Chapter 1

Introduction

General introduction

Sounds are ubiquitous in our living environment. They can stem from different sources, be that biological organisms or mere physical objects. We can detect and select sounds, distill information, and use it to perform appropriate actions in the environment. Here, I am particularly interested in motor behavior and ask: How can sounds provide information to guide our motor system? I also ask: Does the human motor system in turn influence the perception of sound and, if so: How does this work? The studies bundled in this thesis are attempts to square this circle by unraveling the interaction between our auditory and motor systems from different angles.

It is Saturday morning in Amsterdam. It does not rain. You decide to cycle through the city and then go to the center to take a stroll. The closer you get to the city center, the more tourists are walking on the bike path. The tourists do not see you, so the best way to convey your message is to produce a sound, so you ring the bell. For the typical Dutch bell, you gently push or pull the bell’s lever by which discs inside the bell rotate, striking the little bump inside the bell’s hood. This creates noise, and since it is between two caps, it echoes the familiar ‘ding-dong’ sound. That sound thus produced, i.e. a longitudinal wave of air pressure, travels through the air to reach the tourists’ ears. Air pressure changes yield mechanical perturbations at the hair cells in the inner ear that are transformed into electrical signals and eventually reach the auditory cortex and progress further through the brain, where it is analyzed and interpreted as an alert. Once it is altered, the brain evokes the tourist’s motor system that sends signals along his or her corticospinal tract and innervates motor units in the muscles in order to produce movements. And “yes”, if you ring the bell hard enough, the tourist will jump off the bike path.

Before I continue this story, let me step back for a moment and give a brief look at the physical properties of sound (frequency and loudness), the mechanics of the ear, and auditory and motor systems in the human brain.
Sound and human auditory system

The word ‘sound’ refers to both physical stimulus and its perceptual response. The physical definition of sound is: pressure changes in compressible media like air or water; and the perceptual definition is: the experience we have when we hear it. For example, the sentence “The electric guitar plays at a high pitch.” refers to the experience of sound, but “The sound’s intensity level is 60 decibels.” refers to sound as a physical stimulus.

Properties of Sound

When we are close to the source that produces sounds, we perceive them as louder, and quieter when we are further away. Sound intensity is defined as the amount of energy passing through a unit area per unit time and determines the psychological experience we call loudness. The higher the amplitude, the louder the sound perceived. We describe relative sound intensities in units named decibels (dB), a logarithmic scale for measuring the intensity of sound waves. A whisper, for example, generally falls around 30 dB, a normal conversation around 60 dB, while a washing machine produces loud sounds as loud as 80 dBs.

The frequency is the number of cycles in a sound wave per unit of time or the number of times a sound wave repeats itself within a given period. Frequency is most often measured in Hertz (Hz), i.e. cycles per second. Pitch of a sound, on the other hand, is our subjective experience created by sound wave frequency and is directly related to frequency: the higher the frequency of a waveform, the higher the pitch of the sound being heard. The range of hearing in healthy young humans is approximately between 20 Hz and 20,000 Hz; as people grow older, their hearing range reduces.

The auditory system

The human ear is the anatomical structure that is responsible for transducing the fluctuations of air into the neural activity, which ultimately leads to perception of sound. Sound waves have to travel a long way from outside to inside the ear before they are transduced into neural signals and sent to the brain for processing. The pinna is the outer and the only visible part of the ear that collects and directs sound waves. Sound waves travel through the ear canal to the eardrum and cause it to vibrate, which, in turn,
vibrates three tiny bones in the middle ear: malleus, incus, and stapes. These three bones transfer the signals to the oval window, which again transmits them to the inner ear's semicircular canals, the cochlea, which is filled with fluid and nerves (Figure 1).

![Figure 1. The structure of ear. Adapted and edited from (Goldstein, 2010).](image)

The cochlea is a spiral-shaped structure full of tiny hairs that transform fluid waves into electrical impulses. The auditory nerve carries signals generated by the inner hair cells away from the cochlea and toward the cochlear nucleus. From there, they continue to the superior olivary nuclei in the brain stem, the inferior colliculus in the midbrain, and the medial geniculate nucleus in the thalamus. From the latter, there are pathways to the primary auditory cortex (A1) in the temporal lobe of the cortex (Figure 2). The function of A1 is to process sound information such as pitch, intensity, and the location of the sound source.

