BEHAVIORAL AND FMRI MEASURES OF CROSSMODAL PLASTICITY INDUCED BY AUDITORY SENSORY SUBSTITUTION

Thesis by

Noelle R. B. Stiles

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2015

(Defended September 23, 2014)

ACKNOWLEDGMENTS

I would like to thank my advisor, Shinsuke Shimojo, who has been a fantastic mentor throughout my years at Caltech, both pushing me toward excellence and encouraging a deep understanding of the scientific process. I would also like to thank my committee, Henry Lester, Doris Tsao, John O'Doherty, and Carmel Levitan, for their invaluable advice and guidance through the Ph.D. process. I would like to thank Vikram Chib for his patient and helpful fMRI training, advice, and data processing aid.

My collaborators inside and outside of Caltech have also dramatically improved my experience and allowed me to explore scientific questions that are not included in my thesis. In particular, Carmel Levitan, thank you for your scientific fellowship, and I look forward to continuing our work together. Luis Goncalves and Enrico Di Bernardo, thank you for enthusiasm and many excellent scientific discussions. Armand R. Tanguay, Jr., thank you for your collaboration, mentorship, and encouragement; you have added to my success in more ways than I could possibly list.

I also appreciate all of the undergraduate and graduate students who have worked on research projects with me at Caltech (Sepideh Farmani, Yuqian Zheng, Kijun Seo, Charlotte Yang, Juri Minxha, Daisy Lin, Kevin Li, Mason McGill, Elizabeth Lawler, and Allison Ban) and at USC (Ben McIntosh, Jennifer Crisp, and Janet Kim). Thank you for your enthusiasm and hard work.

Finally to my parents, thank you for your endless encouragement and support; I certainly would not be here without you.

Several organizations have supported my research and career at Caltech in many ways, for which I am extremely grateful. I appreciate a fellowship from the National Science Foundation Graduate Research Fellowship Program, the Mary Louise Remy Endowed Scholar Award from the Philanthropic Educational Organization, a Discovery Scholars Prize from the University of Southern California, and timely Travel Awards from the Caltech Graduate Office. I also appreciate research grant support from the Della Martin Fund for Discoveries in Mental Illness; the National Institutes of Health; and the Japan Science and Technology Agency, Core Research for Evolutional Science and Technology.

ABSTRACT

Sensory substitution (SS) aids the blind by encoding information from vision into a tactile or auditory stimulus. The major focus of this thesis is on the study of the crossmodal plasticity engendered through training on SS devices, and on the study and improvement of the rehabilitative potential of SS.

The effectiveness of SS as a rehabilitative device is unfortunately limited. Blind and blindfolded sighted individuals can be trained to interpret SS sounds visually and even learn object localization and identification, as well as depth perception. Despite this, participants require significant training (1 week to 3 months) to learn to use an SS device. Even after that they require significant attentional resources and top-down executive control to perform basic visual tasks. The laborious interpretation of SS is in stark contrast to the effortlessness of visual interpretation even in complex and cluttered environments. Recognition, localization, constancies, and depth perception are attained in vision with ease and a surprising level of automaticity. Therefore, a major focus of this thesis will be studying whether SS devices can be made more intuitive and consequently more vision-like and useful to blind individuals. In particular, we have discovered that certain types of textural patterns provide intuitive associations between auditory and visual interpretation, suggesting both the utility of inherent crossmodal mappings for SS interpretation and a dramatic change in training paradigm. Surprisingly, we found that sounds derived from such textural patterns were correctly matched to images by naïve-sighted participants.

We also studied a crucial element of visual processing, constancies, with SS to determine whether SS users could acquire constancies via crossmodal plasticity and use them effectively. Further, while training the blind and sighted participants to learn constancies, we focused on improving the training procedures to make the device training more effective. Results showed that the sighted and blind participants could learn length and orientation constancy. We also found that spontaneous head-tilting movements while learning length constancy significantly correlated with improved task learning. Moreover, the improvement was transferred to a no-head-tilt condition. Overall, our results indicate that stimuli externalization, vision-like processing, and plastic sensory-motor integration are important and learnable elements of effective SS use.

The second focus of this thesis is the automaticity and topographic mapping of SS through crossmodal plasticity. Previously, auditory or tactile stimuli generated by sensory substitution have been shown to be processed in primary visual regions via crossmodal plasticity in blindfolded sighted and early or late blind individuals. Sensory substitution is therefore intrinsically crossmodal, and has a unique type of crossmodal plasticity between multimodal and unimodal cortical regions. Several studies have recorded this auditory-visual crossmodal interaction with sensory substitution, but very few have attempted to quantify the spatial, temporal, and attentional aspects of this plastic neural network. In this thesis, we will grapple with two of these elements: Spatial representation of sensory substitution stimuli in visual regions (*i.e.,* topographic mapping), and the role of attention (or automaticity) in crossmodal sensory substitution processing.

We used fMRI imaging to investigate the automatic nature of sensory substitution crossmodal plasticity. Sensory substitution has been shown to activate visual cortex with an auditory or tactile stimulus, in a similar pattern to visual processing of objects and locations. However, vision is also bottom-up, perceptual, and automatic. It is still unknown whether sensory substitution, like vision, can also be automatically processed by visual regions. The literature has assumed that automatic processing with SS is not possible, as basic functionality requires extensive training (1 week to 3 months). Unexpectedly, we show in this thesis that sensory substitution activates visual regions in a passive tasks, as well as in tasks distracting attention from the SS stimulus (with fMRI). These results indicate that SS interpretation and crossmodal plasticity is more perceptual, as opposed to cognitive or top-down-controlled, than previously believed.

The topographical mapping of visual space onto visual cortex via sensory substitution was also studied with fMRI imaging. Vision has a retinotopic map such that close regions on the retina are processed by neighboring regions in primary visual cortex. We used fMRI imaging to determine whether this type of mapping holds when the visual image is encoded into sound (with SS) and then activates visual cortex via crossmodal plasticity. Interestingly, we found that topographic mapping with SS is not entirely vision-like, but has similarities.

Overall, this thesis aims to investigate the crossmodal plasticity that enables SS interpretation, and to improve SS as a rehabilitative device for the blind. This thesis investigation shows that SS can be used to learn visual constancies, which are critical to visual rehabilitation of the blind and dependent upon plasticity. It also indicates that crossmodal mappings can be used to intuitively interpret SS with no encoding knowledge and attentive efforts, thereby enabling training design recommendations that may shorten and improve training on SS. Finally, neural imaging studies therein investigated crossmodal plasticity of sensory substitution processing that make SS acquire aspects of vision's automaticity and retinotopic mapping.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1:

INTRODUCTION

Introduction

Sensory substitution (SS) encodes an image into a sound or tactile stimulation, and trained participants have been found not only to utilize the stimulus to coordinate adaptive behavior, but also to process it in early visual areas. Some superusers of a sensory substitution device have further claimed to subjectively experience a vision-like perception associated with device usage (Ward & Meijer, 2010). This chapter will not only go over the technical and historical perspective of SS, but will also more importantly highlight the implications of SS to blind rehabilitation and the potential of SS to reveal crossmodal perceptual organization.

Sensory substitution is processed like vision at cortical levels, but is transduced by audition (or somatosensation) at receptor levels; thus, it should be considered neither pure vision nor audition/somatosensation, but rather a third type of subjective sensation, or “qualia” (the absolute, first-person, quality of sensory experiences). If perceptual experience in sensory substitution is unique, do the same visual primitives hold? Are these visual primitives fundamental to all vision-like processing, or are they dependent on the visual sensory transduction process? Several other questions fundamental to the essential nature of visual experience also become feasible to investigate with this new broader definition of “visual” processing, such as holistic *vs.* local processing, static *vs.* dynamic recognition and depth perception, and perception based on purely sensory *vs.* sensory-motor neural processing. Studies with sensory substitution attempt to aid the blind by understanding these questions and thereby improving both SS devices and the users' quality of life. Further, these investigations advance neuroscience by demonstrating the roles that neural plasticity and sensory integration play in the organization of visual perception. In short, SS provides scientists and philosophers with a new artificial dimension to examine perceptual organization processes.

Overview of Thesis

Sensory substitution studies are bifurcated along rehabilitation and basic neural science objectives. Sensory substitution can aid the blind by enhancing environmental perception, and navigation. Further, as previously mentioned, sensory substitution is a unique crossmodal recombination of modalities, and therefore provides an interesting perspective on crossmodal interactions at the neural and behavioral levels. This thesis serves both of these basic and applied science objectives; each chapter emphasizes one or the other of these aims, as illustrated in Figure 1.1.

The second and third chapters of this thesis use behavioral techniques to improve perception with sensory substitution. In particular, these chapters study making sensory substitution more perceptual and effortless to use, and training participants to learn constancies with the vOICe device, enabling object externalization and adaptive control of behavior. The second chapter discusses the training of sighted and blind participants to learn orientation and length constancy. A surprising outcome of this training is that dynamic interaction with the stimuli was critical to enhanced learning on the length constancy task. The third chapter focuses on shortening vOICe training and improving training outcomes by using innate crossmodal mappings. The results of Chapter 3 may be used to intelligently incorporate crossmodal mappings into training, thereby re-focusing and enhancing it.

The forth and fifth chapters of this thesis use fMRI techniques to understand the neural processing and plasticity that underlie sensory substitution learning. The forth chapter focuses on determining whether attention is required for sensory substitution to activate the visual cortex in blind and sighted participants. The fifth chapter determines whether vOICe perception is contralaterally mapped from visual space to visual perception in the same way that vision is mapped.

This thesis begins to investigate some of the unknown features of sensory substitution. It investigates the similarities and differences between sensory substitution and visual perception. Several experiments within this thesis indicate that sensory substitution may be more automatic and perceptual than shown in previous studies with sensory substitution. Finally, the theme of the unique crossmodal nature of sensory substitution, and how to exploit it for rehabilitative gains, are highlighted in several experiments. Figure 1.2 illustrates several of these thesis themes and how they build toward improving blind participant rehabilitation. Further, Chapter 6 will discuss these thesis themes and their implications for sensory substitution in detail.

