Brain Mechanisms for the Perception of Visual and Auditory Communication Signals – Insights from Autism Spectrum Disorder

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

Doctor rerum naturalium (Dr. rer. nat.)

im Fach Psychologie

eingereicht an der

Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

von **M.Sc., Kamila Borowiak**

Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

Prof. Dr. Bernhard Grimm

Gutachter/innen

1. Prof. Dr. Katharina von Kriegstein
2. Prof. Dr. Isabel Dziobek-Faber
3. PD Dr. Johannes Schultz

Tag der mündlichen Prüfung: 22.06.2020

Abstract

Communication is ubiquitous in our everyday life and appears to be effortless at least to most of us. Yet, individuals with autism spectrum disorder (ASD) can experience difficulties in social communication and interaction and when recognizing socially relevant signals from the face and the voice. Such impairments of social communication can vastly affect the quality of life - a profound understanding of the mechanisms behind these impairments is thus strongly required. Previous research on social communication in ASD has put emphasis on emotion processing and looked at the difficulties mainly from the cognitive perspective. In the current dissertation, I focused on sensory brain mechanisms that underlie the perception of emotionally neutral aspects of communication signals that so far have gained little attention in ASD research. I studied the malleability of voice-identity processing in the human brain using intranasal administration of oxytocin, and thus the potential to alleviate voice-identity recognition impairments in ASD. Furthermore, I investigated brain mechanisms that underlie difficulties in the recognition of visual speech in ASD, because previous evidence on visual-speech recognition in ASD was limited to behavioral findings. For this purpose, I applied methods of functional magnetic resonance imaging (fMRI), eye tracking, and behavioral testing. The contribution of the findings of the present dissertation is twofold. First, the findings provide novel evidence corroborating the view that atypical sensory perception is a critical cornerstone for understanding social communication difficulties in ASD. Dysfunction of visual and auditory sensory brain regions might not only contribute to difficulties in processing aspects of communication signals in ASD, but also determine the efficacy of interventions that aim to improve the behavioral deficits. Second, the findings deliver empirical support for a recent theoretical model of how the typically developing brain perceives moving faces. This improved our current knowledge about brain mechanisms that underlie the processing of visual communication signals from the moving face in the typically developing population. Advanced scientific knowledge about human communication, as provided in the current dissertation, propels not only further empirical research, but also development of clinical interventions that aim to promote communication abilities in affected individuals.

Zusammenfassung

Kommunikation ist allgegenwärtig in unserem Alltag und scheint zumindest für die meisten von uns mühelos zu sein. Jedoch Personen mit einer Autismus-Spektrum-Störung (ASS) können Schwierigkeiten in sozialer Kommunikation und Interaktion zeigen. Sie haben unter anderem Schwierigkeiten beim Erkennen von Kommunikationssignalen, die mittels Gesichts und Stimme dargeboten werden. Da derartige Kommunikationseinschränkungen einen negativen Einfluss auf die Lebensqualität haben können, ist ein tiefgreifendes Verständnis der zugrundeliegenden Mechanismen von großer Bedeutung. Bisherige Forschung zu sozialer Kommunikation in ASS legte ihren Fokus darauf, wie Personen mit ASS Emotionen verarbeiten und betrachtete die sozialen Kommunikationsschwierigkeiten hauptsächlich auf der kognitiven Ebene. In der vorliegenden Dissertation befasste ich mich mit sensorischen Gehirnmechanismen, die der Verarbeitung emotionsneutraler Kommunikationssignale zugrunde liegen und, die in der Forschung zu ASS bisher wenig Beachtung fanden. Erstens untersuchte ich, ob eine intranasale Gabe von Oxytocin die Prozesse der Identitätserkennung von der Stimme beeinflussen kann, und somit die Schwierigkeiten in der Stimmenidentitätserkennung bei Personen mit ASS mildern kann. Zweitens erforschte ich, welche neuronalen Prozesse die Schwierigkeiten in der Wahrnehmung visueller Sprache bei Personen mit ASS erklären können, da bisherige Evidenz nur auf Verhaltensdaten basierte. Um diese Fragestellungen zu beantworten, wendete ich Methoden der funktionellen Magnetresonanztomographie (fMRT), Eyetracking und Verhaltenstestungen an. Die Ergebnisse der Dissertation liefern neuartige Erkenntnisse, die sowohl für Personen mit ASS als auch für typisch entwickelte Personen von hoher Relevanz sind. Zum einen bestätigen sie die Annahmen zur Ätiologie von ASS, dass atypische sensorische Mechanismen einen entscheidenden Grundstein darstellen, um Schwierigkeiten in sozialer Kommunikation und Interaktion in ASS zu verstehen. Sie zeigen, dass eine Dysfunktion visueller und auditorischer sensorischer Gehirnregionen nicht nur den Kommunikationseinschränkungen bei Personen mit ASS zugrunde liegen, sondern auch die Effektivität von Interventionen beeinflussen, die eben jene Schwierigkeiten vermindern sollen. Zum anderen liefern die Ergebnisse empirische Evidenz für aktuelle theoretische Annahmen darüber, wie das typisch entwickelte Gehirn visuelle Kommunikationssignale in einem dynamischen Gesicht verarbeitet. Diese wissenschaftlichen Erkenntnisse zur Wahrnehmung von visuellen und auditorischen Kommunikationssignalen erweitern maßgeblich unser aktuelles Wissen und zukünftige

Forschungsansätze zur zwischenmenschlichen Kommunikation. Außerdem können sie neue Ansätze für Interventionsmöglichkeiten hervorbringen, die auf die Förderung der Kommunikationsfähigkeiten innerhalb klinischer Populationen abzielen.