**The human motor system – a very brief overview**

The motor system enables us to plan, coordinate, and execute a wide range of actions such as breathing, speaking, or cycling. There are several anatomical areas that are engaged in the motor system. The motor areas of the cerebral cortex are divided into primary and secondary motor areas. Each area includes populations of neurons that project from the cortex to the brain stem and the spinal cord. The main output to the spinal cord originates from primary motor cortex (M1). M1 receives input not only from other secondary areas of motor cortex such as premotor areas (PM1), but also from the thalamus and sensory areas.
Primary motor cortex

M1 is located in the frontal lobe of the brain (Figure 3C). According to traditional neuroscience, M1’s major role is to generate neural signals that yield the execution of movement. Signals from one side of M1 cross the body’s midline to stimulate skeletal muscles on the opposite side of the body, meaning that the left M1 controls movement on the right side of the body, and vice versa. The relative area dedicated to any body part indicates the extent of control that the primary motor cortex has over that body part (cf. motor homunculus).

TMS and corticospinal motor tract excitability

Transcranial magnetic stimulation (TMS) is a non-invasive tool for probing brain-behavior relationships. It has numerous applications in the study, diagnosis, and therapy of the brain. The basic principle of this instrument is that it produces brief and focal magnetic fields that induce an electric field, which alters the local field potential field in quite localized parts of the brain. The excitability of the neurons in the vicinity is altered and, if strong enough, the TMS can even elicit action potentials. For instance, TMS over the hand area of M1 can produce a twitch in the corresponding hand muscle.
The twitch, in turn, can be measured using conventional electromyography (EMG). The respective signal is referred to as a motor evoked potential (MEP). By measuring the peak-to-peak amplitude, we can thus quantify the magnitude of corticospinal tract excitability (Figure 3).

![Figure 3](image)

**Figure 3.** (A) Scheme of stimulation of M1 using TMS and recording MEPs. (B) Induction of electrical currents in the brain. (C) Illustration of M1 and motor homunculus. (D) MEP examples; the vertical lines at 0 ms indicates the moment a single pulse of TMS has been generated.

**Premotor cortex and other motor regions**

Other areas of the brain that assist motor function are called the *secondary motor cortices*. These regions consist of PM1 and the supplementary motor area (SMA). Transformation of visual information into motor commands involves the posterior parietal cortex. For instance, the posterior parietal cortex contributes to determining how to grasp objects based on their location in space. This information is sent to both PM1 and SMA. PM1 is anterior to M1 and is involved in both sensory guidance and execution of movements. SMA is involved in planning complex movements such as bimanual actions, mostly in internally generated movements rather than those triggered by external sensory events. Both PM1 and SMA send information to M1 as well as to brainstem motor regions. Next to the brainstem, other subcortical areas contribute to sensorimotor function. For example, the basal ganglia (or basal nuclei) are thought to participate in
various aspects of higher-level motor functions such as action selection. The cerebellum is involved in the integration of sensory and motor information and also plays a major role in motor learning.

**Perception in action – action in perception**

Traditional research on motor system mainly focused on how the central nervous system is involved in generating motor programs. Nikolai Bernstein (1896-1966) was one of the first to recognize that perceptual information is of fundamental importance to understanding movement coordination and control. James Gibson emphasized this issue in more detail. According to Gibson “We must perceive in order to move, but we also must move in order to perceive” (Gibson, 1979: 223), suggesting a close relationship between action and perception when navigating the environment. Although Gibson’s focus was mainly on visual perception, his ideas can be applied to perception in general. The close interplay between action and perception has been further established with the scientific data provided by recent neurophysiological, neuroimaging, and behavioral studies.