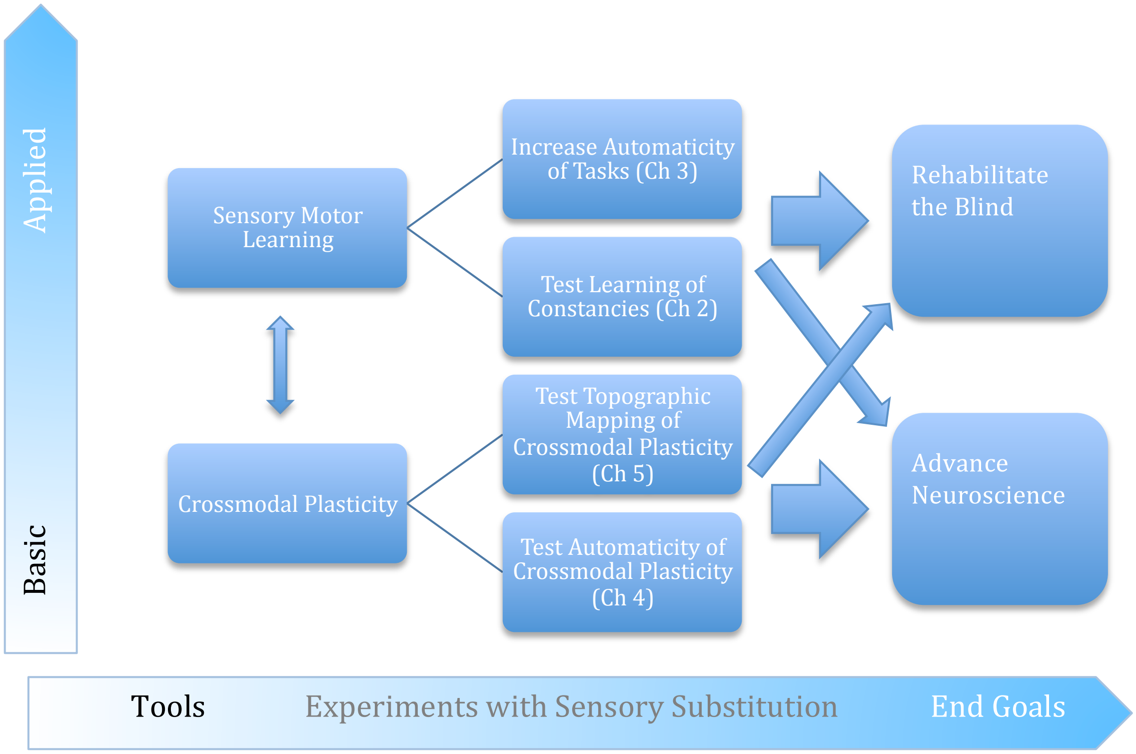


Figure 1.1. Concept web for thesis. This diagram spatially lays out the concepts developed in the thesis, and maps out several interesting inter-connections among concepts. In particular, it maps out the progress from tools to experiments to scientific goals for the thesis. It also shows the range from basic science to more applied science, and various cross-connections among the two. (Note: this figure is repeated as a review in Chapter 6.)

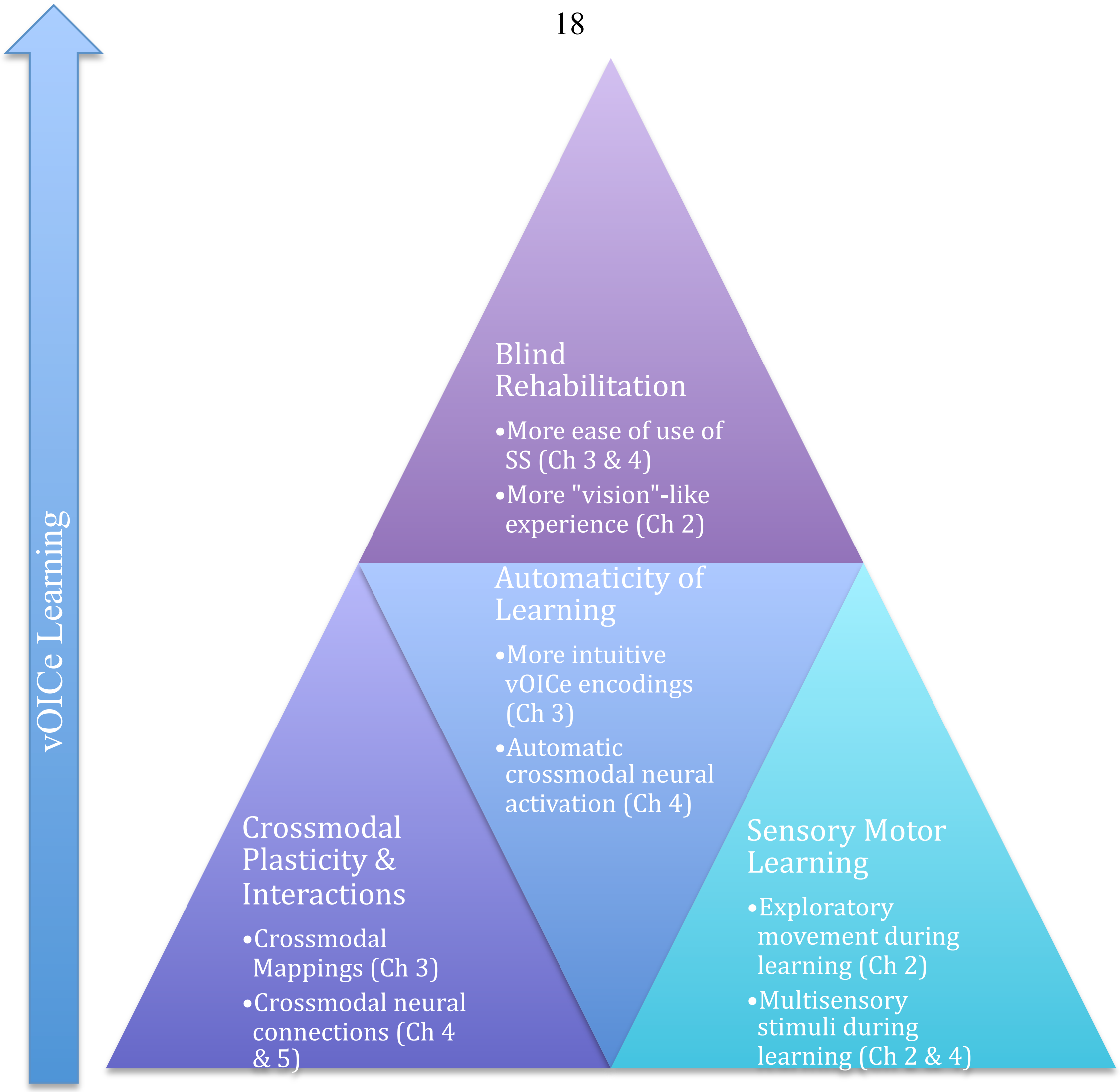


Figure 1.2. Layout of thesis themes. An alternative layout of thesis themes shows the crossmodal plasticity and sensory motor learning at the base of the pyramid, supporting the automaticity of perceptual processing and the rehabilitation of the blind. Each of the pyramid blocks has references to the chapters that relate strongly to those themes. (Note: this figure is repeated as a review in Chapter 6.)

Overview of Chapter

Chapter 1 is organized to include background information on sensory substitution and other relevant perceptual processes. To clarify the information covered in this chapter, a chapter outline is provided in Figure 1.3.

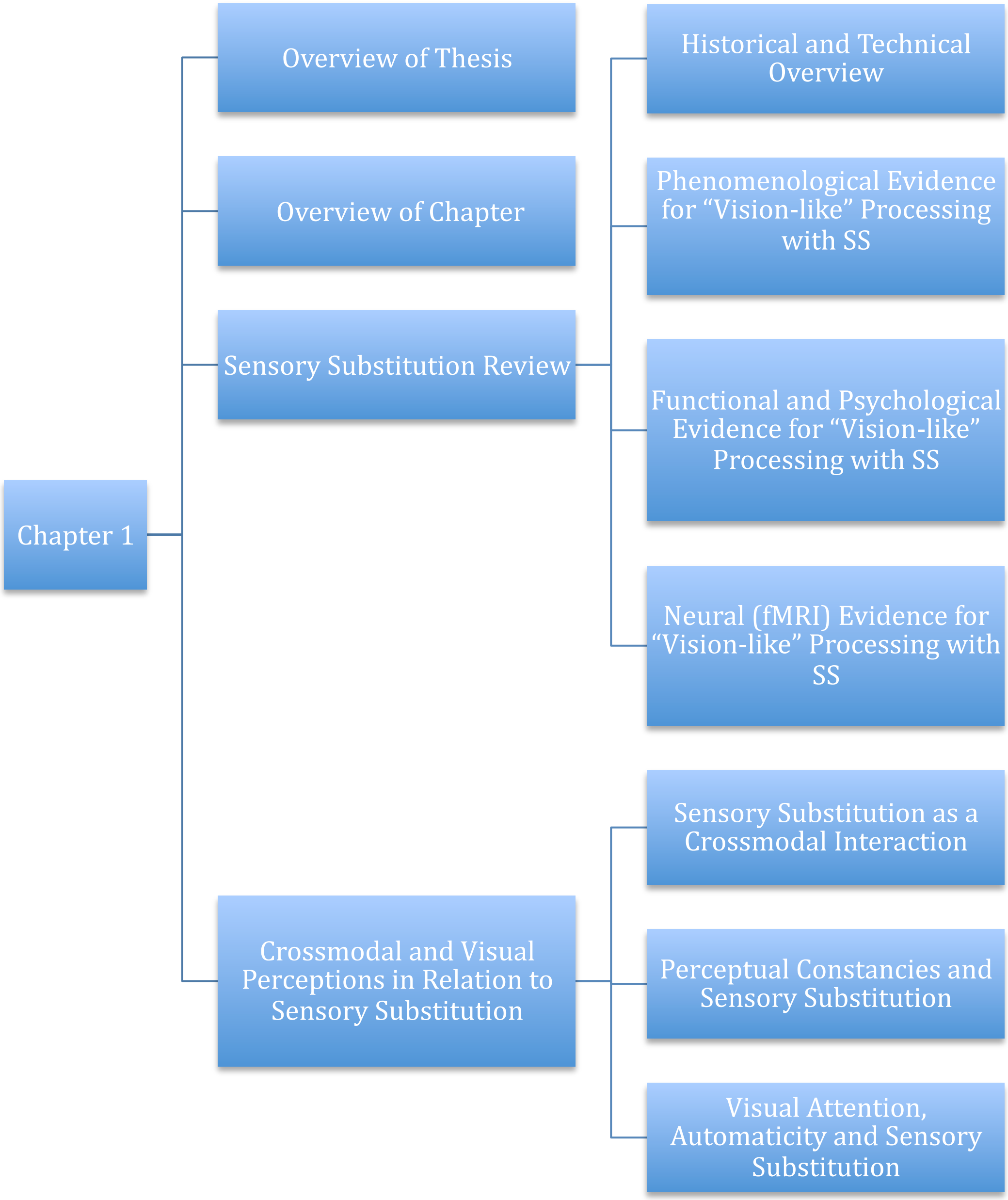


Figure 1.3. Outline of Chapter 1. This figure details the sections of Chapter 1 and their hierarchical structure.