Eingereichte Einzelarbeiten

Borowiak, K., & von Kriegstein, K. (2020). Intranasal oxytocin modulates brain responses to voice-identity recognition in typically developing individuals, but not in ASD. *Translational Psychiatry*, *10*, 1-12. [Study 1] – doi: https://doi.org/10.1038/s41398-020-00903-5

Borowiak, K., Schelinski, S., & von Kriegstein, K. (2018). Recognizing visual speech: Reduced responses in visual-movement regions, but not other speech regions in autism. *NeuroImage: Clinical, 20,* 1078-1091. [Study 2] - doi: https://doi.org/10.1016/j.nicl.2018.09.019

Borowiak, K., Maguinness, C., & von Kriegstein, K. (2019). Dorsal-movement and ventral-form regions are functionally connected during visual-speech recognition. *Human Brain Mapping*, *41*, 952-972. [Study 3] – doi: https://doi.org/10.1002/hbm.24852

Acknowledgments

First and foremost, I would like to thank my supervisor Katharina von Kriegstein for her excellent supervision, guidance and confidence in my work. I am very grateful for the opportunity to learn from and work with her. Many thanks go to all current and former members of the "Neural Mechanisms of Human Communication" group, my Berlin School of Mind and Brain peers, and all the colleagues at the Max Planck Institute and at the Technische Universität Dresden who accompanied me during my PhD. In particular, I would like to thank Claudia Roswandowitz and Stefanie Schelinski not only for being great colleagues, but also for becoming my good and always supportive friends. I hope this will go on like that. I thank Corrina Maguinness for many valuable discussions and her patience with my countless questions about face and voice processing. I am grateful for having met you all.

I would like to express my sincere gratitude to all study participants, especially those with autism spectrum disorder. I appreciate their confiding cooperation and impressive commitment! Without them, this work would not be possible.

I would like to thank my friends, in particular Hanna Kotrba for her endless hospitality and being there for me through thick and thin. Manuel Walter for being my partner in crime and always having an open ear for all the ups and downs during my PhD. Finally and most of all, I would like to thank my parents for their unconditional support and encouragement to study and to explore the world, and for providing me with the means to do so. You never questioned my choices and I love you for that. Dziękuje.

Table of Contents

1 Introduction

Communication is a fundamental component of human nature and ubiquitous in our daily life. We casually talk with friends and family in our leisure time, and we engage in professional discussions or in small talk during coffee breaks with colleagues at work. Thus, our predisposition and ability to communicate significantly affects most of our life domains, including not only our social life, but also our professional choices and success.

Successful human communication relies on the perception of a variety of socially relevant signals that can be conveyed by the whole body, face and voice. In the current dissertation, I focused on the ability to perceive signals from the visual face and the auditory voice (Figure 1). The face and voice allow us to recognize a person that is interacting with us (henceforth termed as face identity and voice identity respectively). We can understand the content of what the person is saying (henceforth referred to as visual speech and vocal speech respectively) and read their current emotional state (henceforth termed as facial emotion and vocal emotion respectively).

Fast and accurate perception of such visual and auditory communication signals is highly relevant to appropriately react in social situations and to adapt our behavior with respect to others. Therefore, we are tuned to perceive them from birth on (for review see Hyde et al., 2016; Streri et al., 2012), while our brain is well equipped with specialized regions dedicated to their processing (e.g., Belin et al., 2000; Kanwisher et al., 1997). A profound understanding of brain processes behind the perception of communication signals is required to elucidate the big question of what human communication is.

Although communication appears to be effortless to most of us, there are individuals who can have difficulties when communicating with others. One of these groups represent individuals with autism spectrum disorder (ASD). ASD refers to a neurodevelopmental disorder that is characterized by core difficulties in social interaction and communication (American Psychiatric Association [APA], 2013). Individuals with ASD have deficits in recognizing most of the visual and auditory communication signals (e.g., for review see Black et al., 2017; Lartseva et al., 2015; Weigelt et al., 2012). However, as ASD represents a spectrum condition characterized by a variable representation of symptoms and intellectual abilities, the presence and severity of such difficulties can vary. For example, perception of

vocal speech can remain relatively typical, at least under listening conditions without background noise (Schelinski et al., 2014; but see Alcántara et al., 2004). Until now, studies aiming to characterize the mechanisms underlying socio-communication deficits in ASD have put most emphasis on emotion processing, while perception of emotionally neutral communication signals, such as voice identity and visual speech have been investigated less intensively.

Figure 1. Schematic overview of some aspects of communication signals that can be extracted from the visual face and the auditory voice during face-to-face communication. For the communication signals highlighted in grey boxes, recognition difficulties have been reported in individuals with ASD. Vocal-speech recognition, highlighted in grey stripes, can remain relatively typical in ASD, at least under good listening conditions.

Individuals with ASD have difficulties recognizing others by voice and these deficits have been related to a dysfunction of an auditory sensory brain region implicated in the perception of voice-identity information (Schelinski et al., 2016; Schelinski et al., 2017). Because deficits in recognizing other people can have significant psychosocial consequences (Yardley et al., 2008; Fine, 2012), it is important to investigate therapeutic approaches aiming for alleviation of the impairments. Intranasal administration of oxytocin represents such an intervention and has recently gained increasing attention due to its efficacy in promoting perception of communication signals in typically developing individuals, as well as in individuals diagnosed with ASD (for review see Kendrick et al, 2017; Yamasue & Domes, 2017). In the first study of this dissertation (Study 1), I investigated the potential of

intranasal oxytocin to promote voice-identity processing in the human brain, particularly in the related auditory sensory brain regions.

