UNIVERSITY OF GENOVA
PHD PROGRAM IN BIOENGINEERING AND ROBOTICS

Building Spatial Skills: Development of Spatial abilities across the sensory systems in children with and without visual impairment

by
Gloria Calafatello

Thesis submitted for the degree of *Doctor of Philosophy* (38° cycle)

June 2026

Prof. Monica Gori	Supervisor
Dr. Silvia Zanchi	Co-supervisor
Prof. Paolo Massobrio	Head of the PhD program

Thesis Jury:

Prof. Marc O. Ernst, *Ulm University*	External examiner
Prof. Marko Nardini, *Durham University*	External examiner
Prof. Silvio Paolo Sabatini, *University of Genova*	Internal examiner

Dibris

Department of Informatics, Bioengineering, Robotics and Systems Engineering

A mio Fratello:
che ha trovato il suo modo di costruire il suo spazio, con poco “input visivo”
To my Brother,
who has found his own way to build his space, with little “visual input”

E ai “*Norman”*: piccoli non possiamo essere.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Gloria Calafatello

June 2026

Abstract

Spatial perception is a fundamental ability that enables humans to orient, navigate, and interact effectively with their environment. The main objective of my Ph.D. project was to investigate how spatial representations develop during early life, and how sensory experience shapes the unisensory and multisensory mechanisms supporting spatial perception. In particular, my Ph.D. aimed at: i) characterising the development of auditory spatial localisation abilities during early childhood; ii) examining how children integrate multisensory spatial information, focusing on audiovisual cues and spatial disparity between objects; iii) investigating the development and neural mechanisms of audio-tactile multisensory integration in early childhood, in both typically developing children and children with severe visual impairment. To achieve these aims, I conducted a series of behavioural and neurophysiological studies across different developmental stages, using age-appropriate experimental paradigms. First, behavioural experiments in preschool children were used to assess the development of auditory spatial localisation. Second, audiovisual spatial object segregation was investigated behaviourally in toddlers using dynamic stimuli, and multisensory performance was evaluated against predictions from computational models of cue combination, including Maximum Likelihood Estimation, to assess whether integration approached optimality. Finally, high-density electroencephalography (EEG) was employed in infants and toddlers to investigate the neural correlates of early audio–tactile multisensory processing, allowing behavioural responses to be directly linked to cortical activity. i) Results on auditory spatial localisation indicated comparable performance between preschool children and adults along the azimuth plane, with residual differences emerging primarily for elevation cues. ii) At the multisensory level, results from the audiovisual segregation study showed that preschool children tended to perceive the spatial discrepancy of objects if the stimulation is bimodal. Substantial inter-individual variability emerged, reflecting differences in sensory weighting strategies during development. iii) The investigation of audio–tactile integration in early childhood demonstrated that multisensory gain is present early in development in both sighted and severely visually impaired children. However, clear differences emerged in neural organisation: while sighted infants recruited integrative cortical networks and maintained multisensory processing under spatial conflict, visually impaired infants relied more strongly

on tactile, body-centred representations, showing reduced engagement of integrative mechanisms when sensory cues were incongruent. Overall, the findings of this Ph.D. project contribute to a comprehensive understanding of how spatial perception develops from infancy to early childhood. They demonstrate that multisensory mechanisms emerge early but are progressively shaped by sensory experience and task demands, with visual input playing a critical role in calibrating spatial representations and supporting flexible multisensory organisation. By integrating behavioural and neural evidence from both typical and atypical developmental trajectories, this work advances our understanding of the mechanisms through which children construct spatial knowledge in a multisensory world.

Key-words:

spatial perception, development, multisensory integration, auditory localisation, audio–visual processing, audio–tactile integration, visual impairment, EEG

Table of Content

List of Figures

List of Tables

Chapter 1

Introduction

The environment that surrounds us acts as a driver, witness, and scaffold for the development of our cognitive and human capacities. Sensing and perceiving are processes that bind us to the space around us, position us within it, and make us active agents of spatial rules that both contain and enable our experience. Understanding the development of spatial abilities thus offers insight into a fundamental human capacity: the ability to establish a meaningful connection between the individual and the surrounding world, enabling engagement with measures, objects, places, shared rules, and conceptual structures. Although spatial abilities are often framed within maturational accounts, developmental findings suggest that spatial performance is shaped by sensory experience and varies across perceptual contexts (Gori et al., 2008; Nardini et al., 2008; Hillock-Dunn & Wallace, 2012).

Across development, particularly during early childhood, cognitive and perceptual abilities do not follow a fixed or uniform trajectory. Instead, development is characterised by substantial inter- and intra-individual variability (De Ribaupierre & Lecerf, 2018), reflecting the dynamic interaction between biological maturation, sensory experience, and environmental factors (Cainelli et al., 2025). The early years of life represent a particularly sensitive window during which knowledge about the world and cognitive abilities rapidly evolve, shaped by multiple interacting influences. This results in a heterogeneous and non-linear developmental process, in which similar functional outcomes may be achieved through different developmental pathways (De Ribaupierre & Lecerf, 2018; Cainelli et al., 2025).

In this context, the present doctoral project aims to explore the development of spatial skills and their emergence in relation to sensory and multisensory integration. This thesis presents a series of studies addressing sensory and multisensory development as applied to different spatial abilities. In Chapter 2, I examined auditory spatial localisation as a unisensory foundation of spatial perception, highlighting that early spatial abilities are functional yet characterised by dimension-specific variability. By characterising localisation accuracy and response dynamics across spatial dimensions and developmental groups, this study demonstrates that auditory spatial representations are functional in early childhood but remain unevenly refined, with dimension-specific limitations that reflect the reliability of acoustic cues. After that, the focus shifted from unisensory processing to dynamic multisensory perception. This transition reflects the broader theoretical perspective that spatial development cannot be fully understood within a single-sensory framework. In Chapter 3, a new device is introduced and

used to deliver controlled audiovisual stimulation, enabling the investigation of how young children perceive spatial separation and localisation when multiple sensory cues unfold over time and space. In Chapter 4, I use this methodological framework to address a central developmental question: whether audiovisual information supports spatial source segregation in preschool-aged children under dynamic conditions. Finally, in Chapter 5, the investigation of multisensory processes focused on the processing of audio-tactile integration, combining behavioural measures with electrophysiological recordings to explore the neural mechanisms supporting early multisensory spatial processing and the role of visual experience in shaping these mechanisms. In particular, this study examines how multisensory gain and conflict resolution differ between sighted and visually impaired infants, providing insight into the developmental organisation of multisensory spatial representations. The final chapter, Chapter 6 integrates these findings, offering a broader discussion on spatial development as a heterogeneous and non-linear process shaped by sensory experience.

1.1 Spatial Abilities

Spatial abilities can be broadly defined as the capacity to detect information from the external world and create a coherent representation of information about locations, configurations, and spatial relationships in the surrounding environment, enabling action and interaction in space (Linn & Petersen 1985; Gori, Amadeo, & Bremner, 2026). While spatial abilities have traditionally been conceptualised as progressively improving with age, developmental evidence suggests that different spatial components may follow partially distinct trajectories, depending on sensory experience and task demands (Gori, Amadeo, & Bremner, 2026). We perceive space from our earliest experiences in life, even though spatial abilities develop gradually over time. Visual perception supports the rapid extraction of spatial configurations, such as object dimensions and distance information within the first months after birth, reflecting both early functional sensitivity and the presence of organized neural substrates for spatial coding in the developing visual system (Baillargeon et al., 1985; Bremner et al., 2006). At the same time, auditory and tactile modalities also convey spatial information from birth: newborns can encode and orient toward the locations of sounds and touches, and neural signatures of spatial processing are observable across these sensory systems in early postnatal life (Muir et al., 1989; Kisilevsky & Muir, 1984). This refinement does not follow a strictly linear progression but reflects heterogeneous developmental patterns across sensory modalities, cross-calibration and spatial dimensions. Over maturation, spatial competences continue to refine throughout the preschool years. In the visual domain, young children progressively improve their use of environmental geometry, landmarks, and relational spatial coding to support reorientation and navigation in space (Dessalegn & Landau, 2008). Parallel developmental changes are observed in auditory spatial perception: localisation accuracy and orienting efficiency to sound sources increase across early childhood, with measurable improvements evident between approximately 3 and 5 years of age (Martin et al., 2015; Eklöf et al., 2022). Tactile spatial processing also shows important maturation during this period. Preschool children demonstrate emerging use of external spatial reference frames for touch, as revealed by developmental changes in crossed-hands localisation performance and visual modulation of tactile spatial coding (Begum Ali et al., 2014).

1.2 Spatial auditory localisation

Spatial auditory localisation provides fundamental information about events occurring beyond the visual field, enabling individuals to orient in space and guide goal-directed action. Unlike

vision, auditory space must be inferred indirectly from acoustic signals shaped by the interaction between sound waves, the listener's body, and the surrounding environment (Blauert, 1997; Kubovy & Van Valkenburg, 2001; Carlini et al., 2024). Spatial hearing operates across three principal dimensions, azimuth, elevation, and distance, each supported by partially distinct acoustic cues and computational mechanisms. Horizontal localisation primarily relies on binaural cues, including interaural time and level differences, which provide relatively robust information about lateral sound position (Makous & Middlebrooks, 1990). In contrast, vertical localisation depends largely on monaural spectral filtering generated by the head and pinnae, and encoded in head-related transfer functions, resulting in greater perceptual ambiguity, interindividual variability, and reduced precision compared with azimuth judgements (Blauert, 1997; Freigang et al., 2015). Although rudimentary orienting toward lateralised sounds is observable at birth, mature auditory spatial perception emerges through a prolonged developmental process characterised by limited early precision, high variability, and strong dependence on attentional and stimulus factors (Muir & Field, 1979; Morrongiello, 1987). Across infancy and childhood, improvements in localisation accuracy and consistency reflect the interaction between neural maturation, growth-related changes in head and pinna morphology, and accumulating sensorimotor experience. Because acoustic spatial cues depend directly on body geometry and acoustic filtering, developmental changes continuously reshape the mapping between acoustic signals and spatial coordinates, requiring ongoing perceptual recalibration (Knudsen & Brainard, 1991; King & Keating, 2015). Consistent with this view, behavioural and neurophysiological evidence demonstrates persistent plasticity in auditory spatial representations and the capacity for experience-dependent recalibration (Hofman et al., 1998; Kumpik et al., 2010). Developmental trajectories are not uniform across spatial dimensions. Sensitivity to binaural cues supporting horizontal localisation emerges relatively early, whereas localisation based on spectral information in the vertical or median plane often remains less accurate and more variable into later childhood (Morrongiello, 1987; Freigang et al., 2015; Otte et al., 2013). This dissociation further supports the view that spatial abilities develop through modality- and cue-specific trajectories rather than through a homogeneous developmental progression**.** Alongside spatial accuracy, developmental refinement is also reflected in the temporal dynamics of localisation behaviour, with progressive reduction in response latency interpreted as markers of increasing processing efficiency and sensorimotor coordination across childhood (Kail, 1991). This developmental trajectory is consistent with evidence from auditory spatial localisation studies showing progressive refinement of spatial hearing abilities and delayed maturation in the absence of

visual input (Cappagli & Gori, 2016; Cappagli et al., 2016).

1.3 Multisensory Integration in Spatial Abilities

Given that natural environments continuously provide information from multiple sensory sources, development during early childhood involves gradual improvements in the use of spatial reference and in combining multiple sources of spatial information together (Hermer & Spelke, 1994), to build a unified world representation. These developmental changes reflect the joint influence of neural maturation, and experiential interactions with the environment (Sluzenski et al., 2004; Woollett & Maguire, 2011). In this perspective, multisensory spatial perception emerges through a gradual transition from early cross-modal calibration processes to later forms of statistically optimal sensory integration, reflecting the progressive coordination between sensory systems during childhood (Gori, Amadeo, Bremner, 2026).

The integration of sensory cues involves determining whether signals from different modalities originate from a common external event and, when appropriate, combining them into a unified perceptual representation. This process relies on sufficient spatial and temporal alignment between signals and follows statistically optimal principles, allowing the perceptual system to reduce uncertainty by weighting redundant sensory information (Alais & Burr, 2010, 2019; Ernst & Banks, 2002). When several modalities provide overlapping evidence about a property such as an object's location, their joint use can reduce uncertainty in the resulting percept, producing what is commonly described as *multisensory gain*, observable through enhanced speed or accuracy in precision of responses. In mature perceptual systems, this improvement often follows the predictions of the Maximum Likelihood Estimation (MLE) framework. Specifically, The MLE model proposed by Ernst and Banks (2002) describes multisensory integration as a statistically optimal process in which estimates from different sensory modalities are integrated by weighting them according to their reliability, defined as the inverse of the signal variance. In the presence of independent noise and Gaussian distributions, the combined estimate corresponds to a weighted sum of the single-sensory estimates, where each weight is proportional to the reciprocal variance of the corresponding modality; this integration produces a final estimate with lower variance than that of each sensory channel considered individually (Ernst & Banks, 2002). In other words, for the MLE model, integration can be described as an adaptive process according to which each sensory signal contributes to the final estimate in proportion to its reliability (Ernst & Banks, 2002).

1.3.1 Neural Basis of Spatial Multisensory Integration

From a neurophysiological perspective, early models described sensory brain organisation as strictly unisensory and hierarchical, with initial processing confined to modality-specific primary cortices and multisensory integration emerging only at later associative stages (Felleman & Van Essen, 1991; Jones & Powell, 1970). However, neuroimaging and electrophysiological evidence have challenged this view, showing that multisensory interactions can arise in subcortical structures and even within primary sensory cortices, indicating that multisensory processing is a distributed property of brain organisation rather than a late-stage function (Alais et al., 2010; Murray et al.,2016). The superior colliculus provides a key model of multisensory integration: neurons in its deep layers receive convergent visual, auditory, and somatosensory inputs and exhibit nonlinear response properties such as superadditivity for congruent stimuli and suppression for discrepant inputs, consistent with the principle of inverse effectiveness (Stein & Meredith, 1993; Stanford et al., 2005). Behaviourally, these mechanisms support faster and more accurate orienting toward congruent multisensory events and develop progressively through multisensory experience and cortical feedback rather than being fully present at birth (Wallace & Stein, 2001; Jiang et al., 2006; Alais et al., 2010). At the cortical level, multisensory integration involves parietal and temporal regions and is shaped by spatiotemporal congruence, perceptual context, and prior experience. Anatomical and functional studies demonstrate cross-modal influences both within primary cortices and through feedback from higher-order multisensory areas, supporting the view that multisensory convergence is intrinsic to cortical organisation rather than restricted to late associative processing (Calvert et al., 1999; Falchier et al., 2002; Ghazanfar & Schroeder, 2006; Kayser et al., 2009; Murray et al.,2016). Consistent with this framework, visual and somatosensory modulations have been observed in auditory cortex and across other primary sensory regions (Musacchia & Schroeder, 2009; King & Walker, 2012; Kuehn & Pleger, 2018; Chanauria et al., 2019).

1.4 Multisensory Development of Spatial Abilities

Multisensory spatial perception develops gradually, rather than being fully operational from birth. Although children are immersed in an environment rich in multisensory stimuli and show early sensitivity to cross-modal correspondences, such as orientation toward speaking faces, detection of audiovisual synchrony, and faster responses to bimodal stimuli, these initial abilities do not yet reflect efficient or statistically optimal integration (Baart, 2015; Bahrick et al., 2012; Lewkowicz, 1996, 2009).

From the earliest stages of life, multisensory stimulation contributes to the construction of initial spatial representations. The tactile system, already functional during gestation, supports early perception of the body and its relationship with the environment in the first months of life (Bremner & Spence, 2017). With development, vision gradually takes on a predominant role, offering immediate, high-resolution access to the spatial properties of the environment and promoting the coherent organization of information from other sensory modalities (Bremner et al., 2013; Tinti et al., 2006). In light of this, vision appears to play a central role in aligning neural representations of space across different modalities, markedly influencing auditory and tactile perceptions under conditions of sensory conflict (Welch & Warren, 1980; Botvinick & Cohen, 1998; Flanagan & Beltzner, 2000; Bertelson & Aschersleben, 2003; Anderson & Zahorik, 2014). Furthermore, during development, visual experience contributes to the calibration of auditory and proprioceptive spatial representation (Loomis et al., 1998; Gori et al., 2012b; Rigato et al., 2014). Studies on animal models further support this perspective, showing that visual adaptation can remodel auditory spatial maps, while early visual deprivation compromises their development (King & Carlile, 1993; Knudsen, 1998). Similarly, even in humans, brief periods of exposure to audiovisual discrepancies can alter auditory spatial representation (Recanzone, 1998; Zwiers et al., 2003). This transition takes several years and depends on the ability to integrate perceptual experiences to construct unified spatial representations (Ruggiero et al., 2016). Behavioural evidence indicates that the stable use of external spatial references emerges around the age of five and continues to refine until around the age of ten (Pagel et al., 2009; Röder et al., 2013).

Within this broader developmental window, particularly during the preschool years and extending into early school age, children undergo a critical transitional stage in the maturation of multisensory spatial processing. While they begin to show measurable behavioural benefits from multisensory stimulation, including faster response times, reduced variability, and improved localisation accuracy when auditory and visual cues are presented together

compared to unimodal conditions (Barutchu et al., 2009), these improvements do not yet reflect statistically optimal cue combination. Instead, multisensory perception during the preschool years appears to rely primarily on intersensory calibration processes, whereby the most accurate modality progressively refines the spatial estimates of the others (Gori et al., 2008; Burr & Gori, 2012). In other words, multisensory integration follows a progressive developmental trajectory in which early multisensory sensitivity observed in infancy evolves into increasingly stable and effective spatial performance during the preschool years, while still remaining distinct from the efficient and statistically optimal integration characteristic of adulthood. This pattern highlights the gradual refinement of multisensory spatial processing across childhood, reflecting age-related improvements in accuracy, response speed, and consistency of behaviour.

1.4.1 Trajectory of Multisensory Neural Development

In most multisensory brain regions, the capacity to integrate inputs from different sensory modalities is not present at birth but instead develops progressively across the lifespan. Evidence from the Superior Colliculus indicates that immature neurons may display rudimentary multisensory responsiveness while lacking genuine integrative properties, such as multisensory enhancement or depression, further supporting the dependence of functional integration on postnatal maturation and sensory experience (Stein et al., 2009; Wallace et al., 2004). The development of descending afferents from extra- primary cortical regions, including the anterior ectosylvian sulcus and the rostral segment of the lateral suprasylvian sulcus, is considered critical for acquiring integrative capacities within the SC (Stein et al., 2009).

At the cortical level, multisensory integration follows a prolonged developmental trajectory. Nonlinear neural interactions, indexed by superadditive or subadditive evoked responses relative to the sum of unisensory activity, represent a key physiological signature of integration (Murray et al., 2005). In adults, multisensory effects emerge both within associative cortices at later latencies and within traditionally unisensory areas at early latencies (< 100 ms), indicating distributed hierarchical processing (Foxe et al., 2000; Fort et al., 2002; Gobbelé et al., 2003; Molholm et al., 2002, 2004; Murray et al., 2004, 2005; Taylor-Clarke et al., 2002; Foxe & Schroeder, 2005). Developmental electrophysiological findings further indicate progressive refinement of these mechanisms: consistent audio-tactile integration is evident in school-age children between approximately 60 and 220 ms across central and associative regions (Brett-Green et al., 2008), with additional temporal-lobe

involvement emerging around 100 ms (Russo et al., 2010). In contrast, mature audiovisual integration within associative cortices appears to develop between approximately 7 and 16 years of age (Brandwein et al., 2011; Molholm et al., 2020; Stefanou et al., 2020), whereas adolescence is characterized by multisensory responses already present at early latencies (< 100 ms) and extending to trimodal interactions (Dwyer et al., 2022). Electrophysiological evidence further indicates that primitive neural coding of bodily boundaries during audio-tactile stimulation emerges shortly after birth (Maitre et al., 2020; Ronga et al., 2021). Within this framework, evoked-potential methodologies provide a crucial tool to characterise the temporal dynamics of these processes (Luck, 2014; Murray et al., 2008; Talsma & Woldorff, 2005). In particular, ERPs enable the identification of both early and late integration signatures, reflecting the hierarchical and distributed organisation of multisensory processing (Foxe & Schroeder, 2005; Ghazanfar & Schroeder, 2006; Kayser & Logothetis, 2007). Complementary analyses of neural oscillations may further elucidate the neurophysiological mechanisms underlying multisensory integration (Senkowski et al., 2008; Keil & Senkowski, 2018).

1.5 Multisensory Development with Sensory Impairment

From the earliest phases of multisensory spatial development, vision provides a fundamental spatial framework through which infants learn to interpret the location of events relative to their body and the external environment (Gori et al., 2021; Rigato et al., 2014; Orioli et al., 2024). Converging evidence indicates that visual input actively drives the maturation of multisensory spatial perception, refining auditory and tactile processing through cross-modal calibration mechanisms that progressively align sensory representations across modalities (Röder et al., 2004; Begum Ali et al., 2015; Thomas et al., 2018; Azañón et al., 2018; Gori et al., 2008, 2014, 2015, 2020). When vision is absent or severely reduced, this calibration process is disrupted, leading to delayed alignment between auditory and tactile spatial maps and broader spatial difficulties extending beyond perception to proprioceptive judgments, spatial imagery, and auditory localisation (Cappagli & Gori, 2016; Cappagli et al., 2017, 2019; Vercillo et al., 2016; Gandhi et al., 2014).

Consistent with this developmental framework, research on spatial processing in visual impairment reveals a complex pattern in which compensatory plasticity and perceptual limitations coexist. Studies in congenitally blind adults report enhanced auditory abilities in several domains of spatial abilities (Röder et al., 1999; Voss et al., 2004; Kolarik et al., 2013), potentially supported by compensatory recruitment of visual cortices by auditory and

somatosensory inputs (Poirier et al., 2005, Collignon et al., 2011). Developmental evidence further shows that visually impaired infants display broader differences and delays across spatial domain relative to sighted peers (Elisa et al., 2002; Cappagli et al., 2016) consistent with less precise emerging spatial representations. At the multisensory level, audio-tactile stimulation produces orienting facilitation relative to unisensory conditions, although with reduced multisensory gain compared to sighted infants, indicating that spatial integration across hearing and touch is present but constrained (Gori et al., 2021). Under spatial conflict, sighted infants preferentially orient toward auditory cues, whereas visually impaired infants rely more strongly on tactile information, revealing the absence of a shared visuo-spatial reference frame for multisensory integration (Gori et al., 2021).

Taken together, findings from both infancy and adulthood demonstrate that the absence of vision does not simply delay spatial abilities but fundamentally reshapes the developmental trajectory of multisensory spatial representations. Visual deprivation leads to task-dependent enhancements and impairments in adulthood, while in infancy, multisensory integration is present yet reduced, unisensory orienting is relatively preserved, tactile reliance is increased, and spatially guided motor development is delayed relative to typical visual development.

1.5.1 Neurophysiological perspective of Multisensory Development with Visual Impairment

From a neurophysiological perspective, studies showed cortical reorganisation in visually deprived individuals, including posterior shifts of auditory processing, recruitment of occipital cortices by nonvisual modalities, and altered multisensory interactions within parietal networks (Röder et al., 1999; Collignon et al., 2015; Ortiz-Terán et al., 2016; Wallace et al., 2004; Carriere et al., 2007; Hotting & Röder, 2009; Occelli et al., 2013). In addition, electrophysiological findings indicate that newborns already display audio-somatosensory responses consistent with early multisensory processing and show stronger responses when sounds occur near the body, suggesting primitive representations of bodily boundaries and peripersonal space (Maitre et al., 2020; Ronga et al., 2021). Together, these results indicate that early visual experience shapes the cortical organisation and spatial strategies underlying multisensory integration, whereas its absence leads to altered developmental trajectories rather than a complete loss of multisensory function.

1.6 MySpace Framework

This doctoral project draws on the theoretical framework of the ERC Horizon 2020 (No.

948349) project "MySpace" (PI Monica Gori). MySpace introduces an integrated approach that connects perceptual development, visual impairment, and rehabilitation, with the aim of understanding the encoding of the space. Specifically, this European project aims to clarify how the brain develops a multisensory representation of space capable of maintaining consistency between the spatial coordinates provided by different sensory modalities. Numerous studies have highlighted the role of visual experience in both the development of multisensory integration (Wallace et al., 2004) and spatial perception (Eimer, 2004; Gori, 2015; Senna et al., 2021). In this context, MySpace specifically investigates the contribution of vision to the integration of sensory information within a unified spatial map.

The theoretical framework of MySpace assumes that vision calibrates other sensory modalities with regard to spatial information. In the absence of visual experience, a position encoded in tactile coordinates is difficult to translate into auditory coordinates. Consistently, blind children show delays in the development of motor and spatial skills (Gori, 2015): the ability to reach for objects emerges around 10–11 months compared to around five months in sighted children (Bigelow, 1986; Tröster & Brambring, 1993), and difficulties in locating sounds can persist into adolescence (Gougoux et al., 2009; Vercillo et al., 2016), despite the integrity of auditory and tactile modalities. A decline in spatial abilities has also been observed in adulthood, contrasting with the hypothesis of generalized sensory compensation which predicts enhanced performance in the remaining sensory modalities following visual deprivation (Champoux et al., 2011; Hötting & Röder, 2004; Röder et al., 2007).

Studies on animal models further highlight the importance of early visual experience in the development of spatial representations. In barn owls, auditory spatial maps adapt to the displacement of visual input induced by prismatic lenses only during a critical period in early life (Knudsen & Knudsen, 1985), whereas this recalibration does not occur in the absence of early visual experience. Similarly, vision dominates multisensory perception in orientation tasks in children (Gori et al., 2008), influencing sensorimotor associations (Gori et al., 2012) and spatial multisensory integration (Calvert, 2004).

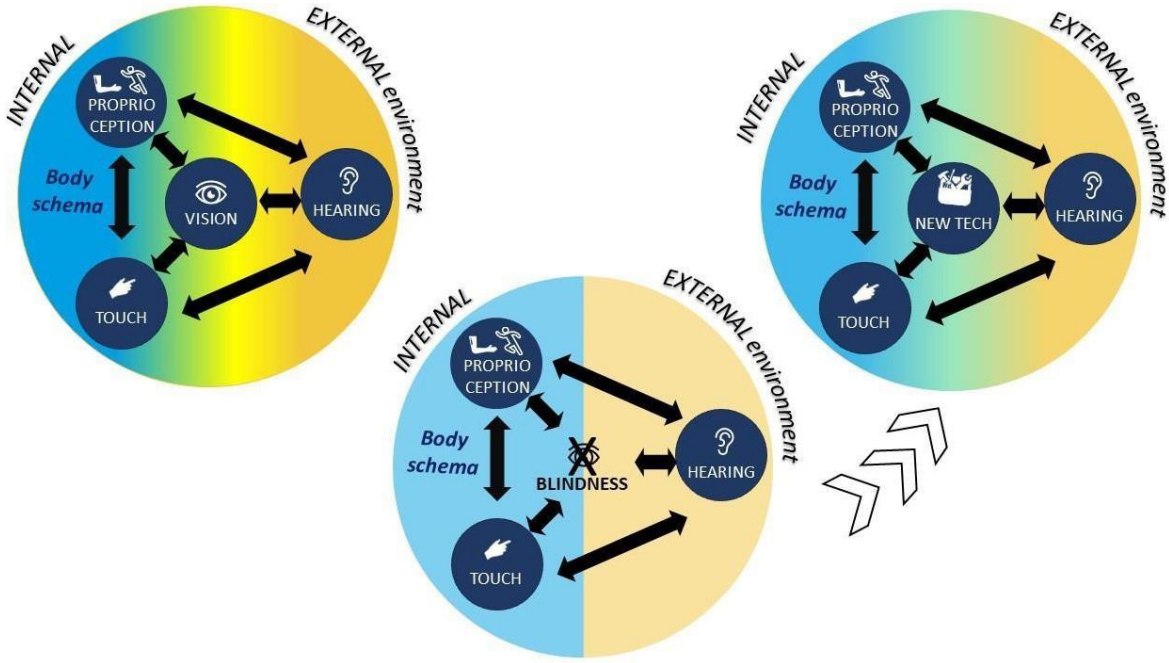

Figure 1.1 Conceptual Framework of MySpace ERC

To address these issues, the MySpace project is organised into four main objectives, articulated across four Work Packages (WPs). The first objective (WP1) aims to clarify the role of vision in the development of auditory and tactile spatial abilities considered separately, at both the behavioural and cortical levels. Building on this foundation, the second objective (WP2) focuses on the development and temporal windows of multisensory spatial representations, specifically in the audio-tactile and audio-visual domains. The third objective (WP3) extends this investigation to the processes of cortical reorganization, particularly within the visual cortex, in supporting spatial processing in the absence of vision. Finally, the fourth and last objective (WP4) represents the translational phase of the project and involves leveraging the findings from previous experiments to develop targeted interventions aimed at improving impaired spatial localisation abilities in blind children.

Within this framework, the present doctoral project primarily contributes to WP1 through the investigation of auditory spatial localisation, with the goal of understanding the role and influence of vision on the development of auditory spatial abilities. It further contributes to WP2 by examining audiovisual multisensory integration processes, and to WP3 by investigating the neurocortical mechanisms underlying multisensory integration in sighted and visually impaired children.

1.7 Objectives of the thesis

The present doctoral thesis aims to provide new insights into the development of spatial abilities during early life, spanning from infancy to preschool age. Given that spatial representations are tightly linked to sensory development, this work investigates spatial perception and spatial skills in both typically developing children and children with visual impairment, in order to clarify the role of vision and the adaptive mechanisms that emerge in its absence. To achieve a comprehensive understanding, both behavioural and neurophysiological (EEG) approaches were employed.

More specifically, the scientific objectives of this thesis are:

1. To characterise the development of auditory spatial localisation abilities in early childhood, identifying their developmental trajectory and limitations across spatial dimensions.
2. To examine how children integrate audiovisual spatial information in the presence of spatial discrepancies and during dynamic object segregation tasks, and to assess whether such integration follows principles of optimal cue combination, such as Maximum Likelihood Estimation (MLE) and to account for inter-individual variability in sensory weighting strategies.
3. To investigate the development and neural mechanisms of multisensory integration, with a focus on audio-tactile processing in typically developing and visually impaired children, and to evaluate how visual experience shapes sensory weighting, spatial representations, and cortical organisation.

Technological Objective

In addition to the scientific aims, this thesis also includes a technological objective:

To design and validate a child-friendly experimental device for the study of spatial abilities, ensuring that it is portable, easy to use, and suitable for young participants, while providing clear and reliable behavioural outputs.

Chapter 2

Development of Auditory Spatial Abilities across Spatial Dimensions

2.1 Introduction

Sound cues are fundamental anchors of the external world, enabling individuals to orient toward events beyond the visual field and guide actions in space. Unlike vision, which provides high spatial resolution, audition relies on temporal information and must infer spatial properties indirectly from acoustic signals that vary with the listener's position (Kubovy & Van Valkenburg, 2001; Blauert, 1997). Accordingly, auditory spatial perception depends on the integration of cues shaped by the interaction between sound waves, the body, and the environment (Blauert, 1997; Carlini et al., 2024).

Auditory spatial perception is inherently multidimensional, involving azimuth, elevation, and distance, each supported by partially distinct cues (Blauert, 1997; Carlini et al., 2024). Horizontal localisation relies on binaural cues, such as interaural time and level differences, which provide relatively reliable spatial information (Makous & Middlebrooks, 1990). In contrast, vertical localisation depends on monaural spectral cues shaped by the head and pinnae, which are more ambiguous and anatomically dependent, resulting in greater uncertainty (Freigang et al., 2015). Consequently, localisation is typically less precise in elevation than in azimuth (Otte et al., 2013), reflecting differences in cue reliability.

These properties make auditory localisation a useful domain for studying development. Across childhood, the auditory system matures while changes in head size and pinna shape continuously alter available cues, requiring ongoing recalibration between acoustic input and spatial representation (Keating & King, 2015; Anbuhl et al., 2017). Although infants can orient to sounds early, spatial precision improves gradually and unevenly across dimensions, with vertical localisation remaining less accurate and more variable than horizontal

localisation (Morrongiello, 1987; Freigang et al., 2015; Otte et al., 2013).

In addition to spatial accuracy, the temporal dynamics of localisation responses provide complementary insights into auditory spatial development. Reaction time or response latency reflects the efficiency with which spatial information is extracted, integrated, and transformed into an overt response, encompassing perceptual, decisional, and sensorimotor components (Mittelstädt et al., 2020). In developmental populations, longer response latencies are commonly observed and are thought to reflect ongoing maturation of attentional, perceptual, and sensorimotor processes, as well as increased demands on evidence accumulation under conditions of uncertainty (Kail, 1991; Eklöf et al., 2022). Importantly, reaction times can reveal differences in processing efficiency even when spatial accuracy differences are subtle or absent, making them a sensitive complementary index of developmental change (Freigang et al., 2015; Eklöf et al., 2022).As such, response latency represents a valuable and complementary index for characterising developmental changes in auditory spatial processing (Eklöf et al., 2022). In this chapter, auditory spatial localisation was examined across the azimuth and elevation planes by comparing performance in children and adults. Using a custom device equipped with speakers and tactile sensors, spatial accuracy and response latency were analysed to characterise developmental differences in localisation performance. In particular, localisation error along the horizontal and vertical axes was considered to assess how cue reliability may shape spatial behaviour. This chapter examines whether auditory localisation differs across spatial planes and developmental groups, and whether these differences are also reflected in response latency. Finally, to explore the role of visual experience in auditory spatial development, particular attention was given to performance in children.