Sensory Substitution Review

Historical and Technical Overview

Sensory substitution was designed as an aid to help the blind recover normal mobility and daily task functionality. Over 250 million people are visually impaired worldwide, with 39 million entirely blind (Visual Impairment and Blindness, 2012). The majority of the blind acquire blindness late in life (Resnikoff et al., 2004), but congenital blindness, or blindness inflicted near birth, still affects 1 out of every 3,300 children in developed countries (Bouvrie & Sinha, 2007). While specialized therapies, surgeries, and medication make most blindness preventable, blindness often cannot be ameliorated after the neural damage is complete. Therefore, several types of electronic prosthetic devices (such as retinal prostheses) have been designed that take over the function of the damaged neural circuitry by stimulating still-functional visual neurons (Humayun et al., 2003; Merabet, Rizzo, Amedi, Somers, & Pascual-Leone, 2005; Stiles et al., 2011; Winter, Cogan, & Rizzo, 2007). However, these devices are invasive, and are still in development. An alternative approach is *sensory substitution*, which encodes visual information into a signal perceived by another still-functional sensory modality, such as somatosensation of the skin or audition. Extensive crossmodal plasticity then enables the brain to interpret the tactile sensations and sounds visually.

Tactile sensation was first used by sensory substitution to transmit visual spatial information. The Tactile Visual Substitution System (TVSS) device used stimulators embedded in the back of a dental chair that were fed video by a camera mounted on a tripod (Bach-y-Rita, Collins, Saunders, White, & Scadden, 1969). With TVSS, six blind participants were anecdotally able to “discover visual concepts such as perspective, shadows, shape distortion as a function of viewpoint, and apparent change in size as a function of distance" (Bach-y-Rita, et al., 1969). TVSS was later modified into the Brainport device that stimulates the tongue surface (Bach-y-Rita, Kaczmarek, Tyler, & Garcia-Lara, 1998) in order to reduce stimulation voltages and energy requirements as well as to utilize the high tactile resolution there.

Audition has also been used for sensory substitution with multiple types of encodings into sound. Early devices, such as the vOICe and PSVA devices, used a direct brightness-to-volume and pixel location to sound frequency transformation. The vOICe device encodes an image by representing vertical position as distinct frequencies, horizontal position as scan time (left to right), and the brightness of individual pixels as volume (Meijer, 1992) (Figure 1.4). The Prosthesis Substituting Vision by Audition (PSVA) device assigns a specific frequency to each pixel, and encodes brightness with volume (Arno et al., 2001; Capelle, Trullemans, Arno, & Veraart, 2002). More recent devices such as the Computer Aided System for Blind People (CASBliP) and the Michigan Visual Sonification System (MVSS) have used 3D sound (encoded with head-related transfer functions) to encode the spatial location of objects (Araque et al., 2008; Clemons, Bao, Savarese, Austin, & Sharma, 2012).

Despite a diverse array of sensory substitution devices, none are currently commercially available or have a large user population. The limited commercial success of sensory substitution is likely due to the long duration (and substantial effort) required to learn a variety of basic visual tasks, and to the limited functionality realized once training is completed. Furthermore, a large part of the training improvement on psychophysical tests appears due to top-down executive control and concentration of attention, even at the intermediate to advanced stages (Browne, 2003; Dunai, 2010; Ward & Wright, 2014). Recent devices, such as the MVSS and CASBliP, hope to increase participant function and decrease training time by changing device encodings from vision-centric to audition-centric. By encoding spatial location in auditory coordinates, these devices exploit existing hardwired processing in auditory cortex while conveying useful information about obstacles. An alternative method to reducing training time and enhancing performance may be improvement of training methods, such as training that exploits intrinsic crossmodal correspondences (Pratt, 1930; Spence, 2011; Stevens & Marks, 1965) to make devices more intuitive, as will be explored in Chapter 3 of this thesis.

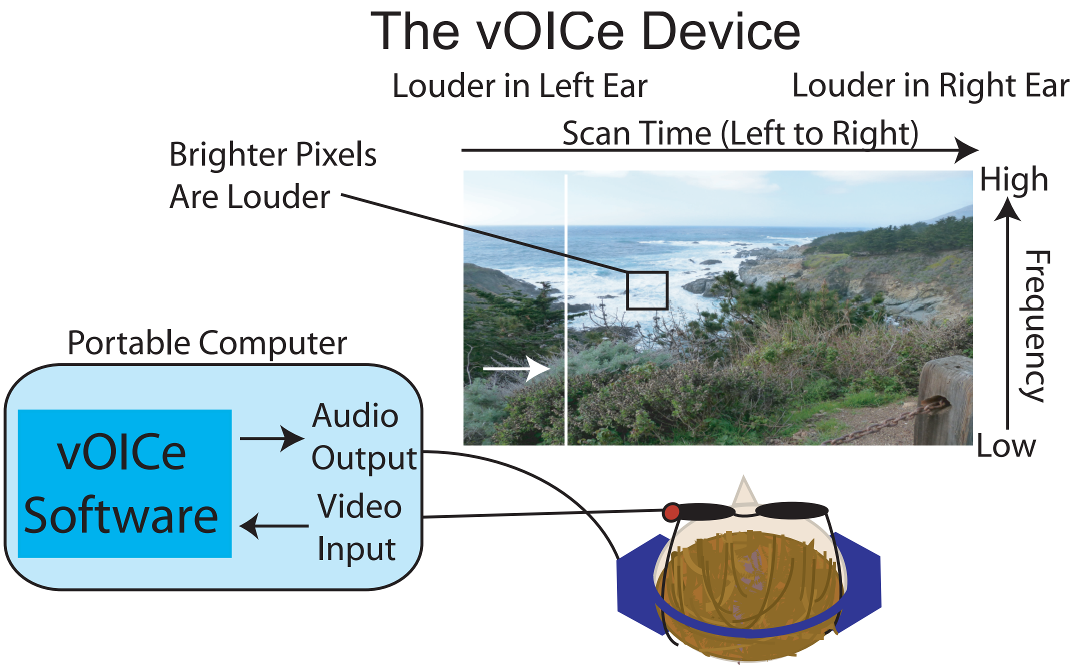


Figure 1.4. Schematic diagram of the vOICe device. A participant wears a pair of glasses with a camera attached that transmits live video to a portable computer. The computer runs the vOICe software, transforming the image into a soundscape by encoding the brightness of pixels into loudness of a sound frequency range that is high for upper pixels and progressively lower for middle and bottom pixels. This column of pixels is scanned across the image at 1 Hz with stereo panning (the scan rate is adjustable). The soundscape representing an image frame is communicated to the user via headphones.

Phenomenological Evidence for "Vision-like" Processing with SS

Sensory substitution generates activation in the primary visual cortex, but may also generate a vision-like perceptual experience, or have visual "qualia" (subjective, conscious quality of perception that can be verbally reported) in select long-term users. (Note that we only refer to the absolute unique quality of subjective perceptual experience here, regardless of whether the neural basis of qualia is a "hard problem" or not, as D. Chalmers has postulated (Chalmers, 1995).) In particular, late-blind vOICe user PF claims to have a visual experience with a sensory substitution device, and to even have color fill-in from previous visual experiences (Ward & Meijer, 2010). PF remembers colors in familiar items such as a strawberry, which she describes as a "red color with yellow seeds all around it and a green stalk"; whereas for unfamiliar objects, her brain "guesses" at the color such as "greyish black" for a sweater, and occasionally reduces the object detail to a line-drawing (Ward & Meijer, 2010). When rTMS was applied to her visual cortex, she claimed to have the visual experience damped, causing her to "carefully listen to the details of the soundscapes" instead of having an automatic "seeing" sensation, qualitatively linking visual activation to "visual" characteristics of the subjective experience (Merabet et al., 2009). The vOICe "visual" experience according to PF:

> *"Just sound?... No, it is by far more, it is sight!... When I am not wearing the vOICe, the light I perceive from a small slit in my left eye is a grey fog. When wearing the vOICe the image is light with all the little greys and blacks... The light generated is very white and clear, then it erodes down the scale of color to the dark black." (Ward & Meijer, 2010)*

Participant PF has not been the only blind user who has reported visual experiences with sensory substitution devices. A study with eighteen blind participants and ten sighted controls found that in the last three weeks of a three-month training period, seven blind participants claimed to perceive phosphenes while using a tactile sensory substitution device (Ortiz et al., 2011). Four out of seven participants with visual experiences retained light perception; they ranged in blindness onset from 1 year old to 35 years old. In most cases, the phosphenes appeared in the shape and angle of the line stimulus tactilely presented; the "visual" perception over time dominated the tactile perception (Ortiz, et al., 2011). The blind group with "visual" experience had activation in occipital lobe regions such as BA 17, 18, and 19 measured via electroencephalography (EEG); in contrast, the non-phosphene blind participants did not have visual activation (Ortiz, et al., 2011).

