CHAPTER 2:

LENGTH AND ORIENTATION CONSTANCY LEARNING WITH AUDITORY SENSORY SUBSTITUTION

Introduction

Chapter 1 discussed that depth illusions (Renier, Laloyaux, et al., 2005), as well as natural and artificial object identification and localization (Amedi, et al., 2007; Bach-y-Rita, et al., 1998; Poirier, De Volder, Tranduy, et al., 2007; Proulx, et al., 2008) including face and word identification (Plaza, et al., 2009; Striem-Amit, Cohen, Dehaene, & Amedi, 2012) have been shown to be learned by SS users. Whether it only indicates sensory-motor learning, or rather adaptive changes of some intrinsically-visual quality/function, remains unsolved. A basic problem with SS is that perceptual constancy, the ability to perceive a feature as constant despite changes in a dynamic visual scene, has not been investigated or shown to be learned with SS. As detailed in Chapter 1, constancies are a critical element of perception that are important to functional task performance as well as accurate environmental perception (*i.e.,* externalization). If constancies can be learned, it would indicate the potential for SS users to attain a high level of functional capability as well as SS to behave as a perceptual modality (like vision or audition alone). Therefore, we used the vOICe auditory SS device to demonstrate that the sighted and blind can learn orientation and length constancy tasks, and that this learning is amplified by dynamic interaction with stimuli, providing further insight into how visual-motor experiences shape perceptual constancies in general.

Visual orientation constancy is the ability to estimate the angle of an object independent of head tilt. This constancy is useful for determining an object's angle relative to gravitation (Palmer, 1999). It would be particularly useful to the blind SS users, as it would allow them to determine object stability via the object's tilt, thus perceiving it as an external object independent of their own locomotion/movement. The stability of tables, chairs, and other furniture as well as natural objects such as rocks, trees, and branches is important mobility information for the blind. It is also useful for obstacle avoidance of leaning objects whose position in space (determined via orientation constancy) is critical to locomotion around them. In particular, orientation constancy allows detection of low-hanging branches that often hang at an angle, and are undetectable by a cane grazing the ground. If a branch angle were misinterpreted by the lack of orientation constancy, than the blind user of a SS device would collide with that branch. Therefore, orientation constancy is particularly relevant to SS users, and is valuable to daily functioning with SS.

Visual orientation constancy is generated by proprioceptive information of head orientation, kinesthetic feedback, and visual frame of reference, allowing for the correction of tilted images due to head tilting (Palmer, 1999). Object orientation is first identified by V1 orientation sensitive cells that detect the angle of lines as they appear on the retina. However, if these cells alone determined perceived orientation, then objects would rotate with head motion. To remove head motion from the perceived angle of objects, vestibular organs sense head angle (via utricle and saccule organs) and changes in head angle (via semicircular canals). Further, kinesthetic feedback provides information about body movement (from sensors in joints) that is used in orientation constancy. Finally, visual frame of reference or context informs orientation constancy. Frame of reference is a set of heuristics of the typical vertical angle of particular objects (such as walls forming a room). Frame-of-reference violations can even generate illusions (such as a tilted room illusion) that override proprioceptive information (Shimojo, 2008). In particular, a rod and frame illusion causes a central line or rod to be perceptually tilted by the context of an external tilted rectangle or frame. Due to frame of reference, in a lit environment orientation constancy is quite robust; however, in a dark room, an approximate error of 10 percent can occur at large head tilts due to the loss of visual context (Palmer, 1999). Overall, both non-visual information and environmental visual context contribute to the accurate perception of orientation constancy.

The cortical processing of orientation constancy has been studied with a variety of imaging techniques and lesion patients. Corbett and colleagues used event-related potentials (ERPs) to measure the neural processing underlying the rod and frame illusion (Corbett, Enns, & Handy, 2009). They found that neural processing represented by P3, characterizing later processing, mediated orientation constancy. A lesion case study indicated a "room tilt illusion" for a patient with posterior cortical atrophy. The scientists hypothesized that the illusion was due to disordered processing of vestibular and visual inputs (Crutch et al., 2011). Finally, Denny and Adorjant showed that electrophysiological responses of cat primary visual cortex were modified by head rotation (Denney & Adorjanti, 1972). Overall, these results indicate later visual processing of orientation constancy as well as possible feedback modulation of earlier visual regions for constancy (in cats).

Length constancy (the ability to estimate length independent of object angle) is a sub-type of shape constancy (the ability to estimate object shape independent of perspective). Length constancy is particularly difficult to accomplish with the vOICe SS device, due to the vOICe image-to-sound encoding of sound frequency in the vertical dimension and scan-time laterally (Meijer, 1992). Horizontal line length is encoded by duration of the sound, whereas the vertical line length is encoded by the range of frequencies in a very brief sound. Inevitably, lines of different angles but the same length are not only perceptually quite different with vOICe, but the computation of length estimation is different as well. In vision, the retina can estimate line length with the same neural computation in any angular dimension. Therefore, it will be particularly interesting if participants can overcome these challenges to learn length constancy with the vOICe device.

Shape and length constancy have been investigated extensively in visual perception. Two-dimensional shapes rotated in three dimensions have been shown to have robust visual shape constancy in both adults and infants (Palmer, 1999). This result indicates that shape constancy is surprisingly innate but also dependent on adequate depth perception. Length constancy of 2D objects rotated in the 2D plane (used in this experiment) is a simpler case, which largely avoids this depth perception dependence. In general, shape constancy is constrained by a tendency toward symmetrical shape perception, which aids in shape constancy of symmetrical shapes and limits constancy of non-symmetrical shapes. Shape constancy of 3D wire objects and unfamiliar opaque 3D objects is surprisingly limited, and is an example of proximal perception (discussed in Chapter 1). Despite this, shape constancy of 3D familiar objects seems to be easy for sighted people. For example, recognition of a banana from 4 different perspectives is straightforward for the typical sighted individual. The ease of daily-life shape constancy could derive from familiarity with the object, or from gradual shape changes as an object moves, or even from identifying axes of symmetry. Overall shape constancy of simple 2D shapes is trivial for the sighted, whereas constancy of 3D arbitrary shapes is difficult and problematic.

fMRI imaging studies of shape constancy have indicated that it is processed in several visual cortical brain regions. Vuilleumier and colleagues used a neural fatigue paradigm with fMRI imaging to show that the left fusiform region decreased in activity in response to repeated stimuli of varying stimuli viewpoints and sizes (Vuilleumier, Henson, Driver, & Dolan, 2002). This fatigue paradigm indicates that the left fusiform is processing visual information from multiple perspectives (*i.e.,* view invariant), a required element of shape constancy. In contrast, the right fusiform activity was only fatigued in response to repeated stimuli from a single vantage point. Kourtzi *et al.* also studied 3D shape perception with a fatigue fMRI design, but their experiment was focused on the lateral occipital complex (LOC), a region known to process shapes (Kourtzi, Erb, Grodd, & Bulthoff, 2003). It was determined that the LOC fatigued in response to objects with the same 3D shape but different 2D shapes. The LOC did not fatigue in response to objects with the same 2D shapes but different 3D shapes. In other words, the LOC was coding for 3D shape in real world coordinates rather than the 2D retinal image shape. Finally, an fMRI study contrasted identity and orientation tasks for objects oriented in depth. Their results "suggested that the parietal/frontal object areas encode view-dependent visual features and underlie object orientation perception" (Niimi, Saneyoshi, Abe, Kaminaga, & Yokosawa, 2011). This object orientation perception is likely an input to the shape constancy that must integrate all views of an object to be view-invariant (performed in a region such as left fusiform region, as detailed above). The neural processing of shape constancy has been shown to involve parietal/frontal areas and the right fusiform region in view-dependent perception that is then processed by the LOC or the left fusiform region for view invariance.