2.2 Methods

2.2.1 Participants

A total of 45 participants took part in the study, including 21 children (mean age = 5.13 years, SD = 0.68) and 24 adults (mean age = 28.89 years, SD = 3.16). Children were recruited from local preschools in the city of Genova (Italy), while adults were tested in our lab. Written informed consent was obtained from their parents or legal guardians prior to participation. Adult participants provided written informed consent themselves. All experimental procedures were conducted in accordance with the Declaration of Helsinki and were approved by the local ethics committee.

2.2.2 Apparatus & Stimuli

The experiment was conducted using a custom-built auditory localization device composed of 25 loudspeakers arranged in a 5 × 5 matrix, 10 cm x 10 cm each (Figure 2.1). These are equipped with 16 tactile sensors or pads (2 cm x 2 cm each) (Figure 2.2). The total size is 50 cm x 50 cm x 10 cm (Setti et al., 2019). Auditory stimuli consisted of pink noise bursts. In order to make the complexity of the task appropriate for children, four loudspeakers out of twenty-five were selected from the array: two corresponding to stimulation along the elevation plane and two corresponding to stimulation along the azimuthal plane (Figure 2.3). On each trial, one of these four loudspeakers was randomly selected to emit the sound.

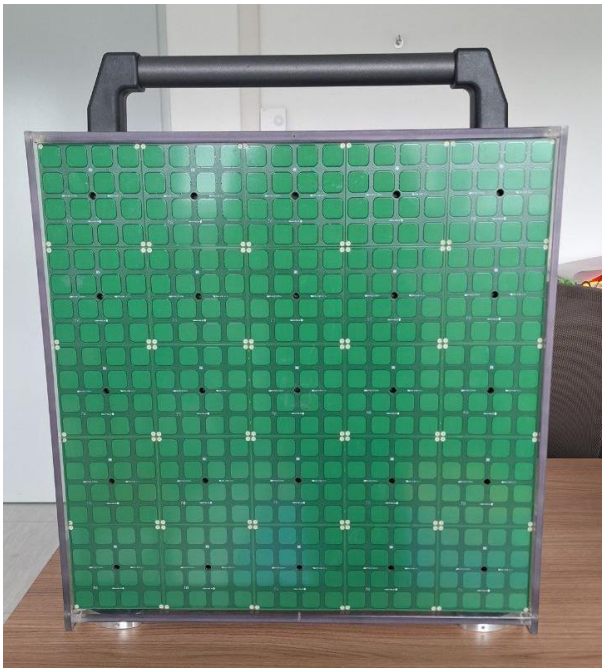

Figure 2.1 ARENA Device (Setti et al., 2019)

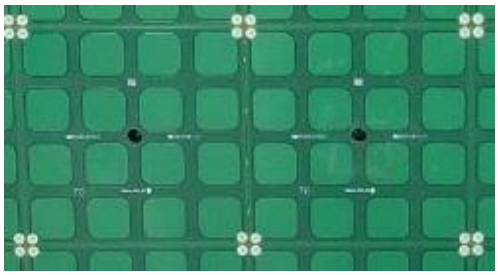

Figure 2.2 Tactile sensors

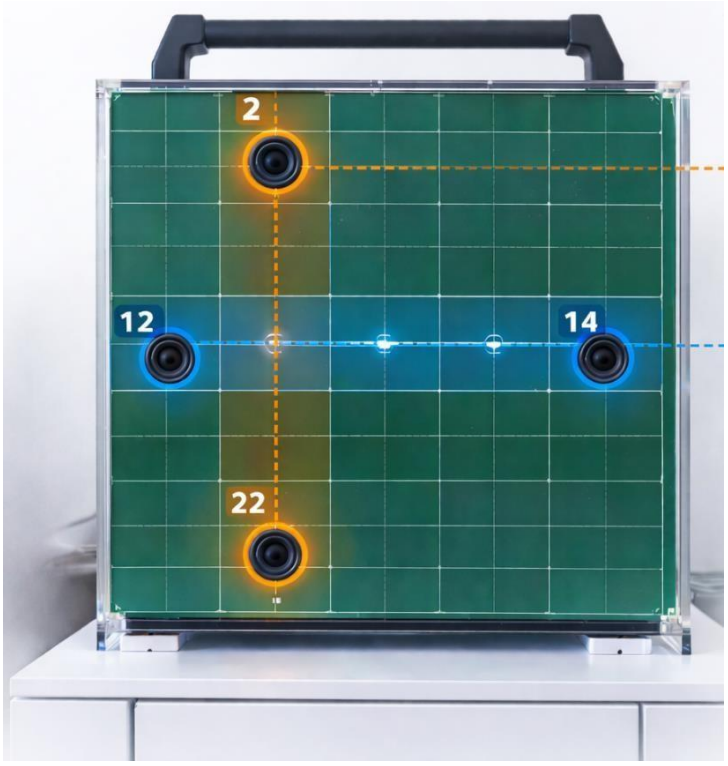

Figure 2.3 Azimuth Plane (Loudspeaker 12 and 14) Elevetion Plane (loudspeaker 2 and 22). This image was edited using AI for visualization purpose.

2.2.3 Procedures

The experiment took place in a quiet room with dim lighting in order to minimise external distractions. Participants were seated in front of the loudspeaker array at a fixed distance of 20 cm. They were positioned centrally with respect to the device, such that the tip of the nose was aligned with the center of the loudspeaker matrix, ensuring a symmetric spatial reference frame across participants. Participants were instructed to keep their eyes closed throughout the task so that localization relied exclusively on auditory information. At the beginning of each trial, a pink noise stimulus was emitted from one of the four selected loudspeakers in a randomized order. Participants were instructed to localize the sound source as quickly and accurately as possible by reaching out and touching the loudspeaker they perceived as the source of the sound. Touch responses were detected through the tactile sensors embedded in the loudspeaker array. Reaction time was measured from sound onset to the moment the participant touched a loudspeaker. Both the response location and the reaction time were recorded for subsequent analyses.

2.3 Data Analysis

2.3.1 Spatial localization accuracy

Spatial localization accuracy was assessed by quantifying the discrepancy between the target sound source and the participant's touch response on the loudspeaker array. For each trial, both the target loudspeaker and the selected response were mapped onto planar coordinates (x, y) using a conversion matrix that assigned a specific spatial position to each loudspeaker. Localization error was computed as the difference between response and target coordinates. Three complementary error measures were derived to characterize localization performance at different levels. First, an absolute spatial error (hereafter referred to as diagonal error) was calculated as the Euclidean distance between the response and the target location, providing a global index of localization accuracy that jointly captures horizontal and vertical deviations. Second, two unidimensional error components were computed. The horizontal error component (X error) quantified the deviation of the response along the left–right axis, whereas the vertical error component (Y error) quantified the deviation along the up–down axis. For the purposes of the present analyses, the absolute values of X and Y errors were

considered, thus indexing the magnitude of localization error independently of its direction. Trials were classified according to the plane of stimulation. Sounds emitted from loudspeakers positioned along the horizontal plane were assigned to the azimuth condition, whereas sounds emitted from loudspeakers positioned along the vertical plane were assigned to the elevation condition. This distinction allowed direct comparison of localization accuracy across spatial planes relying on partially distinct auditory cues. For each participant, localization errors were averaged across trials separately for each plane of stimulation, yielding one mean value per subject and condition for each spatial metric (diagonal error, absolute X error, and absolute Y error). These subject-level averages constituted the dependent variables for subsequent statistical analyses. To assess developmental differences in localization performance, separate mixed-design analyses of variance (ANOVAs) were conducted for each error metric (diagonal, X, and Y errors). In each analysis, Group (Children, Adults) was included as a between-subject factor, and Plane of stimulation (Azimuth, Elevation) as a within-subject factor. This approach allowed us to examine whether localization accuracy differed across age groups and spatial planes, as well as whether these effects varied across different components of spatial error. Post-hoc comparisons between conditions were performed using pairwise t-tests, and p-values were adjusted using Bonferroni correction.

2.3.2 Reaction Times

Reaction times (RTs) were analysed to investigate potential differences in processing speed across spatial planes. To obtain a reliable estimate of perceptual–decisional timing, analyses were restricted to correct trials only, defined as trials in which the participant's response corresponded to the stimulated loudspeaker. This selection criterion ensured that reaction times were not confounded by response uncertainty or post-error adjustments.

For each participant, reaction times were averaged across correct trials separately for each plane of stimulation, yielding one mean RT per subject and condition. These subject-level averages were entered into subsequent statistical analyses. Reaction times were analysed using a mixed-design analysis of variance (ANOVA), with Group (Children, Adults) as a between-subject factor and Plane of stimulation (Azimuth, Elevation) as a within-subject factor. This approach tested whether response speed differed between spatial planes and whether such differences varied as a function of developmental group, while accounting for the repeated-measures structure of the data at the subject level.

2.4 Results

2.4.1 Results for Spatial localization Accuracy

Diagonal Error

For the Diagonal Error, the test revealed a significant main effect of Group, $F(1, 43) = 85.16$, $p < .001$, $\eta^2g = .53$, and a significant main effect of Plane of stimulation, $F(1, 43) = 69.43$, $p < .001$, $\eta^2g = .41$. The Group × Plane interaction was also significant, $F(1, 43) = 8.90$, $p = .005$, $\eta^2g = .08$. Post-hoc pairwise comparisons with Bonferroni correction revealed a significant difference between azimuth and elevation conditions ($p < .001$). This difference was significant in both children ($p < .001$) and adults ($p < .001$). (Figure 2.4A)

Y error

For the Y error, the analysis revealed a significant main effect of Group, $F(1, 43) = 54.91$, $p < .001$, $\eta^2g = .39$, and a significant main effect of Plane of stimulation, $F(1, 43) = 64.19$, $p < .001$, $\eta^2g = .43$. The Group × Plane interaction was also significant, $F(1, 43) = 12.73$, $p < .001$, $\eta^2g = .13$. Post-hoc pairwise comparisons with Bonferroni correction revealed a significant difference between azimuth and elevation conditions ($p < .001$). This difference was significant in both children ($p < .001$) and adults ($p < .001$). (Figure 2.4B)

X error

For the X error, the analysis revealed a significant main effect of Group, $F(1, 43) = 53.02$, $p < .001$, $\eta^2g = .50$. The main effect of Plane of stimulation was not significant, $F(1, 43) = 3.45$, $p = .070$, $\eta^2g = .02$. The Group × Plane interaction was not significant, $F(1, 43) = 0.009$, $p = .927$, $\eta^2g < .001$. No post-hoc comparisons were conducted for this metric, as neither the main effect of Plane nor the interaction reached significance. (Figure 2.4C)

Individual participant data are shown in Figure 2.5.

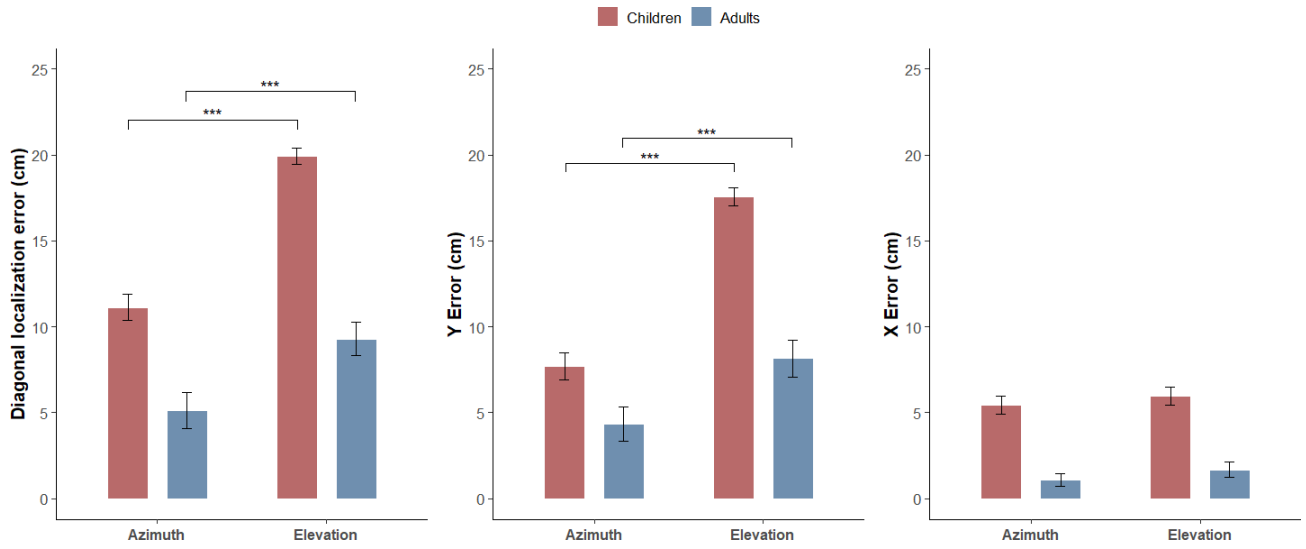

Figure 2.4 Spatial Localization Accuracy in Children and Adults. Diagonal error reflects the overall spatial error combining both dimensions Y error represents vertical deviations (above vs. below the target), X error represents horizontal deviations (left vs. right).

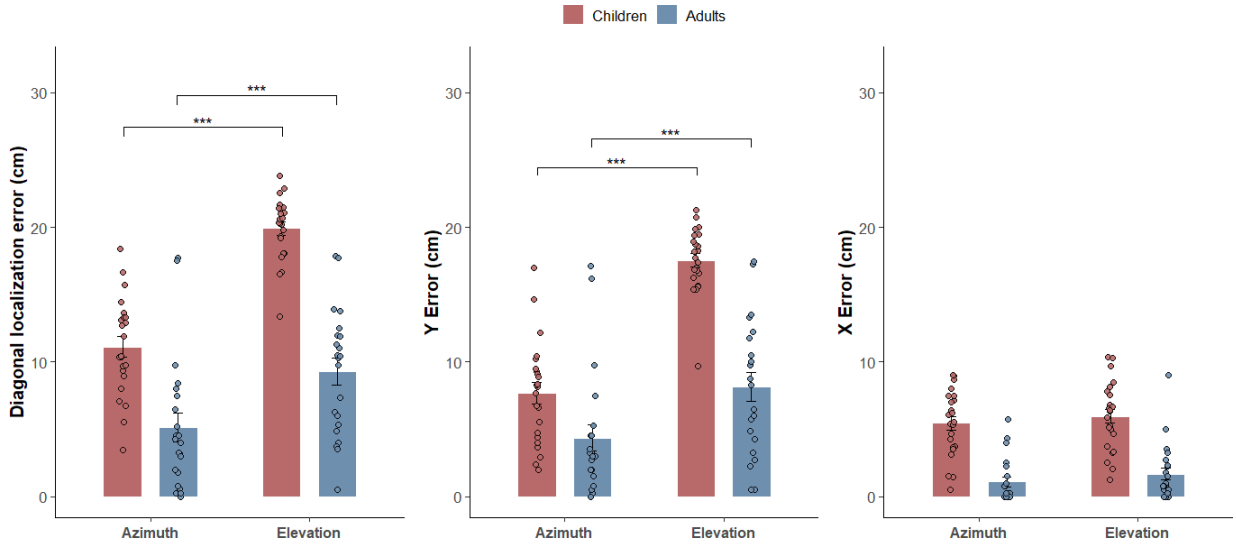

Figure 2.5. Diagonal error reflects the overall spatial error combining both dimensions Y error represents vertical deviations (above vs. below the target), X error represents horizontal deviations (left vs. right). Individual points indicate the distribution of participant's performance.

2.4.2 Results for Reaction Times

Reaction times were analysed using a mixed-design ANOVA with Group (Children, Adults)

as a between-subject factor and Plane of stimulation (Azimuth, Elevation) as a within-subject factor. The analysis revealed a significant main effect of Group, indicating that overall reaction times differed between children and adults, $F(1, 41) = 5.26$, $p = .027$, with children responding more slowly than adults. In contrast, the main effect of Plane of stimulation was not significant, $F(1, 41) = 0.02$, $p = .88$, indicating that reaction times did not differ between azimuth and elevation conditions. The Group × Plane of stimulation interaction was also not significant, $F(1, 41) = 2.42$, $p = .13$, suggesting that the effect of spatial plane on reaction times did not differ between children and adults (Figure 2.6).

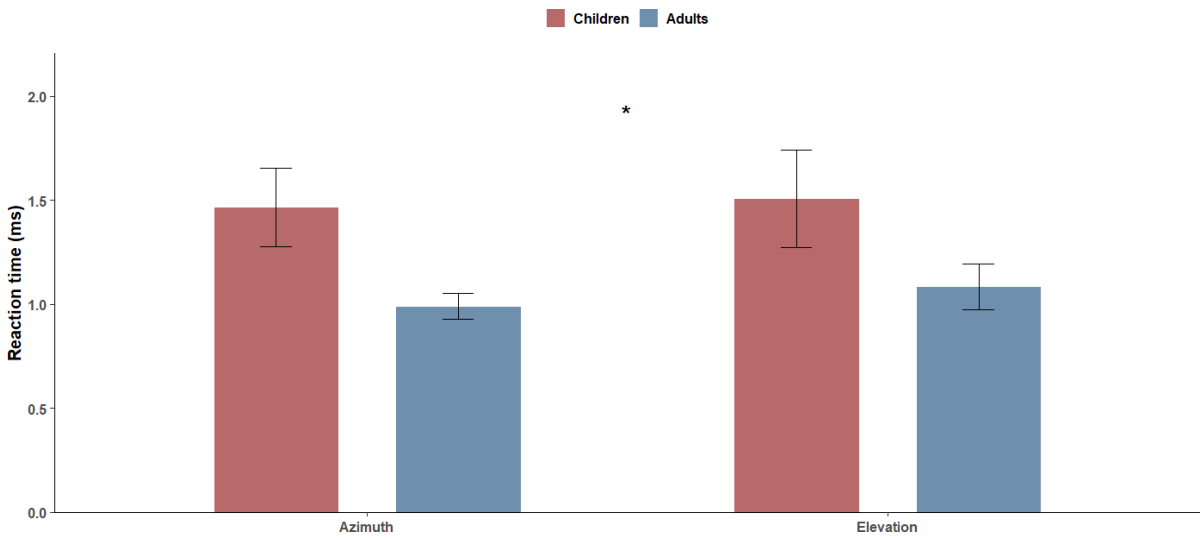

Figure 2.6. Reaction times in children and adults during auditory localisation in azimuth and elevation

2.5 Exploratory Analysis in Several Impaired Children

Exploring how auditory spatial representations develop in cases of visual impairment can provide a useful opportunity to study the mechanisms underlying spatial perception. Vision plays a crucial role in calibrating spatial representations during development, contributing to the alignment of sensory signals within a coherent spatial framework (Burr & Gori, 2012; Gori et al., 2012). From this perspective, studying visually impaired children allows us to distinguish which aspects of spatial perception depend on visual experience and which can develop independently. When it comes to auditory spatial localisation, findings from developmental studies consistently indicate a different pattern compared to adulthood.

Although several studies conducted on blind adults have reported compensatory improvements in auditory processing (Voss et al., 2004; Voss et al., 2014), evidence from developmental populations suggests that visually impaired children show reduced accuracy in auditory spatial tasks compared to their sighted peers, reflecting a delay in the maturation of spatial representations (Cappagli & Gori, 2016; Gori et al., 2014; Vercillo et al., 2015). At the same time, not all components of auditory spatial processing appear to depend equally on vision. For instance, localisation in the horizontal plane, which relies on binaural cues, may be relatively preserved, whereas performance in the vertical plane, which depends on more ambiguous spectral cues, tends to be more variable and less accurate (Blauert, 1997; Lewald, 2002). These findings suggest that different spatial dimensions may follow distinct developmental trajectories depending on the availability of visual input. Starting from these statements, to gain a broader understanding of auditory spatial localisation abilities in early childhood, one of the aims of the present study was to explore potential differences between typically developing children and children with visual impairments, with the goal of clarifying the role of visual experience in the development of spatial representations. Unfortunately, due to the challenges associated with recruiting participants from a sensitive clinical population, it was not possible to obtain a sample comparable in size or age distribution to that of the typically developing group. As a result, the available data do not support formal group-level comparisons and are instead presented as preliminary and exploratory observations.

2.5.1 Data Analysis

Spatial localization accuracy and reaction times were analysed using the same procedures described for the sighted participants. All analyses were replicated on the SVI sample within an exploratory framework, including the computation of subject-level averages and repeated-measures ANOVAs with Plane of stimulation as a within-subject factor. Similarly, to assess the effect of condition on reaction times in SVI participants, a repeated-measures ANOVA was conducted.

2.5.2 Results

To examine the effect of condition on spatial performance in SVI participants, repeated-measures ANOVAs were conducted on diagonal, horizontal (x-axis), and vertical (y-axis) errors. A significant effect of condition emerged for the diagonal error, $F(1, 5) = 118.98$, $p <$

.001, indicating a strong difference between Azimuth and Elevation. Pairwise comparisons confirmed this result ($p < .001$, Bonferroni-corrected), with lower errors observed in the Azimuth condition. A similar pattern was found for the vertical component (y-axis), which also differed significantly between conditions, $F(1, 5) = 124.70$, $p < .001$. In contrast, the horizontal component (x-axis) did not vary across conditions, $F(1, 5) = 0.17$, $p = .70$. (Figure 2.7). Individual participant data are shown in Figure 2.8.

Regarding the Reaction Times, no significant effect of condition emerged, $F(1, 1) = 0.05$, $p = .86$, indicating comparable reaction times across Azimuth and Elevation (Figure 2.9).

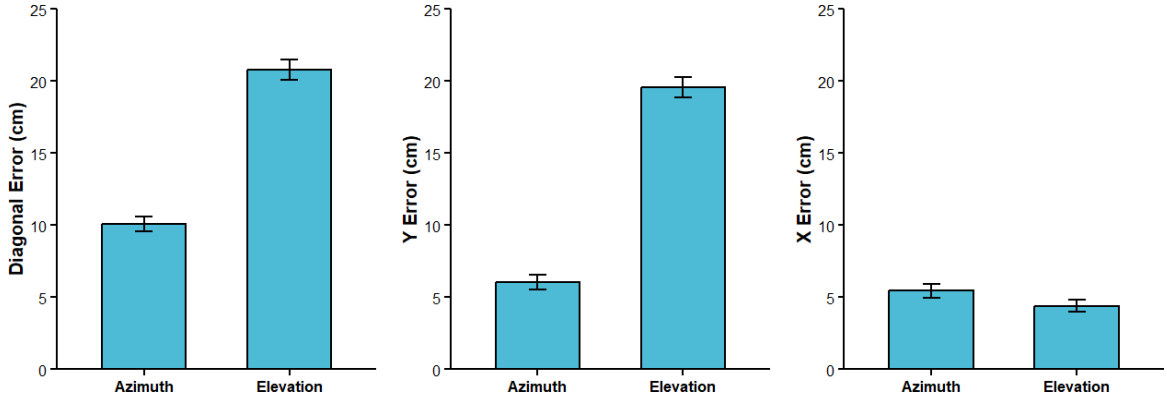

Figure 2.7 Spatial Localization Accuracy in SVI children. Diagonal error reflects the overall spatial error combining both dimensions Y error represents vertical deviations (above vs. below the target), X error represents horizontal deviations (left vs. right).

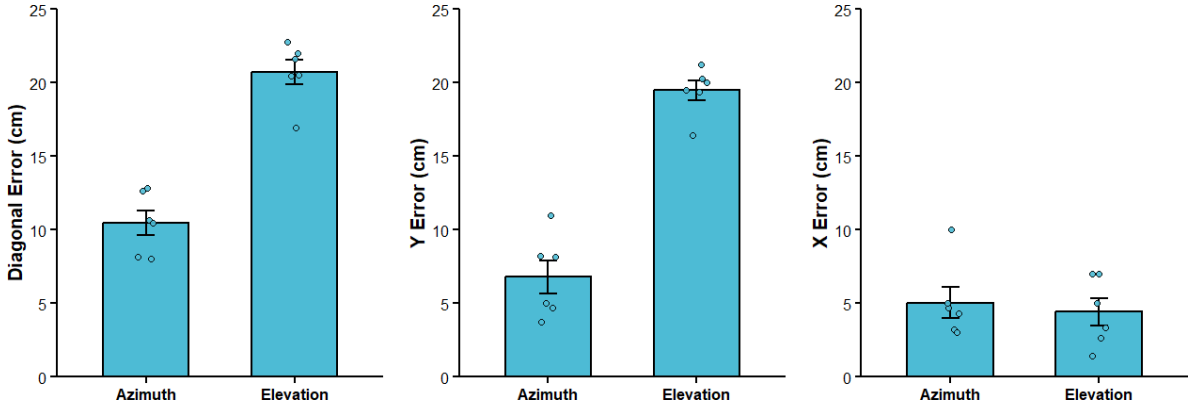

Figure 2.8 Spatial Localization Accuracy in SVI children. Diagonal error reflects the overall spatial error combining both dimensions. Y error represents vertical deviations (above vs. below the target), X error represents horizontal deviations (left vs. right). Individual points indicate the distribution of participants' performance.

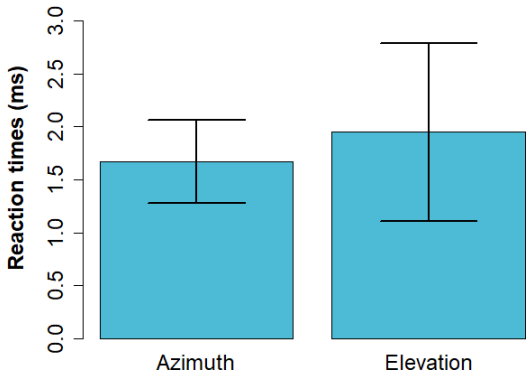

Figure 2.9 Reaction times in SVI children during auditory localisation in azimuth and elevation plane.

2.6 General Discussion

This chapter investigated the development of auditory sound localization on the frontal plane from preschool age to adulthood, comparing azimuth and elevation judgments as well as spatial precision along horizontal and vertical axes. Children showed larger localization errors than adults, consistent with the prolonged maturation of auditory spatial representations reported in developmental research (Morrongiello, 1987; Freigang et al., 2015; Otte et al., 2013). Across groups, localization accuracy differed between spatial planes, with lower performance in elevation compared to azimuth. However, this effect was not uniform across age groups, as indicated by significant interactions between group and plane of stimulation. This suggests that the magnitude of the azimuth–elevation difference varies across development, rather than reflecting a fully shared performance pattern. When examining error components, differences between planes were not observed for the horizontal dimension, whereas the vertical component showed a robust effect of plane and a significant interaction

with group. This indicates that the observed differences in localization performance across planes are primarily associated with the vertical dimension. This pattern is consistent with the known properties of auditory spatial cues, whereby elevation relies on monaural spectral information, which is inherently more ambiguous and variable than the binaural cues supporting horizontal localization (Blauert, 1997; Makous & Middlebrooks, 1990; Freigang et al., 2015). The persistence of reduced vertical precision in adults suggests that these differences reflect modality-specific constraints rather than developmental immaturity alone. Improvements across childhood therefore appear to reflect increased precision, while differences between spatial dimensions remain evident across development. Reaction time findings further indicated slower responses in children, consistent with ongoing perceptual and motor maturation (Kail, 1991; Eklöf et al., 2022), but no differences between azimuth and elevation within groups, suggesting that vertical difficulty primarily affects spatial accuracy rather than response speed. Finally, while vision contributes to developmental spatial calibration (Gori et al., 2008), the persistence of elevation deficits in adults supports the view that auditory spatial localization is constrained by modality-specific processing mechanisms, with vision refining but not fundamentally reshaping auditory spatial coding (Blauert, 1997; Keating & King, 2015). Consistent with this account, preliminary observations from a small sample of visually impaired participants (SVI, $n = 6$) revealed a qualitatively comparable pattern, with reduced accuracy in elevation relative to azimuth, driven by differences along the vertical dimension, and no modulation in the horizontal component or reaction times. Although these findings must be interpreted with caution given the limited sample size, their convergence with the pattern observed in sighted participants tentatively supports the idea that the azimuth–elevation asymmetry primarily reflects intrinsic properties of auditory spatial cues, particularly the greater ambiguity and variability of spectral information underlying elevation, rather than a dependence on visual experience alone. These observations, while preliminary, should be interpreted with caution and remain partly speculative, but point to a more nuanced role of vision: rather than determining the structure of auditory spatial representations, visual input may act to calibrate and stabilise them across development.

Although visual experience does not appear to be essential for the emergence of basic auditory spatial abilities, likely reflecting the maturation of auditory processing and anatomical changes affecting cue reliability, its role becomes more critical when multiple sensory signals must be combined, for example, in the case of audiovisual stimuli. In early

development, it remains unclear whether audiovisual inputs lead to a multisensory benefit, modality dominance, or selective interference. These questions are particularly relevant in dynamic spatial contexts, where information unfolds over time and requires continuous updating of spatial representations. To address these questions, the following chapter introduces a novel experimental approach based on controlled dynamic audiovisual stimulation, designed to investigate multisensory spatial processing in childhood.

Chapter 3

Dynamic multisensory device for investigating perceptual development in young children

As discussed in Chapter 2, auditory spatial localisation emerges early in life but continues to refine throughout development, reaching optimal performance only in adulthood (Carlini et al., 2024). Our findings from the spatial auditory localisation study suggest that spatial representations do not mature uniformly across dimensions. While children show relatively accurate performance in azimuth, they exhibit reduced precision and greater variability in elevation. These results may indicate that, although auditory cues support early interaction with the spatial environment, the auditory system remains progressively tuned during childhood.

Nevertheless, spatial perception in everyday environments rarely relies on a single sensory modality. Instead, it typically arises from the coordinated processing of information across multiple sensory cues. The next step is therefore to examine how multisensory mechanisms contribute to the development of spatial abilities. Investigating these processes requires experimental approaches that are both tightly controlled and suitable for young participants. To meet this challenge, a novel device was developed to present dynamic audiovisual stimulation in a child-friendly and adaptable format. This methodological framework underpins the studies presented in this chapter, which explore the development of multisensory spatial processing in early childhood.

3.1 Introduction

Human perception relies on the brain's ability to extract and combine information coming from multiple sensory modalities; the nervous system often engages in a dynamic integration of cues such as vision, audition, and touch to generate a coherent and stable representation of the surrounding environment (Burr & Gori, 2012; Ernst & Banks, 2002). This process is crucial for spatial cognition: by combining complementary information across senses, adults can form accurate estimates of object location and interact efficiently with the environment (Ernst & Banks, 2002; Burr & Gori, 2012). When the brain processes information from multiple sensory cues, it does not merely determine whether these signals should be integrated. Spatial and temporal coherence provide critical information about whether cues

are related. Beyond this, the brain must preserve the ability to keep sensory signals distinct when they arise from separate objects or events, particularly in dynamic environments (Spence, 2013). This process, commonly referred to as sensory or source segregation, allows individuals to distinguish whether signals belong to the same object or to separate objects (Shams & Beierholm, 2010; Bizley & Cohen, 2013). This dual mechanism, integrating sensory cues when appropriate and segregating them when they originate from distinct sources, reflects an adaptive strategy shaped by the probabilistic nature of perception. When sensory signals are spatially and temporally consistent, integration improves precision and reduces variability. Conversely, when cues deviate beyond expected limits, source segregation allows the brain to maintain distinct representations without necessarily suppressing cross-sensory influences. The ability to flexibly switch between integration and source segregation is considered a hallmark of mature perceptual processing and supports fundamental functions such as navigation, object recognition, and social interaction. Research on multisensory development indicates that truly optimal audio–visual and audio–tactile integration, consistent with statistical models that weight sensory cues according to their relative reliability, typically emerges around seven to eight years of age (Baart, 2015; Gori et al., 2008). Despite this relatively late maturation, several studies indicate that infants exhibit precursors of multisensory competencies much earlier in life. For example, newborns and young infants demonstrate sensitivity to audio-visual correspondences for instance, turning their gaze or head toward a speaking caregiver even before their sensory systems have fully matured (Baart, 2015, Neil et al., 2006). Similarly, localization paradigms have shown that infants younger than one year can detect bimodal audio-visual targets more rapidly than

unimodal ones (Lewkowicz & Ghazanfar, 2009), suggesting that early mechanisms of cross-modal facilitation are present prior to reach statistical optimal integration. When examining a broader developmental window, spanning from early childhood to pre-adolescence, further evidence points to the progressive refinement of these abilities. Children as young as four or five years demonstrate the capacity to integrate visual and auditory information to improve reaction speed and reduce trial-to-trial variability (Burr & Gori, 2012), indicating that the benefits of combining sensory cues are already functionally expressed before integration reaches adult-like optimality. However, this period is also characterized by a wider temporal binding window (TBW) compared to adults, meaning that children are more tolerant of temporal discrepancies between sensory signals while still perceiving them as simultaneous (Lewkowicz & Flom, 2014). This prolonged TBW has been proposed as a hallmark of the gradual tuning processes through which perceptual systems learn to constrain multisensory binding with increasing temporal precision. Taken together, these findings indicate that children aged approximately 4-5 years are situated in a transitional developmental phase: they do not yet exhibit optimal multisensory integration in the strict computational sense, yet they derive clear behavioral advantages from combining sensory inputs. Specifically, multisensory cue presentations tend to enhance both speed and consistency of responses, particularly in tasks characterized by uncertainty or complexity. Nevertheless, the developmental mechanisms underlying multisensory processing remain only partially understood, particularly when sensory signals unfold dynamically over time and space. Most existing paradigms investigating multisensory perception rely on static or discrete presentations of auditory and visual cues and are typically confined to laboratory settings, limiting their ability to capture the continuous and dynamic nature of real-world multisensory experiences. As a result, they do not reflect the continuous and dynamic nature of real- world perception and can limit access to a broader sample of children, especially in early developmental stages where testing opportunities are often tied to preschool and educational settings.