One critical aspect of the subjective visual experience is externalization, *i.e.,* the brain's strong tendency to perceive visual inputs as external objects as opposed to something like visual images attached to the eyes (Palmer, 1999). Tactile devices have been studied for distal attribution of users (*i.e.,* the externalization of the stimulus) as defined by: *1.* the coupling of participant movement and stimulation, *2.* the presence of an external object, and *3.* the existence of "perceptual space" (Auvray, Hanneton, Lenay, & O'Regan, 2005). Distal attribution was tested on 60 participants naïve to the auditory sensory substitution device and its encoding. Participants moved freely with headphones, webcam attached, and a luminous object in hand, and in some conditions were provided an object to occlude the luminous object. A link between participant's actions and auditory stimulation was often perceived; this coupling perception occurred more often than perception of distal object or environmental space. The coupling sensation between action and perception that participants perceived is perhaps another valuable aspect of the "qualia" of visual perception and sensory substitution. In fact, sensorimotor processing has been argued to be critical to visual awareness (O'Regan & Noe, 2001).

Key questions about "visual" sensations with sensory substitution remain. These include the connection between "visual" perception and functionality with the device, showing whether "visual" quality of experience enhances recognition and localization with sensory substitution. The neural mechanisms underlying visual perception with sensory substitution are also still unclear. Is "visual perception" via sensory substitution just mediated by primary visual areas, or do prefrontal and higher visual cortices play a key role? Further, a quantitative rTMS study of Ortiz's participants that have "visual" experience may show whether the visual activation is necessary for their visual perception of sensory substitution stimuli. Deactivation of prefrontal regions (via rTMS) might demonstrate whether those regions are a part of a top-down cognitive network necessary to the distinctively unique subjective experience of "visual" nature with sensory substitution. These are feasible ideas to be tested in the future.

A major complication in visual activation and "visual" perception with sensory substitution is the role of visualization, or visual mental imagination, particularly in the late blind. The late blind have experienced vision and therefore are more familiar with visual principles, but also have the ability to activate visual cortex via visualization, or a mental effort to visually imagine a scene/object. PF is late blind (blindness onset at age of 21 years), and five out of seven of Ortiz's blind participants with "visual" perception had blindness onset at the age of 4 years or later (Ortiz, et al., 2011). Therefore, it is possible that the visual activation in these late-blind participants is due to top-down cognitive visualization rather than an automatic "visual" perception. The major evidence against mere visualization (as an alternative account) was limited to the qualitative claims that (1) the "visual" perception happens automatically, and (2) (in Ortiz's participants) that tactile sensations fade and "visual" perception dominates. A quantitative study of the automaticity of "visual" perception with a sensory substitution device (*i.e.,* does it occur even when top-down attention is distracted) may further clarify the role of visualization in the sensory substitution "visual" experience. It will no doubt provide empirical seeds for theoretical re-consideration of the subjective aspects of perception, including the issue of "qualia." Visualization as it relates to sensory substitution will be discussed in more detail in Chapter 4.

Functional and Psychological Evidence for "Vision-like" Processing with SS

In order for sensory substitution to be visual, it must also mimic the functional and psychological aspects of vision, or the organization and hierarchy of visual processing, that allow people to interact effectively with their environment. Key to visual functionality is depth perception with monocular depth cues such as perspective (parallel lines converge at infinity), relative size of objects, and motion parallax (lateral movement causes object movement to vary with distance) (Palmer, 1999). Furthermore, perceptual illusions are critical probes into vision-like processing, demonstrating the assumptions necessary to disambiguate a 3D world from 2D retinal images. Vision exhibits perceptual constancies that keep our perception of a given object the same despite varying observations of the environment, which may change the ambient brightness (brightness constancy), object distance (size constancy), color of illumination (color constancy), tilt of the head (rotation constancy), and angle of the object (shape constancy), *etc.* (Palmer, 1999). Finally, effortless localization of objects in simple to cluttered environments and recognition of object properties and categories are critical to visual perception.

Recognition of artificial patterns and shapes has been investigated with tactile and auditory sensory substitution devices with positive results. Bach-y-Rita and colleagues tested five sighted participants on simple shape discrimination (such as circles and squares) with a Tongue Display Unit (a tactile sensory substitution device) (Bach-y-Rita, et al., 1998). Recognition performance averaged at 79.8 percent correct across shapes using arrays of 16, 25, 36, or 49 electrodes, and percent correct also improved with object size (Figure 1.5, A.a., line TO). Poirier *et al.'s* study tested pattern recognition with the PSVA (an auditory sensory substitution device) in blindfolded sighted participants (Poirier, De Volder, Tranduy, & Scheiber, 2007). Patterns were simple combinations of vertical and horizontal bars. Six sighted participants performed above 60% correct on element recognition before and after a training of 2 hours, and above 60% correct after training for pattern recognition (Figure 1.5, A.b.). Simple and complex pattern recognition was studied comparatively with auditory sensory substitution device PSVA in Poirier *et al.*'s behavioral analysis; they concluded that participants recognized the element size and spatial arrangement better than the pattern's element features (such as vertical bars and horizontal bars) (Poirier, Richard, Duy, & Veraart, 2006). Overall, sensory substitution studies show that users can recognize patterns and shapes when they are isolated on a plain background. However, recognition of one shape among many (as is most common in natural vision) has significantly less support.

Specialized object recognition has also been studied. In particular, sensory substitution face perception was investigated with PSVA (auditory sensory substitution device) for similar neural correlates to natural visual face perception, but participant recognition performance was not reported (Plaza et al., 2009).

Natural object recognition was tested in Auvray *et al.*'s 2007 study using the vOICe (auditory sensory substitution) (Auvray, Hanneton, & O Regan, 2007). Ten natural objects (such as a plant, shoe, and table) were identified by six sighted participants in an artificial white background (brightness was inverted before sonification) in an average of 42.4 seconds each (Auvray, et al., 2007). Participants listed 1.6 objects on average before choosing the correct object. The time to identification improved over training (from 57.6 seconds to 34.7 seconds), and varied among object type and individual participants. Categories of objects were studied with the 10 natural objects with 9 additional objects in the same category of an original object. Participants performed above chance at recognizing specific objects even within the same category, and participants were more accurate when there were fewer objects in each category.

A majority of the studies on object recognition with sensory substitution have focused on artificial stimuli in simplified environments. Thus far, no studies yet have explored natural objects in natural environments (such as finding a shirt in a closet, or a clock on a nightstand) or the role of distractor objects to object perception (such as recognizing a object in the center of the field of view, with two objects to the left and right). A potential reason is that artificial patterns are easier to identify, and also can be manipulated to test for sensory substitution resolution as well as quantify objects' complexity relatively easily, with a hope that more cluttered scenes would eventually become recognizable in the progress of training. Several key visual questions, such as spatially segregating objects, object recognition independent of point of view (*i.e.,* shape constancy), and differentiation of shadows and reflections from physical objects, remain unanswered.

Vision is to perceive “what is where by looking” (Marr, 1982). Recognition studies investigated the “what” element of perception, and now, localization studies will highlight the “where” element of vision. Clinically, object localization has been most commonly studied with locomotion through a maze of obstacles. Chebat and his collaborators constructed a life-sized maze consisting of white hallway with black boxes, tubes, and bars horizontal (on the floor or partial protruding from the wall) or vertical (aligned with left or right wall)(Chebat, Schneider, Kupers, & Ptito, 2011). Sixteen congenitally blind and eleven sighted controls navigated the maze with a tactile display unit (10 × 10 pixels), and were scored for obstacle detection (pointing at obstacle) and obstacle avoidance (walk past the obstacle without touching it) (Figure 1.5 B.a.). Congenitally blind were able to detect and avoid obstacles significantly more accurately than the sighted controls. Both groups performed the tasks above chance. Larger obstacles were easier to avoid and detect than smaller obstacles, and step-around obstacles were easier to negotiate than step-over obstacles. Other localization studies have investigated artificial maze environments and tracking of stimuli in 2D and 3D space (Chekhchoukh, Vuillerme, & Glade, 2011; Kupers, Chebat, Madsen, Paulson, & Ptito, 2010).

Studies have also investigated localization via a pointing task, and the value to SS learning of SS device use in daily life. A study by Proulx and colleagues (2008) showed that auditory sensory substitution localization was enhanced when participants were allowed to use the SS device in normal life (in addition to device assessments), compared to participants who only used the device during assessments (Proulx, Stoerig, Ludowig, & Knoll, 2008). Auvray and colleagues (2007) used an auditory sensory substitution device to study the accuracy of localization with a pointing task (Figure 1.5, B.b.) and found that 7.8 cm was the mean error for pointing at 4 cm diameter ball (Auvray, et al., 2007). The pointing inaccuracy varied proportionally with distance to the handheld camera (vertically aligned with the participant's elbow).

Depth perception and illusions are also a key part of visual processing. With sensory substitution's monocular camera and low resolution, it can be especially challenging for users to learn. Nevertheless, sighted users have been found to have key illusions of monocular depth perception, and other visual illusions. As described earlier in this chapter, Renier and colleagues have tested for perception of the Ponzo illusion with a auditory sensory substitution device, and found that blindfolded sighted participants could perceive it similarly to the sighted, but early-blind participants could not (Renier, Laloyaux, et al., 2005). Investigation of the vertical-horizontal illusion (vertical lines appear longer than horizontal lines) showed that sighted participants could perceive this illusion with an auditory sensory substitution device, but early blind participants could not perceive it (Renier, Bruyer, & De Volder, 2006). These results may indicate either that previous visual experience is essential for the perception of certain illusions, or that the duration of training may have been too short or superficial. Testing late-blind participants may further elucidate why congenitally blind participants did not perceive these illusions.