As discussed in Chapter 1, constancies (including orientation and length constancy) may be considered a basis for object externalization or distal perception. Orientation constancy allows the object angle to be perceived not as it is on the retina, but instead as an angle of an independent, *external* object in real-world coordinates, thereby enabling adaptive behavior. Learning orientation constancy in SS will therefore be critical to externalizing this new type of sensory input. Length and shape constancy allow an object to be recognized not as changing identity after rotation, but rather as a cohesive single object in the environment. This allows for the object to be externalized out in space as a real singular object. Stimuli externalization is critical for adaptive behavior, because objects are perceived as they are positioned, oriented, and shaped in the external environment rather than on the retinal image. For example, externalization of stimuli will allow for a blind individual to correctly locate a drop cane, approach the cane, and pick up the cane. Without orientation and position constancy, the cane would be jittering in space as the individual moved and would be impossible to reach and grasp. Without shape constancy, the cane would appear to be a different object when the participant tilted their head, making recognition of the cane quite difficult. Overall constancies are critical to the functional use of SS and to improving SS device usage.

Methods

Twelve blindfolded sighted and four blind participants (three late blind, one congenital) were trained on the vOICe device for at least 8 days at approximately 1 hour per day performing three evaluation tasks.

To evaluate orientation constancy, participants were presented with a bar at 6 different angles (0, 90, 45, −45, 22 or −22 degrees relative to vertical) with three potential head positions (vertical, tilted left, or tilted right) and determined the angle of the bar. The experimenter placed the bar on a black felt-covered wall in front of the seated participant, and visually estimated each angle position to be presented to the participant. Participants were permitted to determine the head tilt that they were most comfortable using in each trial, provided their head was stationary. One head position was requested for each trial.

To evaluate the precursor to length constancy, participants were presented with 5 lengths of bars (5AFC: 9, 12,15, 18, 21 cm), while the bar was placed in one of four orientations (0, 90, 45, or −45 degrees relative to vertical). Participants were asked to determine the length of the bar presented independent of the angle it was presented at. Since our primary aim was to explore training style/design, participant training was varied to determine the optimal training procedures. Two sighted participants were directed not to use head tilt during the length constancy task, and were not included in Figure 2.01 or Figure 2.05. One sighted participant was directed to use head tilt intermittently (4 out of 12 sessions with head tilt), and was excluded from Figure 2.05, but included in Figure 2.01. The remaining participants were asked to tilt their head in the initial trials and at the end of the training. Two participants evaluated the bar length without head tilt. Figure 2.05 includes and excludes different participants in a few data points so that the figure can show data for head tilt trials in sessions 0-5, and no-head-tilt data in trials in session 6-7 (sessions 6 data point = 5 participants, session 7 = 7 participants, all other data points = 9 participants). Figure 2.05 was designed with the head-tilt in initial sessions and no head tilt in later sessions to show the retention of the learning gained during the head tilt sessions, which is important to our argument that head-tilting aids learning. The same procedure was used for Figure 2.06 so that it can show data for head tilt in sessions 0-7, and only data for no-head-tilt in session 8-9 (session 8 data point = 3 participants, all other data points = 4 participants). Due to the technical constraints in the usage of the device, the desire for training exploration, and various practical and cognitive limitations in blinds, we could not carry out the experiments just unanimously with one simple and identical procedure.

If time permitted, the participants performed some other tasks such as object recognition and localization, whose data we decided not to include here. Both experimenters followed the same training protocol (training procedures detailed day to day) outlined for all experiment tasks (details in Appendix B).

Participants used a vOICe device to learn the constancy tasks. The vOICe device used a camera embedded in a pair of sunglasses or a webcam attached externally to glasses. Sighted participants were requested to close eyes during training and evaluation and wore opaque glasses and/or mask. The camera provided live video feed of the environment, and used a small portable computer to encode the video into sound in real time. The vOICe device translates the horizontal spatial dimension to scan-time, the vertical spatial dimension to frequency range, and image brightness to sound loudness. The vOICe software was obtained online at seeingwithsound.com and used for the video to sound encoding. All training sessions were recorded for later data analysis and/or presentations; participants were notified of video recording.

Three different vOICe device camera setups were used during training. All setups had a camera attached or embedded in glasses, a small portable computer connected to the camera, and earphones (either separate from the glasses or attached). The camera could be on the side of the glasses, or in the center on the bridge of the nose. Sighted participants' natural vision was obscured with black felt covering the glasses, or a sleeping mask worn under the glasses. In principle, these technical differences would not make any difference in terms of training efficiency and task performance, except a possible minor difference in spatial perception due to the gain of the camera, camera field of view and camera placement. Two blind participants were forced to transition from one camera and device setup to another setup partway through training due to device failure; their data did not indicate any difficulty with this transition.

Data analysis of head tilt and time to decision were performed on the video recordings of training sessions. Head tilt was quantified by counting the number of trials in which the participant used head tilt while exploring the stimulus, and then divided by the total number of trials (160 trials for sighted and 200 trials for blind). Head-tilt was estimated for training sessions with missing video by using the average head-tilt for sessions of the same type (such as head tilt allowed or head tilt not allowed). Two participants were trained without any head tilt permitted in the training, and they are included in the head-tilt correlation and time-to-decision plot but not the main data set on performance, to prevent inhomogeneity in training.

Data analysis on time-to-decision was performed by recording the onset and end of a task during the training session for all training sessions of all training participants. The data was averaged across sighted participants and across blind participants. Training sessions lacking a video recording were omitted from the analysis. One blind participant on the orientation constancy task was omitted from the time-to-decision data due to the lack of three consecutive training session videos; no other participants lacked three or more consecutive training session videos.

ANOCOVA and regression analyses were performed in MATLAB using the aoctool, regstats, and glmfit functions.

Results

Sighted and blind vOICe users were able to classify line angle independent of head tilt (orientation constancy), and to learn to further improve (Figure 2.01 and Figure 2.02). The rate of improvement was significant in both groups (Sighted, 8 training sessions: $p < 0.00$; Blind, 9 training sessions: $p < 0.00$). Blind participants had an average slope of improvement that was not significantly different from that of the sighted ($p_{slope} < 0.195$). However, the intercept of the improvement curves was significantly different between sighted and blind users ($p_{intercept} < 6.54 \times 10^{-9}$) with the sighted starting training at a higher percent correct, likely due to the blind users' diminished spatial perception.

Orientation constancy performance was also evaluated separately at each head position (vertical, tilted left or tilted right) (Figure 2.03 and 2.04). Head vertical position outperformed the head-tilted-left and head-tilted-right conditions for both the sighted and blind participants. The head-vertical position had the advantage of no angle shift calculation; in other words, the angle heard by the participant via vOICe is the angle in the environment, which was not true for the head-tilted-left or right conditions (also, the majority of training experiences had been at this angle). Therefore, with the angle directly perceived and no arithmetic added to the task, the head-vertical task was easier to perform than the head-tilted-left or right task. ANOCOVA analysis indicated that no slope pair between the vertical-head slope, head-tilted-left slope and head-tilted-right slope was significantly different in the blind or sighted participants. Intercepts were also evaluated with ANOCOVA analysis (assuming the slopes are equivalent), and the head-vertical condition was significantly different from the head-tilted-left ($p_{intercept} < 0.0002$) and the head-tilted-right conditions ($p_{intercept} < 1.62 \times 10^{-6}$) for the sighted participants. The blind participants had significantly different intercepts for vertical compared to head-tilted-left ($p_{intercept} < 0.0008$) and vertical compared to head-tilted-right ($p_{intercept} < 0.0036$).