Developing effective therapeutic approaches requires an understanding of the brain mechanisms that underlie the respective behavioral difficulties in the affected populations. Current evidence on visual-speech recognition in ASD is to date limited to behavioral findings (e.g., Foxe et al., 2015; Woynaroski et al., 2013). Therefore, in the next two studies of the present dissertation, I investigated what brain mechanisms contribute to visual-speech recognition deficits in ASD. I studied this question within two theoretical frameworks. First, I specifically referred to an ongoing debate on the contribution of sensory perceptual mechanisms to social communication difficulties in ASD (Baum et al., 2015; Robertson & Baron-Cohen, 2017). Thus, in Study 2, I aimed to disentangle at what level of processing atypical brain mechanisms contribute to visual-speech recognition difficulties in ASD. Second, I considered a current theoretical account for dynamic face perception (Bernstein & Yovel, 2015). This approach may be informative for explaining visual-speech recognition deficits in ASD, because visual speech is naturally embedded in the dynamic face. Thus, in Study 3, I investigated if functions of a brain network for dynamic face perception is altered during visual-speech recognition in ASD.

The following sections of the present dissertation provide a definition of ASD and a systematic overview of previous empirical research on the mechanisms underlying the perception of voice identity and visual speech in the typically developing population and in the ASD population. In addition, I give a short introduction to what oxytocin is. Subsequently, I derive my research questions from the review of the current literature and describe the empirical studies that I conducted during my PhD. Based on the study outcomes I infer implications for the ASD population and for the typically developing population. Finally, I conclude with suggestions for future research and practical applications in clinical populations.

2 Autism Spectrum Disorder

Autism spectrum disorder is a term referring to a group of heterogeneous neurodevelopmental conditions characterized by early-onset social difficulties and restricted behaviors and interests (APA, 2013; Lai et al., 2014). First clinical descriptions of ASD stem from Leo Kanner (1943) and Hans Asperger (1944), who described groups of children demonstrating qualitative impairments in reciprocal social interactions and communication and a restricted repertoire of interests and behaviors. Current characterization of the core symptoms of ASD are still largely based on these seminal delineations.

The 10th edition of the International Classification of Diseases (ICD-10) distinguishes between subtypes of ASD, such as childhood autism and Asperger's Syndrome, which are listed under the category of pervasive developmental disorders (PDD) (World Health Organization [WHO], 2004). Both disorders are commonly diagnosed based on the presence of the core symptom triad including qualitative impairments in reciprocal social interactions and patterns of communication, and a restricted, stereotyped, repetitive repertoire of interests and activities. In contrast to Asperger's Syndrome, the diagnosis of childhood autism requires the presence of impaired or delayed language development. The most recent edition of the Diagnostic and Statistical Manual of Mental Disorders [DSM-5] (APA, 2013) consolidated the subtypes under one single diagnosis of ASD. According to DSM-5, ASD can be characterized by additional specifiers with regard to current symptom severity and presence of intellectual and language impairments or other associated disorders or comorbidities. The triad of symptoms was also collapsed into a dyad, where social and communication difficulties were merged into one single domain.

The prevalence of ASD has been rising in the past years with current worldwide estimates at around 1-2% (e.g., Baron-Cohen et al., 2009; Idring et al., 2012; Kim et al., 2011; Saemundsen et al., 2013), also when considering adults only (Brugha et al., 2011). The increase of prevalence is likely due to increased awareness and recognition of ASD, younger age of diagnosis and changes in diagnostic concepts and criteria (Fisch et al., 2012; Fombonne et al., 2012). ASD is more often diagnosed in males than in females (Baron-Cohen et al., 2011). Twin studies provided strong evidence that genetics have a key role in the

etiology of ASD (Ronald & Hoekstra, 2011; Tick et al., 2016), but the influence of environmental factors and their interaction with genes have also been considered crucial (Grabrucker et al., 2013).

ASD has been also classified with regard to the level of intellectual functioning of the affected individuasl, where one distinguishes between "low-functioning" ASD (LFA) and "high-functioning" ASD (HFA) (Baron-Cohen, 2008). Individuals with LFA have intellectual abilities below a normal range as usually indicated by the intelligence quotient (IQ) below 85, while HFA refers to individuals with intellectual abilities within or above the normal range (IQ > 85) (Baron-Cohen, 2008). This classification has been widely adopted for research purposes. The term HFA usually refers to individuals with childhood autism with an IQ higher than 85 or to individuals with Asperger's Syndrome (Gillberg, 1998). In the current dissertation, study samples included individuals with HFA. From now onwards, I will use the term ASD to refer to individuals from this specific subpopulation.