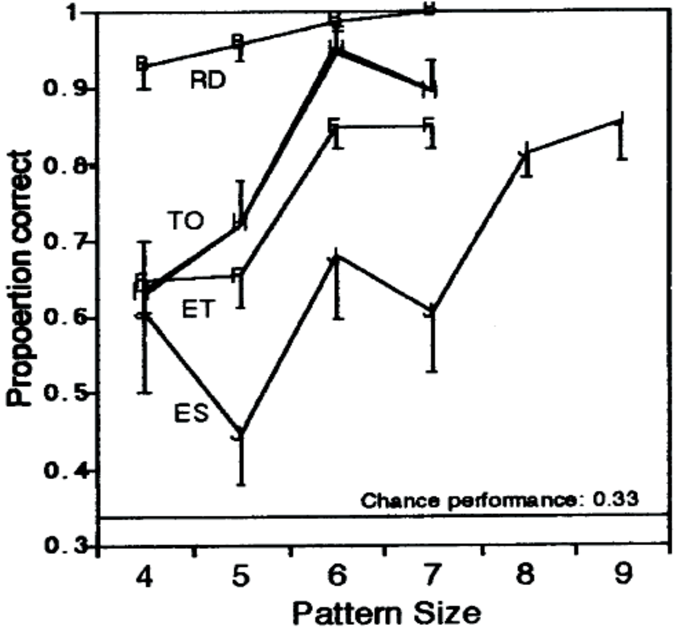
The perceptual organization of sensory substitution perception has many properties yet to be determined. Recognition and localization properties in natural environments are not thoroughly quantified; nor are performances in cluttered environments, or in shadowy and glare-ridden settings. Further questions as to what could be sensory substitution primitives (such as edges or spatial frequencies in vision) have not been answered. Scene perception with sensory substitution is also ambiguous. Questions such as: Can spatial relations of scene be generated with sensory substitution, and How much does it depend on past visual experience and the mode of stimulation (auditory or visual), are still unanswered. The active allocation of attention via gaze is also a critical component of the normal visual function that is entirely absent in sensory substitution encodings. Does the absence of active sensation inhibit the processing of sensory substitution stimuli and the generation of choice? Or instead, would exploration/orienting with the head turn compensate for the gaze shift easily with minimal training? How does the absence of the gaze cascade impact preference in the sensory substitution "visual" experience (Shimojo, Simion, Shimojo, & Scheier, 2003)? Finally**,** Gestalt binding principles of proximity and shared properties may or may not be perceived with sensory substitution, and may be controlled by the transducing modality (somatosensation or audition) or the processing modality (vision).

Neural (fMRI) Evidence for "Vision-like" Processing with SS

Neural imaging and stimulation studies have recently shown visual activation with limited SS device usage in sighted, late blind, and early blind participants. In 2007, Poirier *et al*. reviewed sensory substitution imaging studies, concluding that early blind users use primarily crossmodal plasticity and blindfolded sighted users mainly visual imagery to generate visual activation with sensory substitution use (Poirier, De Volder, & Scheiber, 2007). PET and fMRI studies with tactile and auditory SS devices have shown activation in BA 17, BA 18, and BA 19 with recognition and localization tasks in early and late blind as well as occasionally blindfolded sighted participants (Amedi et al., 2007; Arno, et al., 2001; Kupers, et al., 2010; Merabet, et al., 2009; Poirier, De Volder, Tranduy, et al., 2007; Poirier, De Volder, & Scheiber, 2007; Poirier, De Volder, Tranduy, & Scheiber, 2006; Ptito, Moesgaard, Gjedde, & Kupers, 2005; Renier, Collignon, Poirier, Tranduy, Vanlierde, Bol, Veraart, & De Volder, 2005; Renier & De Volder, 2010; Renier, Laloyaux, et al., 2005). Early PET studies showed activation in occipital cortex for early blind participants, but not for sighted participants (Arno, et al., 2001; Ptito, et al., 2005). fMRI imaging studies later found visual activation with sensory substitution use in sighted participants with pattern recognition and localization, in particular in visual areas within the dorsal and ventral streams (Poirier, De Volder, Tranduy, et al., 2007; Poirier, De Volder, et al., 2006) (Figure 1.6, B).

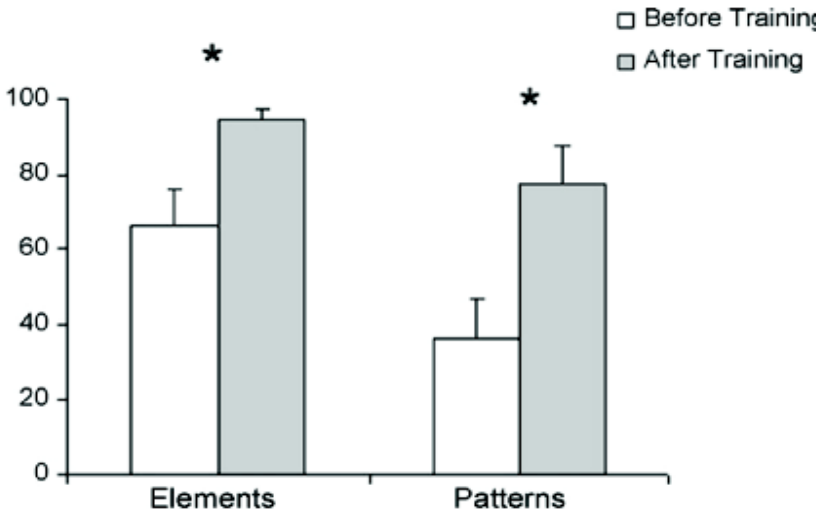
Visual activation due to sensory substitution use has also been shown to be functionally correlated to the task performed during scanning. Amedi and colleagues showed with fMRI imaging that the lateral occipital tactile-visual (LOtv) area known to interpret object shape was also activated by auditory sensory substitution device usage during a shape task (Amedi, et al., 2007) (Figure 1.6, A). Plaza and collaborators in 2009 demonstrated that PSVA could activate the fusiform face area with face stimuli in blindfolded volunteers (Plaza, et al., 2009). Renier *et al.* investigated depth perception with a SS device, and found that blindfolded sighted participants could perceive the Ponzo illusion and had activation in occipito-parietal cortex while exploring 3D images with PET imaging (Renier, Collignon, Poirier, Tranduy, Vanlierde, Bol, Veraart, & De Volder, 2005; Renier, Laloyaux, et al., 2005). Plaza and collaborators in 2012 compared neural activation for an orientation and localization task using images encoded into sound or presented visually (Plaza, Cuevas, Grandin, De Volder, & Renier, 2012). Sighted participants' neural activation was stronger in the right superior parietal lobule (BA 7) for the localization task in comparison to the orientation task in both auditory sensory substitution and vision.

Figure 1.5.A.a. Pattern Recognition, Tactile Sensory Substitution (Bach-y-Rita, et al., 1998)



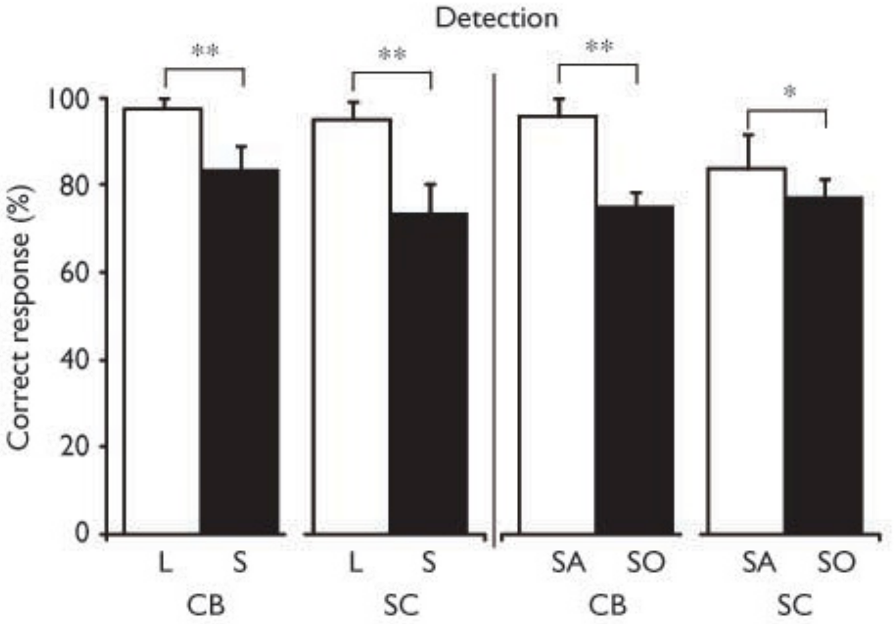
RD: Fingertip perceived raised dots,
TO: Electrotactile tongue discrimination
ET: Fingertip electrotactile discrimination
(participant dynamically modulate current),
ES: Fingertip electrostatic stimulation

Figure 1.5.A.b. Pattern Recognition, Auditory Sensory Substitution (Poirier, De Volder, Tranduy, et al., 2007)



** Statistically significant difference between before and after training (Elements: Wilcoxon test for paired samples: $Z = 1.99, p < 0.05$; Patterns: Wilcoxon test for paired samples: $Z = -2.23, p < 0.03$)*

Figure 1.5.B.a. Object Localization, Tactile Sensory Substitution (Chebat, et al., 2011)



CB: Congenitally Blind, SC: Sighted Controls,
L: Large Object, S: Small Object,
SA: Step-Around Obstacle, SO: Step-Over Obstacle
(*$p \leq 0.05$; **$p \leq 0.001$)

Figure 1.5.B.b. Object Localization, Auditory Sensory Substitution (Auvray, et al., 2007)