In this chapter, I introduce a new device designed to investigate key aspects of multisensory perception through controlled dynamic stimulation. The system allows manual regulation of the distance, speed, and intensity of sensory sources within a child-friendly and highly adaptable experimental setup, suitable for use directly in schools and kindergartens. This approach aims to broaden the feasibility of developmental studies, facilitate data collection in more ecologically relevant contexts, and provide a sensitive framework for examining how

multisensory source segregation emerges in early childhood.

3.1 Multisensory tool for testing perceptual development of multisensory integration with moving stimuli

The device provides a novel, controlled, and child-friendly tool to replicate complex multisensory spatial stimulation and assess how children perceive and process audio, visual, and audio-visual cues (Figure 3.1).

3.1.2 Mechanical Design

The device (Figure 3.1) is composed of a compact support base measuring 60 cm in length, 8 cm in width, and 3.5 cm in height, on which a 56 cm ruler is mounted to guide movement coordination along the horizontal plane. Two independent 7-cm mobile shuttles are positioned on this guide, each equipped with a dedicated mounting to securely host the smartphones delivering the sensory cues. A mechanical transmission system allows the shuttles to be actuated manually while maintaining synchronized, symmetric, and strictly linear motion. Both shuttles start from a central, fully overlapped position and move simultaneously in opposite directions along parallel trajectories. This ensures that the spatial separation increases symmetrically over time, with each shuttle maintaining an equal distance from the central reference point throughout the movement. This behaviour is achieved through an internal mechanism comprising an idle pulley responsible for tension regulation, a transmission belt that transfers motion to the mobile carriers, and a linear guide designed to maintain smooth, controlled translation along a fixed path (Figure 3.2). For visual stimulation, the front section of the device incorporates a semi-transparent plexiglass screen (19.25 cm in height), fully covered with three superimposed diffusion filter sheets of 0.07 mm, 0.1 mm, and 0.25 mm thickness. This multilayer configuration reduces the visibility of the smartphones and homogenizes the light emitted, enabling

controlled scattering and consistent diffusion of the visual stimulus across trials while minimizing extraneous visual cues.

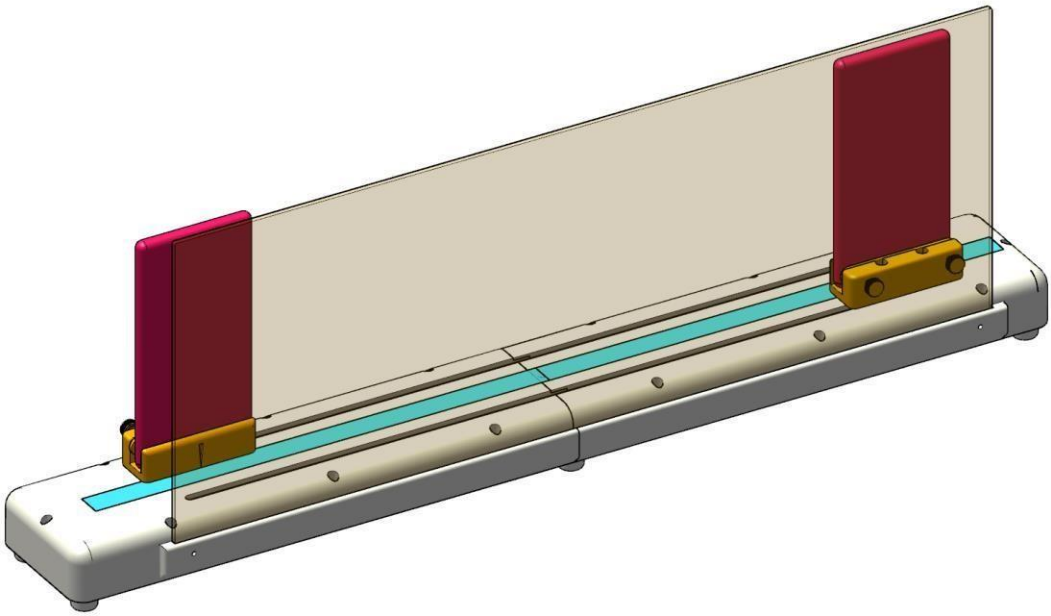

Figure 3.1. *Device (CAD)*

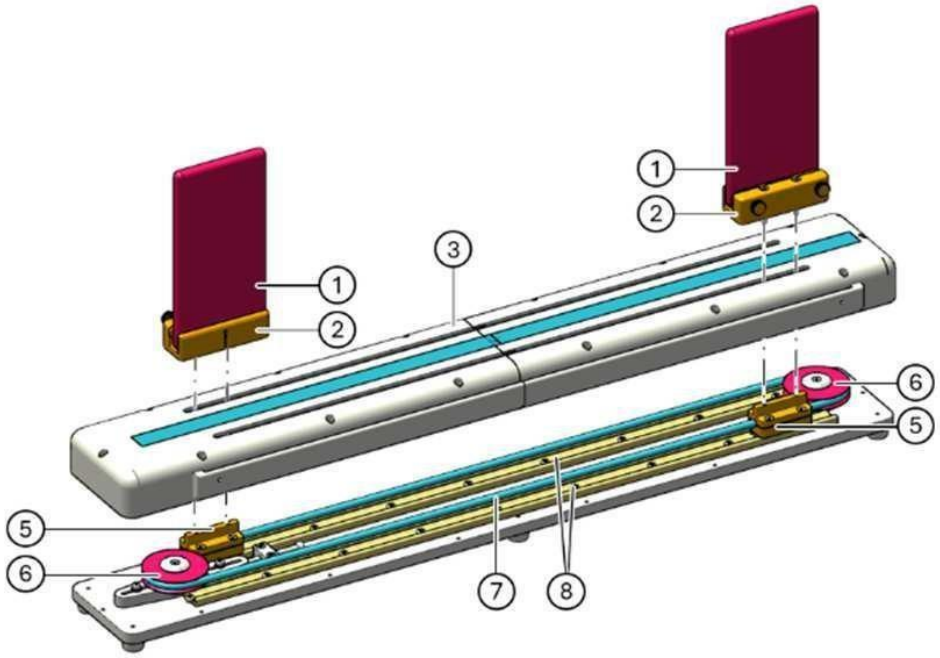

Figure 3.2. *Device components: (1) smartphones (2) moving shuttles (3) top cover (4) semi-transparent screen (5) shuttle base (6) idle pulley (7) transmission belt (8) linear guide*

3.1.3 Software Design

Sensory stimulation is delivered via two smartphones, which operate through a custom-developed mobile application specifically designed for the experimental protocol. The application was built using Flutter, an open-source framework for cross-platform interface development, enabling reliable deployment on different mobile devices. On Android systems, used for all smartphones in this setup, Flutter interfaces with low-level APIs through C/C++ and relies on the Android Native Development Kit (NDK) to access device-level functionalities, while the graphical interface and application logic are implemented in Dart, Google's native programming language for this framework. The experimental configuration includes a DELL computer running a Linux operating system, one XIAOMI POCO M3 Pro 5G (Android 11), and two SAMSUNG Galaxy A40 devices (Android 10). Communication between the computer and smartphones is established through a hotspot network: the workstation transmits data packets to all three devices and assigns them specific roles. One smartphone functions as the "master", managing trial sequencing and acting as the control interface for the experimenter, whereas the remaining two serve as "executors", positioned on the moving shuttles and responsible for delivering the sensory cues. Visual stimulation is produced through the rear-camera LED flash of each smartphone, calibrated at 50 lumens before attenuation and reduced to approximately 45 lumens once diffused through the layered plexiglass filter system. Auditory stimulation consists of two components: broadband white noise and a discrete beep, each lasting 0.1 seconds, emitted at an intensity of approximately 65 decibels. The two auditory signals originate from separate smartphones, ensuring that sound sources remain spatially distinct. Audio and visual cues can be generated independently (visual- only or audio-only) or paired synchronously in an audio-visual condition, with simultaneous onset and identical duration across modalities. The software also manages subject registration and controls the randomized presentation of sensory conditions. Each trial has a duration of approximately 13 seconds, during which stimulation is presented intermittently, with 0.85 seconds of ON time followed by 0.5 seconds OFF, creating a temporally engaging and attention-maintaining sequence for young participants. Throughout the trial, the executor smartphones continuously log acceleration values through their built-in accelerometers. These measurements are streamed in real time to the application, which monitors movement stability and generates visual warnings if acceleration exceeds the predefined threshold of 2 m/s^2. This feature ensures consistency of motion across sessions and reduces unwanted variability linked to manual operation. For each trial, a .csv file is automatically generated, storing the subject identifier, stimulation condition, and the

progressive displacement values of the two smartphones, collected at multiple timestamps for the entire duration of the trial.

3.2 Device Validation

We employed this device in two experimental paradigms targeting different aspects of sensory segregation. The first, a perceptual task (Perception of Segregation), was designed to determine the minimum spatial separation at which two sources are perceived as distinct. The second, a localization task (Localization), assessed whether children's explicit judgments of the distance between the stimuli corresponded to, exceeded, or underestimated the actual physical separation. Together, these experiments demonstrated the feasibility and potential of this child-friendly system for investigating the development of spatial competencies and multisensory integration in early childhood.

3.2.1 Experiment 1 (Perception of Segregation)

The Perception of Segregation task was designed to assess the ability to discriminate between a single sensory source and two spatially distinct sources. Five children participated in the study (mean age ± SD: 3.95 ± 1.49 years). The experimental sessions were conducted in Genoa, Italy. The study protocol was approved by the local ethics committee (Comitato Etico Regione Liguria) and was carried out in accordance with the principles of the Declaration of Helsinki.

Experimental Procedure

Three stimulation conditions were tested: Audio and Visual (unimodal conditions) and Audio-Visual (bimodal condition). Sensory stimulation was delivered by two smartphones mounted on the moving shuttles of the device. Within each condition, both smartphones emitted the same type of stimulation (Audio–Audio, Visual–Visual, or Audio-Visual-Audio-Visual). Each condition was presented three times in a randomized order, resulting in a total of nine trials per participant. Each trial had a fixed duration of approximately 13 seconds. Participants were seated in front of the device at a constant distance and were instructed to observe the stimuli throughout the trial. At trial onset, the two smartphones were aligned at the center of the apparatus, producing stimulation from a single spatial location for the first 2 seconds. Subsequently, the smartphones were progressively displaced along the horizontal axis for the remaining 11 seconds, generating a gradual increase in inter-source distance. The primary outcome measure was the inter-source distance at which participants first exhibited a behavioral response indicative of source segregation. Behavioral responses were operationalized as gaze

shifts between the two stimuli and were recorded using a video camera focused on the child's face. Two independent raters, blind to the experimental conditions and hypotheses, manually coded the video recordings and identified the timestamp corresponding to the first gaze shift between the two sources. A second synchronized camera recorded the motion of the smartphones during each trial. Smartphone positions were quantified using a custom video-tracking pipeline implemented in Python (version; OpenCV library). For each video frame, the distance between each smartphone and a predefined central reference point was computed. The inter-source distance was obtained by summing these two values. Distance data were stored in a .csv file and temporally aligned with a second .csv file containing the timestamps of the behavioral responses. Synchronization of the two datasets was performed using a dedicated Python script, allowing extraction of the inter-source distance at the exact time of the response. This procedure yielded a continuous and objective measure of the spatial separation associated with the perceptual transition from a single source to two distinct sources under each sensory condition.

Data Analysis and Results

The analysis focused on estimating the minimum inter-source distance at which two stimuli were perceived as spatially distinct along the horizontal plane. Based on previous evidence on multisensory processing, we hypothesized that the bimodal Audio-Visual condition would facilitate segregation, leading to shorter perceived segregation distances compared to unimodal conditions.

For each participant and condition, the mean perceived distance and its standard deviation were computed. Prior to inferential testing, the distribution of the distance measures was assessed using the Shapiro–Wilk test, which indicated that the data deviated from normality across all conditions. Consequently, non-parametric statistical analyses were adopted. Differences in mean perceived distance across the three sensory conditions (Auditory, Visual, Audio-Visual) were evaluated using a Friedman test, which revealed no significant effect of condition ($\chi^2 = 1.2$, df = 2, p = 0.5488, Kendall's W = 0.12, 95% CI [0.04, 0.84]).

An additional Friedman test was conducted on the within-condition standard deviations to assess differences in response variability across sensory modalities. This analysis also yielded no significant effect ($\chi^2 = 0.4$, df = 2, p = 0.8187, Kendall's W = 0.04, 95% CI [0.00, 0.76]). Figure 3.3 illustrates the distribution of perceived distances (left panel) and corresponding standard deviations (right panel) for the three conditions.

Although no statistically significant differences emerged, inspection of the data revealed a consistent pattern across participants. Specifically, the Auditory condition tended to be associated with larger perceived segregation distances and greater variability, whereas the Visual condition showed shorter distances and more stable responses, indicating higher spatial precision. The Audio-Visual condition exhibited intermediate values between the two unimodal conditions. This pattern suggests that, even in the absence of statistically reliable effects, children may differentially weight sensory information when judging spatial separation. (Figure 3.3)

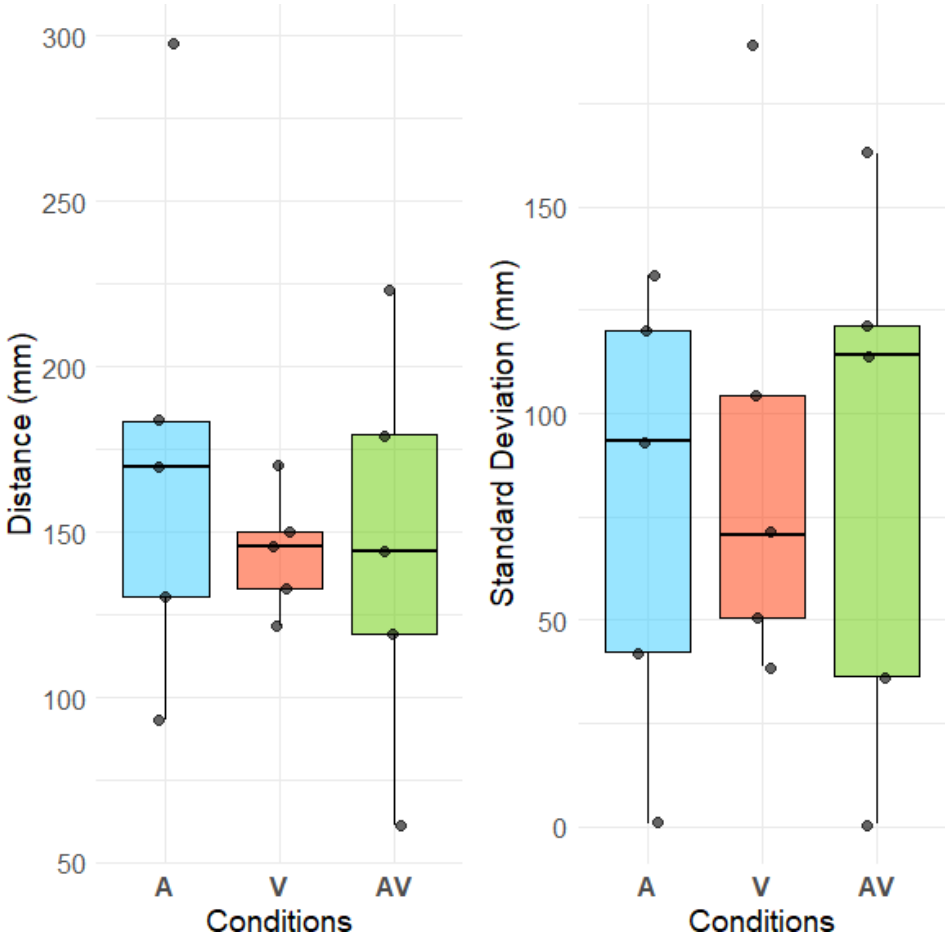

Figure 3.3 Experiment 1 (Perception of Segregation). Boxplots showing the perceived distance (left) and its standard deviation (right) across the three sensory conditions: Auditory (A), Visual (V), and Audio-Visual (AV). Individual data points are over

General Discussion Experiment 1 (Percepition of Segregation)

Despite the absence of statistically significant effects in the Friedman test, additional exploratory analyses were conducted to further characterize potential patterns in the data. Pairwise comparisons between conditions were examined using Wilcoxon effect sizes, with the aim of quantifying the magnitude of observed differences rather than establishing inferential significance. Given the small sample size (n = 5), these analyses should be interpreted as descriptive.

A large effect size was observed for the comparison between Auditory and Audio-Visual conditions (r = 0.784, 95% CI [0.18, 0.93]), while a moderate effect size emerged for the comparison between Auditory and Visual conditions (r = 0.302, 95% CI [0.06, 0.93]). Across comparisons, confidence intervals were wide, reflecting substantial statistical uncertainty.

To further assess the robustness of the dataset, a post hoc power analysis was conducted using G*Power 3.1. Assuming a partial η^2 of 0.08 (equivalent to Cohen's f = 0.30) and $\alpha = 0.05$, the achieved statistical power was estimated at 0.21, indicating that the experiment was underpowered to reliably detect condition differences. A priori power analysis suggested that at least 20 participants would be required to detect an effect of this magnitude with 80% power in a repeated-measures design with three within- subject conditions.

These results suggest that, while no statistically reliable effects emerged, the observed effect size patterns may reflect underlying tendencies that could become more robust with increased sample size. As such, but they provide a useful indication of the potential sensitivity of the paradigm for future investigations.

.

3.2.3 Experiment 2. Localization

The device was also employed to assess spatial localization abilities through a dedicated localization task. The sensory stimuli and experimental structure were identical to those used in the Perception of Segregation task, including the same number of trials, stimulation modalities (Auditory, Visual, and Audio-Visual), and trial duration. Five children participated in the Localization Task (mean age ± SD: 4.4 ± 0.49 years). The experimental sessions were conducted in Genoa, Italy. The study protocol was approved by the local ethics committee (Comitato Etico Regione Liguria) and was carried out in accordance with the principles of the Declaration of Helsinki.

Experimental Procedure

In addition to the previously mentioned Perception of Segregation task, in order to conduct this experiment, we incorporated a spatial reference aid consisting of a visual band mounted on the device, displaying pairs of animals arranged at fixed and equidistant positions along the horizontal axis (Figure 3.4). This visual scale served as a reference system for the children. During each trial, children were instructed to verbally report when the two stimuli were perceived as spatially separate and to indicate the perceived positions of the two sources by naming the corresponding animals on the visual band. The reported positions were then used to compute a localization metric referred to as Position Error, defined as the absolute difference between the perceived inter-source distance and the actual physical distance between the two smartphones at the time of the response. Formally, Position Error was calculated as:

$$\text{Position Error} = |\ \text{Perceived Distance} - \text{Actual Distance}\ |$$

For illustrative purposes we make an example: if a child indicated the position corresponding to 60 mm *(Perceived Distance)* while the true distance *(Actual Distance)* between the smartphones was 40 mm, the resulting *Position Error* would be 20 mm.

The primary hypothesis of this experiment was that bimodal Audio-Visual stimulation would lead to more accurate spatial localization, reflected by smaller Position Error values, compared to the unimodal Auditory and Visual conditions.

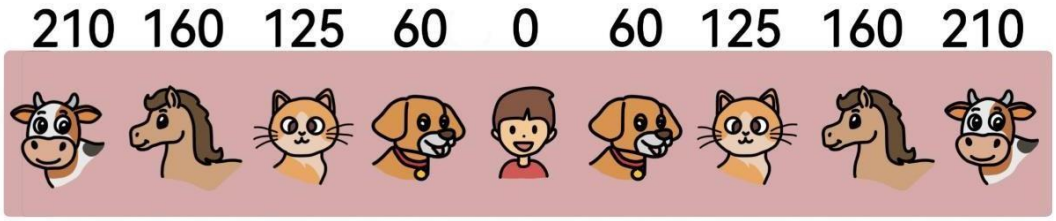

Figure 3.4. Visual Band for reference distance. Each pair of animals is equidistant and corresponds to the signed number.

Data Analysis and Results

For the localization task, spatial accuracy was quantified using Position Error, defined as the absolute difference between the perceived inter-source distance reported by the participant and the actual physical distance between the two smartphones at the time of the response. For each participant and stimulation condition (Auditory, Visual, Audio-Visual), the mean Position Error and its standard deviation were computed. Prior to inferential analysis, assumptions of normality were assessed. Given the small sample size and the non-normal distribution of the data, non-parametric statistical methods were adopted. Differences in mean Position Error across sensory conditions were tested using a Friedman test. An additional Friedman test was performed on the standard deviations to evaluate differences in response variability across conditions. The Friedman test conducted on mean Position Error revealed no significant effect of stimulation condition ($\chi^2 = 1.6$, df = 2, p = 0.45, Kendall's W = 0.16, 95% CI [0.04, 1.00]). Similarly, no significant differences emerged when comparing the standard deviations across conditions ($\chi^2 = 1.2$, df = 2, p = 0.55, Kendall's W = 0.12, 95% CI [0.00, 0.84]). Figure 4 presents boxplots of Position Error (left panel) and corresponding standard deviations (right panel) for the Auditory (A), Visual (V), and Audio-Visual (AV) conditions. Although inferential analyses did not reveal statistically significant differences, descriptive inspection of the data indicates that the Auditory condition tended to be associated with larger localization errors and greater variability. In contrast, the Visual condition showed smaller errors and more consistent responses, whereas the Audio-Visual condition exhibited intermediate values between the two unimodal conditions. (Figure 3.5).

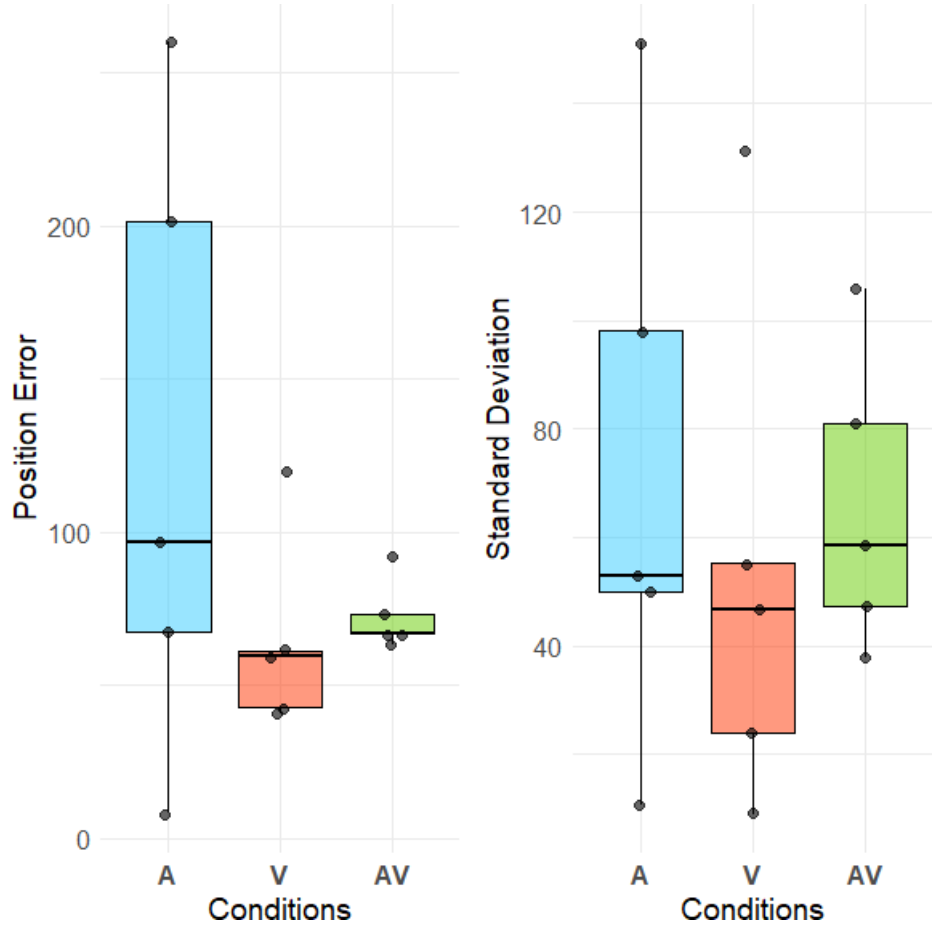

Figure 3.5. *Experiment 2 (Localization). Boxplots showing the Position Error (left) and its standard deviation (right) across the three sensory conditions: Auditory (A), Visual (V), and Audio-Visual (AV). Individual data points are overlaid.*

General Discussion of Experiment 2 (Localization)

Similarly, despite the absence of statistically significant effects in the Friedman test, exploratory analyses were conducted to further describe the data. Pairwise comparisons using Wilcoxon effect sizes indicated a large effect for the Auditory versus Audio-Visual comparison (r = 0.543, 95% CI [0.06, 0.95]). Moderate effect sizes were observed for both the Auditory versus Visual comparison (r = 0.422, 95% CI [0.06, 0.93]) and the Audio-Visual versus Visual comparison (r = 0.302, 95% CI [0.18, 0.95]).

When variability was considered by comparing standard deviations across conditions, moderate effect sizes were found for the Auditory versus Audio-Visual comparison (r = 0.302, 95% CI [0.06, 0.93]) and for the Audio-Visual versus Visual comparison (r = 0.422, 95% CI [0.06, 0.93]), whereas the Auditory versus Visual comparison yielded a small effect size (r = 0.181, 95% CI [0.06, 0.93]). Confidence intervals were consistently wide, indicating high uncertainty.

Taken together, while no statistically reliable effects were observed, the consistency of effect size patterns across comparisons suggests that the dataset may capture meaningful trends,

particularly involving differences between auditory and multisensory conditions. In line with the overall aim of the study, these results should be interpreted cautiously, but they provide a useful indication of the type of variability and sensitivity that the paradigm is able to capture, and they support the need for replication in larger samples.

3.3 General Discussion

In this chapter, I presented the validation of a novel experimental device designed to investigate multisensory spatial perception in preschool-aged children. The system enables a continuous and gradual transition from a single-source configuration to a two-source configuration, allowing precise identification of the spatial point at which perceptual segregation occurs.

Across two experimental paradigms, a segregation perception task and a localisation task, the device proved to be a feasible, child-friendly, and precise tool for studying multisensory processing under controlled yet ecologically plausible conditions. Although no statistically significant effects were observed, the effect size estimates and descriptive patterns suggest a potential visual preference. Importantly, the present results should not be interpreted as evidence against multisensory processing in early childhood, but rather as reflecting the high variability and ongoing refinement that characterise perceptual development at this age. A central contribution of this work lies in the methodological domain. The device combines precise experimental control with an engaging and non-invasive testing format, allowing researchers to study complex perceptual phenomena without requiring young children to perform artificial or cognitively demanding tasks. The use of smartphones as synchronised actuators enables real-time, manually guided motion with accurate temporal alignment between stimulation and behavioural responses. This flexibility makes the system particularly suitable for testing very young participants in non-laboratory environments such as schools and kindergartens. To further contextualise the localisation findings, results were qualitatively compared with those obtained in established bimodal paradigms, such as the study by (Rohlf et al., 2020), despite differences in experimental design, age range, and sample size, and despite the absence of a continuous segregation component in previous work. Similar trends were observed, with auditory stimulation showing the lowest spatial precision, where visual stimulation showing the highest, and bimodal stimulation intermediate performance. While preliminary, these converging patterns may point to a prominent role of visual input in early spatial localization.

Overall, the present study demonstrates the feasibility and potential of the proposed device

for investigating multisensory spatial perception in early development. By enabling controlled investigation of dynamic multisensory phenomena in ecologically valid contexts, this tool provides a promising framework for advancing research on the development of spatial perception and multisensory integration in childhood.

Chapter 2 is partially extracted and adapted from:
Calafatello, G., Zanchi, S., Balzarotti, N., Gori, M., Brmener, A. J., Parmiggiani, A., & Orciari, L. (2025). Testing Perceptual Development with Moving Stimuli: A New Multisensory Tool for Young Children. 2025 IEEE Medical Measurements & Applications (MeMeA), 1–6

Chapter 4

Audiovisual Integration in Dynamic Object Segregation

Building on the findings presented in the previous chapters, the present chapter addresses a more specific and demanding question concerning multisensory spatial development: whether audiovisual information supports spatial source segregation in early childhood, particularly under dynamic conditions. Results presented in Chapter 2 showed that auditory spatial representations in preschool- aged children are functional but not uniformly mature across spatial dimensions, with increased variability and reduced precision depending on the spatial plane. In Chapter 3, a novel experimental device was introduced and validated, enabling the delivery of controlled, dynamic audiovisual stimulation and allowing accurate estimation of spatial reference while remaining suitable for testing children in familiar settings (Calafatello et al., 2025). Together, these findings and methodological advances provide the necessary foundation to investigate not simply whether children process multisensory information, but how such information is used when the perceptual system is required to distinguish between multiple spatial sources.

4.1 Introduction

The sensory and perceptual systems enable individuals to organise information from the environment into coherent spatial representations. Within probabilistic frameworks such as maximum likelihood estimation, sensory cues are combined by weighting them according to their relative reliability, allowing uncertainty to be reduced when multiple signals refer to the same event (Ernst & Banks, 2002; Burr & Gori, 2012). In adults, this process improves perceptual precision and enhances spatial estimates compared to unisensory stimulation (Alais & Burr, 2004; Stein & Stanford, 2008).

Although optimal cue combination is not yet fully established during development (Gori et al., 2008; Nardini et al., 2008; Burr & Gori, 2012; Brandwein et al., 2011), children can

nevertheless, benefit from multisensory information. From early infancy, multisensory stimulation facilitates perceptual responses and speeds up behavioural performance compared to unimodal conditions, indicating the early presence of crossmodal interactions (Lewkowicz, 2009; Neil et al., 2006; Barutchu et al., 2009).

Beyond improving precision, multisensory interactions contribute to the organisation of perceptual input, supporting the segmentation of the environment into distinct objects (Driver & Noesselt, 2008; Lewkowicz, 2014; Spelke, 1990). In spatial contexts, combining information from different sensory modalities can enhance the representation of objects and help determine whether multiple signals originate from a single source or from separate events (Ernst & Banks, 2002; Alais & Burr, 2004). In particular, multisensory cues may facilitate the detection of spatial separation between stimuli, supporting the ability to distinguish between one or multiple objects.

Despite extensive research on multisensory processing in the temporal domain, particularly in relation to audiovisual synchrony (Lewkowicz & Flom, 2014; Hillock-Dunn & Wallace, 2012; Stevenson et al., 2014), much less is known about how children use multisensory information to infer spatial separation, especially in dynamic contexts. Real-world environments are inherently dynamic, requiring continuous tracking of moving objects and ongoing updating of spatial relationships. Understanding how children integrate multisensory information under these conditions is therefore crucial for characterising the development of spatial perception.

The present study aimed to investigate how preschool children use multisensory information in dynamic spatial contexts, focusing on their ability to segregate objects based on spatial disparity. While previous research has shown that children benefit from multisensory stimulation in spatial tasks, it remains unclear whether such facilitation supports the detection of two simultaneously presented sources with different spatial origins, particularly when stimuli are in motion.

To address this gap, we developed a novel device that allowed controlled presentation of audiovisual stimuli while sources gradually moved apart (Calafatello et al., 2025). Children were asked to observe the stimuli and follow them with their gaze, allowing us to assess when they perceived two distinct sources. We hypothesised that audiovisual stimulation would enhance the detection of spatial separation, leading to the perception of two sources at shorter distances compared to unimodal conditions.

4.2 Methods

Participants

Children were recruited from local kindergartens in the city of Genoa. Written informed consent was obtained from the legal guardians of 40 preschool-aged children, of whom 23 were included in the final sample (16 males, 7 females; mean age = 4.09 years, SD = 0.68). The study was approved by the local ethics committee (Comitato Etico Regione Liguria) and conducted in accordance with the Declaration of Helsinki.