**Mirror neurons**

A possible neural basis for the link between perception and action has been put forward by the discovery of so-called mirror neurons. The term ‘mirror neurons’ refers to population of neurons in monkeys’ brains that have been found to fire both when a monkey performs an action (picking a peanut), and when it observes the same action (somebody else picking a peanut: Rizzolatti and Craighero, 2004). Interestingly, mirror neurons respond not only when a specific action is performed and seen by the monkey but also when the monkey hears the sound of the related action (crushing a peanut: Kohler et al., 2002). In the search for these auditory-motor mirror neurons, Parther and colleagues found similar neural firing patterns when a bird was singing itself or when it was listening to songs containing similar sequences sung by other birds (Prather et al., 2008). Mirror neurons may be mediating the comparison of what a bird hears and produces. The presence of such neurons has led to the proposal that this system might have mediated the evolutionary transition from gesticulation to speech (Rizzolatti and Arbib, 1998). However, so far we still wait for convincing evidence supporting the evolution of language from gesture.
As of yet, the existence of auditory mirror neurons in humans is not clear. There are, however, several brain-imaging and neurophysiological data that suggest the existence of a mirror-neuron system in humans during visual and auditory perception (See Rizzolatti and Craighero, 2004 for overview).

A different line of research has provided evidence for the activity in motor regions of the brain in response to auditory stimuli. In listening tasks, functional magnetic resonance imaging (fMRI) has revealed that context-related sounds that trigger a responsive action (e.g. a ringing telephone) provoke stronger activity within PM1 as compared to context-free sounds that do not cue a responsive action (e.g., notes on a piano). Similarly, it has been shown that action-related sounds lead to selective modulation of the left posterior superior temporal and PM1, while the processing of non-action-related sounds results in the bilateral modulation of the temporal pole (Pazzaglia et al., 2008). TMS revealed an increase in corticospinal excitability while subjects were presented with action-related sounds (Aziz-zadeh et al., 2004). It has also been reported that mid-PM1 is activated not only when subjects listened with anticipation and tapped in synchrony to rhythmic sequences, but also when they were passively listening to the musical rhythms (Chen et al., 2008). Another study showed that when subjects listened to a newly-acquired melody, there was activity in cortical motor regions (Broca’s area, the premotor regions, and the left and right inferior frontal gyrus) that was not present during the presentation of untrained-same-notes-music (Lahav et al., 2007). A similar study tested professional musicians and revealed overlapping neural regions (including PM1) both during listening and playing the same piece of music (Bangert et al., 2006). Furthermore, TMS stimulation in musicians showed facilitation of primary motor cortex when listening to a rehearsed piece compared with a non-rehearsed control piece (D’Ausilio et al., 2006). Finally, the modulation of corticospinal excitability has been found in non-musicians who listened to newly-learned action-related sounds (Ticini et al., 2012).

**Emotional sound and action readiness**

Back to our story in Amsterdam, you arrive in the center; you take a break and record all the sounds you hear during the day roaming around. If you replay them afterwards, there will be some sounds that you enjoyed, like the music you heard, the sound of a
baby laughing, or the relaxing sound of the Amstel river. There will also be some sounds that you do not like, e.g., the sound of a siren, a vacuum cleaner, or a crying baby. There are also sounds that are ‘neutral’, such as that of restaurant ambience, walking, and a clock ticking.

What I did here was categorizing the sounds based on their emotional valence. The presence of emotion can be considered to be a vital prerequisite for proper daily actions as it helps qualify information and thus (fine) tunes behavioral responses.

You continue cycling through the town. The emotional content of the sounds you hear can influence your decision of where to go. If you are into music and party, you will consider hearing the sound of music a positive, attractive event. Hearing the sound of people shouting and yelling at each other, by contrast, will be considered a negative event and you will avoid getting too close.

The emotional content of the sound may facilitate the recognition of the emotional intention of the sound and consequently trigger an appropriate reaction. The close interplay between action readiness and emotion has indeed been manifested through different experimental approaches. TMS has been successful in revealing a non-trivial relationship between action preparedness and emotional processing by means of increased corticospinal motor excitability during emotional experiences (Tormos et al., 1997, Oliveri et al., 2003, Baumgartner et al., 2007, Hajcak et al., 2007, Schutter et al., 2008, Coombes et al., 2009, Coelho et al., 2010, Van Loon et al., 2010, Baumert et al., 2011). Despite the numerous reports on motor responses to emotive aspects of visual stimuli, there are surprisingly few reports on corticalspinal tract excitability as a function of auditory processing of sounds carrying emotional contents.