Several theories exist that have aimed to describe mechanisms behind the core symptoms of ASD (for review see Baum et al., 2015; Lai et al., 2014). However, covering the complexity and heterogeneity of ASD symptoms has been a significant challenge for such theoretical frameworks and of little success so far. ASD is characterized by atypical information processing related to structural and functional brain atypicalities, which can be described at the *level of cognition* or at the *level of sensory perception*. So far, previous theories have put their focus on the cognitive perspective. Among these, the most prominent theory is the Theory of Mind (ToM; Baron-Cohen, 1989). It is based on evidence that individuals with ASD have difficulties inferring emotional and mental states of other people (Baron-Cohen et al., 1985; Happé, 1994), and these have been linked to reduced responses in frontal and temporal brain regions implicated in ToM (Castelli et al., 2002; Lombardo et al., 2011). Such atypicalities can be already observed for the primary cornerstones of ToM, such as reduced eye contact, joint attention or emotion perception (Dawson et al., 2004; Fridenson-Hayo, 2016; Senju & Johnson, 2009). Another theoretical account, The Social Motivation Theory, proposes that individuals with ASD are less interested and motivated to engage in social interactions with others (Chevallier et al., 2012). While these two theories are centered on the social domain, other prominent theoretical approaches target more general cognitive mechanisms. One of these theories is the Weak Central Coherence Theory

that refers to a preference and superiority in local information processing, together with diminished integration of information into one global meaning (Frith, 1989; Happé, 1999). Another central theory is the Executive Dysfunction Theory, which focuses on impairments in a broad range of cognitive abilities, such as cognitive flexibility, inhibitory control, planning and attention shifting (Hill, 2004; Ozonoff et al., 1991).

To date, atypical processing at the level of sensory perception has been less explored in comparison to atypicalities at the cognitive level. This is the case even though deficits in sensory processing are very common in ASD across different ages (Leekam et al., 2007) and recognized as a diagnostic criterion in the DSM-5 (APA, 2015). The sensory atypicalities can be observed across different modalities including vision and audition (for review see O'Connor, 2012; Simmons et al., 2009) and are characterized by heterogeneous manifestations of hyper- or hyporeactivity to sensory input (for review see Robertson & Baron-Cohen, 2017). The presence of sensory atypicalities is linked to both socio-communication deficits and repetitive behaviors in the childhood (Boyd et al., 2010; Turner-Brown et al., 2013). One theory considering atypical perceptual processing is The Predictive Coding Hypothesis of ASD (Pellicano & Burr, 2012; Van der Cruys et al., 2014). It assumes that sensory information is less influenced by the prior knowledge of the world making perceptual experience less predictable and more effortful for individuals with ASD (Pellicano & Burr, 2012). On the neural level, neuroimaging studies demonstrated that representations of sensory inputs are less reliable across time in ASD. This has been measured as reduced and more variable response patterns in sensory brain regions (e.g., Haigh et al., 2014; Milne et al., 2011; but see Coskun et al., 2009). Thus, atypical sensory processing can be described in terms of weaker statistical and neural representations.

Recently, there has been an ongoing debate about the contribution of atypical sensory and perceptual processing to socio-communication difficulties at the core of ASD (Baum et al., 2015; Thye et al., 2018). Perceptual representations of the sensory input build a cornerstone for its further processing in terms of cognitive and social mechanisms. Within the scope of this dissertation, I aimed to gain new insights into the contribution of atypical sensory brain mechanisms that underlie the perception of visual and auditory communication signals in ASD.

3 Mechanisms for the Perception of Visual and Auditory Communication Signals

In the following, I will outline brain mechanisms for the recognition of voice identity and their functional atypicalities in ASD. Moreover, I will summarize current evidence indicating that intranasal administration of oxytocin might enhance processing of voice-identity information. Subsequently, I will introduce brain mechanisms that underlie perception of dynamic faces and visual speech in typically developing individuals and provide an overview of behavioral findings on visual-speech recognition difficulties in ASD. I will focus only on those findings that are crucial for the current dissertation.

3.1 Mechanisms of Voice-Identity Perception

The human voice as an acoustic signal is an important source of several aspects of auditory communication signals including not only speech, but also paralinguistic information (Belin et al., 2004; Figure 1). One of the paralinguistic cues that we can extract from the voice is the identity of the person that we communicate with. Because of high relevance of this information for survival and communication success (Sidtis & Kreiman, 2012), we are tuned to attend to voice-identity information already before and right after birth (DeCasper & Fifer, 1980; Kisilevsky et al., 2003). Infants show a preference for their mother's voice in comparison to other unfamiliar female voices or to the voice of their father (Lee & Kisilevsky et al., 2014; Ockleford et al., 1988).

3.1.1 Brain Mechanisms for Voice-Identity Perception

Neuroimaging and lesion studies have provided evidence for the existence of specialized regions in the human brain that are dedicated to voice-identity processing in the typically developing population (for review see Roswandowitz et al., 2019). These regions are typically located in different portions of the bilateral superior temporal sulcus/ gyrus (STS/STG) and characterized by a right-hemispheric dominance (Belin et al., 2000; Belin et al., 2002; Roswandowitz et al., 2018; von Kriegstein et al., 2003; von Kriegstein et al., 2004). Along the right STS/STG, functionally different sub-regions have been related to specific

processing stages of voice-identity recognition. The posterior part of the right STS/STG together with parts of the primary auditory sensory regions, such as the Heschl's gyrus and the planum temporale, have been associated with the perceptual analysis of voice identity (Maguinness et al., 2018). This refers to perception of acoustic voice features that are relevant for voice-identity recognition including vocal pitch and vocal timbre (for review see Mathias & Kriegstein et al., 2014). The posterior STS/STG (pSTS/STG) is more responsive to unfamiliar compared to familiar voices (von Kriegstein et al., 2004; Zäske et al., 2017) presumably because their recognition might require a more elaborate acoustical analysis (Kreiman & Sidtis, 2011; Sidtis & Kreiman, 2012). The mid to anterior portions of the STS/STG are involved in the subsequent recognition of voice identity (Andics et al., 2010; Belin et al., 2002; von Kriegstein & Giraud, 2004). The voice-sensitive posterior and anterior STS/STG (aSTS/STG) share anatomical and functional connections (Blank et al., 2011; von Kriegstein et al., 2004) and are behaviorally relevant for voice-identity recognition (Schall et al., 2015; Schelinski et al., 2016).