Materials

The experiment was conducted using the multisensory device described and validated in Chapter 2, which was specifically designed to assess spatial processing under dynamic conditions in young children. Briefly, the system allows the controlled presentation of auditory, visual, and audiovisual stimuli delivered via two smartphones mounted on moving shuttles that can be displaced symmetrically along the horizontal axis, enabling precise estimation of the spatial distance between sources throughout each trial. Stimulus presentation and trial timing were controlled via a custom application, which also recorded motion data to ensure consistent movement parameters. Visual stimuli consisted of brief light flashes, while auditory stimuli consisted of short sound bursts presented at a comfortable listening level. Stimuli were delivered in auditory-only, visual-only, and audiovisual conditions, with both sources providing the same type of stimulation simultaneously. All experimental parameters and distance values were recorded for offline analysis.

4.2.1 Procedures

Children were tested individually while seated approximately 20 cm from the device, with the screen positioned at eye level in a room within the kindergarten. Before the experimental session, each child completed a brief familiarization phase during which example trials were presented to reduce novelty effects; no data were collected during this phase. At the beginning of each trial, the two smartphones were spatially aligned at the centre of the apparatus behind the semi-transparent Plexiglass screen. Through the layered diffusion filters, the emitted flashes were perceptually blended, such that participants initially perceived the stimulation as originating from a single diffuse visual event rather than from two distinct light sources. The experimenter then gradually moved the two shuttles apart in a symmetrical

manner at a constant speed, progressively increasing the spatial separation between the stimuli over time. As the separation increased, the perceptual transition from one source to two distinct sources emerged gradually rather than through an abrupt onset. Children were not required to provide verbal responses and were simply instructed to observe the stimuli. The session was video-recorded using two synchronized cameras (SONY AX53 Handycam® 4K). One camera focused on the child's gaze behaviour, which served as the primary behavioural measure. A shift in gaze from one stimulus to the other was interpreted as an indication that the child perceived two distinct spatial sources. The second camera recorded the movement of the shuttles, allowing the spatial configuration at the moment of gaze shift to be reconstructed with precision. Each child completed a total of nine trials, with three trials per sensory condition. Trial order was randomized across participants. The full experimental session, including familiarization, lasted approximately ten minutes.

4.3 Analysis

4.3.1 Video-based analysis behavioural measurement

For each participant, two synchronized video recordings were collected: one capturing the child's gaze behaviour and one recording the movement of the shuttles. Gaze recordings were independently analysed by two naïve raters using Avidemux (open-source video editing software). Raters were trained on the coding procedure without being informed about the experimental hypotheses to minimise potential bias.

Each rater identified the video frame corresponding to the first clear shift in gaze from one stimulus to the other following an initial central fixation. The shift in gaze from one source to the other was taken as the behavioural marker of perceived segregation, as it was interpreted as an indication that participants had detected the spatial separation between the two stimuli. Although gaze shift is inherently a temporal measure, it was not analysed in temporal terms. Instead, it was used to identify the moment at which segregation was perceived, and the corresponding spatial distance between the two sources at that time was computed through positional tracking of the moving stimuli. Importantly, , the present task required participants to implicitly signal the moment at which two moving stimuli were perceived as spatially segregated. Consequently, variability in the present task should not be interpreted as a or discrimination threshold, but rather as reflecting the consistency with which perceptual segregation was detected across repeated dynamic trials. Only trials in which gaze was clearly directed toward both sources were considered valid. Inter-rater agreement was high (Spearman's $\rho = .87$, $p < .001$), and the two timing estimates were therefore averaged for each trial. Movement recordings were analysed using a custom Python script based on the OpenCV TrackerMIL algorithm, which tracked shuttle positions frame by frame and converted pixel coordinates into real- world distances using a calibrated ruler. By synchronising gaze and movement recordings, it was possible to compute the exact spatial distance between the two stimuli at the moment of gaze shift. This distance served as the primary dependent variable in all subsequent analyses.

4.3.2 Statistical Analysis

All statistical analyses were performed using RStudio (version 4.2.2; R Core Team, 2022). Data distributions were inspected using the Shapiro-Wilk test and were found to deviate from normality. On this basis, non-parametric statistical approaches were adopted for all subsequent analyses. To examine the influence of sensory condition on spatial segregation, we focused on the distance at which children perceived the two stimuli as originating from separate sources. For each participant and experimental condition, these distances were averaged to obtain the Mean Distance of Perceived Segregation, defined as the mean spatial separation associated with the perceptual report of two distinct sources. In addition, response variability was quantified by computing the standard deviation of the perceived segregation distances for each participant and condition. This measure was taken as an index of Precision, defined as the consistency of perceptual judgments across repeated trials, with lower standard deviation values indicating higher precision. Importantly, this metric reflects variability in the spatial configuration at which segregation was perceived, thus capturing variability in the spatial representation of object separation. Differences in the Mean Distance of Perceived Segregation across sensory conditions, auditory (A), visual (V), and audiovisual (AV), were assessed using a Friedman test. A second Friedman test was conducted on Precision values to determine whether perceptual precision varied as a function of sensory modality.

4.3.3 Best-cue comparison

To determine whether audiovisual stimulation conferred a precision advantage over unisensory processing, we compared variability in the audiovisual condition with that of the most reliable unimodal cue for each participant. Specifically, for each child, the unimodal condition (auditory or visual) associated with the lowest variance was identified as the best cue. Variability in the AV condition was then compared to this best-cue variance using a Wilcoxon signed-rank test. This approach follows current methodological recommendations, according to which multisensory benefit should be evaluated relative to the most informative single modality rather than to the average unisensory performance(Scheller & Nardini 2024).

4.3.4 Maximum Likelihood Estimation Model (MLE)

To further assess whether audiovisual performance reflected statistically optimal cue combination, we applied a Maximum Likelihood Estimation (MLE) model. According to this model, when two sensory cues are combined optimally, the variance of the integrated estimate is reduced as a function of the relative reliability (inverse variance) of each cue (Ernst & Banks, 2002). Within this framework, the MLE model was used as a theoretical benchmark to explore patterns of multisensory cue combination.

For each participant, the predicted audiovisual variance was predicted as:

$$\sigma^2\_\{AV\} = (\sigma^2\_A \times \sigma^2\_V) / (\sigma^2\_A + \sigma^2\_V)$$

where σ^2_A and σ^2_V represent the response variances in the auditory and visual unimodal conditions, respectively.

From these unimodal variances, we also derived predicted auditory and visual weights, reflecting the relative contribution of each modality to the optimal audiovisual estimate:

$w_a = [1 / \sigma_a^2] / [1 / \sigma_a^2 + 1 / \sigma_v^2] = \sigma_v^2 / (\sigma_a^2 + \sigma_v^2)$

$w_v = [1 / \sigma_v^2] / [1 / \sigma_a^2 + 1 / \sigma_v^2] = \sigma_a^2 / (\sigma_a^2 + \sigma_v^2)$

Observed audiovisual variance was compared with the MLE-predicted variance using a Wilcoxon signed-rank test. A non-significant difference was interpreted as evidence that observed audiovisual precision was consistent with MLE predictions.

To complement these analyses, Bayes Factors (BF_{01}) were computed using the *ttestBF* function from the Bayes Factor package (Morey & Rouder, 2018). In line with previous work, BF_{01} values greater than 3 were taken as strong evidence in favour of the null hypothesis (i.e., consistency with MLE predictions), values between 1 and 3 as weaker support, and values below 1 as evidence in favour of a deviation from optimal integration.

4.3.5 Analysis based on sensory-weight similarity

Because the precision gain predicted by the MLE model is maximal when auditory and visual cues are weighted similarly, participants were further classified based on the similarity of their unimodal sensory weights (Bentvelzen et al., 2009; Zanchi et al., 2022). Children whose absolute difference between auditory and visual weights was equal to or smaller than 0.5 were assigned to the *similar-weight* group, whereas those exceeding this threshold were assigned to the *non-similar-weight* group.

All analyses described above were then repeated separately for the two groups. This classification allowed us to examine whether children with more balanced sensory weighting exhibited stronger evidence of multisensory integration, as reflected by reduced audiovisual variability and closer correspondence with MLE predictions.

4.4 Results

4.4.1 Perceived segregation distance

We first examined whether the sensory condition influenced the distance at which children perceived the two stimuli as spatially segregated. A Friedman test conducted on the mean segregation distances across auditory (A), visual (V), and audiovisual (AV) conditions did not reveal a significant effect of condition, $\chi^2(2) = 1.65$, $p = .438$. Pairwise comparisons confirmed that differences between conditions were small, with rank-biserial correlations indicating minimal effects for A vs. V ($r = 0.114$), A vs. AV ($r = 0.127$), and V vs. AV ($r = 0.298$) (Figure 4.1A).

4.4.2 Perceptual precision

To assess perceptual precision, variability in segregation distance was quantified using the standard deviation of responses. A Friedman test on these variability measures did not show a significant effect of sensory condition, $\chi^2(2) = 1.83$, $p = .401$ (Figure 4.1B). Although none of the comparisons reached significance, the contrast between auditory and audiovisual conditions showed a moderate effect size ($r = 0.425$), whereas effects involving the visual condition were small.

4.4.3 Multisensory Benefit

To evaluate whether audiovisual stimulation provided a multisensory benefit beyond unisensory performance, audiovisual precision was assessed using two complementary

criteria: comparison with the most reliable unimodal cue (Best Cue) and comparison with predictions derived from a Maximum Likelihood Estimation (MLE) model.

At the group level, variability in the audiovisual condition did not differ significantly from the variance predicted by the MLE model (Wilcoxon signed-rank test: V = 175, p = .267; Figure 4.1C). Bayesian analysis yielded a BF_{01} of 1.45, indicating weak evidence in favour of equivalence between observed audiovisual variability and MLE predictions. When audiovisual variability was compared to that of the most reliable unimodal cue for each participant, no significant difference was observed (Wilcoxon signed-rank test: V = 148, p = .623, one- tailed; Figure 4.1D). Thus, although audiovisual precision was statistically compatible with MLE predictions, it did not exceed the precision achieved by the best unimodal condition.

4.4.4 Sensory Weighting

Given that the precision gain predicted by the MLE model is maximal when auditory and visual cues are weighted similarly, participants were divided into two subgroups based on unimodal weight similarity. Twelve children were classified as having similar auditory and visual weights, while eleven were classified as having non-similar weights. In the similar-weight group, preferences for auditory and visual cues were evenly distributed, whereas in the non-similar-weight group a greater proportion of participants relied more strongly on visual information.

Similar Weight Group

Within the similar-weight group, a Friedman test on mean segregation distances did not reveal a significant effect of sensory condition, $\chi^2(2) = 4.50$, p = .105; (Figure 4.2A). Despite the absence of a significant main effect, pairwise comparisons indicated moderate-to-large effect sizes for contrasts involving the audiovisual condition, particularly when compared with the visual condition.

Analysis of response variability showed a comparable pattern. A Friedman test on standard deviations did not reveal a significant effect of condition, $\chi^2(2) = 1.17$, p = .558; (Figure 4.2B). Nevertheless, comparisons involving the AV condition showed moderate effect sizes relative to both auditory and visual conditions, whereas variability did not differ between the two unimodal conditions.

Comparison between observed AV variance and MLE predictions revealed no significant

difference (Wilcoxon signed-rank test: V = 32, p = .610; Figure 4.2C). Bayesian analysis yielded a BF_{01} of 3.45, providing substantial evidence in favour of equivalence between observed audiovisual precision and MLE predictions. Comparison between AV variability and the best unimodal cue also did not reveal a significant difference (V = 24, p = .131, one-tailed), although variability in the AV condition was numerically lower than that of the best cue (Figure 4.2D).

Non-similar weight group

In the non-similar-weight group, mean segregation distances did not differ across sensory conditions, $\chi^2(2) = 0.18$, p = .913 (Figure 4.3A), with pairwise comparisons indicating only small effect sizes. Similarly, analysis of response variability revealed no significant effect of condition, $\chi^2(2) = 2.36$, p =

.307 (Figure 4.3B). When observed AV variance was compared with MLE predictions, a significant difference emerged (Wilcoxon signed-rank test: V = 56, p = .041), indicating that audiovisual precision was lower than expected under optimal integration (Figure 4.3C). The corresponding Bayesian analysis yielded a BF_{01} of 1.10, providing weak evidence regarding equivalence with MLE predictions. Finally, comparison between AV variability and the best unimodal cue did not reveal a significant difference (V= 54, p = .973, one-tailed), with variability in the AV condition remaining numerically higher than that of the best cue (Figure 4.3D).

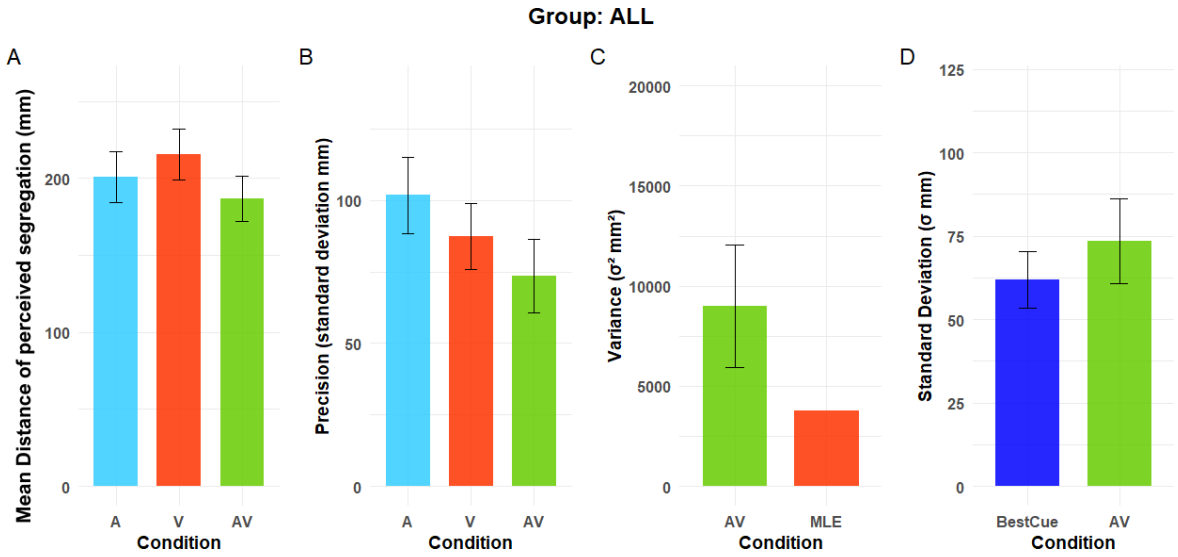

Figure 4.1 Mean Distance Perceived Segregation; Precision; MLE model; Best Cue;

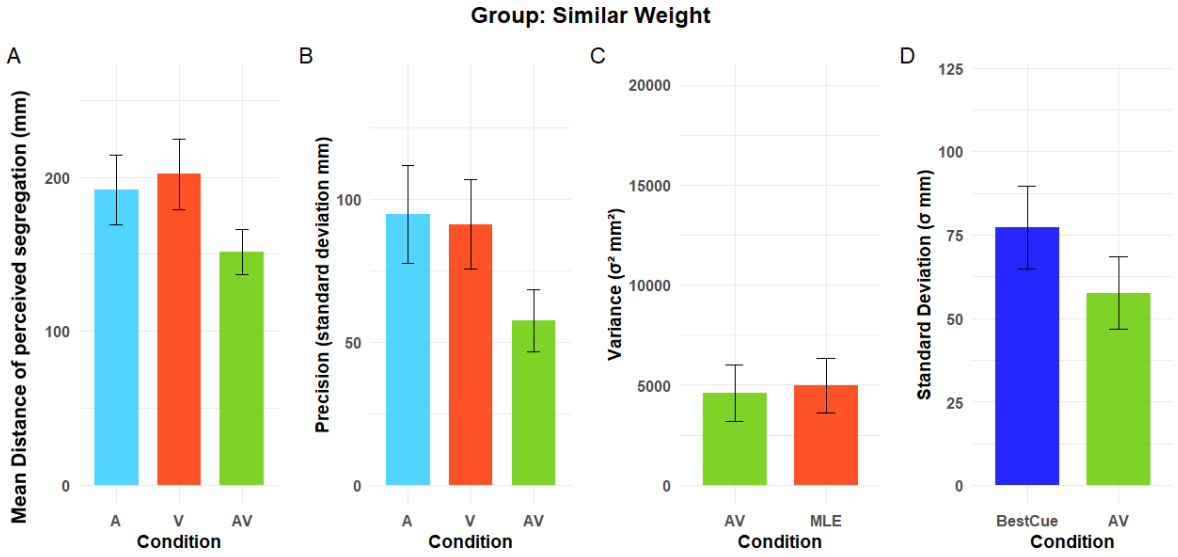

Figure 4.2 Mean Distance Perceived Segregation; Precision; MLE model; Best Cue for Similar Weight Group

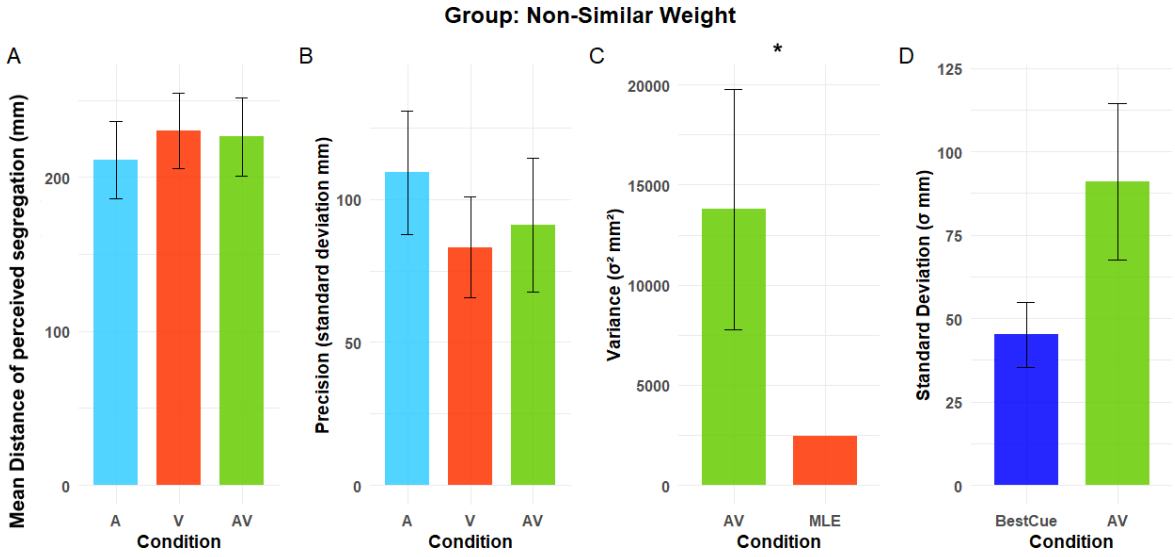

Figure 4.3 Mean Distance of Perceived Segregation; Precision; MLE model; Best Cue of Non-Similar Weight group

4.5 General Discussion

The present study investigated whether audiovisual information supports spatial source segregation in preschool-aged children using a dynamic task involving two moving stimuli. At the group level, no statistically significant differences emerged across sensory conditions, either for the mean distance of perceived segregation or for perceptual precision. Similarly, the comparison between observed AV variance and predictions derived from the Maximum Likelihood Estimation (MLE) model did not reveal a significant difference, and Bayesian evidence in favour of equivalence remained weak. Moreover, no advantage of the AV condition emerged when compared to the best-performing unimodal cue. Therefore, despite a slight descriptive trend toward reduced variability in the AV condition, the overall findings suggest that dynamic audiovisual information may not yet provide a robust behavioural advantage for spatial segregation at this developmental stage. Consistent with developmental evidence, children may not always integrate sensory information to reduce uncertainty but may instead rely on single cues or maintain sensory signals separate (Nardini et al., 2010). Importantly, this absence of group-level effects was accompanied by substantial inter- individual variability. Children differed in their reliance on auditory and visual information, suggesting that multisensory processing during early childhood is not characterised by a single developmental pattern but by the coexistence of multiple strategies.

In this context, averaging across participants may obscure meaningful developmental differences, particularly in tasks involving dynamic spatial inference. To further examine this inter-variability, participants were classified based on the similarity of their unimodal sensory weights. Children with more balanced weights showed a tendency toward improved performance in the AV condition and variability consistent with MLE predictions, although these patterns were not statistically significant and should be interpreted cautiously. In addition, the relatively limited sample size and subdivision into smaller participant groups may have reduced the sensitivity to detect subtle multisensory effects, particularly given the high inter-individual variability typically observed during development. Rather than providing clear evidence of multisensory benefit, these results may reflect early or unstable forms of cue combination. This interpretation is consistent with developmental accounts proposing that multisensory integration emerges gradually and heterogeneously across individuals (Burr & Gori, 2012; Gori, 2015; Rohlf et al., 2020). Previous work has shown that children within the same age range can differ substantially in their sensory strategies, with some combining cues and others relying on a single modality (Negen et al., 2019). Notably, similar inter-individual variability has also been reported in adults (Zanchi et al., 2022; Bentvelzen et al., 2009), suggesting that variability may represent a fundamental feature of multisensory processing rather than a transient developmental limitation. In contrast, children with unbalanced sensory weights did not show any indication of multisensory advantage. Their performance was characterised by greater variability in the AV condition and a deviation from MLE predictions, suggesting a continued reliance on a single dominant modality, most often vision. This is consistent with previous findings indicating that vision typically provides higher spatial reliability during development (Gori et al., 2008; Gori, 2015). Taken together, these findings highlight an important distinction between multisensory integration and spatial segregation. While integration can improve perceptual precision when signals are combined, segregation requires detecting discrepancies between cues and inferring multiple sources. These processes likely rely on partially distinct mechanisms and may follow different developmental trajectories. The present results suggest that, in early childhood, multisensory information alone is not sufficient to support reliable spatial segregation, and that performance remains strongly shaped by individual differences and ongoing sensory calibration.

In the next chapter (Chapter 5) we are going to investigate audio-tactile spatial processing,

examining how visual experience modulates multisensory integration and unisensory preference under conditions of sensory conflict. Using both behavioural and EEG measures, the chapter compares sighted and visually impaired participants to better understand the mechanisms supporting spatial representation in typical and atypical development.

Chapter 5

Neural Mechanisms of Audio-Tactile Integration and Conflict in children with and without visual impairment

The previous chapters examined how spatial perception develops from unisensory auditory processing to audiovisual multisensory integration in early childhood, highlighting both the emergence of evolving competencies and sensory dominance. The present chapter extends this framework to audio-tactile integration, focusing on how spatial information is combined when visual input is absent or limited. By integrating behavioural measures with EEG recordings, this study investigates the neural mechanisms supporting early multisensory spatial processing and explores how visual impairment influences the strategies used to resolve multisensory information, in the case of both spatial alignment and spatial conflict.

5.1 Introduction

Visual experience plays a central role in the early development of multisensory spatial perception, providing children with a reference framework to relate sensory events to their own body and to the external environment (Gori et al., 2021; Rigato et al., 2014; Orioli et al., 2024). A substantial body of research shows that vision contributes to the maturation of multisensory space by calibrating auditory and tactile representations, allowing these signals to be aligned within a coherent spatial framework (Gori et al., 2008; Nardini et al., 2008; Gori et al., 2015; Azañón et al., 2018). When visual input is absent, this calibration is altered: visually impaired children show delayed correspondence between auditory and tactile spatial maps, together with broader difficulties in auditory localisation, proprioceptive processing, and spatial imagery (Cappagli & Gori, 2016; Cappagli et al., 2017; Vercillo et al., 2016; Cappagli et al., 2019).

Behavioural findings from a previous work by our lab are consistent with this framework (Gori et al., 2021). In an audio-tactile localisation task, both sighted (S) and severely visually impaired (SVI) children benefited from spatially congruent multisensory stimulation, indicating that early forms of multisensory integration can emerge even in the absence of visual experience. However, this benefit was reduced in SVI children. More critically, when auditory and tactile cues conveyed conflicting spatial information, the two groups adopted different strategies: sighted children were sensitive to the crossmodal discrepancy, whereas SVI children relied predominantly on tactile cues, suggesting a more body-centred and less flexible mode of spatial coding. These behavioural differences point to potential differences in the neural mechanisms underlying multisensory spatial processing following early visual deprivation.

Neurophysiological evidence indicates that multisensory interactions are supported by cortical mechanisms that emerge early in development. Auditory-somatosensory interactions observed at short latencies in adults (Foxe et al., 2000) and have also been found in typically developing children, though with later latencies (Brett-Green et al., 2008). Evidence from infancy further suggests that multisensory processing is already present at birth and is initially tuned to the body and peripersonal space (Maitre et al., 2020; Ronga et al., 2021). By contrast, the neural organisation of spatial multisensory processing in the absence of early visual input remains poorly understood. Studies in congenitally blind individuals report a posterior reorganisation of spatial processing and altered audio–tactile interactions, indicating that early visual experience plays a critical role in shaping multisensory cortical networks (Röder et al., 1999; Collignon et al., 2015; Occelli et al., 2013).

In this chapter, I explore the development and neural mechanisms underlying audio-tactile multisensory integration during early childhood, focusing on both typically developing children and children with severe visual impairment. This section aims to elucidate how the combination of auditory and tactile cues contributes to early spatial perception, and whether the absence of visual experience shapes the organisation of multisensory processes. To this end, we adopt a multimethod approach combining behavioural measures with neurophysiological recordings. Specifically, high-density electroencephalography (EEG) is employed in children and toddlers to examine the neural correlates of early audio-tactile multisensory processing. This approach allows us to directly link behavioural

responses to underlying cortical activity, providing a more comprehensive understanding of how multisensory integration emerges and is implemented in the developing brain.

5.2 Methods

5.2.1 Participants

Twelve severely visually impaired children (SVI; 6 females; median age = 28.1 months, range = 10–54 months) and twelve sighted children (S; 5 females; median age = 31 months, range = 12–52 months) participated in the study. Written informed consent was obtained from parents prior to participation, and testing was conducted only when children were awake and alert. The study was approved by the Ethics Committee of ASL 3 Genova (Liguria). Visual acuity in the SVI group was assessed using Teller Acuity Cards, a standard method for evaluating visual function in pre-verbal children; further methodological details are reported in Gori et al. (2021). All participants had normal general health, no history of prenatal or perinatal complications, no neurological or metabolic disorders, and no diagnosis of epilepsy. Cognitive and psychomotor development was within the typical range, as confirmed by clinical evaluation and, for visually impaired children, by the Reynell–Zinkin Scale. Cerebral visual impairment was excluded based on medical history and clinical assessment**.** Sample size was informed by previous studies investigating multisensory spatial processing in sighted and severely visually impaired (SVI) infants (Gori and Vitali, 2025), which reported robust behavioural differences between groups using comparable sample sizes. In the present study, we aimed to recruit 12 participants per group in order to increase statistical power relative to previous work and to account for the methodological challenges associated with infant high-density EEG acquisition. Based on the large behavioural effects previously reported in related paradigms, a sample of this size was considered appropriate for detecting group-level differences. Because ERP correlates of multisensory spatial processing in SVI infants have not been extensively investigated, sample size considerations were primarily based on previously observed behavioural effects.

5.2.2 Apparatus & Stimuli

Children were seated either on a child-sized chair or on a parent's lap in front of the testing table. Upper- body movements were recorded using a digital video camera (SONY AX53

Handycam® 4K; 25 fps) for offline behavioural coding conducted by two independent raters. Auditory and audio-tactile stimuli were delivered via custom-built, serial-controlled stimulators (Gori et al., 2019), placed in the infant's palms and secured with a soft bandage Audio-tactile stimuli consisted of a continuous vibration paired with a 112-Hz tone, while auditory-only stimuli consisted of a 926-Hz tone amplitude-modulated at 5 Hz. Each trial lasted 1000 ms and was followed by an 8000-ms response window. Stimulus timing was marked by a visual signal visible only on the video recording, allowing precise synchronisation while keeping raters blind to stimulus laterality.

5.2.3 EEG data Acquisition

EEG activity was continuously recorded with an EGI System by using Hydrogel Geodesic Sensor Net, consisting of 128 silver-chloride electrodes evenly distributed across the scalp. The vertex served as the reference. In order to assess the activation of the primary somatosensory and auditory cortical regions, we considered the area covered by specific clusters of electrodes corresponding to the Centro-Parietal area (CP) for tactile activation, contralateral (channels E 98, E102, E103) and ipsilateral relative to the stimulated hand (channels E41, E46, E47); the Temporo-Posterior area (TP) for audio activation, contralateral (E82, E83, E88, E89, E90, E94, E95, E96, E97, E99, E100, E101, E107, E108) and ipsilateral (channels E45, E50, E56, E57, E58, E63, E64, E65, E68, E69, E70, E73, E74). Then to assess the activation of multisensory associative cortical regions, we considered the Midline- Fronto-Central (MFC) area (channels E4, E5, E6, E11, E12. E13, E19, E20, E112, E118) and the Midline-Central-Parietal (MCP) area (channels E54, E55, E78, E79, E62, E61).

5.2.4 Preprocessing

Signals were band-pass filtered between 0.1 and 30 Hz and subsequently cleaned using Artifact Subspace Reconstruction (ASR; Chang et al., 2018) and Independent Component Analysis (ICA; Delorme & Makeig, 2004). Artifactual components were identified using ICLabel (Pion-Tonachini et al., 2019). Finally, the data were re-referenced to the common average. All preprocessing steps and parameters were consistent with those reported previously**.**

5.2.5 Experimental Design

The experimental design included four sensory conditions: Auditory (A), Tactile (T), Audio–Tactile congruent (AT-congruent; stimulation delivered to the same hand), and Audio–Tactile incongruent (AT- incongruent; stimulation delivered to opposite hands). All stimuli were presented with the hands in a single, uncrossed posture, which was kept constant throughout the experiment. Each condition was presented up to 12 times. Analyses therefore focused on auditory, tactile, and audio-tactile congruent and incongruent trials.

5.2.6 Procedure

Two experimenters were present during each session: one interacted with the child, positioned the hands, and maintained engagement, while the second controlled stimulus delivery via a custom MATLAB script. Each trial began with a brief playful cue "bù" used to position the hands, after which the stimulus was delivered. Trial validity was monitored online, and invalid trials were repeated. Short breaks were provided as needed, and testing continued until all trials were completed or the session was terminated by the child. The first experimenter gently stabilised the child's arms immediately after stimulus delivery and avoided eye contact by looking toward the floor to prevent distraction.

5.3 Data Analysis

5.3.1 Video analysis

Video recordings were independently coded by two raters to identify children's first orienting response on each trial. Raters annotated response direction, latency, and response type for both auditory and tactile conditions. The raters were blind to the experimental hypotheses and to stimulus laterality, but had access to timing markers indicating stimulus onset and offset. Following stimulus onset, lateral eye movements, head turns, or unilateral movements of the hand or arm were classified as orienting responses. Trials characterised by bilateral and symmetrical arm movements were classified as null responses and excluded from further analyses. Limb responses included both fine movements (e.g., finger, wrist, or elbow articulations) and larger withdrawal movements of the arm.

5.3.2 Inter-observer reliability

The video assessments from the two independent raters were evaluated for inter-rater reliability using Cohen's Kappa and weighted Kappa, which quantify agreement for nominal categorical data. Initial reliability was high. For response side (Left, Right, Null), unweighted Kappa was .942 (95% CI: .941–.948) and weighted Kappa was .940 (95% CI: .935–.941). For response modality (Arm/hand,Eye/head, Null), unweighted and weighted Kappa values were .950 (95% CI: .940–.955) and .941 (95% CI: .929–.955), respectively. To maximise statistical power and minimise classification uncertainty, the two raters jointly reviewed and resolved trials on which their initial ratings diverged. After this reconciliation procedure, inter-rater reliability increased further. For response side, unweighted Kappa reached .984

(95% CI: .974–.994) and weighted Kappa .986 (95% CI: .977–.995). For response modality, unweighted Kappa was .990 (95% CI: .982–.998) and weighted Kappa was .993 (95% CI:.988–.999).

5.3.3 Behavioural Analysis

All trials with agreement between raters were included in the behavioural analyses. All analyses were conducted using R (R Core Team, 2012). Analyses were performed separately for congruent and incongruent audio-tactile conditions. Multisensory facilitation was evaluated in the *congruent condition*, defined as a behavioural improvement in performance (e.g., higher accuracy or faster responses) in the bimodal condition relative to unisensory conditions. Multisensory conflict was examined in the *incongruent condition*, reflecting changes in performance when auditory and tactile cues conveyed spatially discrepant information.

Congruent Condition

Gain

To evaluate whether congruent audio-tactile stimulation enhanced spatial localisation beyond unimodal performance, multisensory behavioural gain was assessed by comparing accuracy in the bimodal congruent condition with accuracy in unimodal auditory and tactile conditions. This approach allowed us to quantify the behavioural benefit associated with coherent multisensory input without assuming strict superadditive integration.