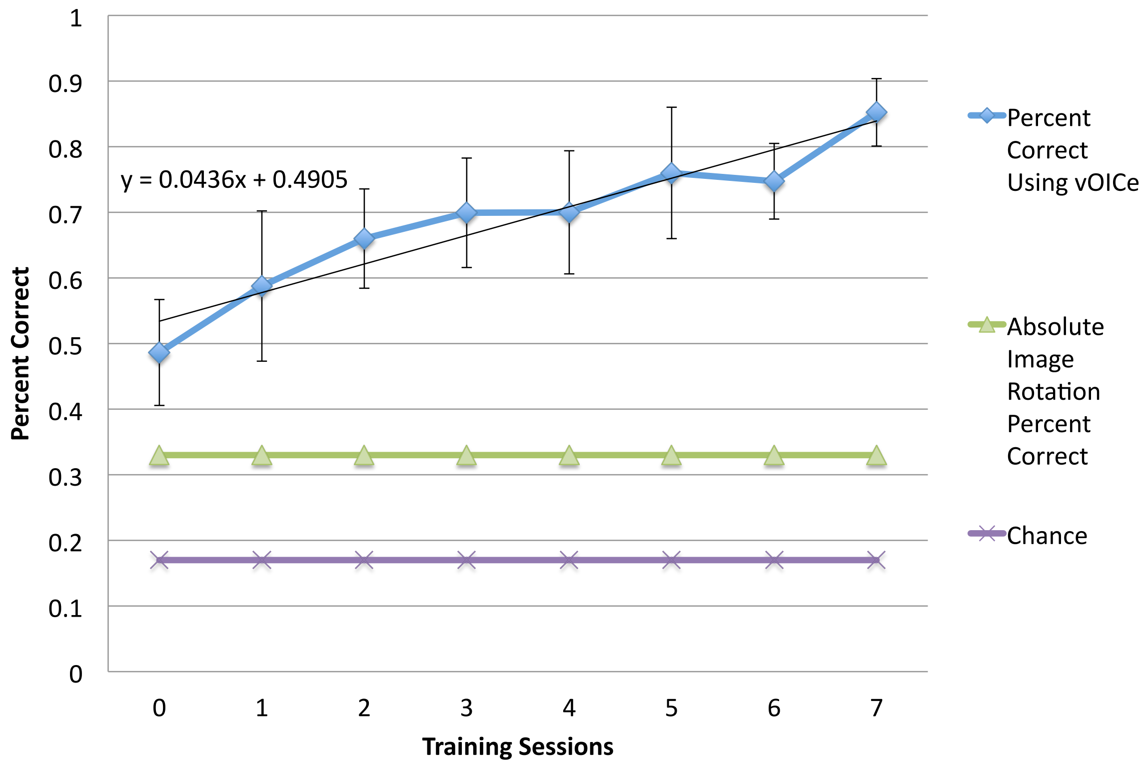

Figure 2.01. Sighted orientation task data. Performance in the orientation constancy task (classification of line angle independent of head tilt) as a function of the number of training sessions in the sighted participants ($N = 10$). Error bars are the standard deviation. Blind participant data are in a Figure 2.02. (Note: Two sighted participants excluded due to differences in training).

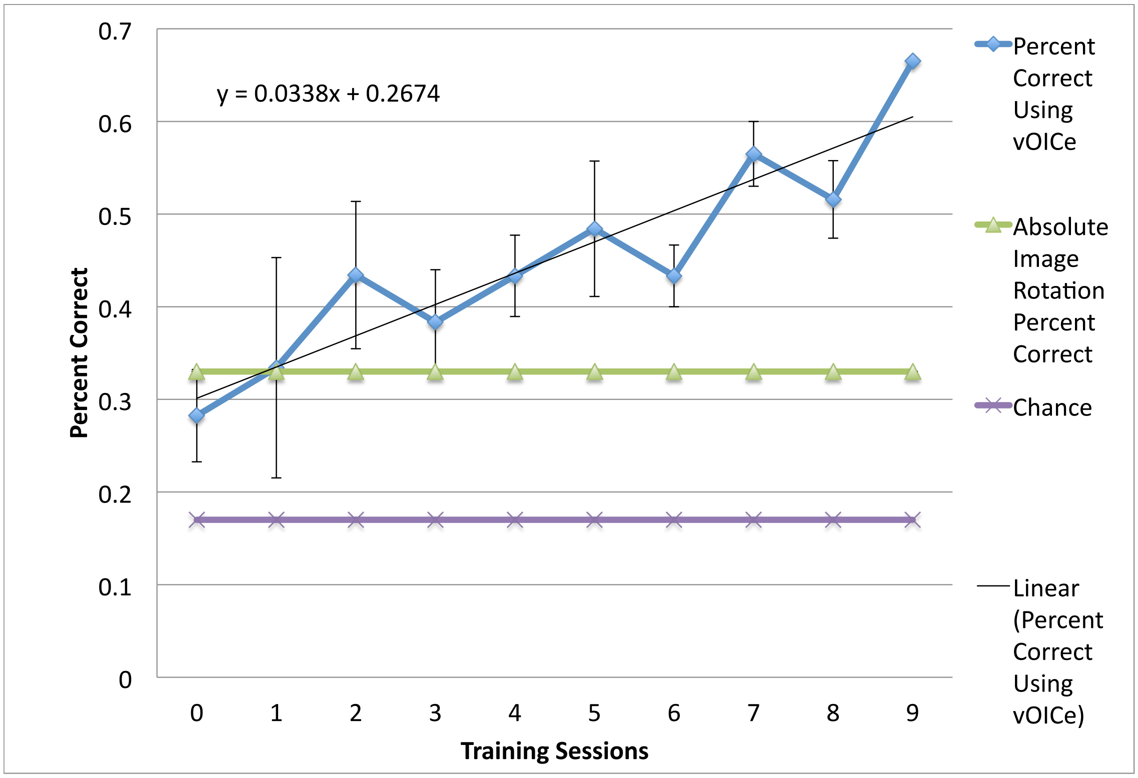

Figure 2.02. Blind orientation task data. Task performance of orientation constancy as a function of the number of training sessions in the blind participants ($N = 4$). Error bars are the standard deviation. The Absolute Image Rotation Percent Correct in Figure 1A is the percent correct if the angle of head tilt is unknown (*i.e.,* only head vertical can be correctly identified, or 1/3 correct).

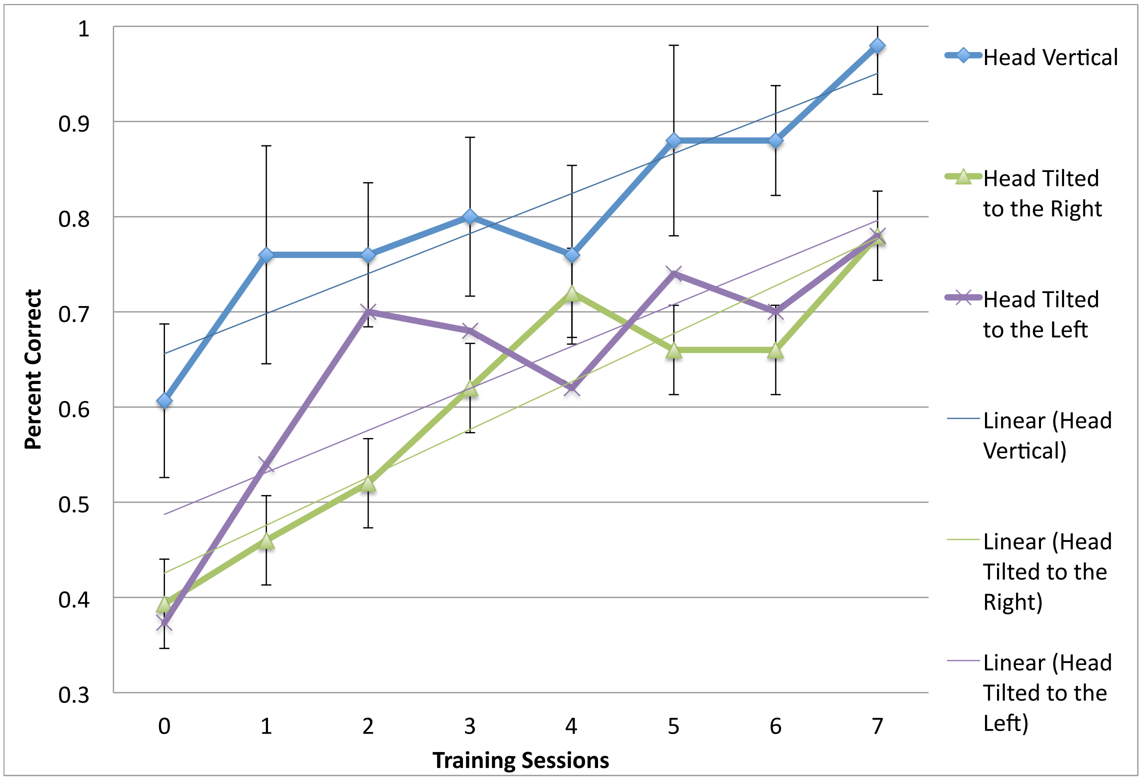

Figure 2.03. Sighted orientation task data divided by head tilt. Task performance of orientation constancy as a function of the number of training sessions in the sighted participants ($N = 10$). Data is separated into the participants' percent correct for each of the potential head positions: Vertical, tilted left, or tilted right. Error bars are the standard deviation.

