carried out the experiment, performed the data analyses and wrote the first draft of the manuscript.

IV. **Werner S**, Noppeney U (2010) The contributions of transient and sustained response codes to audiovisual integration. **Cerebral Cortex** Advanced online publication: published September 1, doi: 10.1093/cercor/bhq161

   I developed the idea for the experiment and designed it together with Uta Noppeney, created the stimuli, programmed and carried out the experiment, performed the data analyses and wrote the first draft of the manuscript.