Accuracy data were analysed using a generalized linear mixed-effects model (GLMM). Condition (Auditory-only, Tactile-only, Audio–Tactile congruent) and Group (S, SVI) were entered as fixed effects, together with their interaction. Age was included as a covariate, and a random intercept for Subject was added to account for individual variability :

$$\textbf{Accuracy} \sim \textbf{Condition} \times \textbf{Group} + \textbf{Age} + (\textbf{1} \mid \textbf{Subject})$$

In parallel, reaction time (RT) data were analysed using a linear mixed-effects model (LMM) with the same fixed-effects structure. Condition and Group were included as fixed effects, Age was entered as a covariate, and a random intercept for Subject assured individual differences in baseline response speed:

RT ~ Condition × Group + Age + (1 | Subject)

Race Model

To determine whether the speeding of responses in the audio–tactile congruent condition reflected genuine multisensory coactivation rather than statistical facilitation, reaction time distributions were compared against the predictions of the race model (Miller, 1982). This approach assesses multisensory redundancy gain by contrasting the empirical cumulative distribution function (CDF) of reaction times in the multisensory condition with the race model bound derived from the unisensory conditions.

Under the assumption of independent processing channels, the probability of a multisensory response occurring within a given time is constrained by the sum of the unisensory CDFs. Therefore, violations of this bound indicate that multisensory responses are faster than predicted by statistical facilitation alone, providing evidence for coactivation.

For each participant, RT CDFs were computed separately for auditory-only, tactile-only, and audio–tactile congruent trials. The race model bound was obtained by summing the unisensory CDFs, and redundancy gain was assessed by comparing this bound with the multisensory CDF across the 10th-90th percentiles. Percentile-wise comparisons were evaluated using Bonferroni-corrected thresholds.

Incongruent Condition

To examine performance in the incongruent audio-tactile condition, the effect of spatial conflict was assessed using a generalized linear mixed-effects model (GLMM). Accuracy was specified as the dependent variable. Group (S, SVI), Sensory Condition (Auditory-only, Tactile-only, Audio–Tactile incongruent), and continuous Age were included as fixed effects, along with all their interactions. A random intercept for Subject was included to account for within-participant variability.

Accuracy ~ Condition× Group + Age + (1|Subject)

Moreover, reaction times were analysed using a linear mixed-effects regression (LMM) to assess whether sensory condition, and group modulated the speed of children' responses during incongruent audio–tactile trials and the corresponding unimodal conditions. The model included Condition (Audio-only, Tactile-only, Audio-Tactile-incongruent), Group (S,

SVI), and Age as fixed effects, with all interactions. A random intercept for Subject accounted for repeated measures across trials:

RT ~ Condition × Group + Age + (1|Subject)

5.4 EEG data analysis

EEG analyses were conducted to characterise multisensory processing under congruent and incongruent audio-tactile stimulation. Multisensory gain was assessed only in the congruent condition, where superadditive responses can be defined. ERP activation patterns and brain-behaviour associations were examined for both congruent and incongruent conditions. To ensure consistent lateralisation across participants, trials presented on the right hand were remapped so that all EEG responses could be expressed relative to the stimulated hand (i.e, ipsilateral vs contralateral).

Gain

To assess whether participants exhibited a multisensory enhancement, we compared the ERP response to the multisensory audio-tactile stimulus (AT) with the linear sum of the corresponding unisensory responses (A + T). This divergence (AT > A + T) is typically interpreted as evidence of superadditivity, indicating that the multisensory response exceeds what would be predicted from independent unisensory inputs. Multisensory gain was quantified specifically within midline regions of interest (fronto-central, MFC, and centro-parietal, MCP) in the 180–220 ms time window, which has been consistently associated with associative multisensory processing in ERP studies. Previous work has shown that multisensory superadditive effects at this latency are reliably expressed over midline associative cortices, rather than primary sensory areas (e.g., Ronga et al., 2021; Murray et al., 2016; Molholm et al., 2002). For each region, we first tested within-group gain significance (gain ≠ 0) using one-sample *t*-tests conducted separately for sighted (S) and severely visually impaired (SVI) participants. We then assessed whether the magnitude of the gain differed between groups by entering gain values into a mixed-design ANOVA. Significant interactions were followed by planned post-hoc *t*-tests comparing: (i) S vs. SVI for each level of the within-subject factor, and (ii) MCP vs. MFC within each group. In addition to the ANOVA, we quantified the overall antero-posterior distribution of multisensory gain by computing the absolute difference between MCP and MFC responses.

Activation

To investigate the neural mechanisms involved in multisensory integration in sighted (S) and severely visually impaired children (SVI), we analysed ERP responses across several regions of interest (ROIs: centro-parietal, temporo-posterior, and midline) and three time windows (45–65 ms, 105–120 ms, and 180–220 ms). These time windows were selected to capture early stages of sensory processing, from early sensory responses (~50 ms) to higher-order multisensory associative processes.

We evaluated differences in ERP mean amplitude as a function of group (S vs SVI) and laterality (ipsilateral vs contralateral hemisphere with respect to the stimulated hand).
To this end, we fitted linear mixed-effects models (LMMs), including group and laterality as fixed effects and subject as a random effect (*ERP mean ~ Laterality× Group + (1 | Subject).* Fixed effects were assessed using likelihood ratio tests (χ^2). Analyses were conducted separately for the congruent and incongruent conditions. Significant effects were further explored using post-hoc contrasts based on t-tests (see Table 1)

Association

To investigate the relationship between behavioural performance and ERP amplitude, separate analyses were conducted for each time window and region of interest (ROI). Linear regression models were first applied to assess the association between ERP mean amplitude and participants' accuracy.

In addition, generalised linear mixed-effects models (GLMMs) were used to predict trial-by-trial response correctness for responses oriented toward the auditory stimulus (reference modality). The models included laterality, group, and ERP mean amplitude as fixed effects, and subject as a random effect (*accuracy ~ laterality × group × ERP mean + (1 | subject)*). Model significance was evaluated using Wald χ^2 tests, and significant effects were further explored with post-hoc contrasts based on Z-tests.

5.5 Multivariate pattern analysis (MVPA)

A multivariate pattern analysis (MVPA) model was applied to test whether distributed EEG activity
patterns reliably distinguished the experimental conditions in each group. Multivariate Pattern Analysis (MVPA) uses supervised machine-learning algorithms to identify reliable patterns

of neural activity associated with specific stimuli or task conditions. Instead of focusing on average amplitudes in single electrodes, MVPA examines distributed patterns across multiple channels, allowing it to detect subtle and spatially distributed information in the EEG signal (Haynes & Rees, 2006). When a classifier can discriminate stimuli or responses above chance level, this indicates that the underlying neural patterns contain information relevant for that distinction. These multivariate decoding approaches, widely used in adult EEG research, are increasingly being applied to infant data (Ashton et al., 2022, Gori &Vitali, 2025). We performed single-trial MVPA using the MVPAlab toolbox (López-García et al., 2022), analysing each group separately. To decode stimulus-related information, we used the same lateralisation procedure described above . To increase statistical power, all trials within each group were pooled. Finally, to decode perceptual responses, we trained classifiers to distinguish trials based on children' behavioural orienting (congruent and incongruent conditions analysed separately). Statistical significance of classification accuracy was assessed using one-sample tests against chance level

5.6 Results

Behavioural Results

Congruent multisensory condition

The model did not reveal a significant main effect of Group ($\chi^2(1)=0.32$, $p=.57$), indicating comparable localisation performance between sighted and visually impaired children. The main effect of Condition was not significant, although it showed a statistical trend ($\chi^2(2)=5.81$, $p=.055$). Age did not significantly predict performance ($\chi^2(1)=3.06$, $p=.08$).

Importantly, no significant interaction between Group and Condition was observed ($\chi^2(2)=0.82$, $p=.66$), indicating that the effect of multisensory stimulation did not differ between groups in the uncrossed posture. No interaction between Group and Age was found ($\chi^2(1)=0.05$, $p=.83$) (Figure 5.2).

The analysis on Reaction Times (RT) revealed a significant main effect of Condition ($\chi^2(2)=553.70$, $p<.001$), indicating that response speed varied as a function of sensory modality. A significant main effect of Group also emerged ($\chi^2(1)=4.63$, $p=.031$), suggesting overall differences in reaction times between sighted children and children with visual

impairment. Age did not significantly predict reaction times ($\chi^2(1)=2.33$, $p=.13$) (Figure 5.3).

Race Model

In the AT condition, both groups showed clear multisensory redundancy gains. For S children, significant violations of the race-model bound were observed from the 10th to the 60th percentile (all ps < .05 after Bonferroni correction), while higher percentiles were not significant. SVI children showed a very similar pattern, with significant redundancy gains from the 10th to the 60th percentile and no significant effects at the 70th–90th percentiles.

Direct group comparisons confirmed the absence of differences between S and SVI children across all percentiles (all ps > .05). Thus, for AT-congruent trials, the magnitude and distribution of multisensory facilitation were statistically indistinguishable across groups (Figure 5.1).

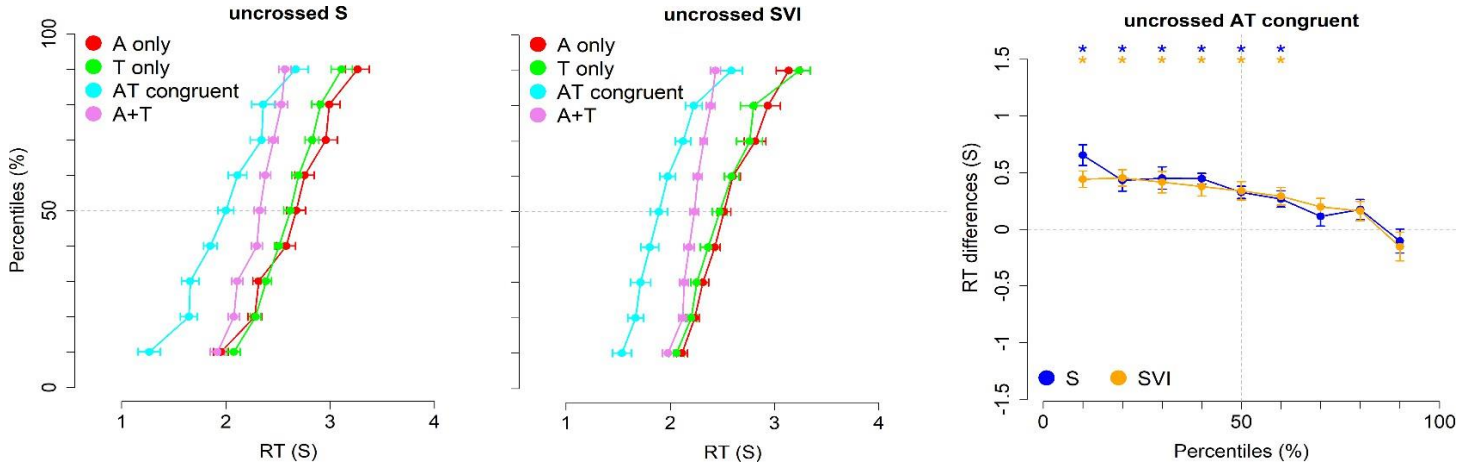

Figure 5.1. Race-model test for multisensory facilitation in the uncrossed condition. Percentile reaction time distributions are shown for unimodal auditory and tactile stimulation, congruent audio-tactile stimulation, and the race-model prediction (summation of unimodal distributions). The right panel illustrates difference waves between observed multisensory performance and race-model predictions. Positive deviations indicate multisensory redundancy gains. Significant violations are marked by asterisks.

Multisensory conflict

To examine performance under multisensory conflict, accuracy in the incongruent audio–tactile condition and the corresponding unimodal trials was analysed using a generalized linear mixed-effects model. The model revealed significant main effects of Group ($\chi^2(1)=32.28$, $p<.001$) and Condition ($\chi^2(2)=13.14$, $p=.001$), as well as a significant Group × Condition interaction ($\chi^2(2)=44.15$, $p<.001$). Age did not significantly predict performance ($\chi^2(1)=1.42$, $p=.23$).

Overall, sighted children were more accurate than visually impaired children (estimate=1.02, SE=0.22, z=4.75, $p<.001$). Collapsing across groups, accuracy was higher in Auditory-only compared to Audio–Tactile incongruent trials (estimate=0.79, SE=0.25, z=3.14, $p=.005$) (Figure 5.2).

Reaction times in the incongruent audio–tactile condition were analysed using a linear mixed-effects model including Condition and Group as fixed effects, Age as a covariate, and a random intercept for Subject.

The analysis revealed significant main effects of Group ($\chi^2(1)=139.7$, $p<.001$) and Condition ($\chi^2(2)=233.8$, $p<.001$), indicating that response speed differed between sighted and visually impaired children and varied across sensory conditions. Age did not significantly predict reaction times.

A significant interaction between Group and Condition was also observed ($\chi^2(2)=457.8$, $p<.001$), indicating that the effect of multisensory conflict on response speed differed across groups (Figure 5.3).

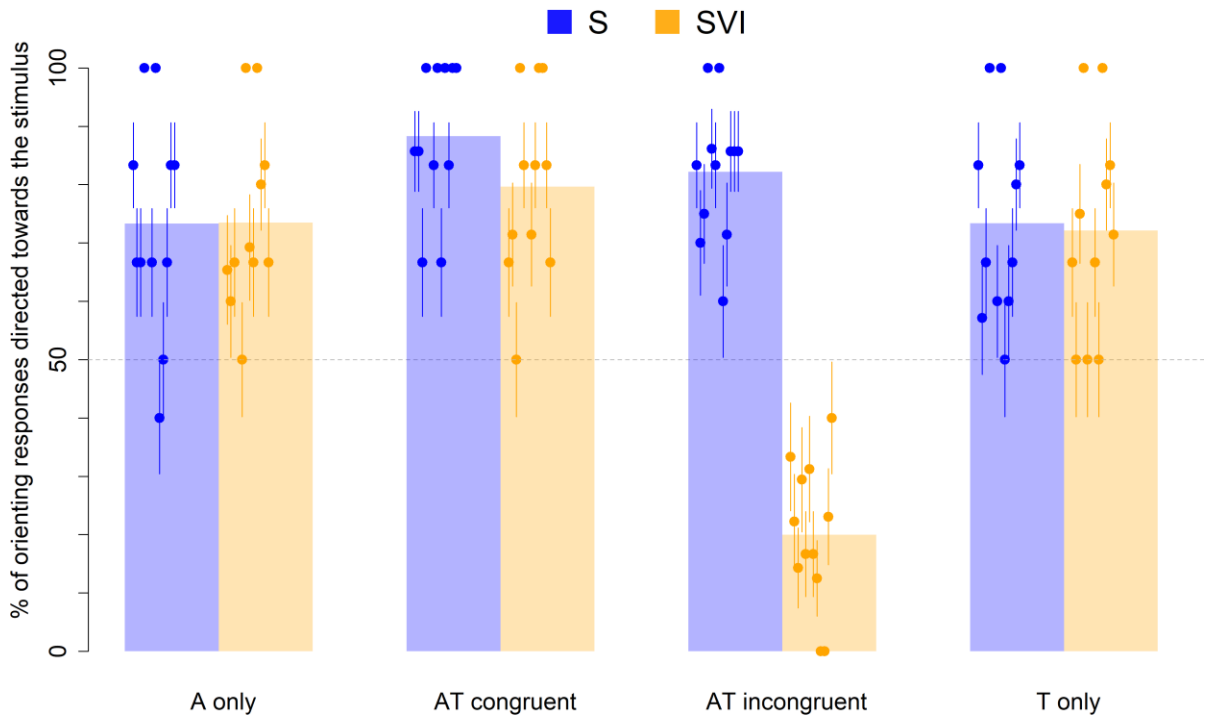

Figure 5.2 Accuracy of spatial orienting responses in sighted (S) and visually impaired (SVI) children across sensory conditions (A only, AT congruent, AT incongruent, T only). Bars represent the mean percentage of responses directed towards the target stimulus, while dots indicate individual participants. Error bars around the dots represent within-participant variability (± standard error). In the incongruent audio–tactile condition, accuracy reflects orienting towards the auditory stimulus, consistent with the definition adopted in the statistical analyses. The dashed horizontal line at 50% indicates chance-level performance.

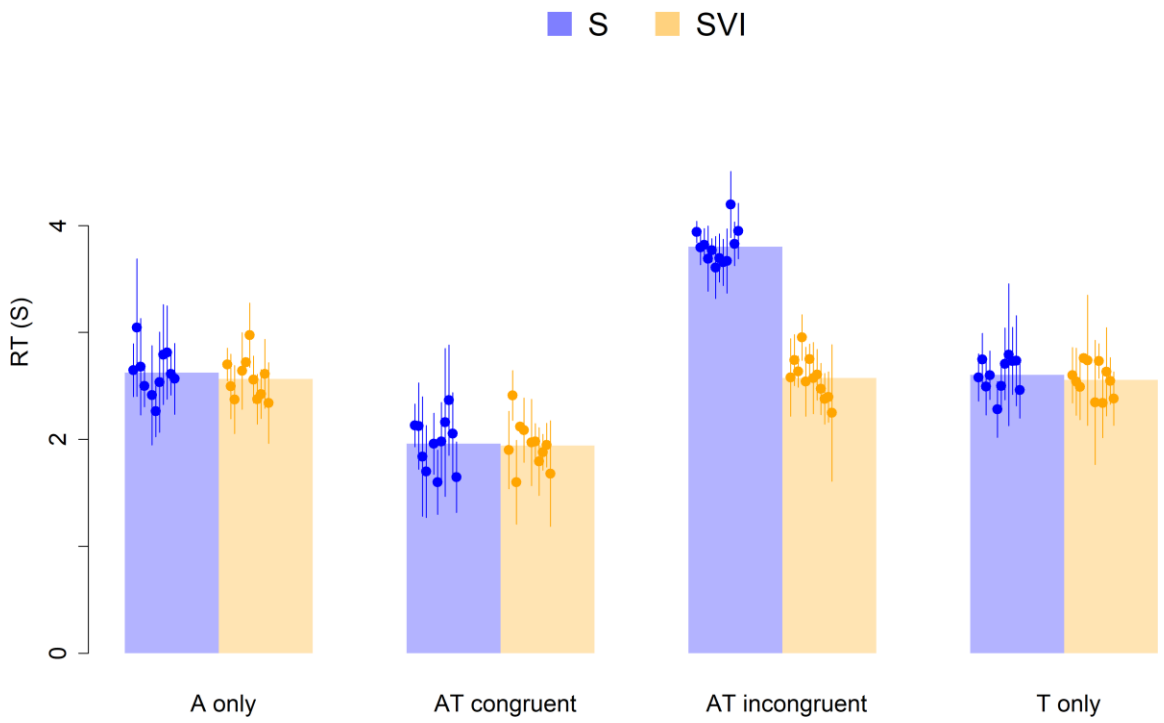

Figure 5.3 Reaction times (RT) of spatial orienting responses in sighted (S) and visually impaired (SVI) children across sensory conditions (A only, AT congruent, AT incongruent, T only). Bars represent mean RTs, while dots indicate individual participants. Error bars represent ± standard error.

5.5.3 Results EEG

Congruent Condition

Below, we report the results for congruent and incongruent conditions separately, for each relevant ROI and time window, focusing only on statistically significant effects. Unisensory conditions are not discussed here, as the present analyses specifically target multisensory interactions.

GAIN

Centro-Parietal Region (105-120 ms)

For congruent audio-tactile stimulation, a significant ERP gain emerged only in the contralateral hemisphere for both groups (S: $t(9)=9.47$, $p<0.001$; SVI: $t(8)=17.88$, $p<0.001$), whereas ipsilateral responses did not differ from zero in either group. The ANOVA on ERP

gain revealed robust main effects of laterality ($F(1,17)=100.5$, $p<0.001$) and group ($F(1,17)=40.74$, $p<0.001$), as well as a significant laterality × group interaction ($F(1,17)=7.30$, $p=0.015$). Within-group contrasts indicated that both groups exhibited greater contralateral than ipsilateral activation (S: $t(9)=-5.17$, $p<0.001$; SVI: $t(8)=-9.38$, $p<0.001$). Post-hoc comparisons showed that contralateral gain was significantly larger in SVI than in S ($t(17)=-7.48$, $p<0.001$) (Figure 5.4, Table 5.1 at the end of the chapter).

Temporo-Posterior Area (105-120 ms)

One-sample t-tests indicated that ERP gain was significant in the contralateral hemisphere in both groups (S: t(9)=12.73, p<.001; SVI: t(8)=6.20, p=.001), whereas ipsilateral responses did not differ

from zero. The ANOVA revealed significant main effects of laterality (F(1,17)=141, p<.001) and group (F(1,17)=23.19, p<.001), as well as a significant laterality × group interaction (F(1,17)=33.71, p<.001).

Post-hoc comparisons showed that contralateral responses were greater than ipsilateral ones in both groups (S: t(9)=−9.72, p<.001; SVI: t(8)=−8.23, p<.001), and that contralateral gain differed between groups (t(17)=6.55, p<.001) (Figure 5.4, Table 5.1 at the end of the chapter).

Midline Area (180-220 ms)

One-sample t-tests showed that ERP gain was significant at mid-frontocentral (MFC) sites in sighted children (t(9)=9.65, p<.001) and at mid-centroparietal (MCP) sites in SVI children (t(8)=6.54, p<.001). The ANOVA revealed significant main effects of antero-posteriority (F(1,17)=25.94, p<.001) and group (F(1,17)=11.44, p=.0035), as well as a significant interaction (F(1,17)=130, p<.001). Post-hoc comparisons showed that S children differed between MFC and MCP (t(9)=−9.31, p<.001), and SVI children differed between sites (t(8)=7.58, p<.001). Additional post-hoc comparisons indicated differences between groups at MCP (t(17)=−6.38, p < .001) and MFC (t(17)=8.62, p < .001), as well as a significant difference in the antero–posterior contrast between groups (t(17)=4.49, p < .001), reflecting differences in the distribution of multisensory gain along the midline. (Figure 5.4, Table 5.1 at the end of the chapter).

Multisensory ERP Gain

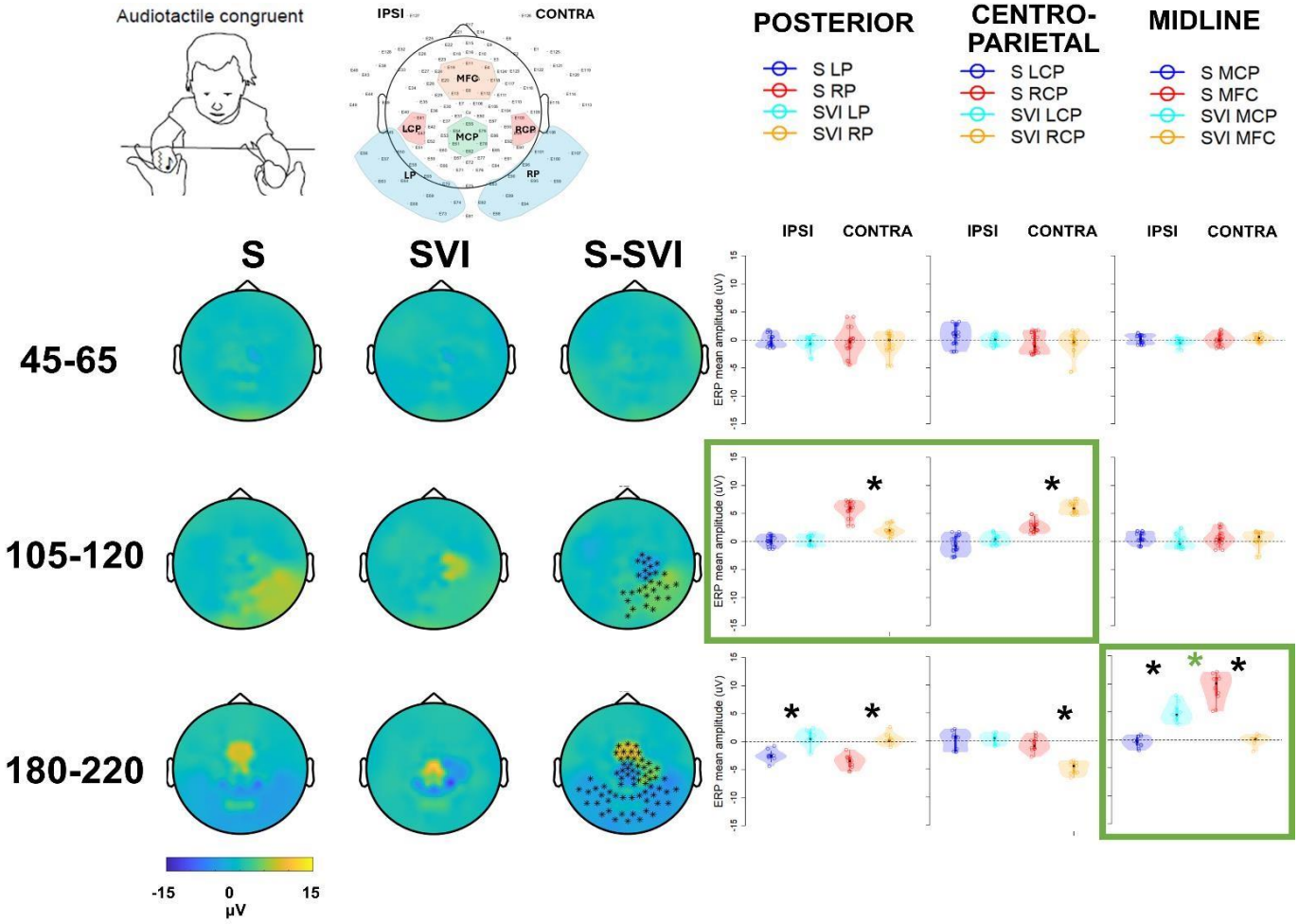

Figure 5.4 Multisensory ERP gain during congruent audio-tactile stimulation in sighted (S) and visually impaired (SVI) children. Scalp maps depict the spatial distribution of ERP gain across three temporal windows (45-65 ms, 105-120 ms, 180-220 ms), along with group differences (S–SVI). Plots on the right show mean ERP gain in posterior, centro-parietal, and midline ROIs, separately for ipsilateral and contralateral hemispheres. Individual dots represent participants, and error bars denote standard errors. Asterisks indicate significant effects.

5.5.3.2 ACTIVATION ERP

Centro-Parietal area (105-120 ms)

A linear mixed-effects model revealed a significant main effect of laterality ($\chi^2(1)=913.4$, $p < .001$), a significant main effect of group ($\chi^2(1)=71.84$, $p < .001$), and a significant laterality × group interaction ($\chi^2(1)=105$, $p < .001$). Post-hoc comparisons revealed a significant difference between groups in the contralateral condition ($t(62)=-13.19$, $p < .001$), whereas no significant difference emerged in the ipsilateral condition. Within-group comparisons showed that both S and SVI children exhibited greater contralateral than ipsilateral responses (S: $t(476)=-14.12$, $p < .001$; SVI: $t(476)=-28.62$, $p < .001$) (Figure 5.5, 5.6; Table 5.2).

Temporo-posterior area (105–120 ms)

A linear mixed-effects model revealed a significant main effect of laterality ($\chi^2(1)=289.4$, $p < .001$), a significant main effect of group ($\chi^2(1)=39.02$, $p < .001$), and a significant laterality × group interaction ($\chi^2(1)=55.83$, $p < .001$).

Post-hoc comparisons revealed a significant difference between groups in the contralateral condition ($t(37)=9.40$, $p < .001$), whereas no significant difference emerged in the ipsilateral condition.

Within-group comparisons showed that both S and SVI children exhibited greater contralateral than ipsilateral responses (S: $t(475)=-17.31$, $p < .001$; SVI: $t(475)=-6.75$, $p < .001$) (Figure 5.5, 5.6; Table 5.2).

Midline area (180-220 ms)

A linear mixed-effects model revealed a significant main effect of antero-posteriority ($\chi^2(1)=58.86$, $p < .001$), a significant main effect of group ($\chi^2(1)=11.4$, $p < .001$), and a significant antero-posteriority × group interaction ($\chi^2(1)=156.6$, $p < .001$).

Post-hoc comparisons revealed significant differences between groups at both MCP ($t(60)=-6.44$, $p < .001$) and MFC ($t(60)=11.19$, $p < .001$).

Within-group comparisons showed that both S and SVI children exhibited significant differences between MCP and MFC sites (S: $t(227)=-14.27$, $p < .001$; SVI: $t(227)=3.42$, $p < .001$) (Figure 5.5, 5.6; Table 5.2).

.

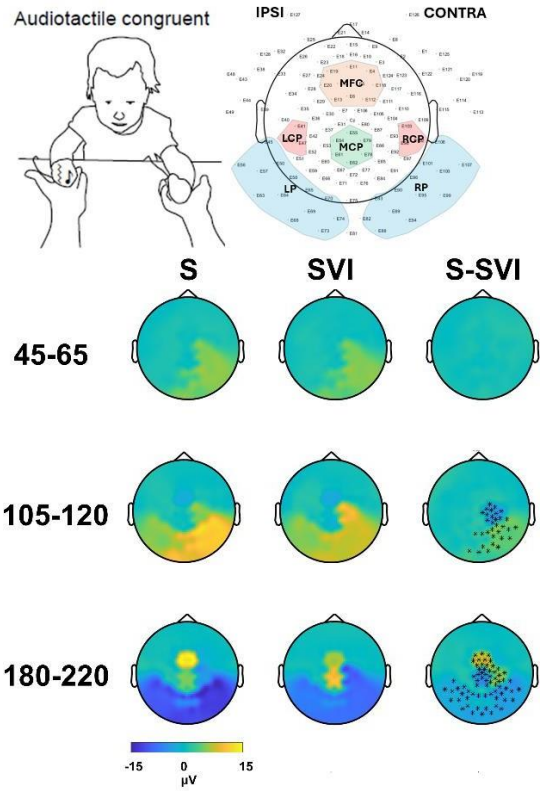

Figure 5.5 Scalp topographies of ERP mean activation during audio-tactile congruent trials in sighted (S) and visually impaired (SVI) children across three time windows (45-65 ms, 105-120 ms, 180-220 ms). The S-SVI panel displays the between-group difference. Significant group differences are observed in the 105-120 ms and 180-220 ms time windows.

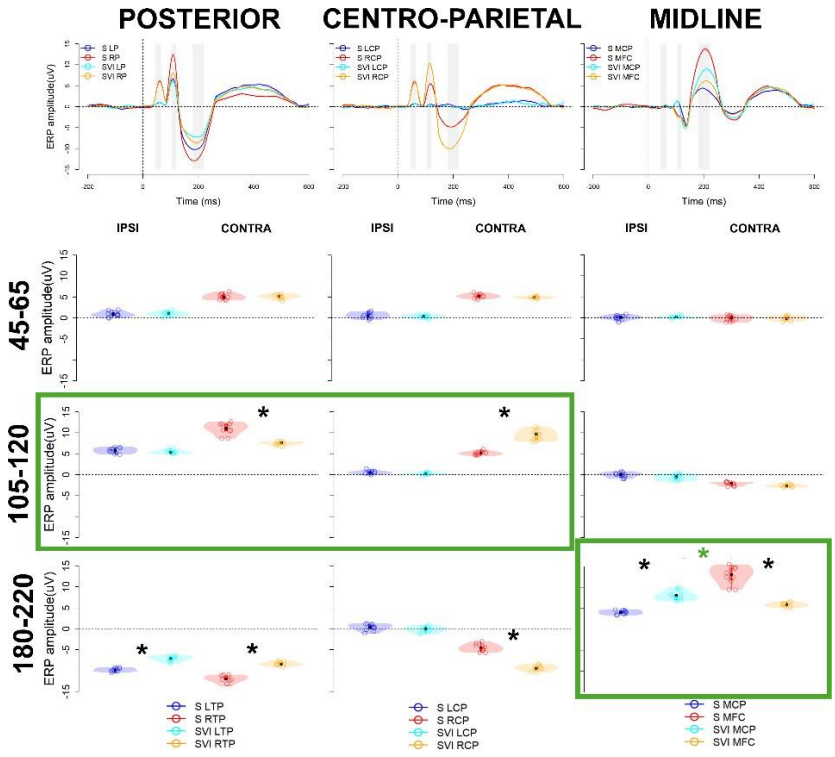

Figure 5.6 ERP mean activation during the audio-tactile congruent condition in sighted (S) and visually impaired (SVI) children. Top panels show grand-average ERP waveforms in posterior, centro-parietal, and midline regions of interest. Lower panels display ERP mean amplitudes (μV) across three time windows (45-65 ms, 105-120 ms, and 180-220 ms), separately for ipsilateral and contralateral responses. Individual data points and group distributions are shown. Significant effects (highlighted in green) are observed in centro-parietal and temporo-posterior regions during the 105-120 ms time window, and in the midline region during the 180-220 ms time window.

5.5.3.3 Behavioural Association

We assessed how ERP amplitudes are related to behavioural performance using linear and mixed-effects models, testing effects of group, laterality, and ERP amplitude on accuracy at both aggregated and trial- by-trial levels. Results are summarised below for each region of interest and reported separately for the congruent and incongruent conditions, focusing only on statistically significant effects.