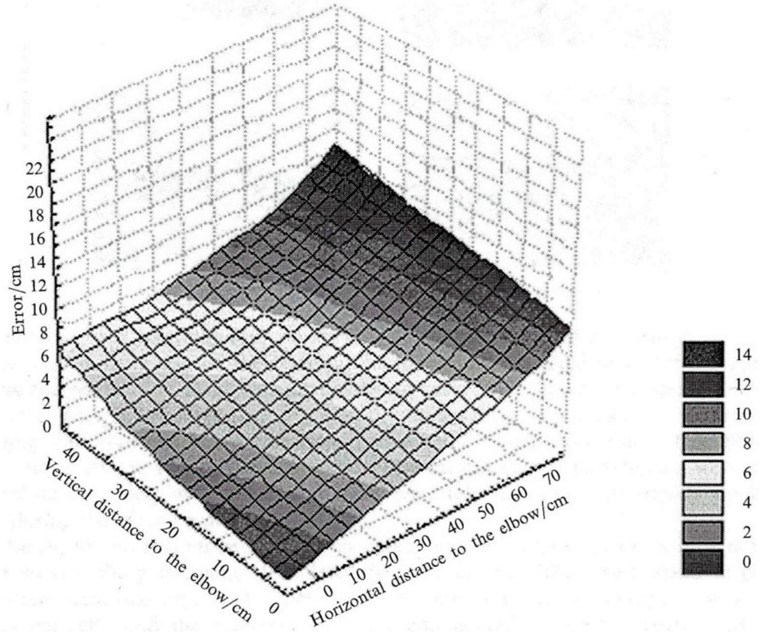
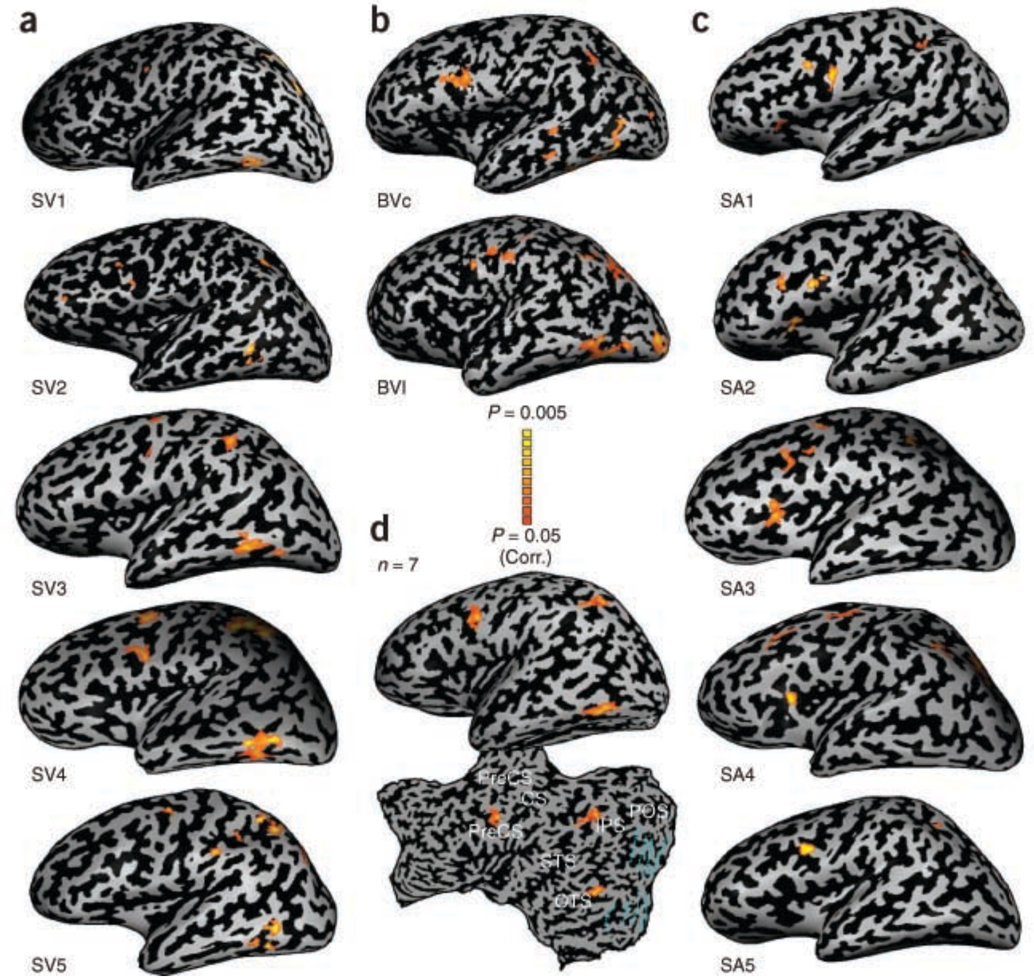


Figure 1.5. Behavioral outcomes of sensory substitution training. Psychophysical testing with tactile and auditory sensory substitution devices has had similar outcomes. Object recognition testing with the Tongue Display Unit (A.a.) has shown a correlation between the pattern size and the proportion correct; all participants exceeded the chance performance. Pattern recognition with an auditory device (A.b.) significantly improved with training, and had a similar average percent correct as tactile pattern recognition (between 0.6 and 0.8 proportion correct). Obstacle localization in an uncluttered maze environment with a tactile device (B.a.) had between 0.8 and 1 proportion correct for most object types. Localization of a 4 cm diameter ball with an auditory device showed that inaccuracy increased with distance to the object (webcam to view environment was held in the right hand and aligned with the elbow). (Auvray, et al., 2007; Bach-y-Rita, et al., 1998; Chebat, et al., 2011; Poirier, De Volder, Tranduy, et al., 2007)

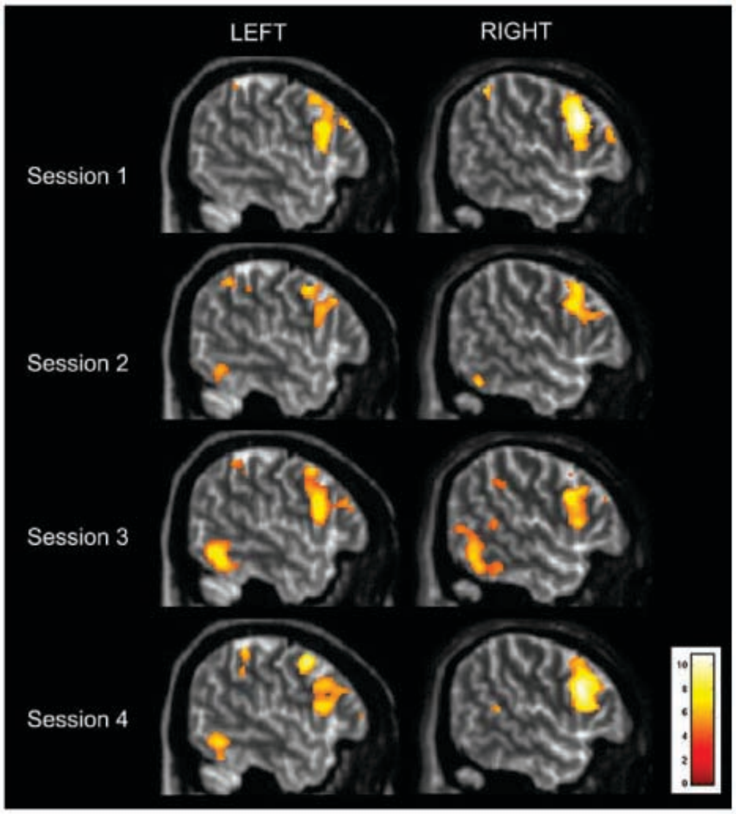
Even non-sensory-substitution binding of crossmodal stimuli can generate visual activation from unimodal stimuli. Zangenehpour and Zatorre found that training on the spatial and temporal congruence of beeps and flashes activated visual cortex even in the auditory-only condition (Zangenehpour & Zatorre, 2010). Therefore, visual cortex can be trained to respond to audition if the participants are taught to associate temporally and spatially co-located beeps and flashes. This indicates that a critical part of training-induced plasticity is simultaneous stimulation of sensory substitution (audition or somatosensation) and vision (for sighted participants), potentially due to Hebbian learning. Hebbian learning can also be potentially extended to the blind if stimuli are felt by the hand simultaneously with stimulation by sensory substitution.

Figure 1.6A. Activation in Blind and Sighted with a Shape Estimation Task (Amedi, et al., 2007)



a. Single sighted participant's neural activation, b. Blind participant neural activation, c. Single sighted participant activation from auditory control task, d. Average across seven vOICe trained users (participants in a and b).

Figure 1.6B. Sighted Participant Activation as a Function of Training Session on a Pattern Recognition Task (Poirier, De Volder, et al., 2006)

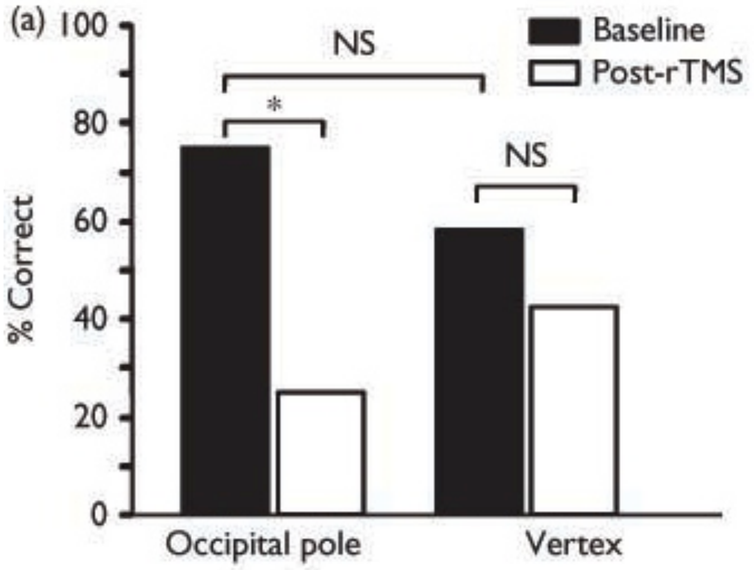


Voxels corrected for multiple comparisons in the whole brain and threshold exceeding $p < 0.05$. *Six sighted participants.*

Figure 1.6. Imaging with sensory substitution. Neural activation was shown on the left occipito-temporal cortex in all sighted and blind expert users during sensory substitution shape classification (A.a. – A.b.), whereas sighted users did not have visual activation with auditory control task (A.c.). Averaged results show activation in several multimodal regions (A.d.). During a sensory substitution pattern recognition task, six sighted participants showed a progressive increase in occipital activation with training on an auditory sensory substitution device (B.) (Amedi, et al., 2007; Poirier, De Volder, et al., 2006).