In Chapter 2, I summarize a study asking whether (i) emotional processing of nonverbal auditory stimuli would lead to increased corticospinal tract excitability, (ii) this modulation is lateralized in response to the valence of the stimuli, and that (iii) delivering the sounds to the left ear, right ear, or both ears may yield lateralization in motor facilitation.
Speech

Depending on the environment we live in, the sounds we hear during the day have many different sources and different messages to convey. Speech and music are two examples of how we interact with sounds through action.

After all the cycling through Amsterdam, you do need a drink. As soon as you start ordering, you produce speech sounds. The bartender (hopefully) understands the sound you made, which involves ‘speech perception’. – If you order the usual, bartenders are quite good at that.

The movements we make with our vocal organs cause vibrations in the air, which travel across to the ear as series of sound waves, which activate a network of tiny bones and cells to enable us to eventually hear the sound. These steps are the central parts of the speech chain and constitute the main constituents of the field called phonetics, which is concerned with describing the speech sounds that occur in languages and is divided into three types:

- **Articulatory phonetics** asks how speech sounds are produced. In almost all speech sounds, the respiratory system provides the basic source of power. It pushes the air out of the lungs that goes up the windpipe and into the larynx. At this point, it must pass through vocal cords, which modify and modulate the airstream by means of several articulators. This combination of anatomical properties and the airstream enables phoneticians to precisely define all speech sounds, or as is sometimes called segments that are used in natural languages.

- **Auditory phonetics** investigates the processes through which speech is perceived by the listener. The starting point for any auditory speech is the study of the auditory system, which concerns the anatomy and physiology of both the ear and the brain. Since the auditory system cannot react to all features present in the sound wave, it is essential to determine what we perceive and how we perceive it. This enormously complex field is referred to as speech perception, which is not only of interest to phonetics, but also to cognitive neuroscience and experimental psychology.
- Unlike the first two, which were concerned about how speech sounds are produced and heard, *acoustic phonetics* deals with describing speech sounds in terms of what we can hear. The description would enable *acoustic phoneticians* to explain, for example, why certain sounds are confused with one another, or to specify sounds (such as vowels) that are hard to describe as far as articulatory movements are concerned.

**Theories of speech perception**

There is consensus that the ear amplifies incoming signals and transfers them to auditory areas where a primary analysis takes place, filtering out noise and other non-speech-related features of sound. Auditory patterns are then distilled using a sort of cognitive recognition device. Current theories differ in what this recognition device may be. ‘*Active theories*’ assume that the process of speech perception involves some sort of internal speech production that the listener applies to their articulatory knowledge when they analyze incoming signals. The listeners do not only act when the speaker produces the speech, but also when they perceive it. ‘*Passive theories*’, by contrast, emphasize the sensory aspect of the perceptual process and relegate the process of speech production to a minor role. They postulate the use of a stored neural pattern, which may even be innate.

Importantly, the motor system appears to be involved in speech perception (Liberman et al., 1967, Liberman and Mattingly, 1989). Listeners may perceive spoken language by generating forward models in the motor system and by activating articulatory phonetic gestures used to produce acoustic speech signals rather than by identifying the sound patterns generated in speech. Supporting evidence has been provided through altered M1 and PM1 activity during speech perception. Passive listening to speech sounds results in an increase in M1 activity (Wilson and Iacoboni, 2006) even in a somatotopic manner (Pulvermüller et al., 2006). Fadiga and colleagues showed that listening to words requiring strong tongue movement results in facilitation of MEPs recorded from the tongue muscles (Fadiga et al., 2002). Similarly, Watkins and co-workers reported an increase in MEP amplitude of lip muscles involved in speech production when observing speech-related lip movements or listening to speech sounds (Watkins et al., 2003). Repetitive TMS over the left PM1 or M1 causes the capacity of phonetic discrimination to be significantly reduced (Meister et al., 2007, Möttönen and Watkins, 2009, Sato et al., 2009), indicating even a causal relationship between the mo-
tor system and speech perception. Furthermore, functional lesion studies suggest that the involvement of motor areas during speech production does not necessarily contribute to speech perception (for review, see e.g., Hickok and Poeppel, 2000, Scott et al., 2009).