Centro parietal area (105-120 ms)

A significant association between ERP amplitude and behavioural performance was observed in SVI participants, with contralateral ERP amplitude strongly correlating with accuracy ($R^2 = 0.98$, $p < .001$). No significant relationship was observed in sighted children. The GLMM revealed a significant main effect of ERP amplitude ($\chi^2(1) = 4.51$, $p = .034$), which was further qualified by a significant laterality × group × ERP mean interaction ($\chi^2(1) = 12.5$, $p < .001$). Post-hoc comparisons indicated a significant difference between groups in the contralateral condition ($Z = -3.95$, $p < .001$), suggesting that the relationship between ERP amplitude and behavioural performance was specific to SVI participants and depended on hemispheric laterality (Figure 5.7, 5.8; Table 5.3).

Temporo-Posterior area (105-120 ms)

A significant association between ERP amplitude and behavioural accuracy emerged in sighted children, with contralateral ERP responses correlating with accuracy ($R^2 = 0.92$, $p < .001$). No significant association was observed in SVI participants.

Consistent with this pattern, the GLMM revealed a significant interaction between group and ERP amplitude ($\chi^2(1) = 13.49$, $p < .001$), as well as a significant laterality × group × ERP mean interaction ($\chi^2(1) = 5.6$, $p = .01$), indicating that the relationship between ERP amplitude and behavioural performance differed across groups and depended on hemispheric laterality (Figure 5.7, 5.8; Table 5.3).

Midline area (180-220 ms)

In sighted children, ERP amplitude at midline fronto-central (MFC) sites was strongly associated with behavioural accuracy ($R^2 = 0.92$, $p < .001$), whereas no significant association was observed at centro-parietal (MCP) sites. Conversely, in SVI participants, accuracy was strongly associated with ERP amplitude at MCP sites ($R^2 = 0.95$, $p < .001$), with no significant relationship at MFC sites. The GLMM revealed a significant interaction between antero-posteriority and group ($\chi^2(1) = 7.77$, $p = .005$), indicating that the spatial distribution of ERP–behaviour relationships differed between groups. This pattern reflects a shift in the topographical organisation of multisensory processing, with sighted children relying more on anterior regions and SVI participants on posterior regions (Figure 5.7, 5.8; Table 5.3).

Consistent with these findings, MVPA results showed that classification accuracy was significantly above chance in the 105–120 ms and 180–220 ms time windows in both groups (all ps < .001).

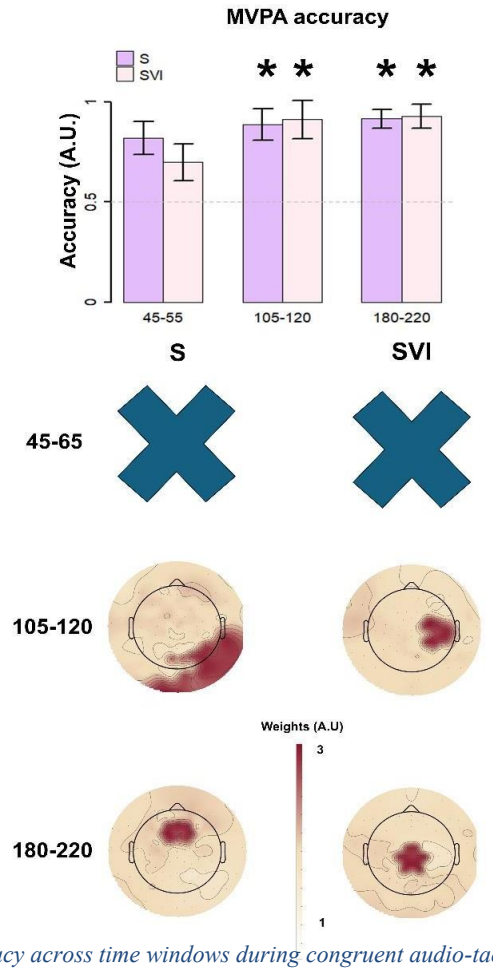

Figure 5.7 MVPA classification accuracy across time windows during congruent audio-tactile stimulation in sighted (S) and visually impaired (SVI) children. Bars represent mean classification accuracy (± SE) at 45-65 ms, 105-120 ms, and 180-220 ms. Scalp maps show the spatial distribution of decoding weights for each group and time window. No significant decoding was observed in the earliest window, whereas above-chance classification emerged at 105-120 ms and 180-220 ms in both groups.

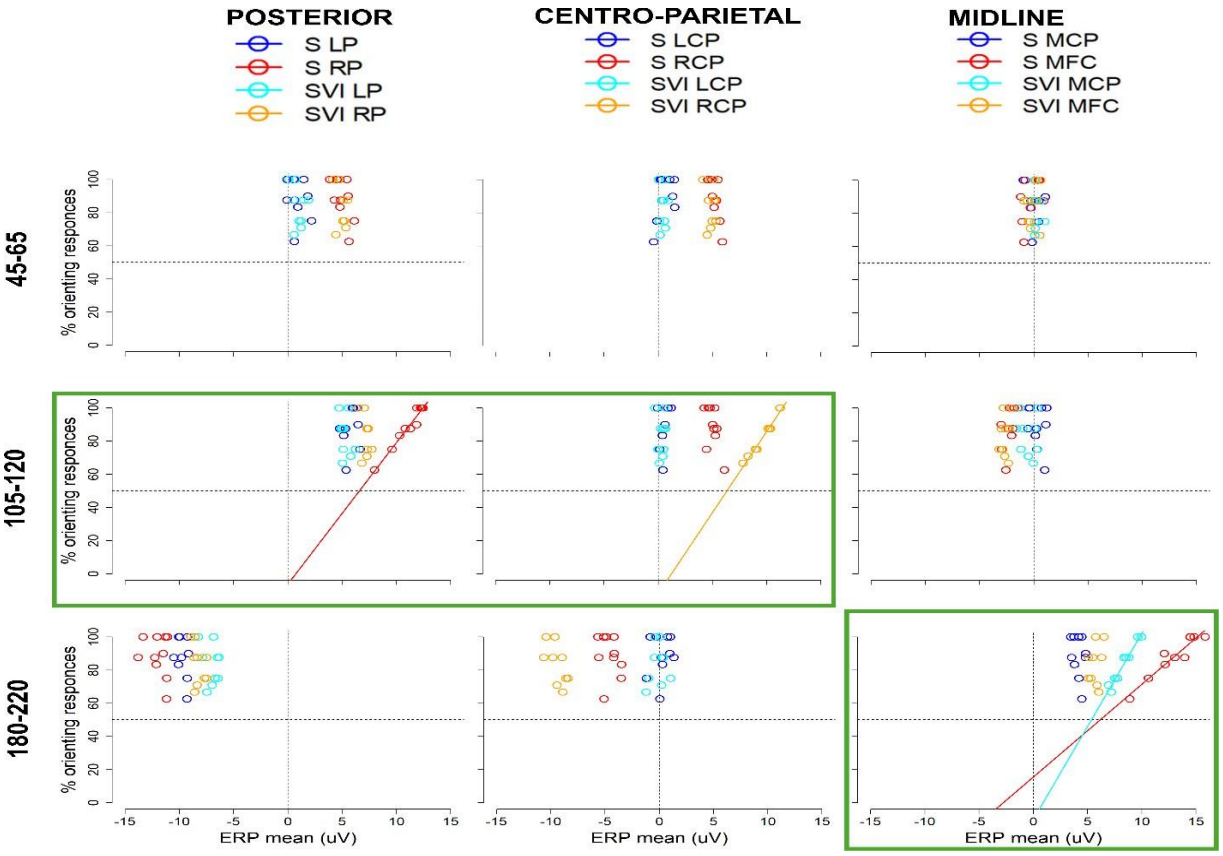

Figure 5.8 Relationship between ERP mean amplitude and behavioural orienting accuracy across posterior, centro-parietal, and midline regions in sighted (S) and visually impaired (SVI) children. Scatter plots show individual participants' data across three time windows (45–65 ms, 105-120 ms, and 180-220 ms). Solid lines indicate significant correlations between neural responses and behavioural performance. Panels highlighted in green mark the regions and time windows showing significant brain-behaviour associations.

EEG RESULTS INCONGRUENT CONDITION
ERP ACTIVATION

Centro-Parietal area (105-120 ms)

A linear mixed-effects model (ERP mean ~ laterality × group + (1|subject)) revealed a significant main effect of laterality ($\chi^2(1)=116$, $p < .001$) and a significant main effect of group ($\chi^2(1)=7.14$, $p = .0075$). The interaction between laterality and group was not significant ($\chi^2(1)=0.05$, $p = .822$).

Post-hoc comparisons revealed no significant differences between groups in either ipsilateral or contralateral conditions. Within-group comparisons showed significant differences between ipsilateral and contralateral responses in both groups (S: $t(958)=7.20$, $p < .001$; SVI: $t(958)=6.98$, $p < .001$) (Figure 5.9, 5.10; Table 5.4).

Temporo-Posterior (105-120 ms)

A linear mixed-effects model (ERP mean ~ laterality × group + (1|subject)) did not reveal a significant main effect of laterality ($\chi^2(1)=0.96$, $p = .326$) or group ($\chi^2(1)=0.41$, $p = .524$). The interaction between laterality and group was not significant ($\chi^2(1)=0.29$, $p = .588$).

No significant differences were observed between groups or between ipsilateral and contralateral responses in either group (Figure 5.9, 5.10; Table 5.4).

Midline area (180–220 ms)

A linear mixed-effects model (ERP mean ~ antero-posteriority × group + (1|subject)) revealed a significant main effect of antero-posteriority ($\chi^2(1)=8.03$, $p = .0046$) and a significant main effect of group ($\chi^2(1)=5.00$, $p = .025$). The interaction between antero-posteriority and group was also significant ($\chi^2(1)=19.26$, $p < .001$).

Post-hoc comparisons revealed significant differences between groups at both MCP ($t(25)=11.25$, $p < .001$) and MFC sites ($t(25)=16.75$, $p < .001$). Within-group comparisons showed a significant difference between MCP and MFC in sighted children ($t(472)=-5.03$, $p < .001$), whereas no significant difference was observed in SVI participants (Figure 5.9, 5.10; Table 5.4).

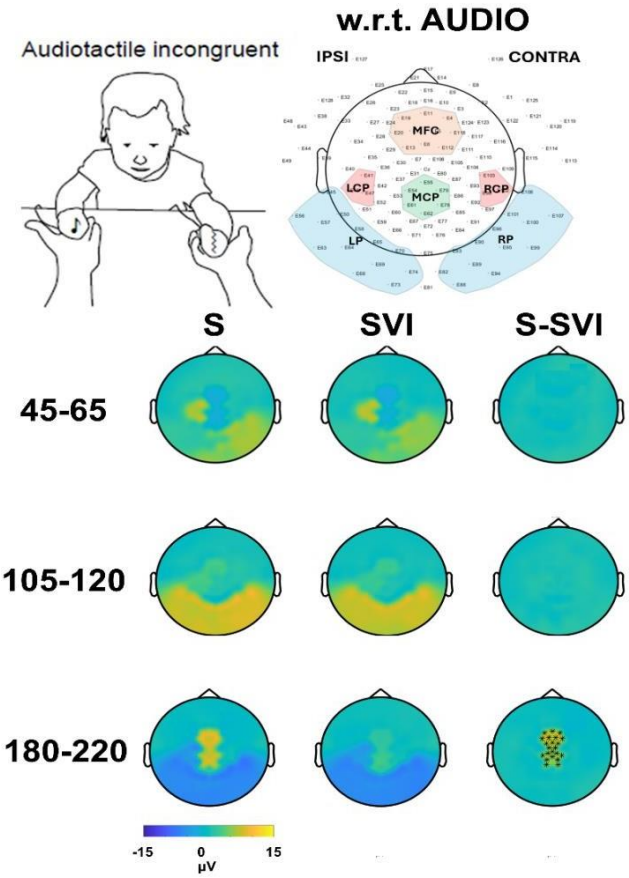

Figure 5.9 Scalp topographies of ERP mean activation during audio-tactile incongruent trials (with respect to the auditory stimulus (w.r.t)) in sighted (S) and visually impaired (SVI) children across three time windows (45-65 ms, 105-120 ms, 180-220 ms). The S–SVI panel displays the between-group difference. Significant group differences are observed only in the 180-220 ms time window.

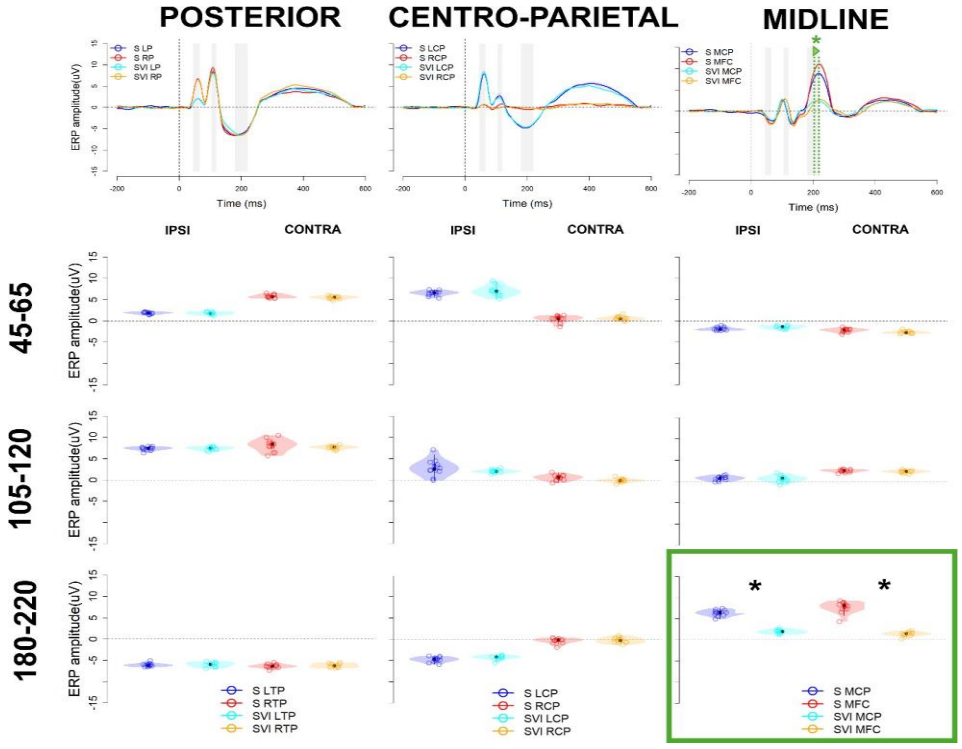

Figure 5.10 ERP mean activation during the audio-tactile incongruent condition (reference: auditory stimulus) in sighted (S) and visually impaired (SVI) children. Top panels show grand-average ERP waveforms in posterior, centro-parietal, and midline regions of interest. Lower panels display ERP mean amplitudes (μV) across three time windows (45–65 ms, 105-120 ms, and 180-220 ms), separately for ipsilateral and contralateral responses. Individual data points and group distributions are shown. Significant effects are observed in the midline region during the 180-220 ms time window (highlighted in green), reflecting differences between MCP and MFC sites and their interaction with group.

5.5.4.2 Behavioural Association

Centro-parietal area (105–120 ms)

A linear regression analysis showed that, in sighted children, ERP mean amplitude significantly predicted behavioural accuracy in the contralateral condition ($R^2 = 0.9798$, $p < .001$), whereas no significant association was observed in the ipsilateral condition or in children with severe visual impairment (SVI).

A generalized linear mixed-effects model revealed a significant main effect of group ($\chi^2(1) = 51.13$, $p < .001$), but no significant main effects of laterality ($\chi^2(1) = 1.10$, $p = .294$) or ERP mean amplitude ($\chi^2(1) = 0.72$, $p = .395$).

A significant interaction emerged between laterality and group ($\chi^2(1) = 9.27$, $p = .002$), as well as a significant three-way interaction between laterality, group, and ERP mean amplitude ($\chi^2(1) = 5.23$, $p = .022$), indicating that the relationship between neural responses and behavioural accuracy differed across groups depending on spatial laterality.

Post-hoc comparisons revealed a significant difference between groups in the ipsilateral condition ($Z = 2.29$, $p = .022$). Within the sighted group, behavioural accuracy also differed significantly between ipsilateral and contralateral responses ($Z = -2.29$, $p = .022$) (Figure 5.11, 5.12; Table 5).

Temporo-Posterior area (105-120 ms)

A linear regression analysis showed that, in the contralateral condition, ERP mean amplitude significantly predicted behavioural accuracy in sighted children ($R^2 = 0.8976$, $p < .001$), whereas no significant association was observed in the ipsilateral condition or in SVI participants. A generalized linear mixed-effects model (accuracy ~ laterality × group × ERP mean + (1|subject)) revealed no significant main effect of laterality ($\chi^2(1) = 0.14$, $p = .708$), but a significant main effect of group ($\chi^2(1) = 64.29$, $p < .001$). The main effect of ERP mean was not significant ($\chi^2(1) = 2.01$, $p = .156$). The interaction between laterality and group was significant ($\chi^2(1) = 8.481$, $p < .001$), whereas the interaction between laterality and ERP mean was not significant ($\chi^2(1) = 0.03$, $p = .867$). The interaction between group and ERP mean was significant ($\chi^2(1) = 19.96$, $p < .001$). A significant three-way interaction between laterality, group, and ERP mean was also observed ($\chi^2(1) = 31.33$, $p < .001$).

Post-hoc comparisons revealed a significant difference between groups in the contralateral condition

($Z = 5.59$, $p < .001$), whereas no significant difference was observed in the ipsilateral condition. Additional post-hoc comparisons showed a significant difference between ipsilateral and contralateral responses in the sighted group ($Z = -5.59$, $p < .001$), whereas no significant difference was observed in the SVI group (Figure 5.11, 5.12; Table 5)

Midline area (180-220 ms)

Linear regression analyses showed that ERP mean amplitude significantly predicted behavioural accuracy in sighted children at both mid-centro-parietal (MCP: $R^2 = 0.767$, $p < .001$) and mid-fronto-central (MFC: $R^2 = 0.674$, $p = .001$) sites. No significant associations were observed in the SVI group. Mixed-effects modelling revealed a significant main effect of group ($\chi^2(1) = 4.16$, $p = .041$), whereas no main effects were found for antero-posteriority ($\chi^2(1) = 0.003$, $p = .955$) or ERP mean amplitude ($\chi^2(1) = 1.01$, $p = .315$).

A significant interaction between group and ERP mean amplitude was observed ($\chi^2(1) = 19.33$, $p < .001$), indicating that the neural–behavioural relationship differed between sighted and visually impaired children.

Post-hoc comparisons confirmed significant group differences at both MCP ($Z = 3.40$, $p < .001$) and MFC ($Z = 2.87$, $p = .004$) sites (Figure 5.11, 5.12; Table 5). Consistent with these findings, MVPA results showed above-chance classification accuracy in different time windows across groups: SVI participants showed significant decoding in the early 45–55 ms window, whereas sighted children showed significant decoding in the 105–120 ms and 180–220 ms windows.

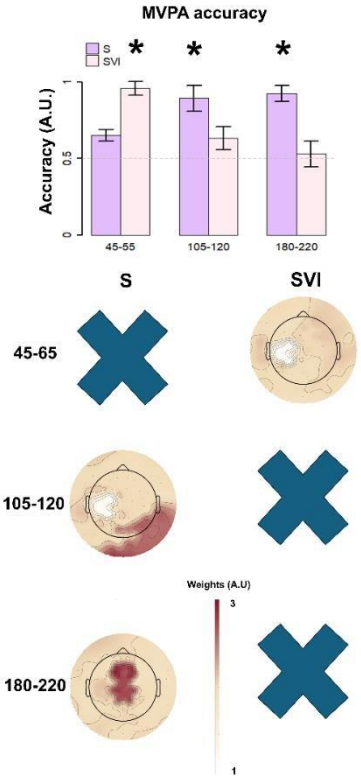

Figure 5.11 Multivariate pattern analysis (MVPA) accuracy across time windows. Bar plots show classification accuracy (mean ± SE) for sighted (S) and visually impaired (SVI) children in three temporal windows (45–65 ms, 105–120 ms, 180–220 ms). Asterisks indicate above-chance decoding performance. Scalp maps illustrate the spatial distribution of classifier weights for significant time windows in each group.

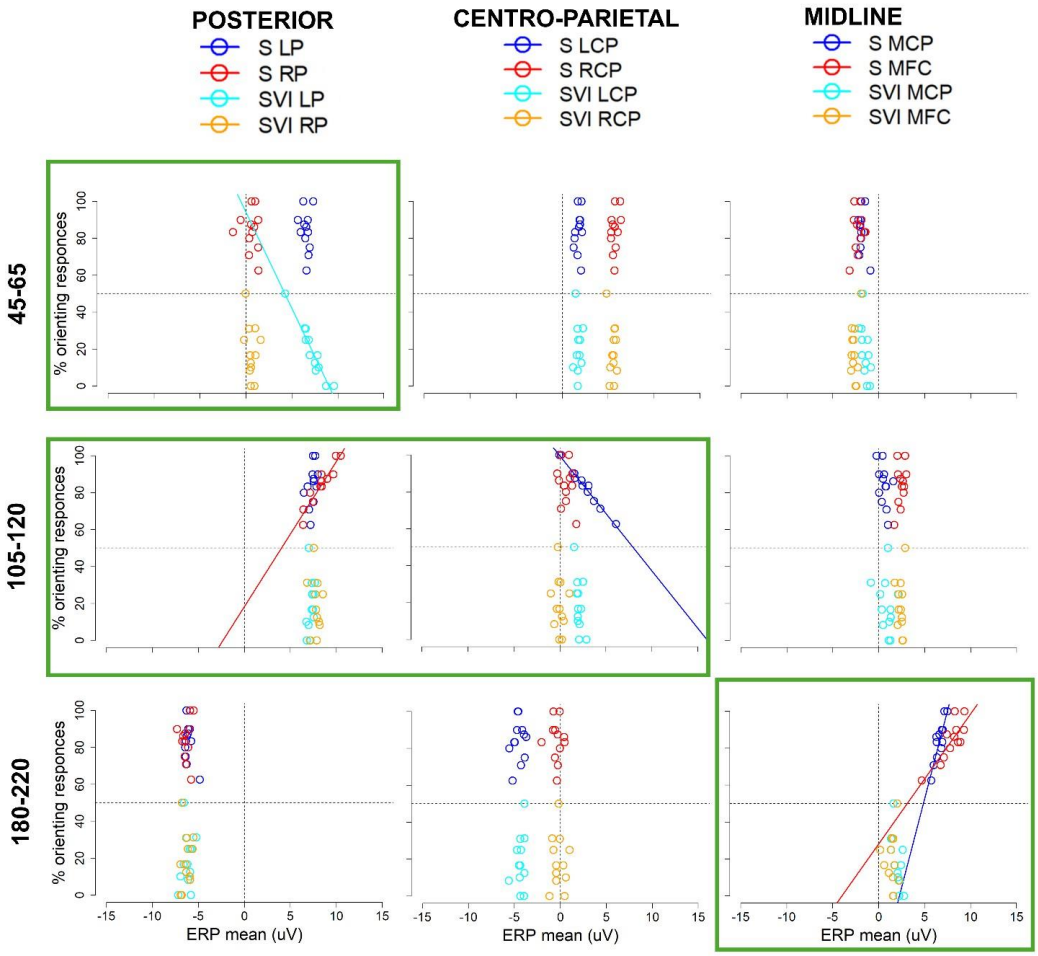

Figure 5.12 Relationship between ERP amplitude and behavioural accuracy across regions of interest and time windows.
Scatter plots show the association between mean ERP amplitude and percentage of orienting responses for sighted (S) and visually impaired (SVI) children in posterior, centro-parietal, and midline regions. Each row corresponds to a specific temporal window (45–65 ms, 105–120 ms, 180–220 ms). Regression lines are shown only for statistically significant associations. RIQUADRI VERDI

Tables

Congruent Condition

Gain

Centro-Parietal area [105-120]ms

	Congruent stimuli			
ERP gain significance	**S**		**SVI**	
	t(df)	P	t(df)	P
Ipsi	-	-	-	-
Contra	9.467(9)	<0.001*	17.88(8)	<0.001*
ANOVA of the ERP gain	*ERP gain~laterality*group*			
	$F(df_{num}, df_{den})$		**P**	
laterality	100.5(1,17)		<0.001*	
group	40.74(1,17)		<0.001*	
laterality*group	7.298(1,17)		0.01513*	
Post hoc				
Comparisons for laterality	**Ipsi**		**Contra**	
	t(df)	P	t(df)	P
S vs SVI	-	-	-7.479(17)	<0.001*
Comparisons for group	**S**		**SVI**	
	t(df)	P	t(df)	P
Ipsi vs Contra	-5.169(9)	<0.001*	-9.384(8)	<0.001*

Posterior area [105-120]ms

	Congruent stimuli			
ERP gain significance	**S**		**SVI**	
	t(df)	P	t(df)	P
Ipsi	-	-	-	-
Contra	12.73(9)	<0.001*	6.197(8)	0.001041*
ANOVA of the ERP gain	*ERP gain~laterality*group*			
	$F(df_{num}, df_{den})$		**P**	
laterality	141(1,17)		<0.001*	
group	23.19(1,17)		<0.001*	
laterality*group	33.71(1,17)		<0.001*	
Post hoc				
Comparisons for laterality	**Ipsi**		**Contra**	
	t(df)	P	t(df)	P
S vs SVI	-	-	6.551(17)	<0.001*
Comparisons for group	**S**		**SVI**	

	t(df)	P	t(df)	P
Ipsi vs Contra	-9.721(9)	<0.001*	-8.226(8)	<0.001*
Midline area [180-220]ms				
	Congruent stimuli			
ERP gain significance	**S**		**SVI**	
	t(df)	P	t(df)	P
MCP	-	-	6.539(8)	<0.001*
MFC	9.646(9)	<0.001*	-	-
ANOVA of the ERP gain	*ERP gain~anteroposteriority*group*			
	$F(df_{num}, df_{den})$		**P**	
anteroposteriority	25.94(1,17)		<0.001*	
group	11.44(1,17)		0.003538*	
anteroposteriority*group	130(1,17)		<0.001*	
Post hoc				
Comparisons for anteroposteriority	**MCP**		**MFC**	
	t(df)	P	t(df)	P
S vs SVI	-6.379(17)	<0.001*	8.624(17)	<0.001*
Comparisons for group	**S**		**SVI**	
	t(df)	P	t(df)	P
MCP vs MFC	-9.309(9)	<0.001*	7.579(8)	<0.001*

Table 5.1

ERP Activation

Centro-Parietal area [105-120]ms				
	Congruent stimuli			
Linear mixed effect models of the ERP mean	*ERP mean~laterality*group+(1|subject)*			
	χ^2(df)		**P**	
laterality	913.4(1)		<0.001*	
group	71.84(1)		<0.001*	
laterality*group	105(1)		<0.001*	
Post hoc				
Comparisons for laterality	**Ipsi**		**Contra**	
	t(df)	P	t(df)	P
S vs SVI	-	-	-13.19(62)	<0.001*
Comparisons for group	**S**		**SVI**	
	t(df)	P	t(df)	P

Ipsi vs Contra	-14.12(476)	<0.001*	-28.62(476)	<0.001*

Temporo-posterior area [105-120]ms

	Congruent stimuli			
Linear mixed effect models of the ERP mean	*ERP mean~laterality*group+(1|subject)*			
	χ^2(df)		**P**	
laterality	289.4(1)		<0.001*	
group	39.02(1)		<0.001*	
laterality*group	55.83(1)		<0.001*	
Post hoc				
Comparisons for laterality	**Ipsi**		**Contra**	
	t(df)	P	t(df)	P
S vs SVI	-	-	9.402(37)	<0.001*
Comparisons for group	**S**		**SVI**	
	t(df)	P	t(df)	P
Ipsi vs Contra	-17.31(475)	<0.001*	-6.745(475)	<0.001*

Midline area [180-220]ms

	Congruent stimuli			
Linear-mixed effect models of the ERP mean	*ERP mean~anteroposteriority*group+(1|subject)*			
	χ^2(df)		**P**	
anteroposteriority	58.86(1)		<0.001*	
group	11.4(1)		<0.001*	
anteroposteriority*group	156.6(1)		<0.001*	
Post hoc				
Comparisons for anteroposteriority	**MCP**		**MFC**	
	t(df)	P	t(df)	P
S vs SVI	-6.436(60)	<0.001*	11.19(60)	<0.001*
Comparisons for group	**S**		**SVI**	
	t(df)	P	t(df)	P
MCP vs MFC	-14.27(227)	<0.001*	3.424(227)	<0.001*

Table 5.2

Association

Centro-Parietal area [105-120]ms

	Congruent stimuli	
Linear regression analysis		
	S	**SVI**

	R^2	**p**	**R^2**	**p**
Ipsi - Accuracy~ ERP ean	-	-	-	-
Contra - Accuracy~ERP mean	-	-	0.9796	<0.001*
Generalized linear mixed models	*accuracy~laterality*group*ERP mean+(1|subject)*			
	χ^2(df)		**P**	
laterality	0.9571(1)		0.3279	
group	1.217(1)		0.27	
ERP mean	4.513(1)		0.03364*	
laterality*group	0.419(1)		0.517	
laterality*ERP mean	0.8778(1)		0.3488	
group*ERP mean	3.783(1)		0.05179	
laterality*group*ERP mean	12.5(1)		<0.001*	
Post hoc GLMM				
Comparisons for laterality		**Ips**		**Contra**
	Z	**p**	**Z**	**p**
S vs SVI	-	-	-3.947	<0.001*
Comparisons for group		**S**		**SVI**
	Z	**p**	**Z**	**p**
Ipsi vs Contra	-	-	3.06	0.00221*

Temporo-Posterior area [105-120]ms

	Congruent stimuli			
Linear regression analysis				
		S		**SVI**
	R^2	**p**	**R^2**	**p**
Ipsi - Accuracy~ERP mean	-	-	-	-
Contra - Accuracy~ERP mean	0.9176	<0.001*		
Generalized linear mixed models	*accuracy~laterality*group*ERP mean+(1|subject)*			
	χ^2(df)		**P**	
laterality	0.095(1)		0.757	
group	0.0001(1)		0.992	
ERP mean	1.815(1)		0.178	
laterality*group	0.2763(1)		0.599	
laterality*ERP mean	0.052(1)		0.818	
group*ERP mean	13.49(1)		0.0002*	
laterality*group*ERP mean	5.6(1)		0.01*	

Post hoc GLMM				
Comparisons for laterality	**Ipsi**		**Contra**	
	Z	**p**	**Z**	**p**
S vs SVI	2.513	0.01198*	3.664	<0.001*
Comparisons for group	**S**		**SVI**	
	Z	**p**	**Z**	**p**
Ipsi vs Contra	-2.319	0.02038*	-	-

Midline area [180-220]ms

	Congruent stimuli			
Linear regression analysis				
	S		**SVI**	
	R^2	**p**	**R^2**	**p**
MCP - Accuracy*ERP mean	-	-	0.95	<0.001*
MFC - Accuracy*ERP mean	0.9209	<0.001*	-	-
Generalized linear mixed models	*accuracy~anteroposteriority*group* ERP mean+(1|subject)*			
	χ^2(df)		**P**	
anteroposteriority	0.3135(1)		0.5756	
group	0.012(1)		0.9128	
ERP mean	0.04542(1)		0.8312	
anteroposteriority*group	7.765(1)		0.005326	
anteroposteriority*ERP mean	2.373(1)		0.1234	
group*ERP mean	0.7755(1)		0.3785	
anteroposteriority*group*ERP mean	14.37(1)		0.0001498	
Post hoc GLMM				
Comparisons for anteroposteriority	**MCP**		**MFC**	
	Z	**p**	**Z**	**p**
S vs SVI	-2.927	0.003425*	2.604	0.009218*
Comparisons for group	**S**		**SVI**	
	Z	**p**	**Z**	**p**
MCP vs MFC	-2.244	0.02484	3.45	<0.001*

Table 5.3

Incongruent Condition

Activation ERP

Centro-Parietal area [105-120]ms

	Incongruent stimuli
Linear mixed models of the ERP mean	*ERP mean~laterality*group+(1|subject)*

	χ^2**(df)**		**P**	
laterality	116(1)		<0.001*	
group	7.14(1)		0.007536*	
laterality*group	0.05078(1)		0.8217	
Post hoc				
Comparisons for laterality	**Ipsi**		**Contra**	
	t(df)	P	t(df)	P
S vs SVI	-	-	-	-
Comparisons for group	**S**		**SVI**	
	t(df)	P	t(df)	P
Ipsi vs Contra	7.203(958)	<0.001*	6.982(958)	<0.001*

Temporo-posterior area [105-120]ms

	Incongruent stimuli			
Linear -mixed effects models of the ERP mean	*ERP mean~laterality*group+(1|subject)*			
	χ^2**(df)**		**P**	
laterality	0.9643(1)		0.3261	
group	0.4061(1)		0.524	
laterality*group	0.2936(1)		0.5879	
Post hoc				
Comparisons for laterality	**Ipsi**		**Contra**	
	t(df)	P	t(df)	P
S vs SVI	-	-	-	-
Comparisons for group	**S**		**SVI**	
	t(df)	P	t(df)	P
Ipsi vs Contra	-	-	-	-