3.1.2 Atypical Perception of Voice Identity in ASD

Already in the very first clinical descriptions of ASD, Kanner (1943) mentioned that children with ASD tend to orient less to vocal sounds. Later reports described that children with ASD lack preference for their mother's voice compared to sounds of noise created from superimposed voices (Klin et al., 1991). There is evidence that individuals with ASD have deficits recognizing paralinguistic voice information, such as vocal emotion and voice identity (Lartseva et al., 2015; Schelinski et al., 2017). In contrast, recognition of vocal speech can be relatively typical, at least under a good signal-to-noise ratio (Alcántara et al., 2004; Schelinski et al., 2014; Schelinski et al., 2016). For the purpose of this dissertation, I will only describe processes of voice-identity recognition in more detail.

Difficulties in voice-identity recognition can be observed in individuals with ASD across different ages (Boucher et al., 1998; Schelinski et al., 2017). Interestingly, in the typically developing population, the ability to recognize voice identity was shown to be related to autistic traits, where individuals with a higher level of autistic traits had a poorer performance on voice-identity recognition tasks (Skuk et al., 2017). A most recent study reported that adults with ASD compared to typically developing controls have more

difficulties discriminating between voices, and learning and recognizing unfamiliar voices (Schelinski et al., 2017; but see Lin et al., 2015). The deficits are accompanied by a selectively impaired discrimination of vocal pitch, but intact perception of vocal timbre and non-vocal pitch (Schelinski et al., 2017). In contrast, recognition of famous voices is comparable between the groups (Schelinski et al., 2017).

Schelinski et al. (2016) demonstrated that individuals with ASD compared to typically developing individuals have significantly decreased brain responses in the right pSTS/STG during the recognition of voice identity compared to speech. In contrast, the right aSTS/STG responses to the same contrast were comparable between the groups. Thus, atypical functions of the right pSTS/STG likely contribute to the behavioral deficits in voice-identity recognition in ASD.

The behavioral and the neural profile of the voice-identity recognition deficits in ASD converge to point towards a dysfunction at the perceptual level of voice-identity processing (Schelinski et al., 2016; Schelinski et al., 2017).

3.1.3 Potential of Intranasal Oxytocin to Enhance Voice-Identity Perception

Oxytocin is a nine-amino acid neuropeptide that is synthesized in the paraventricular nucleus and in the supraoptic nucleus of the hypothalamus in the human brain (Landgraf & Neumann, 2004; Sofroniew, 1983). It is released via the posterior pituitary acting both as a hormone in the peripheral bloodstream and as a neurotransmitter in the central nervous system (Churchland & Winkielman, 2012). Peripheral release of oxytocin is crucial for birth-related processes, such as contraction of uterus muscles during parturition and lactation (Gimpl & Fahrenholz, 2001). Central release of oxytocin has been related to socio-cognitive processes including affiliative behavior, stress regulation and social perception (for review see Bartz et al., 2011; Meyer-Lindenberg et al., 2011).

Oxytocin can also be administered externally. First findings stem from animal research where an injection of synthetic oxytocin into a rat brain elicited maternal behavior towards foster pups (Pedersen & Prange, 1979). A groundbreaking study by Born et al. (2002) provided a non-invasive method for human application by demonstrating that intranasal administration of neuropeptides allows them to enter the brain. Thus, intranasal administration is a useful approach to study central effects of oxytocin in humans. Since then

multiple human studies have administered oxytocin in form of a nasal spray to test its influence on behavior and brain mechanisms (for review see Grace et al., 2018; Stevens et al., 2013; Zink & Meyer-Lindenberg, 2012). Its beneficial influence was demonstrated for several social functions including anxiolytic effects, stress reduction, increased trust and improved mental attribution. The promising state of evidence has led to an intensive evaluation of oxytocin as a therapeutic intervention for clinical disorders, especially those characterized by social difficulties such as ASD (Parker et al., 2017; see also Neumann et al., 2016; van Zuiden et al., 2017). Individuals with ASD might be particularly sensitive to oxytocin treatment due to functional atypicalities of their oxytocinergic system (LoParo & Waldman., 2015; Modahl et al., 1998; but see Miller et al., 2013; Parker et al., 2014).

Intranasal oxytocin can promote social interaction and communication by enhancing the perception of visual and auditory communication signals (for review see Kanat et al., 2014; Kendrick et al, 2017; Yamasue & Domes, 2017). Most of the current evidence comes from studies on face processing. In typically developing individuals, intranasal oxytocin can facilitate recognition of facial emotion and face identity (Bate et al., 2014; Lischke et al., 2012; Rimmele et al., 2009) and elevated responses to face perception in visual sensory regions (Domes et al., 2014; Kanat et al., 2015). Similar oxytocin-related influences have been reported in ASD, where oxytocin improved the ability to infer the emotional state of others from the face (Domes et al., 2014; Guastella et al., 2010) and boosted responses in face-sensitive visual regions (Andari et al., 2016; Domes et al., 2014). In the auditory domain, current evidence stems only from studies conducted in ASD. Intranasal administration of oxytocin enhanced comprehension of vocal emotion (Hollander et al., 2007) and increased brain responses to emotional voices in auditory sensory regions including the bilateral STS/STG (Gordon et al., 2016).