Speech and gesture

Back to the bar where you ordered a drink. When the bartender brings your order, you pay immediately. “Keep the change”, you say and hear the bartender’s “Thank you”. He/she shows you a ‘thumbs up’ to indicate both appreciation and agreement. This is speech/gesture production. Of course, you need to perceive what is being said and what the gesture really means.

How speech sound might be perceived has been sketched above. But how do we perceive the gestural signals even when they are not accompanied with speech sounds? The strong coupling between gestures and speech suggests the presence of a unified communication system. Neuropsychological and neurophysiological evidence indeed support the idea that speech and gestures share common neural substrates. For example, imaging studies revealed activities in language areas (e.g., Broca’s area) during gestural observation (MacSweeney et al., 2002).

Möttönen and co-workers reported that MEPs elicited by stimulating hand representation in M1 do not differ when participants observed signs with known versus signs with unknown meanings (Möttönen et al., 2010). If M1 hand area is insensitive to the distinction between action associated and not associated with words, then other regions in M1 like the tongue or lip areas might be better candidates. It remains unclear, however, whether motor representations of tongue and lips are capable of distinguishing between those actions that symbolically represent words (e.g., an object or a state) and those that do not.

In Chapter 3, I discuss whether observation of newly learned hand gestures paired and not paired with words may or may not result in changes in the excitability of the hand and tongue areas of motor cortex.
Speech and multisensory perception – the McGurk effect

The speech signal is multisensory in nature. It has long been known that the speaker’s facial movements contribute to speech perception in noisy conditions.

*It is much easier to perceive the bartender’s appreciation when not only hearing the “Thank you”, but also looking at his or her face, especially if you picked a noisy bar in the center of Amsterdam.*

The by now classic auditory illusion addressing this is named after Harry McGurk. He and John MacDonald employed incongruent audiovisual videotape stimuli on which acoustic syllable /ba/ was spoken over the facial motion of /ga/ (McGurk and MacDonald, 1976). When listeners were asked what consonant they heard, they reported hearing the fused response /da/. This is a terrific example showing that the accurate perception of information might engage the contribution of more than one sensory system in order to help us perceive the world around us better.

Several neuroimaging studies demonstrated the involvement of different brain areas during the McGurk illusion, including superior temporal sulcus (STS) (Calvert et al., 1998, Sekiyama et al., 2003, Beauchamp et al., 2010). It is the same STS that has been considered the cortical region for the integration of auditory and visual stimuli (Beauchamp et al., 2004). Interestingly, Beauchamp and colleagues showed that TMS delivered over the STS reduced the likelihood of the McGurk illusion, which presents evidence for a causal role of STS in multisensory integration (Beauchamp et al., 2010).

One way of studying the involvement of motor processes in speech perception is to examine the contribution of speech motor system in speech perception. Similar to McGurk effect (McGurk and MacDonald, 1976), it has been shown that the perception of auditory syllables can be disturbed or improved when the listener silently articulates incongruent or congruent syllables, respectively (Sams et al., 2005, Mochida et al., 2013). In Chapter 6, I go beyond the presence of audiomotor integration in speech and address brain mechanisms underlying this particularly important multisensory capacity of humans.
Sensory guidance of action using sound (action execution)

You pass Dam Square. People are talking, a band is playing music, and some street artists are performing a show. You decide to stop for a while and listen to the music. The band is playing blues music, the guitarist is blind and is standing right in front of you; the bass is to your left, and the saxophone to your right. If you close your eyes, you experience an ‘auditory space’ in which you can easily localize the band’s members at different positions.