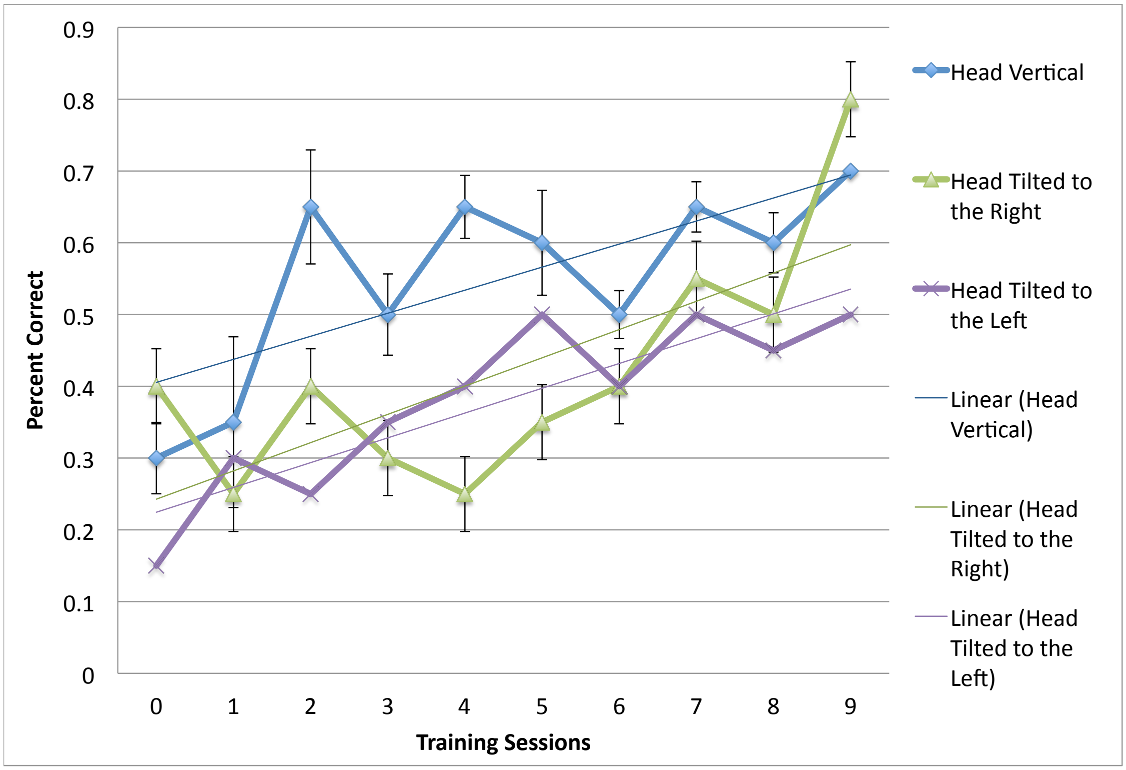

Figure 2.04. Blind orientation task data divided by head tilt. Task performance of orientation constancy as a function of the number of training sessions in the blind participants ($N = 4$). Data is separated into the participants' percent correct for each of the potential head positions: Vertical, tilted left, or tilted right. Error bars are the standard deviation.

Sighted and blind vOICe users were able to classify line-length independent of angle (length constancy), and to learn to further improve (Figure 2.05, Figure 2.06). The rate of improvement was significant in both groups (Sighted, 8 training sessions: $p < 0.01$ Blind, 10 training sessions: $p < 0.03$). Nonetheless, blind participants had an average slope that was not significantly different from the sighted, while the intercepts were significantly different ($p_{slope} < 0.179$, $p_{intercept} < 0.0014$). During head-tilt allowed sessions, head-tilting was encouraged. Head tilting frequency during the task correlated significantly with improved line length classification (Figure 2.07). The sighted participants correlated head tilt with length constancy task improvement with a coefficient of 0.6560 ($p < 0.03$), whereas sighted and blind participant data combined had a coefficient of 0.6024 ($p < 0.02$). The blind-only correlation is seemingly lower partly because the participants were fewer, and they have a wider range of capabilities and spatial perception. (For further comparison of blind to sighted, see Figure 2.08). Head tilt frequency was not significantly correlated with participants' initial performance at the length constancy task (*i.e.,* intercept) (for sighted and blind combined $rho = 0.0786$, $p < 0.78$).

Blind participants were divided into late and early blind categories, and their slope of improvement was compared to the sighted participants (Figure 2.08). In both the shape constancy and orientation constancy tasks, the early blind participant ($N = 1$) improved the slowest (*i.e.,* smallest slope). The late blind participants ($N = 3$) improved the second slowest, and the sighted ($N = 9$-10) improved the fastest (*i.e.,* largest slope). As discussed in Chapter 1, the sighted participants have the advantage of familiarity with visual principles (such as relative size), and visuomotor skills, due to daily visual experience. These existing skills can be used to advantage when learning to process the vOICe visually or re-learn a constancy. The late blind have less visual experience than the sighted, as they have been visually deprived for years if not decades. Finally, the early blind have no visual experience. Therefore, the rate of learning seems to correlate with visual experience; however, no definitive statement can be made due to the low number of late blind ($N = 3$) and the extremely low number of early blind participants ($N = 1$).

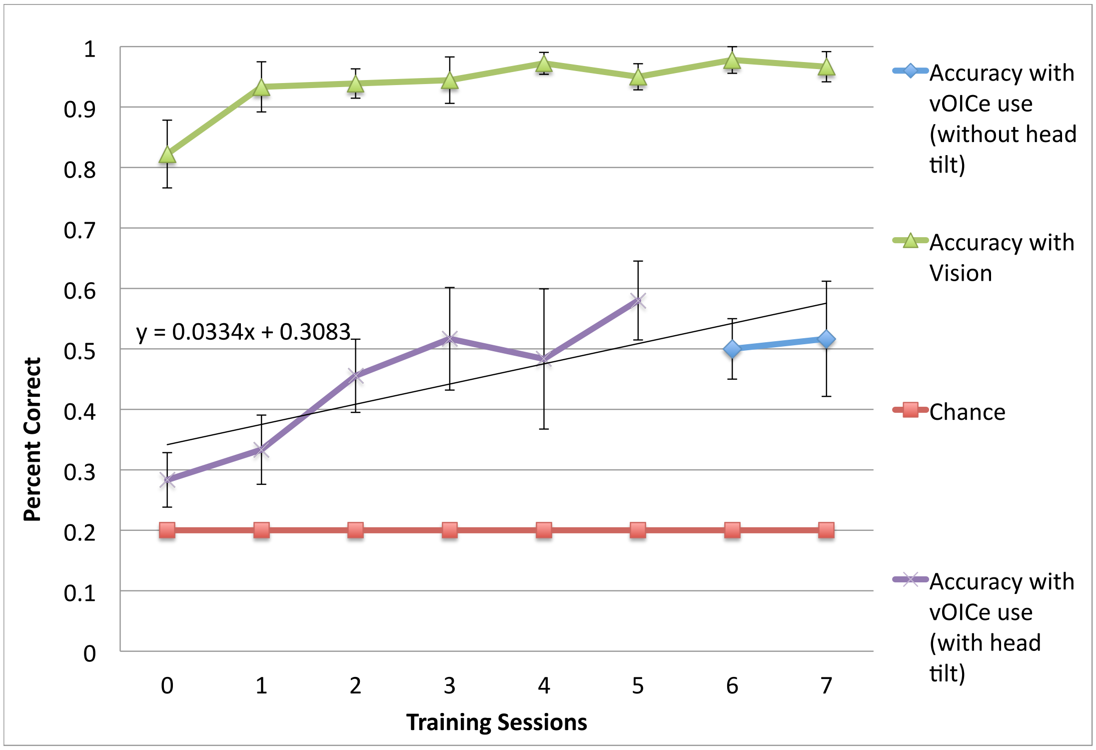

Figure 2.05. Sighted length constancy data. Performance in the length constancy task (classification of line length independent of angle) as a function of the number of training sessions in sighted participants ($N = 9$). Error bars are the standard deviation. Blind participant data are in a Figure 2.06 (Note: Three sighted participants were excluded due to no head tilt used, or intermittent use of head tilt as directed by the experimenter).