To date, it is an open question whether oxytocin can facilitate processing of voice identity. Such an effect of oxytocin would especially benefit individuals with ASD to alleviate their voice-identity recognition difficulties and to boost the reduced responses in the right pSTS/STG (Schelinski et al., 2016; Schelinski et al., 2017). However, it is uncertain if intranasal oxytocin can elevate brain responses in the atypical right pSTS/STG in ASD. Thus, it remains unknown if the atypical sensory mechanisms for voice-identity recognition in ASD might modulate the efficacy of intranasal oxytocin administration.

3.2 Mechanisms of Visual-Speech Perception

Although speech perception is a predominantly auditory experience, speech information can be also extracted from the visual face (Figure 1). Visual speech refers to speech information that is inherent in the visible articulatory movements of the face. It is most effectively conveyed by movements of the lips, jaw, tongue and teeth, but movements in the extra-oral face parts, such as the cheeks and the larynx can be informative as well (Thomas & Jordan, 2004). The visual-speech signal contributes significantly to acquisition and comprehension of vocal speech (McGurk & MacDonald, 1976; Sumby & Pollack, 1954) and our ability to recognize it develops already during early infancy (Dodd, 1979; Kuhl & Meltzoff, 1984). Perception of visual speech supports speech development so that infants around the age of 6 to 10 months show a preference for looking at talker's mouth, likely to prepare for subsequent speech production (Hunnius & Geuze, 2004; Lewkowicz & Hansen-Tift, 2012). Moreover, visual speech compensates for hearing in deaf individuals and individuals with hearing impairments (Bernstein et al., 2000; Giraud et al., 2001; Rouger et al., 2007) and enhances comprehension of vocal speech in normal-hearing individuals (Schroeder et al., 2008; Van Wassenhove et al., 2005). Such an effect is especially beneficial under noisy conditions (Ross et al., 2007; Sumby & Pollack, 1954), but we tend to use the visual-speech signal also when the vocal speech is clear (Arnold & Hill, 2001; McGurk & MacDonald, 1976).

3.2.1 Brain Network for Dynamic Face Perception

The human face is dynamic by nature. When we talk, laugh or raise our eyebrows, our face changes rapidly. Fast and accurate perception of facial movement is essential for successful communication. Visual speech represents one of multiple social cues that can be conveyed by the moving face (Figure 1). Previous theoretical models have attempted to explain how the human brain perceives dynamic faces. For the purpose of this dissertation, I will delineate the theoretical approaches by Haxby et al. (2000) and by Bernstein & Yovel (2015). However, it is important to note that other theoretical models also exist that described brain mechanisms underlying face perception (e.g., Duchaine & Yovel, 2015; Haxby et al., 2000; O'Toole et al., 2002; Pitcher et al., 2014; Rossion et al., 2011).

Haxby et al. (2000) proposed a brain network that consists of a core system and an extended system. The core system is dedicated to the visual analysis of a face, while the extended system performs further analysis of the visual information extracted from the face. Here, I will only describe the core system. It consists of three face-sensitive visual sensory regions: (i) the occipital face area (OFA) located in the inferior occipital gyrus (Gauthier et al., 2000; Haxby et al., 1999), (ii) the pSTS/STG (Hoffman & Haxby, 2000; Kanwisher et al., 1997) and (iii) the fusiform face area (FFA) in the posterior portions of the fusiform gyrus (Kanwisher et al., 1997; McCarthy et al., 1997). The OFA likely facilitates an early stage of face perception, including the perceptual analysis of facial structure and features like the eyes and the nose (Liu et al., 2002; Liu et al. 2010). The OFA outputs to both the pSTS/STG and the FFA. The pSTS/STG is associated with processing of facial emotion, eye gaze and visual speech (Campbell et al., 2001; Engell & Haxby, 2007). The FFA is thought to form complex face representations that are relevant for discrimination and recognition of identity (Andrews & Ewbank, 2004, Grill-Spector et al., 2004). Thus, the core system includes one pathway extending into the dorsal temporal cortex (i.e., OFA and the pSTS/STG) and one pathway including regions in the ventral temporal cortex (i.e., OFA and FFA). The dorsal pathway is dedicated to processing of changeable (i.e., variant) face cues, while the ventral pathway is involved in processing of invariant information. However, previous empirical findings contradict this functional division. The pSTS/STG has been shown to be involved in the perception of invariant information (e.g., Anzellotti & Caramazza 2017, Dobs et al., 2018; Fox et al., 2011) while the FFA can be recruited for the perception of changeable information (e.g., LaBar et al., 2003; Sato et al., 2004; Schultz & Pilz, 2009).

The theoretical account by Bernstein and Yovel (2015) proposed a brain network that can be also divided into a dorsal and a ventral pathway. However, contrary to the model by Haxby et al. (2000), both pathways are concurrently recruited during perception of dynamic faces regardless of what information is exactly processed (i.e., changeable, invariant). The dorsal pathway is located along the visual motion area 5 (V5/MT) and the pSTS/STG. The V5/MT is an extrastriate visual area that is responsive to non-human and human movement (Grossman et al., 2000; Zeki et al., 1991), while the pSTS/STG is selectively responsive to human movement including facial movement (e.g., facial emotion, visual speech) (Grosbras et al., 2012; Grossman et al., 2005). The V5/MT and the pSTS/STG share structural and

functional connections and thus might be suitable to build one processing pathway (Bernstein et al., 2018; Ethofer et al., 2011; Furl et al., 2014). The ventral pathway includes the occipital face area (OFA) and the fusiform face area (FFA) that are also known to be structurally and functionally well connected to each other (Ethofer et al., 2011; Fairhall & Ishai, 2006; Gschwind et al., 2012; Pyles et al., 2013). The dorsal pathway is assumed to be important for processing movement information in the face, while the ventral pathway is involved in processing structural form information of the face. Henceforth, I will refer to these pathways as the dorsal-movement pathway and the ventral-form pathway, respectively (Figure 2A).