The sense of hearing can work together with that of vision to guide orientation to stimuli and movement in our environment. Bats and owls are very specialized in sound localization and use it for hunting prey and navigating. Positioning and navigation are also multisensory in nature, and therefore one needs to understand how senses (auditory, vestibular, proprioceptive) are integrated to understand auditory localization.

Intercepting movement (sound localization)

Since the two ears are separated by the width of the head, sound waves arrive in one ear slightly earlier than the other ear. This is technically referred to as the interaural time difference (ITD). The head also creates a barrier -“sound shadow”- that makes differences in the amplitude of the sound waves entering the two ears and is known as interaural level difference (ILD). Another known binaural localization cue is the head-related transfer function (HRTF), which is a unique way in which a sound spectrum is filtered by the physical properties of an individual’s pinna, head and torso. It has been shown that some mammals (including humans) use all the three auditory cues, ITD, ILD, and HRTF to localize sounds (for a review, see Grothe et al., 2010). Visually impaired people have the potential not only to echolocate objects, but they are also able to interact with the moving sound source in space.

The guitarist is wearing a baseball cap, and you may wonder whether he really likes baseball or if he plays the ‘beep baseball’. The game is similar to baseball with the difference that the players are visually impaired and use sounds to dive onto the ground to stop a beeping ball and run at full speed toward the sound of a buzzing base to score a run.
Guiding movements to intercept a sound source requires reliable information about the direction along which the sound is travelling. It has been suggested that the increase in the binaural rate of intensity (or loudness growth) of a sound is the primary information source used to determine the trajectory and velocity of an approaching sound emitting source (Shaw et al., 1991). Although ILD seems to be an advantageous auditory cue for the localization of an approaching sound, sound localization in the horizontal plane (azimuth) involves both ITD and ILD binaural cues. In Chapter 4, I show which strategies humans could adopt in order to intercept a low-frequency rotating sound that has been created by varying the ITD.

**Sensorimotor synchronization (music and movement)**

*You are still in the Dam Square. You are enjoying the music and spontaneously start to synchronize your body movements with the rhythm’s beat such as nodding the head or tapping a foot. Here you feel the power of music.*

Music is a powerful stimulus for evoking timely movements. The skill to synchronize body movements with music often occurs naturally and does not require previous musical training. Already from the first months of life, infants move their body more naturally under the presence of musical rhythm than speech rhythm (Zentner and Eerola, 2010).

Synchronizing movements with musical rhythms is indeed one the most natural and instinctive ways humans use to interact with their auditory environment. Infants are not only able to automatically synchronize their movements with different musical tempos, but they are also selectively sensitive to melodies presenting different pitch structures (Zentner and Kagan, 1998). Adults use a different walking strategy under the guidance of music than under a metronome beat (Styns et al., 2007, Wittwer et al., 2013). Using a finger-tapping paradigm, it has been shown that synchronization error was significantly less when tapping with music cues than doing so with a metronome (Thaut, 1997). What emerges from these studies is that in addition to the timing cues, music conveys additional properties that help guide the coordination of movement. Hence, investigating whether and how non-temporal cues, such as pitch and harmony, influence movement synchronization seems crucial for understanding the inseparable connection between action and perception.
Musical chords are either **consonant** (sounding pleasant or stable) or **dissonant** (sounding unpleasant or instable). Although composers make use of both intervals to evoke diverse feelings of ‘tension’ and ‘resolution’, consonant intervals in tonal music occur more often than dissonant ones (Vos and Troost, 1989). Interestingly, consonant intervals are more preferred by human infants (Zentner and Kagan, 1998, Trainor et al., 2002, Masataka, 2006). The preference of consonance over dissonance seems to be cross-cultural, as it has been reported among native African populations who did not have prior experience with Western music (Fritz et al., 2009). Overall, it seems that some characteristics of musical pitch interval perception might be innate and represent a by-product of fundamental biological properties.