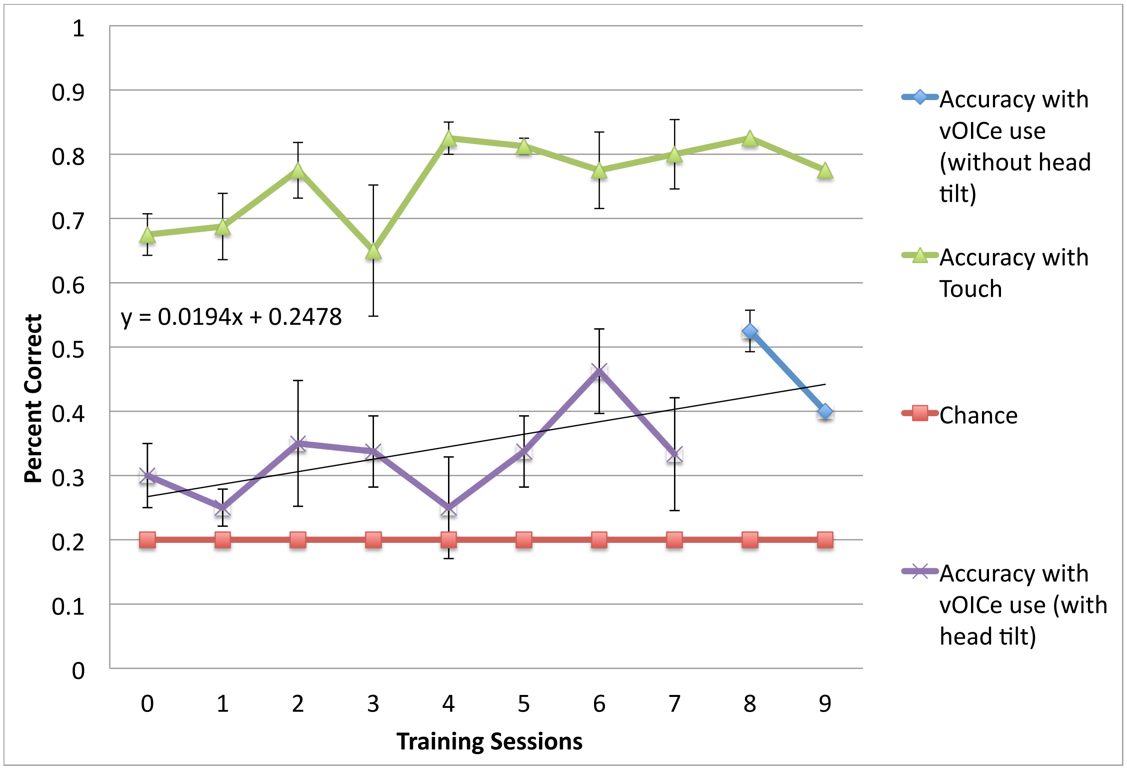

Figure 2.06. Blind length constancy data. Task performance of length constancy as a function of the number of training sessions in the blind participants ($N = 4$). Error bars are the standard deviation.

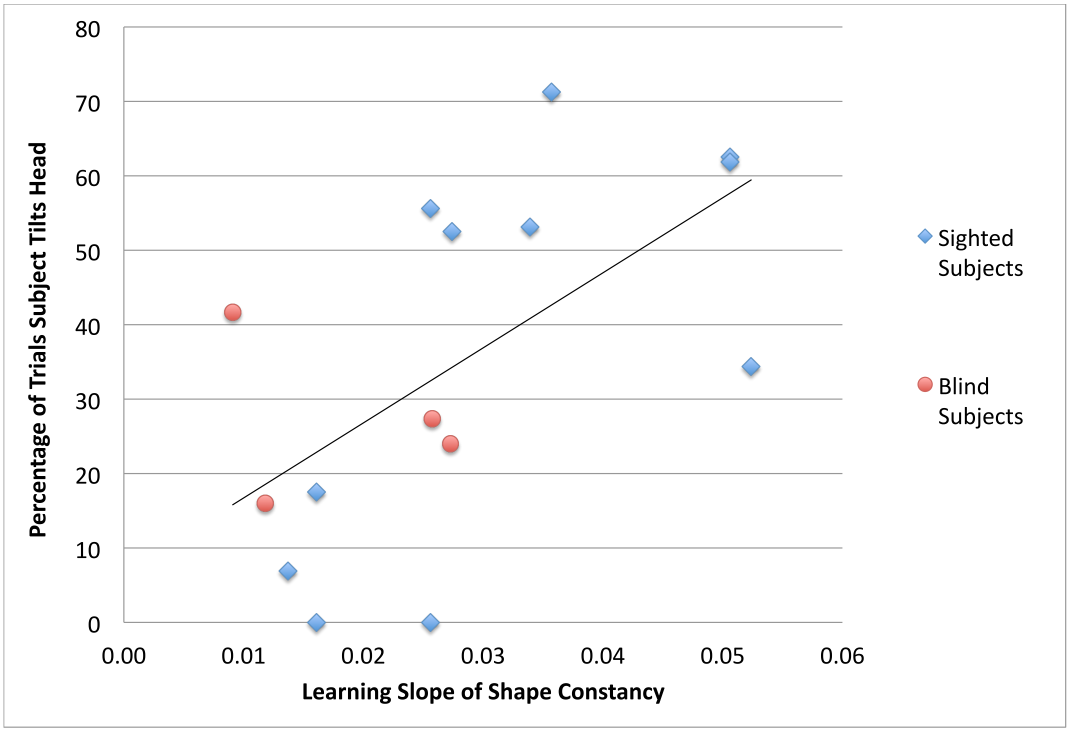

Figure 2.07. Length constancy head tilt and performance improvement correlation. Significant correlation between head tilt and performance improvement in the length constancy task ($N = 16$) ($rho = 0.6024$, $p < 0.02$). The number of trials that participants tilted their head was counted for all task sessions from video recordings (the average number of head tilt trials were used for sessions with missing video). The percent of trials that head-tilt was used was plotted against the slope of the participants length constancy improvement (from the interpolated slope in individual participant plots similar to Figure 2.05). (Note: One sighted participant was excluded due to intermittent use of head tilt, and two extra sighted participants were included who did not use head tilt at all, and were thus excluded from the analyses for Figure 2.01 and Figure 2.05).