Movement and form information are facial cues that are both behaviorally relevant for successful perception of dynamic faces and their integration is required to yield a coherent percept of facial movement (Giese & Poggio, 2003; Maguinness & Newell, 2015). So far, it is unclear how movement and form information is integrated in the human brain. Network connectivity between the dorsal-movement pathway and the ventral-form pathway could provide such a suitable mechanism. Two studies demonstrated that the dorsal-movement regions and the ventral-form regions were functionally connected during facial emotion perception (Foley et al., 2012; Furl et al., 2014). To date, it remains unknown whether a similar network connectivity might be recruited for other face information such as visual speech.

3.2.2 Brain Network for Visual-Speech Perception

Visual speech represents one of many communication signals conveyed by the visual face. Therefore, besides the more general brain network for the perception of the dynamic face, the human brain has also developed a brain network that is more specifically dedicated to the processing of visual speech (for review see Bernstein & Liebenthal, 2014). This brain network can be broadly divided into visual sensory brain regions in the dorsal part of the temporal cortex and into brain regions in the frontal lobe. The first brain regions are associated with perception of visual movement in the face, while the latter are related to perception and production of speech. Henceforth, I will refer to them as dorsal-movement regions and speech-related regions, respectively (Figure 2B).

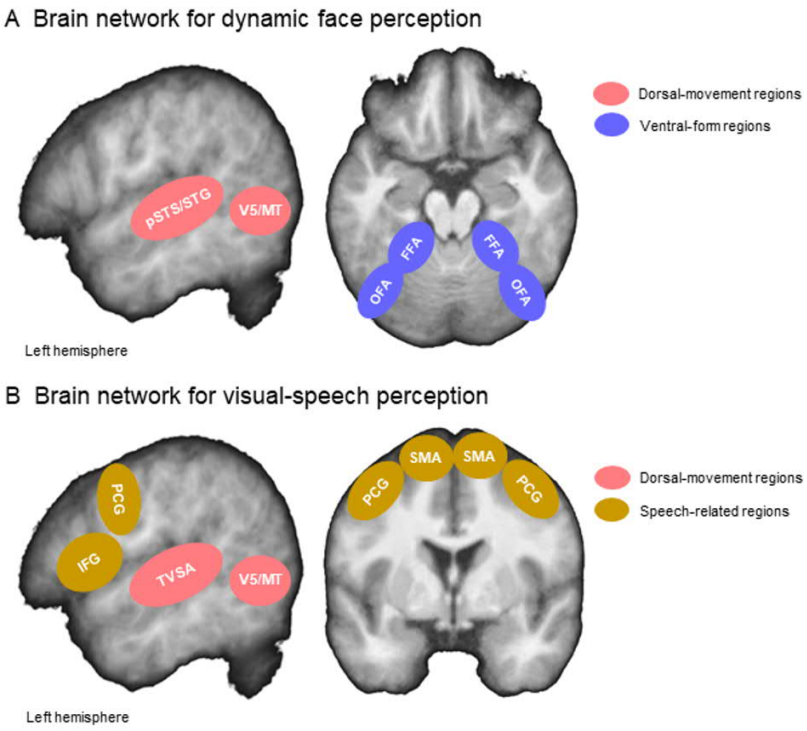

Figure 2. Schematic illustration of brain networks considered in the investigation of brain mechanisms underlying visual-speech recognition. pSTS/STG = posterior Superior Temporal Sulcus/ Gyrus; V5/MT = Visual Motion Area 5; OFA = Occipital Face Area; FFA = Fusiform Face Area; TVSA = Temporal Visual Speech Area; IFG = Inferior Frontal Gyrus; PCG = Precentral Gyrus; SMA = Supplementary Motor Area.

The dorsal-movement regions refer to the V5/MT and the pSTS/STG described in the previous section. The V5/MT and the pSTS/STG have been shown to be causally related to visual-speech recognition (Campbell et al., 1997; Riedel et al., 2015). Neuroimaging studies corroborated that the V5/MT is involved in visual-speech perception (Callan et al., 2014; Calvert & Campbell, 2003; Paulesu et al., 2003; Campbell et al., 1997). However, the V5/MT responses were not specific to visual speech, as they were lower to visual-speech movements compared to non-speech movements (Bernstein et al., 2011). In contrast, the pSTS/STG contains portions that are selectively sensitive to visual speech, which have been labeled as the so-called temporal visual speech area (TVSA; Bernstein et al., 2011). The TVSA in the left hemisphere might be particularly relevant for visual speech, because its responses were associated with behavioral recognition accuracy (Hall et al., 2005; Lee et al., 2007). Taken together, the V5/MT likely performs a general analysis of the movement in the face and then

forwards the information to the TVSA for a more complex and specific analysis of visual-speech information (Ethofer et al., 2011; Furl et al., 2014).

Visual-speech perception has been shown to elicit responses also in portions of the bilateral supplementary motor area (SMA), bilateral precentral gyrus (PCG) and the left inferior frontal gyrus (IFG) (Blank & von Kriegstein, 2013; MacSweeney et al., 2000; Okada & Hickok, 2009). These brain regions correspond to speech-related regions that are involved in both perception and production of speech (for review see Hickock & Poeppel, 2015; Tryfon et al., 2018). The PCG and the SMA are proposed to map the phonological representation of the perceived speech sound onto its articulatory representation, while the IFG is involved in mapping the speech percept onto its semantic meaning (for review see Friederici, 2011; Hickok & Poeppel, 2015). This is likely facilitated by anatomical connections between the speech-related regions and the pSTS/STG (for review see Friederici, 2015).