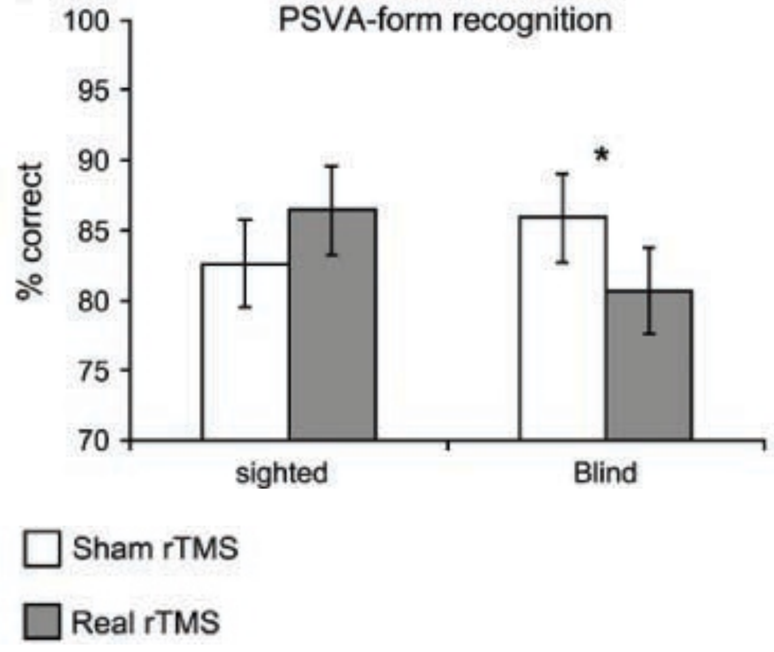
fMRI and PET studies have demonstrated that visual cortex activation correlates with sensory substitution use, but cannot prove causality. Repetitive Transcranial Magnetic Stimulation (rTMS) deactivates a region of cortex, examining the possible causal link between neural activation and participant performance. Collignon and colleagues applied rTMS to the right dorsal extrastriate occipital cortex of seven sighted and seven early blind participants (both trained on the PSVA auditory sensory substitution device) preceding sensory substitution pattern recognition (Collignon, Lassonde, Lepore, Bastien, & Veraart, 2007). Early blind participants had longer reaction times and lower accuracies with rTMS applied as compared to a sham rTMS condition; sighted participants had no performance change (Collignon, et al., 2007) (Figure 1.7, B). Merabet *et al.* also deactivated with rTMS occipital peristriate regions of a late blind sensory substitution superuser, PF, and demonstrated a decrement in recognition accuracy relative to pre-rTMS and post-sham rTMS conditions (Merabet, et al., 2009) (Figure 1.7, A). In the tactile domain, TMS applied to occipital cortex elicited somatotopic tactile sensations in blind but not blindfolded sighted users of a tactile sensory substitution device (Kupers et al., 2006). Overall, rTMS studies indicate that the blind users of sensory substitution devices functionally and causally recruit the occipital cortex, potentially due to long-term crossmodal plasticity from visual deprivation.

Figure 1.7A. rTMS on a Late Blind Auditory Sensory Substitution Expert (Merabet, et al., 2009)



*NS: Not Significant, *: $p < 0.05$*

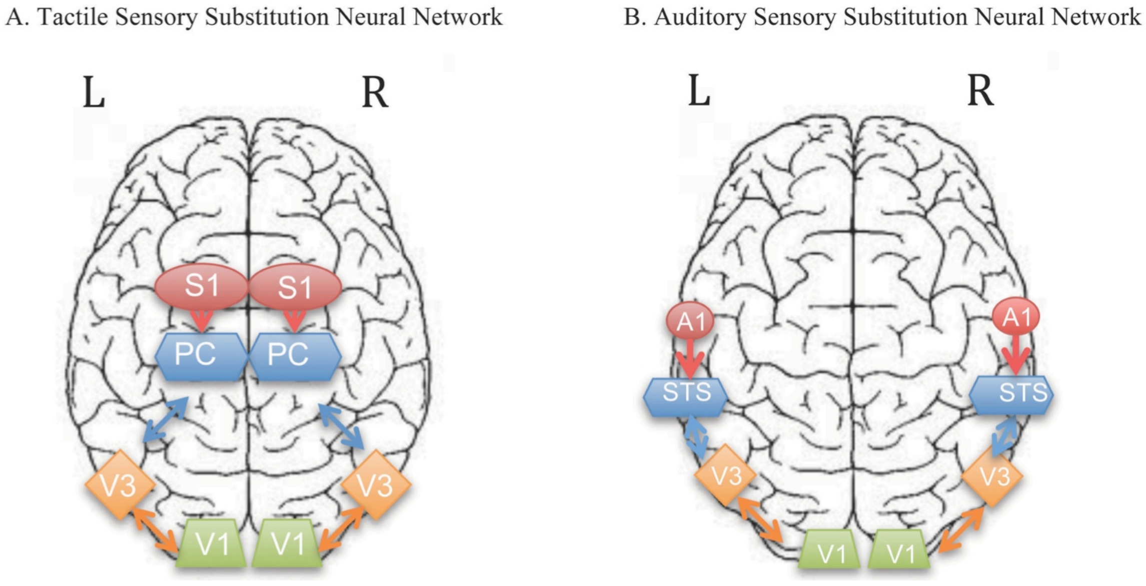
Figure 1.7B. rTMS on Early Blind Auditory Sensory Substitution Users (Collignon, et al., 2007)



*: $p < 0.05$, *Error bars indicate standard errors.*

Figure 1.7. rTMS with sensory substitution. Repetitive Transcranial Magnetic Stimulation (rTMS) decreases neural activation and influences behavior, thereby generating a causal link between behavioral outcomes and neural region activation. rTMS of an occipital region significantly reduced percent correct at object identification in an expert vOICe user, PF (A.). PF's recognition was not significantly impaired by rTMS of a vertex location. Seven early blind participants were also impaired at the sensory substitution pattern recognition task with rTMS to right dorsal extrastriate occipital cortex (B.). Seven sighted participants' performance was not significantly affected by rTMS (B.) (Collignon, et al., 2007; Merabet, et al., 2009).

Dynamic Causal Modeling (DCM) studies in the blind have constructed a crossmodal network for auditory and somatosensory connections to the visual cortex (Fujii, Tanabe, Kochiyama, & Sadato, 2009; Klinge, Eippert, Roder, & Buchel, 2010). It remains to be shown whether these networks are used in blind participants with sensory substitution, and whether the crossmodal network in the sighted is similar to, or different from blind participants. Nevertheless, literature on functional connectivity of sensory substitution "stimuli" and dynamic causal modeling of the blind can be used to generate several neural network possibilities (Figure 1.8, A and B) with feedforward and feedback connections. The network likely includes the primary sensory region of the transducing modality (somatosensation or audition), which connects to a multimodal region that further connects to primary visual regions (V3, V2, or V1). The filtering of stimuli as sensory substitution stimuli or natural stimuli could occur at the primary region of transducing modality (A1 or S1) or the multimodal region. More studies on the specificity of the plasticity would be required to elucidate this. The role of prefrontal regions in top-down cognitive processing of the crossmodal stimulus has yet to be shown. More critically, which specific regions in the network are casually linked to performance, and therefore the roles that the regions play in stimulus processing, have yet to be fully determined. Feedback between visual regions and the multimodal regions may play a significant role in stimulus processing, yet the degree of feedback in sensory substitution processing is unclear. Motor regions and other primary sensory regions may also play an important role in plastic changes in the sensory substitution neural network.



A. Tactile Sensory Substitution Neural Network

B. Auditory Sensory Substitution Neural Network

Figure 1.8. Neural network with sensory substitution. Visual, auditory, and tactile regions generate a neural network in blind and sighted sensory substitution users that processes sensory information within a feedforward and feedback hierarchy (A. for tactile devices and B. for auditory devices) (Poirier, De Volder, & Scheiber, 2007). The sensory information is first filtered by primary sensory regions (A1 or S1 for auditory and tactile devices, respectively). Sensory information is then communicated to multimodal regions (such as STS or Parietal Cortex [PC]) and forwarded to primary visual regions (V3, V2 [not shown] or V1). It is also likely that feedback and reiterative processing play a role in the perception of the sensory substitution stimuli.

Structural connectivity between and within sensory regions has also been measured in the blind and sighted using MRI Diffusion Tensor Imaging (DTI) and Diffusion Tensor Tractography (DTT). Shimony *et al.* found that early blind individuals maintained their white matter tracts between visual cortex and orbital frontal and temporal cortices (Shimony et al., 2006). Shu *et al.* used DTT and found that early blind participants had reduced connectivity compared to sighted controls (Shu et al., 2009). Overall, white matter neural imaging indicates interesting similarities and differences between the sighted and blind populations.

Crossmodal and Visual Perceptions in Relation to Sensory Substitution

Sensory Substitution as a Crossmodal Interaction

Regardless of the specific encoding employed, sensory substitution is intrinsically crossmodal, as the information from the transducing modality is communicated to visual cortex for processing by means of either intrinsic mapping (such as matching of high pitch and high spatial location) or neural plasticity engendered through training. The crossmodal interactions utilized by sensory substitution exist as both hardwired developmental connections and plasticity-induced changes in adulthood.

As an example of a more hardwired crossmodal interaction, the Illusory Flash or Double Flash Illusion (in which a single flash accompanied by two short sounds is perceived to be doubled) seems to be lower-level-sensory, since the illusion is relatively immune to at least certain cognitive factors, such as a feedback and reward (Andersen, Tiippana, & Sams, 2004; Mishra, Martinez, Sejnowski, & Hillyard, 2007; Rosenthal, Shimojo, & Shams, 2009; Shams, Kamitani, & Shimojo, 2000). This illusion demonstrates that the modality carrying the more discontinuous (and therefore salient) signal becomes the influential or modulating modality (Shams, Kamitani, & Shimojo, 2002; Shimojo & Shams, 2001). It has also been shown that a wide variety of crossmodal information is combined such that the resulting variance is minimized, thereby mimicking maximum likelihood estimation (MLE) models (Ernst & Banks, 2002). Ernst and Banks were able to conclude from MLE that the modality that dominates in crossmodal information integration is the one with the lowest variance.

Crossmodal intrinsic mappings also bridge modalities with learned spatial, linguistic, or emotional connections (Spence, 2011). Many visual and auditory mappings have been studied for over 50 years, including relating high visuo-spatial position with high pitch and vice versa (Pratt, 1930). Other audio-visual mappings include the matching of loudness with brightness (Stevens & Marks, 1965) and the matching of visual spatial frequency and auditory amplitude modulation rate (Guzman-Martinez, Ortega, Grabowecky, Mossbridge, & Suzuki, 2012). Mappings also exist between vision and tactile sensation (Spence, 2011), including between darkness and weight (Walker, Francis, & Walker, 2010), and color and temperature (Morgan, Goodson, & Jones, 1975). These are likely prior to any associative learning via experiences, considering that these correspondences are common across various natural languages (from different origins) with very few exceptions. In fact, preverbal infants of only 3 to 4 months have been shown to have correspondences between high auditory pitch and high spatial position as well as visual “sharpness” (Walker et al., 2010). Nevertheless, new relations between senses can be learned with training (Ernst, 2007).