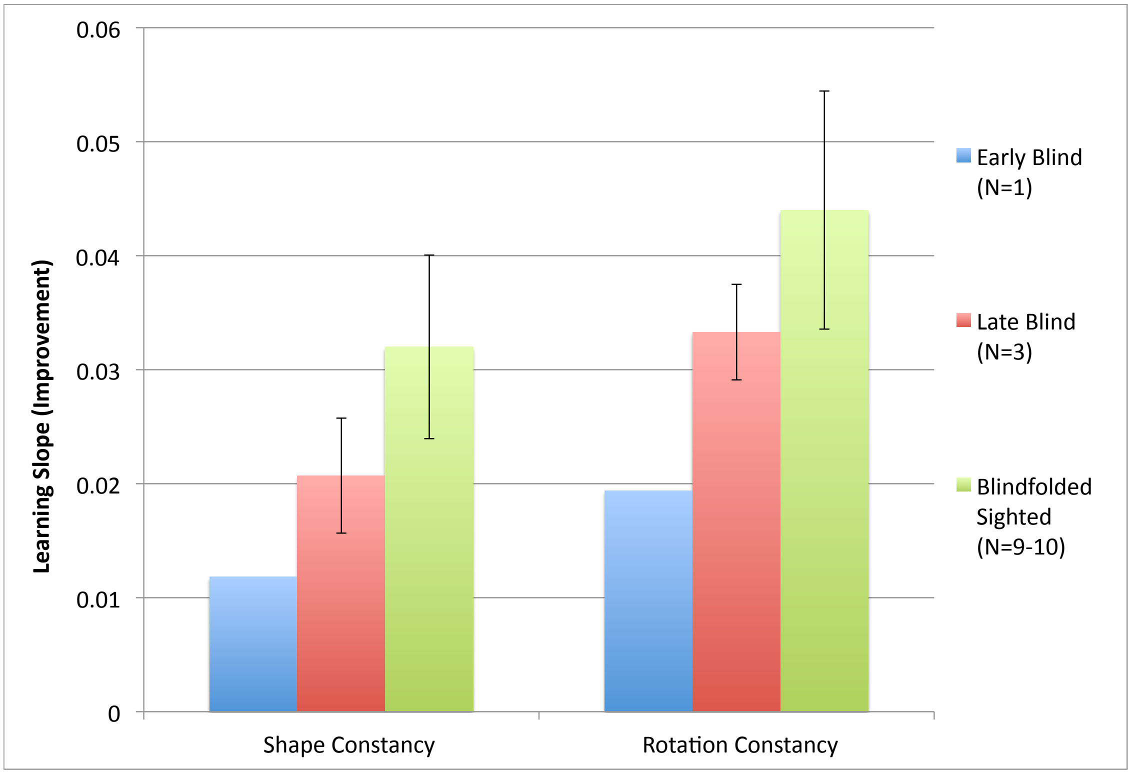

Figure 2.08. Vision and blind task improvement comparison. The slope of task performance as a function training session was determined for each of the blind and sighted participant groups. Slopes were calculated for each individual's data, and then averaged into a group.

The performance time in both tasks decreased as training sessions progressed for sighted and blind participants (Figures 2.09 and 2.10). The decrease in time to perform the training task indicates a tendency toward task automaticity and away from extensive top-down attention, thereby beginning to mimic the intuitive and automatic nature of perceptual constancies in the sighted.

It is revealing that head-tilt significantly correlated with the improved length constancy performance (Figure 2.07). This is the most critical and core finding of this study, as it indicates an improved SS training technique with additional sensorimotor interaction. It also indicates a key method for learning of constancy in vision as well as with SS.

The benefit of head-tilt with length constancy can be described in mathematical and psychophysical terms. As a participant spontaneously tilts their head, they alter the tilt of the camera attached to glasses on the head, thereby altering the angle of the line heard. As the angle of a line is rotated, the length and the width also change according to $L*Sin(\theta)$ and $L*Cos(\theta)$, where L is the length of the line and theta is the tilt, in the head or frame of reference. The change of vertical length (range of pitch in the SS device) and horizontal width (duration of sound) with head tilt are plotted in Figure 2.11 and 2.12, for a line placed vertically (Figure 2.11) or horizontally (Figure 2.12). In these plots, the radius of each half circle represents the line length, which remains the same (*i.e.,* constant) across all the different head tilts or line angles. The brain learns from association of different points on each line to identify each line as one entity (*i.e.,* a white bar of a particular length), and to separate the different lines as separate bars of different length. Learning general length constancy (not just bars used, but all bars at different angles) requires similar exploration with head-tilt and similar association of different input patterns to be identified as the same real-world object (and it applies to the real world, natural seeing as well as via the sensory substitution device). Further, participants can begin to associate all of the stimuli types (*i.e.,* 0, 90, 45, −45 degrees) for a given line length, which in effect provide orientation-invariance and correspond to the object identity (line length). Obviously, active head-tilting and its sensory feedback have a critical role in such a dynamic associative learning of length constancy, specifically in sensory substitution but more generally in vision. Further, due to the significant head-tilt correlation, memorization has been shown to not be a successful learning strategy for length constancy (*i.e.,* if memorization were used by participants, head tilt would make the task more difficult, whereas head tilt improved performance at the task).

By tilting their head, the observer receives dynamic yet systematic changes of input parameters, as illustrated in Figure 3-A and -B. In effect, learning aims to identify all the data points within each curve as an "identical horizontal line," whereas discriminating across different curves as "different length." Quite intuitively, it would be much easier if the brain compared an entire curve *vs.* another in the graph using head-tilt to move along the curve, as opposed to a point-by-point comparison in a set of (static) parameters. One may easily implement this more computationally in terms of S/N ratios in a Bayesian or a MLE framework.

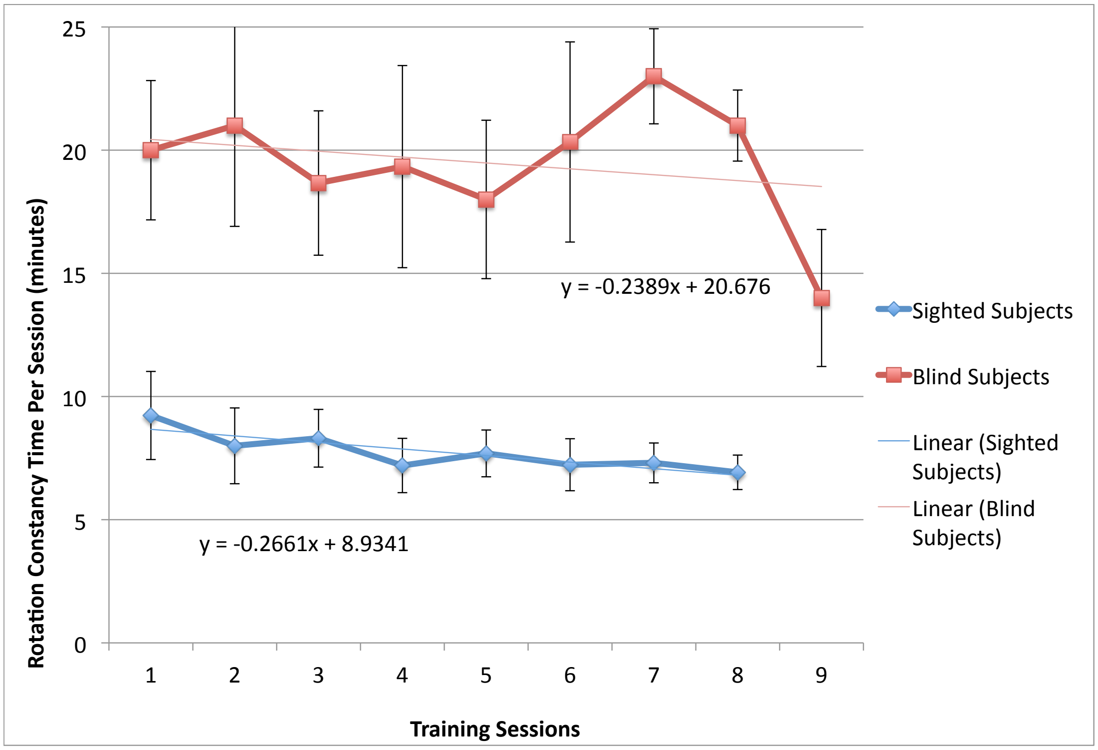

Figure 2.09. Orientation constancy task duration. Orientation constancy task duration for sighted ($N = 12$) and blind participants ($N = 3$). The duration of all trials for each participant was determined from video recordings of training sessions, and then averaged across participants. This orientation constancy task duration is plotted as a function of the training sessions. One blind participant was omitted from the data due to lack of video from three consecutive sessions. Error bars are the standard deviation.