Midline area [180-220]ms

	Incongruent stimuli			
Linear mixed-effect models of the ERP mean	*ERP mean~anteroposteriority*group+(1|subject)*			
	χ^2**(df)**		**P**	
anteroposteriority	8.033(1)		0.004593*	
group	5.00(1)		0.025*	
anteroposteriority*group	19.26(1)		<0.001*	
Post hoc				
Comparisons for anteroposteriority	**MCP**		**MFC**	
	t(df)	P	t(df)	P
S vs SVI	11.25(25)	<0.001*	16.75(25)	<0.001*

Comparisons for group	**S**		**SVI**	
	t(df)	P	t(df)	P
MCP vs MFC	-5.033(472)	<0.001*	-	-

Table 5.4

Association

Centro-Parietal area [105-120]ms

	Incongruent stimuli			
Linear regression analysis				
	S		**SVI**	
	R^2	**p**	**R^2**	**p**
Ipsi - Accuracy~ERP ean	-	-	-	-
Contra - Accuracy~ERP mean	0.9798	<0.001*		
Generalized linear mixed models	*accuracy~laterality*group*ERP mean+(1|subject)*			
	χ^2(df)		**P**	
laterality	1.103(1)		0.2936	
group	51.13(1)		<0.001*	
ERP mean	0.7239(1)		0.3949	
laterality*group	9.273(1)		0.002326*	
laterality*ERP mean	0.02171(1)		0.8829	
group*ERP mean	3.588(1)		0.0582	
laterality*group*ERP mean	5.229 (1)		0.0222 *	
Post hoc GLMM				
Comparisons for laterality	**Ips**		**Contra**	
	Z	**p**	**Z**	**p**
S vs SVI	-2.29	0.0220*	-	-
Comparisons for group	**S**		**SVI**	
	Z	**p**	**Z**	**p**
Ipsi vs Contra	-1.388	0.0226 *		

Temporo-Posterior area [105-120]ms

	Incongruent stimuli			
Linear regression analysis				
	S		**SVI**	
	R^2	**p**	**R^2**	**p**
Ipsi - Accuracy~ERP mean	-	-	-	-
Contra - Accuracy~ERP mean	0.8976	<0.001*		
Generalized linear mixed models	*accuracy~laterality*group*ERP mean+(1|subject)*			
	χ^2(df)		**P**	

laterality	0.1405(1)		0.7077	
group	64.29(1)		<0.001*	
ERP mean	2.014(1)		0.1558	
laterality*group	8.481(1)		<0.001*	
laterality*ERP mean	0.02804(1)		0.867	
group*ERP mean	19.96(1)		<0.001*	
laterality*group*ERP mean	3.133(1)		<0.001*	
Post hoc GLMM				
Comparisons for laterality	**Ipsi**		**Contra**	
	Z	**p**	**Z**	**p**
S vs SVI	-	-	5.599	<0.001*
Comparisons for group	**S**		**SVI**	
	Z	**p**	**Z**	**p**
Ipsi vs Contra	-5.597	<0.001*	-	-

Midline area [180-220]ms

	Incongruent stimuli			
Linear regression analysis				
	S		**SVI**	
	R^2	**p**	**R^2**	**p**
MCP - Accuracy~ERP mean	0.767	<0.001*	-	-
MFC - Accuracy~ERP mean	0.6738	0.001066*		-
Generalized linear mixed models	*accuracy~anteroposteriority*group* ERP mean+(1|subject)*			
	χ^2(df)		**P**	
anteroposteriority	0.003251(1)		0.9545	
group	4.161		0.04138*	
ERP mean	1.01(1)		0.315	
anteroposteriority*group	0.01296(1)		0.9094	
anteroposteriority*ERP mean	0.06077(1)		0.8053	
group*ERP mean	19.33(1)		<0.001*	
anteroposteriority*group*ERP mean	0.03039(1)		0.8616	
Post hoc GLMM				
Comparisons for anteroposteriority	**MCP**		**MFC**	
	Z	**p**	**Z**	**p**
S vs SVI	3.399	<0.001*	2.874	<0.001*
Comparisons for group	**S**		**SVI**	
	Z	**p**	**Z**	**p**

MCP vs MFC	-	-	-	-

Table 5.5

5.6 General Discussion

In this study, we investigated how audio-tactile information is combined during early childhood, and how this process is shaped by the presence or absence of visual experience. By combining behavioural measures with electrophysiological recordings, we were able to characterise not only whether multisensory gain is present, but how it operates under different spatial conditions. A key finding of this work is that multisensory processing is already present in both sighted (S) and severely visually impaired (SVI) children at this early developmental stage. Under congruent conditions, both groups showed clear evidence of audio-tactile gain, suggesting that basic multisensory interactions can emerge even in the absence of vision. This is consistent with previous behavioural findings (Gori et al., 2021) and supports the idea that early multisensory processing can rely on relatively low-level mechanisms that do not require fully calibrated spatial representations. However, the critical insight of this study emerges under conditions of spatial conflict. When auditory and tactile signals were incongruent, the two groups diverged markedly. Sighted children maintained stable behavioural performance and showed robust associations between neural activity and behaviour, indicating that they were able to flexibly resolve crossmodal discrepancies. In contrast, SVI children showed a reduction in performance together with weaker and less consistent brain–behaviour relationships, suggesting a different mode of multisensory processing under conflict.

Importantly, these findings do not indicate an absence of multisensory gain in SVI children. Rather, they suggest that integration operates differently when sensory signals are inconsistent. While sighted children appear to flexibly combine and reweight sensory information depending on the context, SVI children seem to rely more strongly on tactile input as a stable spatial reference. This strategy may be effective when signals are aligned, but becomes less adaptive when they conflict. This difference can be understood in light of the role of early visual experience in calibrating spatial representations (Burr & Gori, 2012; Ernst & Banks, 2002). In sighted children, vision likely supports the alignment of sensory reference frames and enables flexible multisensory processing across conditions. In contrast, in the

absence of vision, spatial representations may remain more closely tied to the body, leading to a greater reliance on somatosensory information and reduced flexibility in resolving crossmodal conflict (Röder et al., 2002; Hötting & Röder, 2009; Gori et al., 2014). Taken together, these findings suggest that multisensory gain is not simply present or absent in early development, but is critically shaped by how the brain deals with spatial conflict. Vision does not appear to be necessary for multisensory interactions to emerge, but plays a fundamental role in enabling flexible, context-dependent integration, allowing sensory information to be dynamically reweighted when signals are inconsistent. Although the present findings should be interpreted in light of the relatively limited sample size and broad developmental age range, the number of participants was informed by previous studies using comparable audio-tactile paradigms in sighted and severely visually impaired infants, where large behavioural effects were successfully detected with similar sample sizes. Nevertheless, future studies involving larger and more age-homogeneous samples will be important to further clarify the developmental mechanisms underlying multisensory spatial processing and crossmodal conflict resolution.

Chapter 6

General Discussion

The objective of this doctoral project was to investigate the development of spatial abilities, with a particular focus on how sensory systems and multisensory integration contribute to the emergence and organisation of spatial representations during early childhood, within a dynamic developmental framework. This work was conducted within the broader framework of the ERC-funded MySpace project, which aims to understand how sensory experience shapes spatial cognition across development in children with and without visual impairment. Through behavioural and neurophysiological measures, these skills were explored both in conditions of typical development and in conditions of sensory deprivation, involving participants with visual impairment. This approach has enabled a deeper understanding of the role of vision in the development of multisensory integration and spatial representations, as well as of how its absence influences behavioural performance and neural organization. The investigation of spatial abilities begins with the study of auditory spatial localisation presented in Chapter 2. By comparing preschool children and adults in tasks assessing accuracy and reaction times, these results suggested that auditory spatial representations are already functional in early childhood. However, localisation performance differs between azimuth and elevation, with a more pronounced discrepancy observed in children. This pattern suggests that the maturation of experience-dependent processes, together with body and behavioural development, contributes to the progressive refinement of auditory spatial localisation. Subsequently, multisensory audiovisual integration abilities were investigated in preschool children, with the aim of determining whether such processes provide a perceptual advantage in object segregation; in Chapter 3, a novel experimental device was introduced and validated, specifically designed to assess this ability in young children, while Chapter 4 presented the spatial object segregation study conducted using this device. The results did not reveal a clear multisensory advantage at the group level for this age range and task. Nevertheless, a subgroup of preschool children showed a tendency to combine auditory and visual cues, perceiving spatial discrepancies between objects at shorter distances under bimodal stimulation. These findings highlighted substantial inter-individual variability and heterogeneous sensory weighting strategies across participants, although the relatively limited sample size suggests that further studies with larger cohorts will be important to better characterise these developmental patterns.

Finally, Chapter 5 investigated audio-tactile multisensory integration and spatial conflict in sighted and visually impaired infants using both behavioural measures and EEG recordings. The findings showed that a multisensory gain for audio-tactile stimulation is already present early in development in both groups, but acts differently in spatial conflict. Also, differences emerged in neural organisation: whereas sighted infants recruited integrative cortical networks and maintained multisensory processing under spatial conflict, visually impaired infants relied more on tactile, body-centered representations and showed reduced engagement of integrative mechanisms when sensory cues were incongruent. These findings suggest that spatial abilities do not develop along a uniform or linear trajectory, but rather emerge through a dynamic and heterogeneous developmental process, in which different components evolve at different rates and are shaped by sensory experience, task demands, and inter-individual variability before progressively stabilising into adult-like patterns.

6.1 Spatial Abilities in Sound Localization

The first spatial skill investigated is auditory unisensory perception. The interpretation of sound cues from the environment occurs after the encoding of information. In fact, unlike vision, which encodes space directly through retinotopic mapping, hearing lacks an equivalent "focusing" mechanism and must reconstruct the position of sound sources from temporal, level, and spectral cues introduced by the interaction between sound, the head, and the outer ear. Auditory localisation is therefore a prime example of perceptual inference: the horizontal dimension is based mainly on binaural comparisons, interaural time differences (ITDs) and interaural level differences (ILDs), whereas elevation depends largely on monaural spectral cues linked to head-related transfer functions (HRTFs) (Middlebrooks & Green, 1991). Consequently, adults typically show greater accuracy for azimuthal than elevation judgments. In our study, children exhibited larger localization errors than adults across conditions, consistent with the protracted maturation of auditory spatial representations reported in previous developmental work (Morrongiello, 1987; Freigang et al., 2015; Otte et al., 2013). Despite these age-related differences in overall accuracy both children and adults showed a similar performance pattern across spatial dimensions, with significantly poorer accuracy in elevation than in azimuth. This consistent elevation disadvantage across age groups suggests that, although spatial precision improves with development, the underlying organisation of auditory space is broadly shared. This pattern aligns with behavioural

evidence showing that vertical localisation relies primarily on monaural spectral cues, which are inherently more ambiguous and variable than the binaural cues supporting horizontal localisation (Blauert, 1997; Makous & Middlebrooks, 1990; Freigang et al., 2015). The persistence of greater vertical errors in adults therefore suggests that reduced elevation precision reflects intrinsic constraints of the auditory system rather than immature sensory processing alone. From a developmental perspective, improvements across childhood appear to involve increased precision and reduced variability, while the relative difficulty associated with different spatial dimensions remains stable. Although developmental changes in head and pinna morphology alter the available spectral cues, experience-dependent sensorimotor calibration can partially refine their functional use over time (Keating & King, 2015). Nevertheless, elevation judgments remain less reliable than azimuth judgments even in adulthood (Otte et al., 2013; Carlini et al., 2024), reinforcing the interpretation of a modality-specific limitation rather than a purely developmental delay.

A key aspect is the pronounced contribution of errors on the vertical axis. Localisation inaccuracies were disproportionately determined by deviations along the Y dimension, indicating a systematic difficulty in estimating the elevation of the sound. It is important to note that this variability was not limited to elevation assessments alone, but also affected responses during azimuth localisation, where sounds were often mislocalised slightly above or below their actual horizontal position, while the error on the X-axis (more to the right or more to the left) remained minor. This uncertainty could similarly be justified by the nature of the physical properties of sound and how it is encoded. These results suggest that although auditory localisation could provide spatial information alone in an adequate manner already during childhood, and improves with experience, they raise the possibility that visual spatial cues may provide an important source of calibration and refinement. Indeed, previous research has shown that combining auditory and visual information enhances spatial accuracy in both adults and children, supporting the view that vision contributes to optimizing auditory spatial representations across development (e.g., Gori et al., 2008; Burr & Alais, 2006). Also, in our study, preliminary observations in children with severe visual impairment (SVI), although limited by the small sample size, revealed a similar pattern of results, with larger localisation errors in elevation than in azimuth and no differences in reaction times across conditions. This convergence with the pattern observed in sighted participants is consistent with previous evidence showing that basic auditory spatial abilities can be preserved in the absence of vision, particularly for

simpler spatial judgments (Röder et al., 1999; Voss et al., 2004). Although these findings should be interpreted with caution, they suggest that the reduced accuracy in vertical localisation may reflect intrinsic constraints of auditory processing that are preserved even in the absence of vision. In particular, the reliance on monaural spectral cues for elevation, which are inherently more ambiguous and variable, may represent a fundamental limitation of the auditory system that is not fully compensated by experience alone. At the same time, developmental studies indicate that when spatial tasks become more demanding, visually impaired children show reduced accuracy compared to sighted peers, highlighting the role of visual experience in calibrating auditory spatial representations (Cappagli & Gori, 2016; Cappagli et al., 2019). Within this framework, vision appears to support the refinement of spatial precision rather than the initial establishment of auditory spatial organisation. When examining reaction times, children responded more slowly than adults overall, consistent with the continued maturation of perceptual and sensorimotor processes (Kail, 1991; Eklöf et al., 2022). However, reaction times did not differ between azimuth and elevation within either group, indicating that the increased difficulty associated with vertical localisation primarily affected spatial accuracy rather than response speed. This dissociation suggests that both children and adults were able to detect and respond to sound sources with comparable temporal efficiency across spatial dimensions, even when spatial estimates, particularly along the vertical axis, were less precise.

Finally, these results respond to the broader question of whether auditory spatial localisation is fundamentally “autonomous” or shaped by visual experience. Although vision contributes to the calibration of spatial representations during development (Gori et al., 2008), the persistence of reduced elevation accuracy in adulthood indicates that visual input does not fully compensate for the intrinsic ambiguity of auditory spectral cues. The present findings therefore support accounts proposing modality-specific constraints in auditory spatial localisation, particularly for vertical space, with vision refining but not fundamentally altering the cue structure underlying auditory spatial coding (Blauert, 1997; Keating & King, 2015).

6.2 Audiovisual integration for object segregation

I investigated whether multisensory integration of auditory and visual cues supports young children's ability to spatially segregate two moving sensory sources. The hypothesis was that, if a multisensory benefit were present, children would perceive the two sources as distinct at a shorter separation distance in the bimodal condition. Behavioural comparisons across conditions did not reveal significant group-level differences. However, given the descriptive trends observed in the audiovisual condition, we further explored whether audiovisual performance showed patterns broadly compatible with multisensory cue combination by applying a Maximum Likelihood Estimation (MLE) framework focused on variance reduction. While predicted and observed performance did not differ significantly, Bayesian analyses provided only weak evidence in favour of the null hypothesis.

While this does not support the presence of statistically optimal integration at the group level, it opens a broader perspective on the developmental processes underlying multisensory abilities. Substantial inter-individual variability in the use of multisensory cues is expected during development, with children differing in the extent to which they rely on single cues or combine information across modalities (Gori et al., 2008; Nardini et al., 2008; Burr & Gori, 2012). Importantly, similar variability has also been observed in adulthood, even when group-level behaviour approximates optimal integration (Ernst & Banks, 2002; Zanchi et al., 2022). In light of this variability, participants were further examined according to the similarity of their sensory weights. Within the subgroup of children showing more balanced auditory and visual weighting (similar-weight group), performance patterns suggested a relative advantage for audiovisual processing compared to unimodal conditions. Although these effects were not statistically significant, their consistency may be compatible with the possibility of emerging multisensory benefits in spatial processing. Consistent with this interpretation, the MLE model showed that the observed audiovisual variance was relatively close to the variance predicted by the model, suggesting that these children may display response patterns broadly compatible with audiovisual cue combination. Similarly, comparison between the audiovisual condition and the best-performing unimodal cue (Best Cue) did not reveal significant differences, although the audiovisual condition showed a consistent numerical advantage. Taken together, these findings may suggest that some children could begin to show more stable audiovisual processing strategies, even in the absence of fully optimal integration. These patterns may reflect emerging integrative tendencies that are not yet fully stabilised but could nonetheless contribute to perceptual performance in dynamic spatial tasks. In contrast,

within the non-similar-weight group, the evidence does not support the presence of multisensory integration. No statistically significant differences emerged across conditions in terms of either accuracy or precision, but rather a tendency to rely more strongly on visual stimuli. This interpretation is reinforced by the MLE model analysis: the observed variance in the audiovisual condition significantly differed from the variance predicted by the model, indicating a deviation from optimal integration. However, the associated Bayes Factor provided only weak evidence in favour of the null hypothesis, suggesting that this deviation should be interpreted with caution. Although the observed audiovisual variance was higher than the MLE prediction, substantial inter-individual variability may have diluted the statistical signal, limiting the conclusiveness of the Bayesian evidence. Consistently, the comparison between the audiovisual condition and the best-performing unimodal cue (Best Cue) did not reveal significant differences, while the audiovisual variance remained numerically higher than that of the Best Cue. This pattern further supports the absence of a multisensory benefit in this group, suggesting that, unlike children with more balanced weighting, these participants may not yet integrate sensory cues in a way that enhances spatial precision. In the absence of balanced sensory weighting, children are known to rely on the modality that provides greater spatial reliability, typically vision (Gori et al., 2008; Nardini et al., 2008; Gori, 2015), due to its higher spatial resolution (Bruns & Röder, 2023; Alais & Burr, 2004). Consistent with this, we observed a clear preference for visual input within this subgroup. This finding is not surprising, as it aligns with previous evidence showing that, when multisensory integration is not yet established, children tend to rely on a single, task-dominant modality rather than combining multiple sensory cues (Burr & Gori, 2012). Beyond their developmental relevance, these individual differences have also been observed in adulthood across different types of multisensory tasks (Bentvelzen et al., 2009; Zanchi et al., 2022). For instance, in spatial navigation tasks based on audiovisual landmarks, adults show heterogeneous integration strategies, with only individuals assigning similar weights to auditory and visual cues benefiting from multisensory information in terms of increased precision (Zanchi et al., 2022). Similarly, in dynamic perceptual tasks such as

audiovisual speed discrimination, integration strategies depend on the relative reliability of sensory inputs: when cue weights are balanced, performance approximates optimal integration, whereas large discrepancies between cue reliabilities lead to alternative strategies, such as probability summation rather than true integration (Bentvelzen et al., 2009). These findings suggest that variability in multisensory processing is not limited to development but reflects a general property of the perceptual system, whereby integration depends on task demands and the relative weighting of available sensory cues. From a developmental perspective, this variability likely originates early in life, as children show heterogeneous and task-dependent trajectories in the emergence of multisensory integration (Burr & Gori, 2012; Gori, 2015; Negen et al., 2019). However, the present study focused on a single preschool age range, which limits the possibility of directly characterising developmental changes across different stages of childhood. Although the paradigm was specifically designed for young children and proved suitable for investigating multisensory spatial processing at this developmental stage, future studies involving broader age ranges and larger cohorts will be important to better clarify how audiovisual segregation abilities evolve across development.

In this context, some children may begin to combine sensory cues, whereas others rely predominantly on a single modality and may not yet integrate sensory information to reduce uncertainty (Gori et al., 2008; Nardini et al., 2010), possibly reflecting a developmental stage in which perceptual systems are not yet optimised for uncertainty reduction. Overall, these findings do not indicate a complete absence of multisensory processing, but rather suggest that audiovisual cue combination at this developmental stage may not yet be stable or consistently expressed across individuals. While some preschool children showed response patterns potentially compatible with emerging multisensory benefits, these effects were not consistently observed at the group level and appeared to depend strongly on individual sensory weighting strategies. Importantly, the relatively limited sample size, particularly following subgroup subdivision, warrants caution in interpreting these patterns. Rather than reflecting a stable or uniform process, multisensory processing in this task appears to be heterogeneous and task-dependent, with children differing in the extent to which they rely on single sensory cues or combine information across modalities.

6.3 Audio-tactile multisensory development in sighted and visually impaired infants

A substantial body of research has highlighted the central role of vision in guiding spatial

representations across other sensory modalities, including audition and touch (Cappagli et al., 2015; Thinus-Blanc & Gaunet, 1997; Vercillo et al., 2016). The development of sensory and spatial abilities in the absence of vision represents a clear example of a non-linear developmental process, in which typical trajectories are reorganised rather than simply delayed. Indeed, visual deprivation has been shown to affect the development of spatial abilities (Cattaneo et al., 2008; Ungar et al., 1995), providing a valuable framework for understanding how and to what extent sensory experience shapes the emergence of spatial representations.

The last study descripted in this thesis, I investigated the emergence of audio-tactile integration in early childhood by comparing sighted (S) and severe visually impaired (SVI) infants in a spatial localisation task, combining behavioural orienting responses with electrophysiological measures. Building on previous behavioural findings (Gori et al., 2021), the results provide converging evidence that multisensory processing is already present in the first years of life, while critically revealing that its underlying organisation is strongly shaped by early visual experience. At a general level, both groups exhibited clear markers of multisensory processing. Behavioural findings showed that in the bimodal congruent condition, infants increased accuracy and presented faster orienting responses, while electrophysiological data revealed enhanced neural activity within centro-parietal and fronto-central regions in temporal windows classically associated with early multisensory interactions (Hötting & Röder, 2009; Brandwein et al., 2011; Murray et al., 2005; Foxe & Molholm, 2009). These findings indicate that basic multisensory coactivation mechanisms can emerge even in the absence of visual experience. However, despite this shared capacity, the results revealed a difference in how multisensory processing is organised across groups. Crucially, these differences emerged at early stages of processing, suggesting that visual experience influences multisensory development from the initial phases of sensory encoding. Sighted infants showed stronger engagement of temporo-posterior regions associated with auditory processing, whereas SVI exhibited enhanced activity in centro-parietal regions linked to somatosensory processing. This early dissociation suggests that differences from its earliest stages shape multisensory integration. Within this framework, vision appears to support the calibration role in organising multisensory development. In sighted infants, early visual experience supports the calibration of auditory spatial representations and promotes the alignment of sensory inputs within a shared external reference frame (Gori et al., 2012; Cappagli et al., 2017; Zwiers et al., 2003). In contrast, in the absence of such calibration, auditory spatial representations remain less precise, leading to a shift toward tactile-dominant

spatial coding (Röder et al., 1999; Gori et al., 2014; Vercillo & Gori, 2015). This interpretation is consistent with evidence showing reduced crossmodal alignment and increased reliance on tactile information in visually deprived individuals (Occelli et al., 2012; Occelli et al., 2013). These differences were also reflected in distinct patterns of neural organisation within the multisensory processing time window (180–220 ms). While sighted infants showed greater recruitment of fronto-central regions, SVI infants exhibited a more posterior distribution of activity centred over centro-parietal areas, suggesting partially different neural dynamics underlying multisensory processing. This pattern is broadly consistent with previous evidence from congenitally blind individuals showing enhanced recruitment of occipito-parietal regions and crossmodal plasticity following early visual deprivation (Hötting & Röder, 2009; Fiehler & Rösler, 2010; Zhang et al., 2019; Collignon et al., 2015). Additional studies suggest that reduced or absent visual input during development may contribute to long-lasting reorganisation of cortical networks and altered recruitment of posterior regions during non-visual processing (Collignon et al., 2015; Duymuş et al., 2024). However, these interpretations should be considered cautiously, as the present cohort consisted of severely visually impaired (SVI) infants rather than totally blind participants, and the degree of visual experience likely varied across individuals.

Importantly, these neural differences were accompanied by distinct behavioural tendencies. Sighted infants appeared to maintain multisensory representations even when sensory inputs were less consistent, potentially reflecting the early emergence of more externally oriented spatial coding mechanisms (Bremner et al., 2008). In contrast, SVI infants showed a stronger tendency to rely on tactile input, which may indicate a greater reliance on body-centred spatial representations in situations where one modality provides more stable spatial information (Cattaneo et al., 2008; Occelli et al., 2013; Thinus-Blanc & Gaunet, 1997). This interpretation is broadly compatible with previous findings reporting enhanced somatosensory processing and reduced influence of external spatial coordinates following early visual deprivation (Röder et al., 2002; Röder et al., 2004; Hötting & Röder, 2009), although further work will be needed to clarify how these mechanisms operate across different levels of visual impairment. Another key difference between the two groups is the distinct patterns that emerged under conditions of spatial conflict, when auditory and tactile stimuli conveyed incongruent spatial information. In sighted infants, behavioural performance remained relatively stable and neural responses continued to involve multisensory processing regions, suggesting that multisensory representations were maintained even when sensory inputs were spatially inconsistent (Bremner et al., 2008). In contrast, SVI infants showed a stronger

tendency to rely on tactile information under conflict conditions, together with reduced evidence of sustained multisensory involvement. This pattern may reflect a greater reliance on body-centred spatial representations in situations where tactile input provides the most stable spatial reference (Cattaneo et al., 2008; Occelli et al., 2013; Thinus-Blanc & Gaunet, 1997). Such an interpretation is broadly compatible with previous findings reporting enhanced somatosensory processing and reduced influence of external spatial coordinates following early visual deprivation (Röder et al., 2002; Röder et al., 2004; Hötting & Röder, 2009), although further work will be needed to clarify how these mechanisms operate across different levels of visual impairment. Taken together, these findings indicate that while multisensory gain emerges early in both groups, early visual experience plays a fundamental role in shaping how sensory signals are weighted, aligned, and integrated. Similarly, multisensory development therefore does not follow a uniform trajectory but unfolds as a dynamic and experience-dependent process, characterised by different developmental pathways depending on the availability and reliability of sensory inputs.

Concluding Remarks

During my Ph.D., I investigated how spatial abilities develop across early life, focusing on the contribution of sensory systems and multisensory integration in both typical and atypical development. To address this question, I designed a series of behavioural and neurophysiological experiments targeting different aspects of spatial perception, including auditory localisation, audiovisual spatial segregation of moving stimuli, and audio-tactile integration in infancy.

The integration of multiple methodological approaches, ranging from unisensory to multisensory paradigms and from behavioural measures to EEG recordings, allowed for the characterisation of spatial representations across different sensory modalities, task demands, and developmental stages. This framework enabled the investigation of both relatively simple and more computationally demanding spatial processes, while also allowing the direct examination of the role of visual experience through the inclusion of children with severe visual impairment. The findings illustrated in this thesis allow for the following conlcusions:

1. Spatial perception emerges early and presents a relatively stable organisational structure across development, although its precision appears to be progressively refined over time. Auditory spatial localisation is already functional in early childhood, with broadly similar patterns of performance across age groups, but with higher variability and reduced precision in younger children. Persistent limitations in vertical localisation are consistent with the presence of modality-specific constraints that may not be fully overcome during development.

2. The findings of this thesis suggest that multisensory integration may not emerge as a unitary mechanism uniformly enhancing spatial performance across different tasks and developmental stages. While clear multisensory benefits were observed in audio-tactile spatial localisation during infancy, suggesting that integration mechanisms can operate early when sensory signals directly support spatial estimation, audiovisual performance in the segregation of dynamically moving stimuli during preschool age did not show consistent facilitation at the group level. Importantly, the multisensory paradigms investigated in this thesis capture distinct phases of developmental organisation, thereby providing complementary insight into the temporal unfolding of multisensory spatial processing across early life. This dissociation suggests that multisensory processing during development is shaped not only by task difficulty, but also by the nature of the spatial computations required. In particular, dynamic object segregation may rely on more

complex processes, such as the continuous evaluation of sensory correspondence over time and the inference of multiple causal sources, which may follow a more protracted developmental trajectory than simpler forms of cue combination. Consequently, multisensory behaviour in early childhood appears highly heterogeneous and reflects substantial inter-individual variability in sensory weighting strategies and developmental trajectories.

3. 2The findings are consistent with the idea that visual experience contributes substantially to the calibration and refinement of audio-tactile spatial representations during development, rather than enabling the emergence of spatial abilities per se. In the absence of vision, spatial processing may rely on partially different organisational strategies, characterised by a greater reliance on body-centred representations and modality-specific processing. More broadly, these findings are consistent with the hypothesis that vision contributes to the alignment of sensory information within common external spatial reference frames and may facilitate flexible multisensory processing under conditions of uncertainty or conflict.

4. 4Future research should further investigate how spatial abilities evolve across childhood, with particular attention to how multisensory integration mechanisms interact with perceptual learning, environmental experience, and task complexity. It will be especially important to characterise how different sensory modalities contribute to spatial representations under varying ecological and computational demands, as well as how individual variability shapes the developmental emergence of multisensory competence.

Taken together, the work presented in this thesis provides both methodological and empirical contributions to the understanding of spatial development within a multisensory framework. In particular, this research introduces a novel experimental approach for investigating spatial behaviour using dynamic and ecologically relevant multisensory stimulation in young children. This methodological perspective allows the study of spatial processing under conditions that more closely approximate real-world sensory environments while preserving experimental control. Furthermore, the findings contribute to a more refined account of the developmental organisation of spatial abilities, suggesting that aspects of spatial maturation may follow partially dimension-specific trajectories and that multisensory processing is characterised by substantial inter-individual variability and heterogeneous patterns of organisation. By jointly examining unisensory and multisensory mechanisms across different developmental stages and clinical populations, this work advances a more nuanced

understanding of how spatial representations emerge through the interaction between sensory experience, perceptual learning, and task demands.

More broadly, the results of this thesis provide a conceptual and methodological framework for future investigations aimed at characterising the development of multisensory spatial processing across diverse contexts and populations. These findings highlight the importance of conceptualising multisensory development as a dynamic and inherently heterogeneous process, shaped by adaptive reorganisation and individual variability rather than reflecting a fully uniform progression toward adult-like performance.

Bibliography

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. Current Biology, 14(3), 257–262.

Alais, D., Newell, F. N., & Mamassian, P. (2010). Multisensory processing in review: from physiology to behaviour. Seeing and perceiving, 23(1), 3.

Azañón, E., Camacho, K., Morales, M., & Longo, M. R. (2018). The sensitive period for tactile remapping does not include early infancy. Child development, 89(4), 1394-1404.

Baart, M., Bortfeld, H., & Vroomen, J. (2015). Phonetic matching of auditory and visual speech develops during childhood: Evidence from sine-wave speech. Journal of Experimental Child Psychology, 129, 157–164.

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. Cognition, 20(3), 191-208.

Barutchu, A., Crewther, D. P., & Crewther, S. G. (2009). The race that precedes coactivation: Development of multisensory facilitation in children. Developmental Science, 12(3), 464–473. https://doi.org/10.1111/j.1467-7687.2008.00782.x

Beatch, J. A. (2020). Audiovisual integration in 4-and 5-year-old children: Behavioural and neural responses. https://ucalgary.scholaris.ca/items/44b10cfb-23a3-4007-a315-af56cdada596

Begum Ali, J., Cowie, D., & Bremner, A. J. (2014). Effects of posture on tactile localization by 4 years of age are modulated by sight of the hands: Evidence for an early acquired external spatial frame of reference for touch. Developmental Science, 17(6), 935–943. https://doi.org/10.1111/desc.12184

Bentvelzen, A., Leung, J., & Alais, D. (2009). Discriminating Audiovisual Speed: Optimal Integration of Speed Defaults to Probability Summation When Component Reliabilities Diverge. Perception, 38(7), 966–987. https://doi.org/10.1068/p6261

Bentvelzen, A., Leung, J., & Alais, D. (2009). Discriminating audiovisual speed: optimal integration of speed defaults to probability summation when component reliabilities diverge. Perception, 38(7), 966-987.

Bigelow, A. E. (1986). The development of reaching in blind children. British Journal of Developmental Psychology, 4(4), 355–366. https://doi.org/10.1111/j.2044-835X.1986.tb01031.x

Bizley, J. K., & Cohen, Y. E. (2013). The what, where and how of auditory-object perception. Nature Reviews Neuroscience, 14(10), 693–707.

Blauert, J., & Butler, R. A. (1985). Spatial hearing: The psychophysics of human sound localization by Jens Blauert. Acoustical Society of America. https://pubs.aip.org/asa/jasa/article-abstract/77/1/334/654073

Brandwein, A. B., Foxe, J. J., Russo, N. N., Altschuler, T. S., Gomes, H., & Molholm, S. (2011). The development of audiovisual multisensory integration across childhood and early adolescence: A high-density electrical mapping study. Cerebral Cortex, 21(5), 1042–1055.

Bremner, A. J., & Spence, C. (2017). The development of tactile perception. Advances in child development and behavior, 52, 227-268.

Bremner, A. J., Hill, E. L., Pratt, M., Rigato, S., & Spence, C. (2013). Bodily illusions in young children: developmental change in visual and proprioceptive contributions to perceived hand position. PLoS One, 8(1), e51887.