Modalities are also plastic after development and can generate learned (or trained) relations across senses, as witnessed in visual activation during echolocation, sound localization, and braille reading in the blind (late blind *vs*. early blind) (Bavelier & Neville, 2002; Cohen et al., 1997; Collignon, Voss, Lassonde, & Lepore, 2009; Sadato et al., 1996). Braille reading activated primary visual cortex (BA 17) and extrastriate cortices bilaterally in blind participants (Sadato, et al., 1996). Repetitive Transcranial Magnetic Stimulation (rTMS) was used to deactivate visual cortical regions in blind braille experts, and generated errors in braille interpretation (Cohen, et al., 1997). In fact, an early-blind individual lost the ability to read braille after experiencing a stroke in the occipital cortex (Hamilton, Keenan, Catala, & Pascual-Leone, 2000). These results demonstrate a functional and causal link between visual activation and the ability to read braille in the blind. Other studies provide even more evidence for plasticity in the handicapped, such as enhanced visual ERPs (Event Related Potentials) in early-onset deaf (Neville & Lawson, 1987; Neville, Schmidt, & Kutas, 1983), auditory ERPs in the posterior (occipital) region in early and late blind (Kujala et al., 1995), and posterior DC potentials in blind by tactile reading (Uhl, Franzen, Lindinger, Lang, & Deecke, 1991).

Further, due to plasticity-induced changes, it has been proposed that the brain, including the visual cortex, may be “metamodal,” such that brain regions are segregated by processing of different types of information, and not by stimulus modality (Pascual-Leone & Hamilton, 2001). The metamodal theory of the brain was supported by the activation of the shape-decoding region, Lateral Occipital tactile-visual area (LOtv), by audition when shape was conveyed by vOICe encoded sounds (Amedi, et al., 2007).

Crossmodal perceptual organization usually refers to Gestalt principles, such as proximity-based (both in space and time) grouping/segregation, regularity, and Prägnanz (good shape). Vision, audition, and somatosensation have partly the same, but partly different (unique) perceptual organization rules. For example, segregation or chunking rules operate across modalities in the same way at the most abstract level, but indeed, it could be spatial in vision but temporal in audition (Bregman & Campbell, 1971; Neri & Levi, 2007; Vroomen & de Gelder, 2000). SS provides an opportunity to investigate what would happen to such perceptual organization rules when between-modality connectivity is enhanced by training. To be more specific, questions including: A) would the auditory or the tactile modality acquire vision-like perceptual organization rules and B) would crossmodal combinations themselves self-organize and generate new crossmodal organization principles, can be investigated in detail with sensory substitution.

Existing literature on crossmodal interactions is a guide to understanding and interpreting the visual nature of sensory substitution processing. Sensory substitution also requires plastically generating new learned relationships across modalities, but it may also rely on existing developmental connections. In fact, SS might modulate the strength of existing developmental connections, and thereby alter crossmodal perception, even in sighted participants. Ideally, the training of participants can exploit these existing crossmodal interactions and mappings to enable effortless training and signal interpretation. In addition, training on SS devices should take into account crossmodal interaction variance across both functional and experimental participant groups, including the early blind with no visual experience, the late blind who have limited visual experience, and the sighted with normal visual perception (Bavelier & Neville, 2002; Poirier, De Volder, & Scheiber, 2007).

Perceptual Constancies and Sensory Substitution

Constancies (the perception of a unchanging environment despite various observation conditions such as head movement, lighting changes, and distance alteration) are critical to the perception of stimuli as distal (within the external environment) rather than proximal (on the retinal surface) (Palmer, 1999). This distal nature of perception conferred by constancies is equivalent to the externalization of perception. It is critical to perceive the world distally, as it is useful for the functional engagement of the objects in the environment. For example, if a person would like to pick up a dropped coin, it is easier to recognize the coin if its color and brightness is constant whether it is under the shadowy desk or brightly illuminated in your hand. Further, it is easier to pick up the coin if the size of the coin appears to be constant independent of whether it's closer (such as in your hand) or farther away (such as under the desk). There are several types of visual constancies such as size constancy (constant size independent of distance), shape constancy (constant shape independent of object rotation), orientation constancy (constant object orientation independent of head tilt), position constancy (constant object position independent of head and body movement), and brightness constancy (constant brightness independent of external illumination) (Palmer, 1999). Auditory perceptual constancies exist as well, such as: Loudness constancy (constant loudness independent of distance, like size constancy) (Zahorik & Wightman, 2001) and rotation constancy (constant sound location independent of head rotation, similar to position constancy) (Day, 1968).

Two general ideas dominate the visual perceptual constancy theory: the indirect idea originated by Helmholtz, and the direct idea generated by Gibson. The indirect approach claims that constancy is constructed by using the retinal image information in combination with other additional information (optical and nonoptical), such as accommodation in size constancy (von Helmholtz, 1925). Additional information for constancy processing can include head angle for orientation constancy, or external lighting properties for brightness constancy. In fact, for position constancy, it has been argued that an efferent copy of movement commands is sent to brain regions for constancy computation. In contrast, the direct approach uses image properties that are invariant to determine constancy (Palmer, 1999). For example, position constancy can be generated using optic flow calculations based on the retinal image. In particular, Gibson argues that some constancy properties such as size and distance in size constancy are unavoidably linked and inseparable (Gibson, 1950a).

While people can perceive object constancies, they can also perceive object changes. These two types of perception (constancies and changes) are known as the distal and proximal modes (respectively) (Palmer, 1999). The distal mode (*i.e.,* using constancy) is useful for functional tasks and interaction with environmental objects (as discussed above). Therefore, the distal mode frequently dominates in daily life tasks. The proximal mode, or retinal image perception, can be occasionally useful for select activities such as realistic painting or the detection of an illusion due to distal (or constancy) assumptions. Occasionally, individuals use the proximal mode of vision for these tasks, and can even perceive constancies and changes simultaneously.

Constancies processing has yet to be tested with vOICe or with sensory substitution devices. If it can be proven that participants can learn constancy with sensory substitution, it will be an important indication of perceptual externalization and distal perception with sensory substitution devices. Further, constancy will be important to the comparison of sensory substitution processing to purely visual and purely auditory processing capabilities.

Visual Attention, Automaticity, and Sensory Substitution

Vision is well known to be automatic, requiring limited attention and cognitive processing for complex tasks such as visual search (A. Treisman, 1985) or the processing of faces (Palermo & Rhodes, 2007). Overt visual attention is often expressed with eye and head orienting movements that bring the object of interest into view for inspection, such as with facial feature evaluation. However, attention can also be exerted by internal mechanisms of focus that highlight different object properties or spatial regions, called property selection and spatial selection, respectively (Palmer, 1999). In addition, attention can be oriented based on exogenous cues (cues from the environment, such as a loud sound) or on endogenous cues (internally generated cues) (Posner, 1980).

Attention selection and processing without attention (automaticity) has been studied in audition and vision with distraction paradigms. The goal of distraction is to focus attention on an unnecessary task, thereby limiting the attention that can be used in the task of interest. In audition, distraction tasks were performed using repetition of words communicated through the left or right side of a pair of headphones (Cherry, 1953; Palmer, 1999). Then participants were tested for the information heard in the unattended headphone. Participants were able to detect general sensory information in the unattended ear, such as the voice gender or if it was speech, but not specifics such as the language or content. Further experiments showed that personal information could also be detected in the unattended ear, such as the individual's name. The combined early and late processing required for general sensory information detection as well as detailed personal information detection generated the attenuator theory of auditory processing. This theory claims that attention is allocated for early processing information of interest as well as highly salient representations of words (Palmer, 1999; A. M. Treisman, 1960). The threshold for conscious awareness of words processed is then dynamically regulated based on an individual's value of that word. Visual distraction paradigms were designed with a similar distraction task (such as relative length of two presented lines) and with an unattended visual stimulus of interest. When questioned about the unattended object, participants could identify the location, color, and number of objects, but not the shape (Palmer, 1999; Rock, Linnett, Grant, & Mack, 1992). However presentation of the object's name did allow for most participants to identify it although it was unattended (Palmer, 1999). Therefore, similarly to audition, a combination of early and late processing is allocating attention in vision.

Several automaticity-imaging studies have indicated that visual stimuli are still processed in the brain despite being cognitively ignored, via a distraction paradigm. Schwartz and colleagues showed that the attention load of the central task modulated the visual activation of the ignored peripheral stimuli (Schwartz et al., 2005). Further, Schwartz *et al.* indicated that peripheral regions neighboring the central visual regions (processing the attentive task) were suppressed more by high attention load than distant peripheral cortex. Several neuroscientists have investigated emotional processing of face stimuli with limited attention, and have obtained mixed results (Pessoa, 2005; Pessoa, McKenna, Gutierrez, & Ungerleider, 2002; Vuilleumier, Armony, Driver, & Dolan, 2001). Pessoa claims that these mixed results are in fact logical, if cognitive load is considered, with the high cognitive load reducing the emotional vision processing the most, and vice versa (Pessoa, 2005). However, in the words of R. Palermo, "although neural responses in face-selective cortex are reduced for ignored compared with attended faces, they are certainly not eliminated" (Palermo & Rhodes, 2007). Results seem to indicate that unattended visual stimuli have reduced activation in visual cortex relative to attended stimuli, but that nevertheless unattended stimuli are still processed in visual regions.

While vision automaticity (processing without active attention) has its limitations, it does occur for simple image/sound properties, and complicated properties of personal value. These visual automaticity properties are important to the hierarchy of visual processing and will be important for the comparison of vOICe to visual processing. Unfortunately, the current sensory substitution imaging studies all use tasks with active attention and cognitive processing. No distraction task paradigms with SS have been tested with psychophysics or brain imaging. Therefore, the intensity of attention required and the different roles attention may play in interpreting vOICe, or other sensory substitution devices remains unknown. While it has been assumed that vOICe is processed entirely cognitively with active attention resources, this may not entirely be the case. Elements of vOICe processing, mediated by existing crossmodal interactions and connections, may be processed without attention in a similar manner to that in which some elements of vision are processed. Automaticity and its relevance to sensory substitution will be discussed further in Chapters 3 and 4.