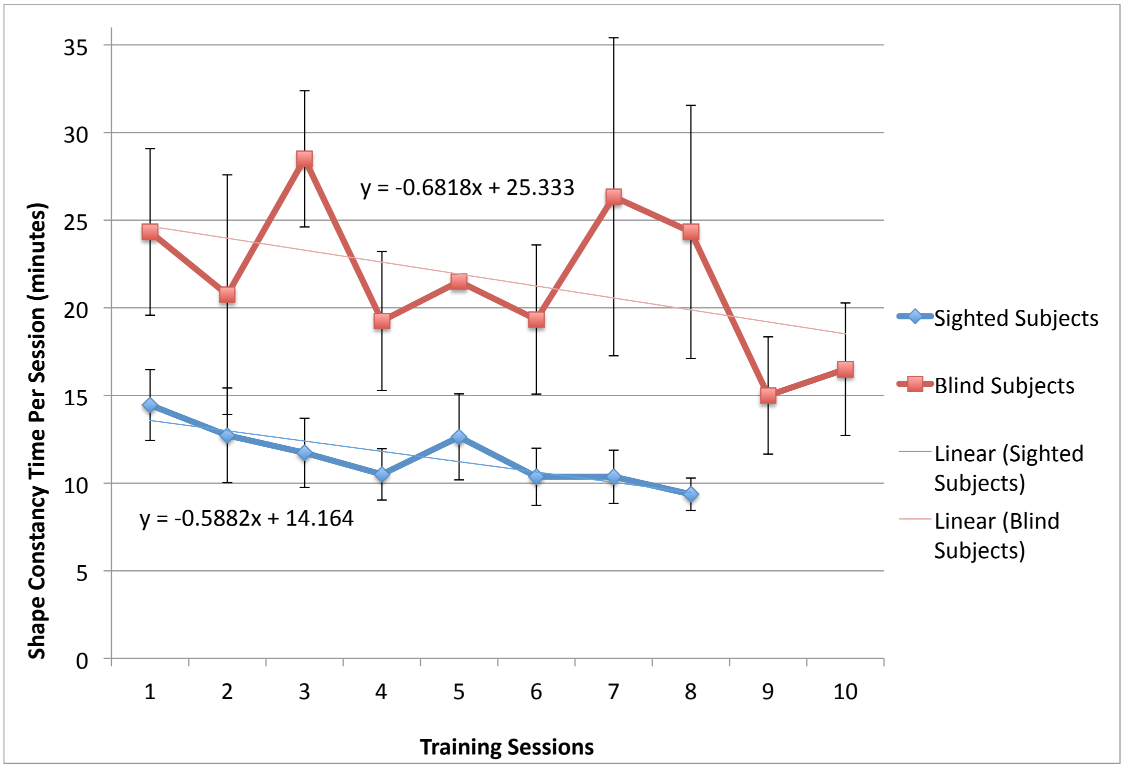

Figure 2.10. Length constancy task duration. Length constancy task duration for sighted ($N = 11$) and blind participants ($N = 4$). The duration of all trials for each participant was determined from video recordings of training sessions, and then averaged across participants. This length constancy task duration is plotted as a function of the training sessions. Error bars are the standard deviation. One sighted participant was omitted due to four different transitions between head-tilt allowed trials and no-head tilt allowed trials.

From smallest radius to largest in order: 9 cm bar, 12 cm bar, 15 cm bar, 18 cm bar, 21 cm bar

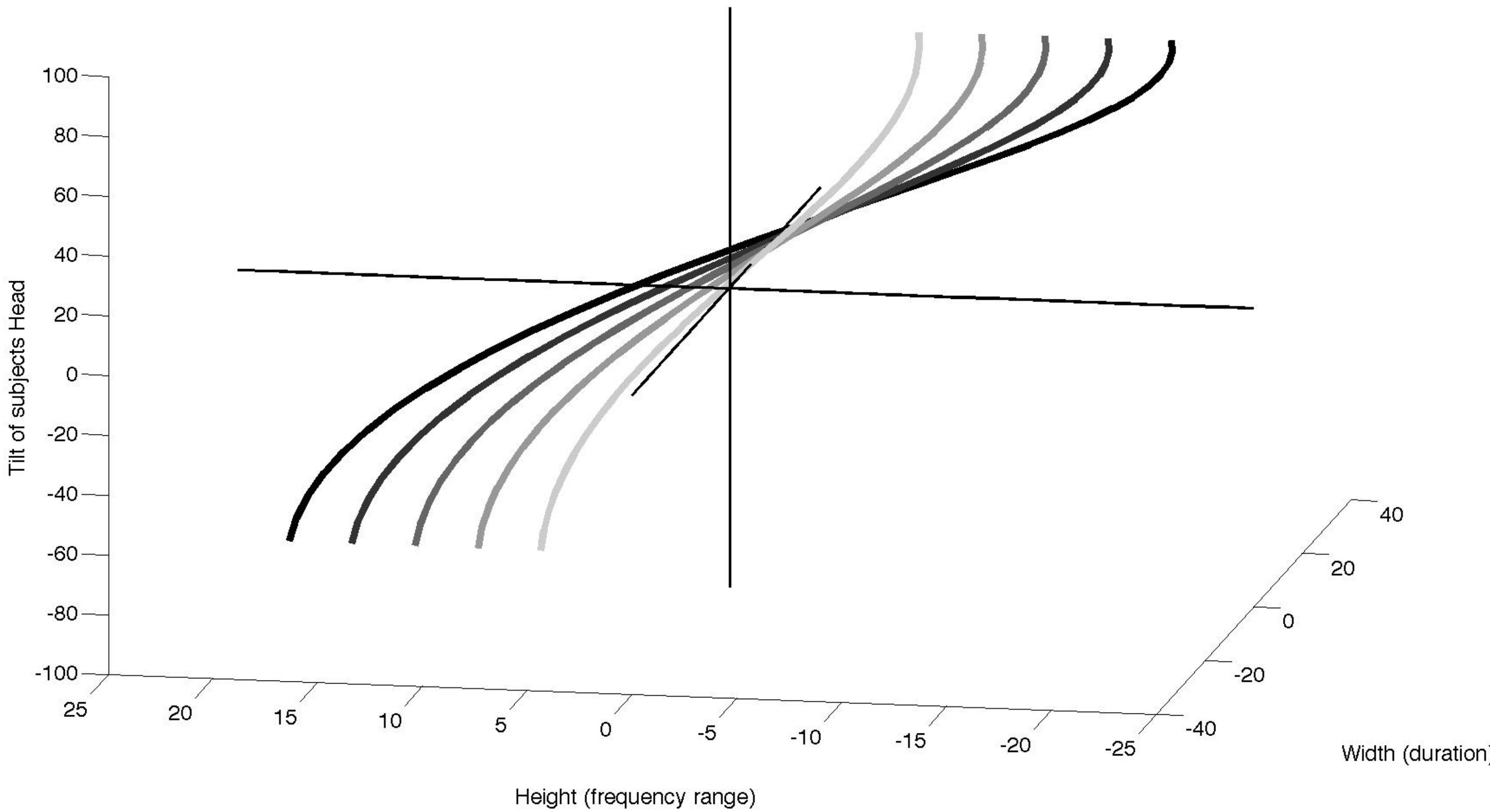

Figure 2.11. Head tilt and length constancy with a horizontal line. A horizontal line's dynamic change as participant tilts their head from vertical (no head tilt, tilt = 0) to 90 degrees left (negative tilt) or right (positive tilt). Each line represents a different line length ranging from 9 cm to 21 cm in length. Thus, in effect, learning aimed to identify all the data points within each curve as an "identical horizontal line," while discriminating across different curves as a "different length."