Pythagoras defined consonance by the fact that the length of string segments form simple integer ratios (e.g., 3:2, 2:1) and dissonance when string length ratios are more complex (e.g., 16:15, 243:128). Galileo echoed this idea and proposed that simple integer ratios are the ones that do not “keep the ear drum in perpetual torment” (Galilei, 1914: 104). Von Helmholtz (1863) extended this concept and argued that consonance occurs not only as a consequence of simple frequency ratio relationships, but also as a result of the interference between overtones of slightly different frequencies – a phenomenon known as beating (Helmholtz, 1954). When overtone frequencies are close, the beating gets faster and an unpleasant sensation, called roughness, is formed (Helmholtz, 1877/1954).

Neuroimaging studies revealed marked differences in the processing of musical intervals at both cortical (e.g., PM1: Minati et al., 2009) and subcortical levels (e.g., brainstem: Bidelman and Krishnan, 2009, 2011), indicating the involvement of networks for sensory and cognitive processing. The preferential encoding of consonants might be rooted in a more robust and coherent neuronal synchronization (Fishman et al., 2001, Mckinney et al., 2001, Tramo et al., 2001).

In view of the ecological relevance of consonance versus dissonance, the harmonic structure of sounds may affect the spatiotemporal characteristics of movements when using them to guide timed actions. In Chapter 5, I address this issue in a synchronization-continuation paradigm, in which the participants were asked to synchronize their
movements with auditory tones differing in the degree of dissonance and then to maintain the same movement patterns in the absence of the auditory stimuli.
Research questions and outline of the thesis

With the research described in my thesis, I aim to tackle the interaction between auditory and motor system from different points of view. The research has been motivated by the following research questions:

Does valence of sound affect the excitability of the corticospinal motor tract?

– Chapter 2.

Does speech perception through the observation of gestures alter M1 excitability?

– Chapter 3.

What strategies can be used to intercept moving sounds?

– Chapter 4.

Do non-temporal aspects of musical sound affect movement timing?

– Chapter 5.

What are the cortical correlates of audiomotor and audiovisual integration?

– Chapter 6.

The remainder of this thesis is divided into six chapters summarizing different experimental studies followed by an epilogue containing a general discussion.

The study described in Chapter 2 addressed to what degree non-verbal emotionally characterized sounds presented separately to the left and right ear yield differential and possibly lateralized excitability of the corticospinal motor tract. While subjects were listening to sounds (monaurally and binaurally), single TMS pulses were delivered to either left or right M1, and EMG activities were recorded from the contralateral abductor pollicis brevis muscle.

In Chapter 3, the mere observation of hand gestures, paired and not paired with words, is reported as it may result in changes in the excitability of the hand and tongue areas of M1. Once again, I used single-pulse TMS, while the participants viewed video sequences of bimanual hand movements associated or not-associated with nouns.

To gain more insight into the interaction between sound perception and the motor system, I further investigated how people use auditory information to control their action.
In the experiment given in Chapter 4, I examined subjects’ ability to intercept a laterally moving virtual sound object by controlling the displacement of a sliding handle so that it reached its endpoint at the same time as the sound object. I tested whether the ITD of an arriving sound could be the main source of perceptual information used by individuals to enable them to successfully intercept the object.

Synchronizing movements with musical rhythms are the most natural and instinctive ways through which humans interact with their auditory environment. In Chapter 5, I summarize the effects of musical consonance/dissonance on sensorimotor timing in a synchronization-continuation paradigm during which participants performed reciprocal aiming movements. This was assessed by measuring timing accuracy and variability along with parameters defining the shape of the movement trajectory.

Another way of studying the involvement of motor processes in auditory perception is to examine the contribution of the motor system in speech perception. In Chapter 6, I sketch a search for brain mechanisms underlying audiovisual and audiomotor multisensory integration during speech perception. During the experiment, listeners were asked to recognize the auditory syllables while silently articulating congruent/incongruent syllables (motor condition) or observing videos of a speaker’s face articulating the syllables (visual condition). EEG responses were recorded during all the conditions.

As mentioned, I reflect on the main results of all studies reported in the thesis and revisit the leading research questions in the final Chapter 7, the Epilogue. And for the impatient: That final part also contains the denouement of the stroll through Amsterdam 😊