Bremner, A. J., Holmes, N. P., & Spence, C. (2008). Infants lost in (peripersonal) space?. Trends in cognitive sciences, 12(8), 298-305.

Bremner, A. J., Mareschal, D., Lloyd-Fox, S., & Spence, C. (2008). Spatial localization of touch in the first year of life: early influence of a visual spatial code and the development of remapping across changes in limb position. Journal of Experimental Psychology: General, 137(1), 149.

Brett-Green, B. A., Miller, L. J., Gavin, W. J., & Davies, P. L. (2008). Multisensory integration in children: a preliminary ERP study. Brain Research, 1242, 283-290.

Bruns, P., & Röder, B. (2023). Development and experience-dependence of multisensory spatial processing. Trends in Cognitive Sciences, 27(10), 961-973.

Burr, D., & Gori, M. (2012). Multisensory integration develops late in humans. https://europepmc.org/article/nbk/nbk92864

Cainelli, E., Stramucci, G., & Bisiacchi, P. (2025). A light in the darkness: Early phases of development and the emergence of cognition. Developmental cognitive neuroscience, 72, 101527.

Calafatello, G., Zanchi, S., Balzarotti, N., Gori, M., Brmener, A. J., Parmiggiani, A., & Orciari, L. (2025). Testing Perceptual Development with Moving Stimuli: A New Multisensory Tool for Young Children. 2025 IEEE Medical Measurements & Applications (MeMeA), 1–6. https://ieeexplore.ieee.org/abstract/document/11066292/

Calvert, G. A., & Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. Journal of Physiology-Paris, 98(1-3), 191-205.

Cappagli, G., & Gori, M. (2016). Auditory spatial localization: Developmental delay in children with visual impairments. Research in Developmental Disabilities, 53, 391–398.

Cappagli, G., Cocchi, E., & Gori, M. (2017a). Auditory and proprioceptive spatial impairments in blind children and adults. Developmental Science, 20(3), e12374. https://doi.org/10.1111/desc.12374

Cappagli, G., Cocchi, E., & Gori, M. (2017b). Auditory and proprioceptive spatial impairments in blind children and adults. Developmental Science, 20(3), e12374. https://doi.org/10.1111/desc.12374

Carlini, A., Bordeau, C., & Ambard, M. (2024). Auditory localization: a comprehensive practical review. Frontiers in Psychology, 15, 1408073.

Carriere, B. N., Royal, D. W., Perrault, T. J., Morrison, S. P., Vaughan, J. W., Stein, B. E., & Wallace, M. T. (2007). Visual Deprivation Alters the Development of Cortical Multisensory Integration. Journal of Neurophysiology, 98(5), 2858–2867. https://doi.org/10.1152/jn.00587.2007

Champoux, F., Collignon, O., Bacon, B. A., Lepore, F., Zatorre, R. J., & Théoret, H. (2011). Early- and Late-Onset Blindness Both Curb Audiotactile Integration on the Parchment-Skin Illusion. Psychological Science, 22(1), 19–25. https://doi.org/10.1177/0956797610391099

Chanauria, N., Bharmauria, V., Bachatene, L., Cattan, S., Rouat, J., & Molotchnikoff, S. (2019). Sound induces change in orientation preference of V1 neurons: audio-visual cross-influence. Neuroscience, 404, 48-61.

Chang, C. Y., Hsu, S. H., Pion-Tonachini, L., & Jung, T. P. (2018, July). Evaluation of artifact subspace reconstruction for automatic EEG artifact removal. In 2018 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC) (pp. 1242-1245). IEEE.

Chen, Y. C., & Spence, C. (2017). Assessing the role of the 'unity assumption'on multisensory integration: A review. Frontiers in psychology, 8, 445.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. Psychological bulletin, 70(4), 213.

Collignon, O., Dormal, G., De Heering, A., Lepore, F., Lewis, T. L., & Maurer, D. (2015). Long-lasting crossmodal cortical reorganization triggered by brief postnatal visual deprivation. Current Biology, 25(18), 2379-2383.

Collignon, O., Vandewalle, G., Voss, P., Albouy, G., Charbonneau, G., Lassonde, M., & Lepore, F. (2011). Functional specialization for auditory–spatial processing in the occipital cortex of congenitally blind humans. Proceedings of the National Academy of Sciences, 108(11), 4435–4440. https://doi.org/10.1073/pnas.1013928108

De Ribaupierre, A., & Lecerf, T. (2018). On the importance of intraindividual variability in cognitive development. Journal of Intelligence, 6(2), 17.

De Ribaupierre, A., & Lecerf, T. (2018). On the importance of intraindividual variability in cognitive development. Journal of Intelligence, 6(2), 17.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of neuroscience methods, 134(1), 9-21.

Denervaud, S., Gentaz, E., Matusz, P. J., & Murray, M. M. (2020). Multisensory gains in simple detection predict global cognition in schoolchildren. Scientific reports, 10(1), 1394.

Dessalegn, B., & Landau, B. (2008). More Than Meets the Eye: The Role of Language in Binding and Maintaining Feature Conjunctions. Psychological Science, 19(2), 189–195. https://doi.org/10.1111/j.1467-9280.2008.02066.x

Dessalegn, B., & Landau, B. (2008). Vision and language: Recoding of visual representations. Journal of Vision, 8(6), 419-419.

Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on 'sensory-specific'brain regions, neural responses, and judgments. Neuron, 57(1), 11–23.

Duymuş, H., Verma, M., Güçlütürk, Y., Öztürk, M., Varol, A. B., Kurt, Ş., ... & Farooqui, A. A. (2024). The visual cortex in the blind but not the auditory cortex in the deaf becomes multiple-demand regions. Brain, 147(10), 3624-3637.

Dwyer, P., Takarae, Y., Zadeh, I., Rivera, S. M., & Saron, C. D. (2022). Multisensory integration and interactions across vision, hearing, and somatosensation in autism spectrum development and typical development. Neuropsychologia, 175, 108340.

Eklöf, M., Asp, F., & Berninger, E. (2022). The Development of Sound Localization Latency in Infants and Young Children with Normal Hearing. Trends in Hearing, 26, 23312165221088398. https://doi.org/10.1177/23312165221088398

Ellis, B. J., Sheridan, M. A., Belsky, J., & McLaughlin, K. A. (2022). Why and how does early adversity influence development? Toward an integrated model of dimensions of environmental experience. Development and psychopathology, 34(2), 447-471.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. Nature, 415(6870), 429–433.

Fairweather-Schmidt, A. K., Lee, C., & Wade, T. D. (2015). A longitudinal study of midage women with indicators of disordered eating. Developmental Psychology, 51(5), 722.

Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. Journal of Neuroscience, 22(13), 5749-5759.

Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. Journal of Neuroscience, 22(13), 5749-5759.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. Cerebral cortex (New York, NY: 1991), 1(1), 1-47.

Fiehler, K., & Rösler, F. (2010). Plasticity of multisensory dorsal stream functions: Evidence from congenitally blind and sighted adults. Restorative Neurology and Neuroscience, 28(2), 193–205. https://doi.org/10.3233/RNN-2010-0500

Fiehler, K., & Rösler, F. (2010). Plasticity of multisensory dorsal stream functions: evidence from congenitally blind and sighted adults. Restorative Neurology and Neuroscience, 28(2), 193-205.

Flanagan, J. R., & Beltzner, M. A. (2000). Independence of perceptual and sensorimotor predictions in the size–weight illusion. Nature neuroscience, 3(7), 737-741.

Fort, A., Delpuech, C., Pernier, J., & Giard, M. H. (2002). Early auditory–visual interactions in human cortex during nonredundant target identification. Cognitive brain research, 14(1), 20-30.

Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. Frontiers in psychology, 2, 154.

Gandhi, T. K., Ganesh, S., & Sinha, P. (2014). Improvement in Spatial Imagery Following Sight Onset Late in Childhood. Psychological Science, 25(3), 693–701. https://doi.org/10.1177/0956797613513906

Getzmann, S., & Lewald, J. (2007). Localization of moving sound. Perception & Psychophysics, 69(6), 1022–1034. https://doi.org/10.3758/BF03193940

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory?. Trends in cognitive sciences, 10(6), 278-285.

Gobbelé, R., Schürmann, M., Forss, N., Juottonen, K., Buchner, H., & Hari, R. (2003). Activation of the human posterior parietal and temporoparietal cortices during audiotactile interaction. Neuroimage, 20(1), 503-511.

Gori, M. (2015). Multisensory integration and calibration in children and adults with and without sensory and motor disabilities. Multisensory Research, 28(1–2), 71–99.

Gori, M., Amadeo, M. B., & Bremner, A. J. (2026). The development of spatial perception with and without visual experience. Nature Reviews Psychology, 1-13.

Gori, M., Campus, C., Signorini, S., Rivara, E., & Bremner, A. J. (2021). Multisensory spatial perception in visually impaired infants. Current Biology, 31(22), 5093-5101.

Gori, M., Del Viva, M., Sandini, G., & Burr, D. C. (2008). Young children do not integrate visual and haptic form information. Current Biology, 18(9), 694–698.

Gori, M., Vitali, H., Bremner, A. J., Tonelli, A., Amadeo, M. B., Setti, W., ... & Campus, C. (2025). Vision drives the neural construction of a two-stage hierarchy of spatial processing in infancy. Iscience, 28(11).

Gouteux, S., & Spelke, E. S. (2001). Children's use of geometry and landmarks to reorient in an open space. Cognition, 81(2), 119-148.

Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. Nature reviews neuroscience, 7(7), 523-534.

Heed, T., & Azañón, E. (2014). Using time to investigate space: a review of tactile temporal order judgments as a window onto spatial processing in touch. Frontiers in psychology, 5, 76.

Hermer, L., & Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. Nature, 370(6484), 57-59.

Hillock-Dunn, A., & Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. Developmental Science, 15(5), 688–696. https://doi.org/10.1111/j.1467-7687.2012.01171.x

Hofman, P. M., & Van Opstal, A. J. (1998). Spectro-temporal factors in two-dimensional human sound localization. The Journal of the Acoustical Society of America, 103(5), 2634-2648.

Hötting, K., & Röder, B. (2004). Hearing cheats touch, but less in congenitally blind than in sighted individuals. Psychological science, 15(1), 60-64.

Hötting, K., & Röder, B. (2009). Auditory and auditory-tactile processing in congenitally blind humans. Hearing research, 258(1-2), 165-174.

Hötting, K., Friedrich, C. K., & Röder, B. (2009). Neural correlates of cross-modally induced changes in tactile awareness. Journal of Cognitive Neuroscience, 21(12), 2445-2461.

Jiang, W., Jiang, H., & Stein, B. E. (2006). Neonatal cortical ablation disrupts multisensory development in superior colliculus. Journal of neurophysiology, 95(3), 1380-1396.

Jones, E. G., & Powell, T. P. S. (1970). An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. Brain, 93(4), 793-820.

Kail, R. (1991). Development of processing speed in childhood and adolescence. Advances in Child Development and Behavior, 23, 151–185.

Kayser, C., Petkov, C. I., & Logothetis, N. K. (2009). Multisensory interactions in primate auditory cortex: fMRI and electrophysiology. Hearing research, 258(1-2), 80-88.

Keating, P., & King, A. J. (2015). Sound localization in a changing world. Current Opinion in Neurobiology, 35, 35–43.

Keil, J., & Senkowski, D. (2018). Neural Oscillations Orchestrate Multisensory Processing. The Neuroscientist, 24(6), 609–626. https://doi.org/10.1177/1073858418755352

King, A. J., & Carlile, S. (1993). Changes induced in the representation of auditory space in the superior colliculus by rearing ferrets with binocular eyelid suture. Experimental Brain Research, 94(3). https://doi.org/10.1007/BF00230202

Knudsen, E. I. (1998). Capacity for Plasticity in the Adult Owl Auditory System Expanded by Juvenile Experience. Science, 279(5356), 1531–1533. https://doi.org/10.1126/science.279.5356.1531

Knudsen, E. I., & Brainard, M. S. (1991). Visual Instruction of the Neural Map of Auditory Space in the Developing Optic Tectum. Science, 253(5015), 85–87. https://doi.org/10.1126/science.2063209

Knudsen, E. I., & Brainard, M. S. (1991). Visual instruction of the neural map of auditory space in the developing optic tectum. Science, 253(5015), 85-87.

Knudsen, E. I., & Knudsen, P. F. (1985). Vision Guides the Adjustment of Auditory Localization in Young Barn Owls. Science, 230(4725), 545–548. https://doi.org/10.1126/science.4048948

Kolarik, A. J., Cirstea, S., Pardhan, S., & Moore, B. C. (2014). A summary of research investigating echolocation abilities of blind and sighted humans. Hearing research, 310, 60-68.

Koustriava, E., & Papadopoulos, K. (2010). Mental Rotation Ability of Individuals with Visual Impairments. Journal of Visual Impairment & Blindness, 104(9), 570–575. https://doi.org/10.1177/0145482X1010400910

Koustriava, E., & Papadopoulos, K. (2010). Mental rotation ability of individuals with visual impairments. Journal of Visual Impairment & Blindness, 104(9), 570-575.

Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. Cognition, 80(1–2), 97–126.

Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. Cognition, 80(1-2), 97-126.

Kuehn, E., & Pleger, B. (2018). How Visual Body Perception Influences Somatosensory Plasticity. Neural Plasticity, 2018, 1–12. https://doi.org/10.1155/2018/7909684

Kuehn, E., & Pleger, B. (2018). How visual body perception influences somatosensory plasticity. Neural plasticity, 2018(1), 7909684.

Kumpik, D. P., Kacelnik, O., & King, A. J. (2010). Adaptive reweighting of auditory localization cues in response to chronic unilateral earplugging in humans. Journal of Neuroscience, 30(14), 4883-4894.

Lewald, J. (2002). Vertical sound localization in blind humans. Neuropsychologia, 40(12), 1868–1872.

Lewald, J. (2002). Vertical sound localization in blind humans. Neuropsychologia, 40(12), 1868-1872.

Lewkowicz, D. J. (2012). Development of multisensory temporal perception. https://europepmc.org/article/nbk/nbk92854

Lewkowicz, D. J. (2014). Early experience and multisensory perceptual narrowing. Developmental Psychobiology, 56(2), 292–315. https://doi.org/10.1002/dev.21197

Lewkowicz, D. J., & Flom, R. (2014). The audiovisual temporal binding window narrows in early childhood. Child development, 85(2), 685-694.

Lewkowicz, D. J., & Flom, R. (2014a). The audiovisual temporal binding window narrows in early childhood. Child Development, 85(2), 685–694.

Lewkowicz, D. J., & Flom, R. (2014b). The audiovisual temporal binding window narrows in early childhood. Child Development, 85(2), 685–694.

Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. Trends in cognitive sciences, 13(11), 470-478.

Lewkowicz, D. J., & Ghazanfar, A. A. (2009a). The emergence of multisensory systems through perceptual narrowing. Trends in Cognitive Sciences, 13(11), 470–478.

Lewkowicz, D. J., & Ghazanfar, A. A. (2009b). The emergence of multisensory systems through perceptual narrowing. Trends in Cognitive Sciences, 13(11), 470–478.

Loomis, J. M., Klatzky, R. L., Philbeck, J. W., & Golledge, R. G. (1998). Assessing auditory distance perception using perceptually directed action. Perception & Psychophysics, 60(6), 966–980. https://doi.org/10.3758/BF03211932

López-García, D., Peñalver, J. M., Górriz, J. M., & Ruz, M. (2022). MVPAlab: A machine learning decoding toolbox for multidimensional electroencephalography data. Computer Methods and Programs in Biomedicine, 214, 106549.

Luck, S. J. (2014). An introduction to the event-related potential technique. MIT press.

Maitre, N. L., Key, A. P., Slaughter, J. C., Yoder, P. J., Neel, M. L., Richard, C., Wallace, M. T., & Murray, M. M. (2020). Neonatal Multisensory Processing in Preterm and Term Infants Predicts Sensory Reactivity and Internalizing Tendencies in Early Childhood. Brain Topography, 33(5), 586–599. https://doi.org/10.1007/s10548-020-00791-4

Makous, J. C., & Middlebrooks, J. C. (1990). Two-dimensional sound localization by human listeners. The Journal of the Acoustical Society of America, 87(5), 2188–2200.

Martin, K., Johnstone, P., & Hedrick, M. (2015). Auditory and visual localization accuracy in young children and adults. International Journal of Pediatric Otorhinolaryngology, 79(6), 844-851.

Molholm, S., Murphy, J. W., Bates, J., Ridgway, E. M., & Foxe, J. J. (2020). Multisensory audiovisual processing in children with a sensory processing disorder (I): behavioral and electrophysiological indices under speeded response conditions. Frontiers in integrative Neuroscience, 14, 4.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. Cognitive brain research, 14(1), 115-128.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. Cognitive brain research, 14(1), 115-128.

Morey, R. D., & Rouder, J. N. (2018). BayesFactor: Computation of Bayes factors for common designs.

Morey, R. D., & Rouder, J. N. (2018). BayseFactor: computation of bayes factors for common designs.

Morrongiello, B. A., & Rocca, P. T. (1987). Infants' localization of sounds in the horizontal plane: Effects of auditory and visual cues. Child Development, 918–927.

Muir, D., & Field, J. (1979). Newborn infants orient to sounds. Child development, 431-436.

Muir, D., & Field, J. (1979). Newborn infants orient to sounds. Child development, 431-436.

Muir, D., Engvall, E., Varon, S., & Manthorpe, M. (1989). Schwannoma cell-derived inhibitor of the neurite-promoting activity of laminin. The Journal of cell biology, 109(5), 2353-2362.

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP Analyses: A Step-by-Step Tutorial Review. Brain Topography, 20(4), 249–264. https://doi.org/10.1007/s10548-008-0054-5

Murray, M. M., Lewkowicz, D. J., Amedi, A., & Wallace, M. T. (2016). Multisensory processes: a balancing act across the lifespan. Trends in neurosciences, 39(8), 567-579.

Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C., ... & Foxe, J. J. (2005). Grabbing your ear: rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. Cerebral cortex, 15(7), 963-974.

Murray, M. M., Thelen, A., Thut, G., Romei, V., Martuzzi, R., & Matusz, P. J. (2016). The multisensory function of the human primary visual cortex. Neuropsychologia, 83, 161-169.

Murray, M. M., Thelen, A., Thut, G., Romei, V., Martuzzi, R., & Matusz, P. J. (2016). The multisensory function of the human primary visual cortex. Neuropsychologia, 83, 161-169.

Musacchia, G., & Schroeder, C. E. (2009). Neuronal mechanisms, response dynamics and perceptual functions of multisensory interactions in auditory cortex. Hearing research, 258(1-2), 72-79.

Musacchia, G., & Schroeder, C. E. (2009). Neuronal mechanisms, response dynamics and perceptual functions of multisensory interactions in auditory cortex. Hearing research, 258(1-2), 72-79.

Nardini, M., Bedford, R., & Mareschal, D. (2010). Fusion of visual cues is not mandatory in children. Proceedings of the National Academy of Sciences, 107(39), 17041–17046. https://doi.org/10.1073/pnas.1001699107

Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. Current Biology, 18(9), 689–693.

Negen, J., Chere, B., Bird, L.-A., Taylor, E., Roome, H. E., Keenaghan, S., Thaler, L., & Nardini, M. (2019). Sensory cue combination in children under 10 years of age. Cognition, 193, 104014.

Neil, P. A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., & Shimojo, S. (2006a). Development of multisensory spatial integration and perception in humans. Developmental Science, 9(5), 454–464. https://doi.org/10.1111/j.1467-7687.2006.00512.x

Neil, P. A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., & Shimojo, S. (2006b). Development of multisensory spatial integration and perception in humans. Developmental Science, 9(5), 454–464. https://doi.org/10.1111/j.1467-7687.2006.00512.x

Orioli, G., Dragovic, D., & Farroni, T. (2024). Perception of visual and audiovisual trajectories toward and away from the body in the first postnatal year. Journal of Experimental Child Psychology, 243, 105921.

Ortiz-Terán, L., Ortiz, T., Perez, D. L., Aragón, J. I., Diez, I., Pascual-Leone, A., & Sepulcre, J. (2016). Brain plasticity in blind subjects centralizes beyond the modal cortices. Frontiers in Systems Neuroscience, 10, 61.

Otte, R. J., Agterberg, M. J., Van Wanrooij, M. M., Snik, A. F., & Van Opstal, A. J. (2013). Age-related hearing loss and ear morphology affect vertical but not horizontal sound-localization performance. Journal of the Association for Research in Otolaryngology, 14(2), 261-273.

Pagel, B., Heed, T., & Röder, B. (2009). Change of reference frame for tactile localization during child development. Developmental Science, 12(6), 929–937. https://doi.org/10.1111/j.1467-7687.2009.00845.x

Pasqualotto, A., Spiller, M. J., Jansari, A. S., & Proulx, M. J. (2013). Visual experience facilitates allocentric spatial representation. Behavioural Brain Research, 236, 175-179.

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. NeuroImage, 198, 181-197.

Poirier, C., Collignon, O., DeVolder, A. G., Renier, L., Vanlierde, A., Tranduy, D., & Scheiber, C. (2005). Specific activation of the V5 brain area by auditory motion processing: an fMRI study. Cognitive Brain Research, 25(3), 650-658.

Poirier, C., Collignon, O., DeVolder, A. G., Renier, L., Vanlierde, A., Tranduy, D., & Scheiber, C. (2005). Specific activation of the V5 brain area by auditory motion processing: an fMRI study. Cognitive Brain Research, 25(3), 650-658.

R Core Team (2022). R: A language and environment for statistical computing.

Recanzone, G. H. (1998). Rapidly induced auditory plasticity: The ventriloquism aftereffect. Proceedings of the National Academy of Sciences, 95(3), 869–875. https://doi.org/10.1073/pnas.95.3.869

Recanzone, G. H. (1998). Rapidly induced auditory plasticity: the ventriloquism aftereffect. Proceedings of the National Academy of Sciences, 95(3), 869-875.

Rigato, S., Ali, J. B., Van Velzen, J., & Bremner, A. J. (2014). The neural basis of somatosensory remapping develops in human infancy. Current Biology, 24(11), 1222-1226.

Rigato, S., Bremner, A. J., Mason, L., Pickering, A., Davis, R., & van Velzen, J. (2013). The electrophysiological time course of somatosensory spatial remapping: vision of the hands modulates effects of posture on somatosensory evoked potentials. European journal of neuroscience, 38(6), 2884-2892.

Röder, B., Kusmierek, A., Spence, C., & Schicke, T. (2007a). Developmental vision determines the reference frame for the multisensory control of action. Proceedings of the National Academy of Sciences, 104(11), 4753–4758. https://doi.org/10.1073/pnas.0607158104

Röder, B., Kusmierek, A., Spence, C., & Schicke, T. (2007b). Developmental vision determines the reference frame for the multisensory control of action. Proceedings of the National Academy of Sciences, 104(11), 4753–4758. https://doi.org/10.1073/pnas.0607158104

Röder, B., Ley, P., Shenoy, B. H., Kekunnaya, R., & Bottari, D. (2013). Sensitive periods for the functional specialization of the neural system for human face processing. Proceedings of the National Academy of Sciences, 110(42), 16760–16765. https://doi.org/10.1073/pnas.1309963110

Röder, B., Stock, O., Bien, S., Neville, H., & Rösler, F. (2002). Speech processing activates visual cortex in congenitally blind humans. European Journal of Neuroscience, 16(5), 930–936. https://doi.org/10.1046/j.1460-9568.2002.02147.x

Rohe, T., & Noppeney, U. (2015). Cortical hierarchies perform Bayesian causal inference in multisensory perception. PLoS biology, 13(2), e1002073.

Rohe, T., & Noppeney, U. (2015). Sensory reliability shapes perceptual inference via two mechanisms. Journal of vision, 15(5), 22-22.

Rohlf, S., Li, L., Bruns, P., & Röder, B. (2020). Multisensory integration develops prior to crossmodal recalibration. Current Biology, 30(9), 1726–1732.

Ronga, I., Galigani, M., Bruno, V., Noel, J.-P., Gazzin, A., Perathoner, C., Serino, A., & Garbarini, F. (2021). Spatial tuning of electrophysiological responses to multisensory stimuli reveals a primitive coding of the body boundaries in newborns. Proceedings of the National Academy of Sciences, 118(12), e2024548118. https://doi.org/10.1073/pnas.2024548118

Ruggiero, G., D'Errico, O., & Iachini, T. (2016). Development of egocentric and allocentric spatial representations from childhood to elderly age. Psychological research, 80(2), 259-272.

Russo, N., Foxe, J. J., Brandwein, A. B., Altschuler, T., Gomes, H., & Molholm, S. (2010). Multisensory processing in children with autism: high-density electrical mapping of auditory–somatosensory integration. Autism research, 3(5), 253-267.

Scheller, M., & Nardini, M. (2023). Correctly establishing evidence for cue combination via gains in sensory precision: Why the choice of comparator matters. Behavior Research Methods, 56(4), 2842–2858. https://doi.org/10.3758/s13428-023-02227-w

Senkowski, D., Schneider, T. R., Foxe, J. J., & Engel, A. K. (2008). Crossmodal binding through neural coherence: implications for multisensory processing. Trends in neurosciences, 31(8), 401-409.

Senna, I., Cuturi, L. F., Gori, M., Ernst, M. O., & Cappagli, G. (2021). Spatial and temporal perception in sensory deprivation. Frontiers in Neuroscience, 15, 671836.

Setti, W., Cuturi, L. F., Maviglia, A., Sandini, G., & Gori, M. (2019, June). ARENA: a novel device to evaluate spatial and imagery skills through sounds. In 2019 IEEE international symposium on medical measurements and applications (MeMeA) (pp. 1-6). IEEE.

Shams, L., & Beierholm, U. R. (2010). Causal inference in perception. Trends in Cognitive Sciences, 14(9), 425–432.

Sluzenski, J., Newcombe, N. S., & Satlow, E. (2004). Knowing where things are in the second year of life: Implications for hippocampal development. Journal of cognitive neuroscience, 16(8), 1443-1451.

Spelke, E. S. (1990). Principles of object perception. Cognitive science, 14(1), 29-56.

Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. Annals of the New York Academy of Sciences, 1296(1), 31-49

Spence, C., & Ho, C. (2015). Multisensory information processing. https://psycnet.apa.org/record/2014-29521-027

Stanford, T. R., & Stein, B. E. (2007). Superadditivity in multisensory integration: putting the computation in context. Neuroreport, 18(8), 787-792.

Stanford, T. R., Quessy, S., & Stein, B. E. (2005). Evaluating the operations underlying multisensory integration. Journal of Neurophysiology, 93, 298–309.

Stefanou, M. E., Dundon, N. M., Bestelmeyer, P. E., Ioannou, C., Bender, S., Biscaldi, M., ... & Klein, C. (2020). Late attentional processes potentially compensate for early perceptual multisensory integration deficits in children with autism: evidence from evoked potentials. Scientific Reports, 10(1), 16157.

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. Nature reviews neuroscience, 9(4), 255-266.

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. Nature reviews neuroscience, 9(4), 255-266.

Stein, B. E., Stanford, T. R., & Rowland, B. A. (2009). The neural basis of multisensory integration in the midbrain: its organization and maturation. Hearing research, 258(1-2), 4-15.

Talsma, D., & Woldorff, M. G. (2005). Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. Journal of cognitive neuroscience, 17(7), 1098-1114.

Taylor-Clarke, M., Kennett, S., & Haggard, P. (2002). Vision modulates somatosensory cortical processing. Current biology, 12(3), 233-236.

Tinti, C., Adenzato, M., Tamietto, M., & Cornoldi, C. (2006). Visual Experience is not Necessary for Efficient Survey Spatial Cognition: Evidence from Blindness. Quarterly Journal of Experimental Psychology, 59(7), 1306–1328. https://doi.org/10.1080/17470210500214275

Ungar, S., Blades, M., & Spencer, C. (1995). Mental Rotation of a Tactile Layout by Young Visually Impaired Children. Perception, 24(8), 891–900. https://doi.org/10.1068/p240891

Ungar, S., Blades, M., & Spencer, C. (1995). Mental rotation of a tactile layout by young visually impaired children. Perception, 24(8), 891-900.

Vasilyeva, M., & Lourenco, S. F. (2010). Spatial Development. In R. M. Lerner, M. E. Lamb, & A. M. Freund (Eds), The Handbook of Life-Span Development (1st edn). Wiley. https://doi.org/10.1002/9780470880166.hlsd001020

Vercillo, T., & Gori, M. (2015). Attention to sound improves auditory reliability in audio-tactile spatial optimal integration. Frontiers in integrative neuroscience, 9, 34.

Vercillo, T., Burr, D., & Gori, M. (2016). Early visual deprivation severely compromises the auditory sense of space in congenitally blind children. Developmental psychology, 52(6), 847.

Voss, P., Lassonde, M., Gougoux, F., Fortin, M., Guillemot, J.-P., & Lepore, F. (2004). Early-and late-onset blind individuals show supra-normal auditory abilities in far-space. Current Biology, 14(19), 1734–1738.

Voss, P., Tabry, V., & Zatorre, R. J. (2015). Trade-off in the sound localization abilities of early blind individuals between the horizontal and vertical planes. Journal of Neuroscience, 35(15), 6051–6056.

Wallace, M. T., Carriere, B. N., Perrault Jr, T. J., Vaughan, J. W., & Stein, B. E. (2006). The development of cortical multisensory integration. The Journal of neuroscience, 26(46), 11844-11849.

Wallace, M. T., Roberson, G. E., Hairston, W. D., Stein, B. E., Vaughan, J. W., & Schirillo, J. A. (2004). Unifying multisensory signals across time and space. Experimental brain research, 158(2), 252-258.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. Psychological bulletin, 88(3), 638.

Woollett, K., & Maguire, E. A. (2011). Acquiring “the Knowledge” of London's layout drives structural brain changes. Current biology, 21(24), 2109-2114.

Zanchi, S., Cuturi, L. F., Sandini, G., & Gori, M. (2022). Interindividual differences influence multisensory processing during spatial navigation. Journal of Experimental Psychology: Human Perception and Performance, 48(2), 174.

Zwiers, M. P., Van Opstal, A. J., & Paige, G. D. (2003). Plasticity in human sound localization induced by compressed spatial vision. Nature neuroscience, 6(2), 175-181.

Scientific Production

Papers

Calafatello, G., Zanchi, S., Balzarotti N., Bremner, A.J, Parmiggiani A., Orciari L., Gori, M. (2025). Testing perceptual development with moving stimuli: A new multisensory tool for young children. In Proceedings of the IEEE International Symposium on Medical Measurements and Applications (MeMeA) (pp. 1–6). https://doi.org/10.1109/MeMeA65319.2025.11066292 (published)

Calafatello, G., Zanchi S., Balzarotti N., Bremner A.J, Gori, M. **"**The role of multisensory integration in spatial object segregation in early childhood" (under review)

Calafatello G.,* Gori, M.,* Zanchi S., Vitali H., Bremner A.J, Campus C. "Unveiling the neural mechanisms of integration and conflict of audio-tactile stimuli in early childhood with and without visual impairment" (in preparation)

Calafatello G., Zanchi S., Tonelli A., Setti W., Tammurello C., Gori, M. "Comparison of sound spatial localisation skills and reaction times between adults and children on the frontal plane" (in preparation)

Tonelli A., Esposito E., **Calafatello G**., Gori, M. Early visual experience shapes temporal but not statistical priors in auditory spatial perception (in preparation)

Conferences Contribution

Talk Presentation

Calafatello G., Zanchi S., Tonelli A., Setti W., Tammurello C., Gori M. "*Spazio sensoriale: Esplorando le dimensioni e lo sviluppo della percezione del corpo e dell'ambiente con e senza disabilità sensoriale"*
Associazione Italiana Psicologi (AIP), Noto, 2024

Calafatello G., Zanchi S., Balzarotti N., Bremner A.J, Parmiggiani A., Orciari L., Gori M. *"Testing Perceptual Development with Moving Stimuli: A New Multisensory tool for young children"* at IEEE Medical Measurments & Application (MeMeA) 2025, Chania, Greece

Calafatello G, Zanchi S., Balzarotti N., Bremner A.J, Gori M. "*Integrazione e segregazione multisensoriale: segnali percettivi emergenti nei bambini in età pre-scolare".* Associazione Italiana Psicologi (AIP), Torino, 2025

Poster Presentation

Calafatello G., Zanchi S., Tonelli A., Setti W., Tammurello C., Gori M. *"Audiovisual Spatial Localization in Children and Adults in the frontal plane. Comaprison of Localization Accuracy in the Azimuth and Elevation Plane".* IMRF conference (Internation Mulisensory Research Forum), Bruxelles, 2023

Calafatello G., Zanchi S., Balzarotti N., Bremner A.J, Gori, M. "Investigating Audiovisual spatial integration in infacy". Internation Conference on Infants Studies (ICIS), 2024

Calafatello G., Zanchi S., Vitali H., Bremner A.J, Campus C., Gori, M. "*Neural correlates of multisensory spatial perception in visually impaired children"* International Multisensory Research Forum, (IMRF), Durham, 2025.