From smallest radius to largest in order: 9 cm bar, 12 cm bar, 15 cm bar, 18 cm bar, 21 cm bar

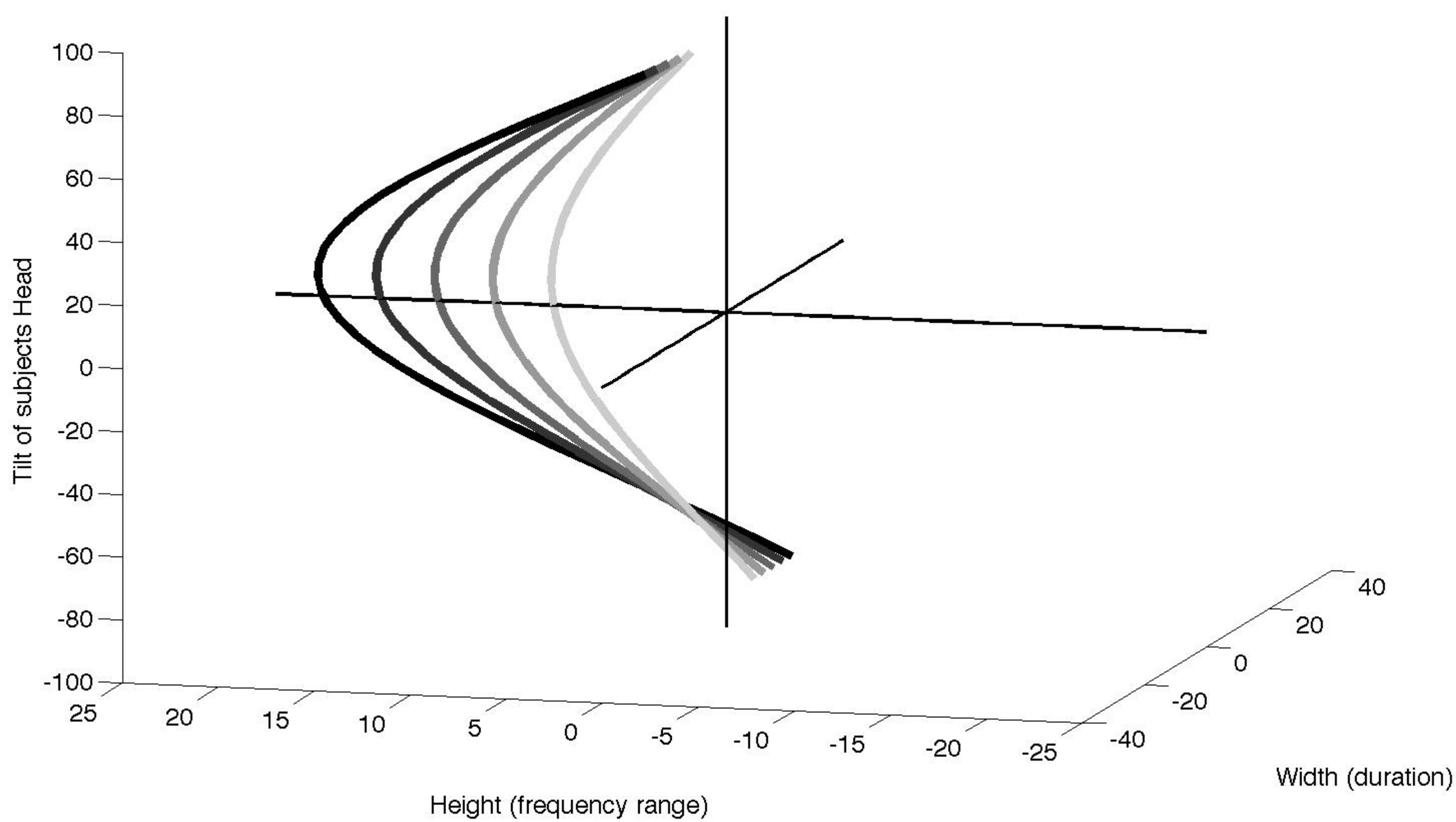

Figure 2.12. Head tilt and length constancy with a vertical line. A vertical line stimulus dynamically changes in horizontal width (duration of the sound) and vertical height (range of sound pitch) as the participant tilts their head from vertical (no head tilt, tilt = 0) to 90 degrees left (negative tilt) or right (positive tilt). Each line represents a different line length ranging from 9 cm to 21 cm in length. Thus, in effect, learning aimed to identify all the data points within each curve as an "identical vertical line," while discriminating across different curves as a "different length."

The enhanced learning due to head-tilt (retained in no-head-tilt trials) is consistent with a neural network using supervised training. In supervised learning, a neural network improves at classification with more training images and correct answer pairs presented[17]. In length constancy training, with no-head-tilt in each trial, the number of unique training images presented to the neural network is correlated with number of trials. But when the participant uses head-tilt, the number of unique training stimuli presented increases by a factor of at least *three*, one factor for each head tilt (vertical, left tilt, and right tilt), because each head position is a unique training stimulus. Each head position presents different parameters to calculate length, effectively tripling the training stimuli during supervised training and making classification more accurate (Changizi, Hsieh, Nijhawan, Kanai, & Shimojo, 2008). Further, while this calculation is based on snapshots at each head tilt, dynamic feedback would be even more "educational" to the network.

Active, as opposed to passive, interactions with the environment have been proven to be more effective for sensory-motor learning (Held & Hein, 1963). Some argue even more strongly that active interaction is crucial to visual awareness (O'Regan & Noe, 2001). Reynolds and Glenny showed that interactive two-participant training generated better performance at a vOICe device localization task than typical active or passive training with one participant (Reynolds & Glenney, 2012). J. Gibson's classical concepts such as "dynamic, direct perception" or "picking up higher-order invariance from the input affordance" may be relevant (Gibson, 1950b). Modern neural computation modals indicate that the brain is associative (Dayan, Abbott, & Abbott, 2001), using synapse weighting to correlate related properties between neurons that “recognize” objects (Quiroga, Reddy, Kreiman, Koch, & Fried, 2005). Reafferent signals or feedback from motor commands have been hypothesized by V. Mountcastle to provide a memory-based prediction to optimize sensory-motor learning (Mountcastle, 1978). Such re-afferent signals seem to critically enhance SS spatial perception and, therefore, constancy. More critically, the parametric analyses of the constancy with regard to the head tilt (Figure 2.11 and 2.12) not only reveal how the brain learns it with the SS device, but also in principle capture the learning of constancies in natural viewing. Thus, the implications of the current results/analyses go beyond just the SS learning.

Constancy learning with the vOICe demonstrates the dramatic plasticity of the adult brain. Tactile and auditory sensory substitution learning functionally recruits visual regions via extensive plasticity in blind and sighted users (Amedi, et al., 2007; Kupers, et al., 2010; Merabet, et al., 2009; Poirier, De Volder, & Scheiber, 2007). In particular, SS face stimuli activates the FFA (fusiform face area), SS shape discrimination activates the LOtv (Lateral Occipital tactile visual) area, and SS reading activates the VWFA (visual word form area)(Amedi, et al., 2007; Plaza, et al., 2009; Striem-Amit, et al., 2012). In addition, repetitive TMS had shown that congenital and late-blind users causally recruit visual regions for SS processing (Collignon, et al., 2007; Merabet, et al., 2009). Although no previous studies have brain-scanned the constancy learning with SS, a broad network of regions from sub-cortical auditory areas, primary auditory regions, multimodal regions and then visual regions may play a role. Further, active sensory feedback between these regions and motor areas likely improves multisensory network efficacy. If multisensory experiences and feedback shape such a neural network, active, as opposed to passive or static, learning procedures may enhance network shaping.

Length and orientation constancy are critical for rehabilitative use of SS, especially since it will enable “vision”-like processing and stimulus externalization. Constancies allow for the neural association of stimuli that have different proximal properties but represent the same object or feature in real 3D space, thus making a cohesive representation of external objects. SS devices have yet to aid the large population of blind people still limited in their daily-life functionality. Externalization of objects via length and orientation (and other) constancies could be the first critical stepping stone in the training process towards higher functionality at more complicated tasks in cluttered natural environments.

In sum, critical perceptual properties such as constancy and externalization can be achieved with current sensory substitution devices. Dynamic interaction with stimuli is shown to be critical to learning with sensory substitution owing to sensory-motor engagement and additional training information provided to the cognitive neural network.