Previous study by Chu et al. (2013) shed light on the temporal dynamics of the brain network for visual-speech recognition indicating that the speech-related regions respond subsequently to the dorsal-movement regions during visual-speech perception. Such network dynamics suggest that visual-speech recognition could be broadly divided into two processing stages: (i) an earlier stage of facial-movement perception followed by (ii) a later stage where speech information is extracted from the movement.

3.2.3 Atypical Perception of Visual Speech in ASD

The ability to recognize visual speech is highly variable across normal-hearing individuals (Tye-Murray et al., 2014; Auer & Bernstein, 2007). Depending on the complexity of the visual-speech signal (i.e., syllable, isolated words, words in sentences), the recognition accuracy of normal-hearing individuals can vary between 0 and 75% (Bernstein et al., 2000). Impairments in visual-speech recognition have been reported in clinical conditions including ASD (Smith & Bennetto, 2007; see also Mohammed et al., 2006; Surguladze et al., 2001). Deficits in recognizing visual speech in ASD can be observed across different ages (Foxe et al., 2015; Schelinski et al., 2014). Individuals with ASD compared to typically developing controls show lower performance when asked to recognize syllables, single words and words embedded in sentences from visual-only articulatory movements (Iarocci et al., 2010;

Irwin et al., 2011; Woynaroski et al., 2013; but see also Saalasti et al., 2012). One of the potential mechanisms contributing to the recognition deficits in ASD could be atypical gaze behavior representing one of the core characteristics of ASD (for review see Guillon et al., 2014; Nomi & Uddin, 2015). Individuals with ASD show reduced gaze behavior to the face and to its most informative mouth area during visual-speech recognition (Irwin et al., 2011; Irwin & Brancazio, 2014). However, visual-speech recognition difficulties have been also observed, when gaze behavior in ASD was comparable to typically developing individuals (Foxe et al., 2015; Irwin et al., 2011).

Yet, it remains an open question what brain mechanisms underlie visual-speech recognition in ASD. On the one hand, individuals with ASD might be less able to recognize visual speech because they have difficulties in perceiving movement in the face due to atypical processing in the dorsal-movement regions. On the other hand, they might be challenged to extract speech information from the facial movement due to atypical mechanisms in the speech-related regions. Thus, according to the first scenario visual-speech recognition deficits in ASD would arise at the level of sensory perception, while according to the latter at the level of cognition. Considering findings about atypical gaze behavior in ASD, distinguishing between the two scenarios requires recording eye movements to assess if participants attend to visual-speech signal (Marassa & Lansing, 1995).

4 Questions Addressed in the Dissertation

The aim of the current dissertation was to examine the role that sensory brain regions that are related to perception of communication signals from the face and the voice, have for our understanding of some aspects of social communication difficulties in ASD. For this purpose, I tested the malleability of brain responses to voice-identity recognition using intranasal administration of oxytocin and investigated brain mechanisms behind visual-speech recognition difficulties in ASD. In the following, I will outline the scientific questions addressed in the three empirical studies that I conducted within the scope of this dissertation.

4.1 Does Oxytocin Modulate Voice-Identity Processing in the Human Brain?

Intranasal administration of oxytocin can augment behavioral and brain mechanisms for face and voice perception in typically developing individuals (e.g., Rimmele et al., 2009; Domes et al., 2014) and in individuals with ASD (e.g., Andari et al., 2016; Gordon et al., 2016). Currently, it is unknown whether recognition of voice identity and its underlying brain responses can be also boosted by intranasal oxytocin. Considering impairments of voice-identity recognition in ASD (Schelinski et al., 2016; Schelinski et al., 2017), such an oxytocin-related enhancement would be particularly beneficial for this population.

The aim of **Study 1** was to disclose whether intranasal administration of oxytocin can increase right pSTS/STG responses to voice-identity recognition and its behavioral accuracy in typically developing individuals and in individuals with ASD. However, in ASD, it was unclear if intranasal oxytocin would be able to increase the right pSTS/STG responses to voice-identity recognition due to atypical functioning of this brain region during voice-identity recognition.

4.2 What Brain Mechanisms Underlie Visual-Speech Recognition Deficits in ASD?

Individuals with ASD have difficulties with recognizing visual speech (e.g., Foxe et al., 2015; Schelinski et al., 2014). To date, it remains unknown what brain regions or networks

might contribute to the behavioral deficits. One of important factors to consider are decreased eye movements to the face in ASD, particularly to the mouth region (Irwin et al., 2011; Irwin & Brancazio, 2014; but see Foxe et al., 2015).

Study 2 aimed to disentangle at what processing level the visual-speech recognition deficits in ASD emerge. Considering the brain network for visual-speech recognition delineated in typically developing individuals, the deficits might be due to atypical brain mechanisms at the level of visual-movement perception ("dorsal-movement regions") or at the level of speech processing ("speech-related regions"). In addition, gaze behavior was recorded using eye tracking to assess if participants attended to the visual-speech information.

4.3 Are Dorsal-Movement Regions and Ventral-Form Regions Functionally Connected during Visual-Speech Recognition?

The theoretical account for dynamic face perception by Bernstein and Yovel (2015) proposed that visual sensory regions in the dorsal-movement pathway and in the ventral-form pathway are concurrently involved in processing of both changeable and invariant information in the dynamic face. Recent studies demonstrated that the dorsal-movement and the ventral-form regions are functionally connected during facial emotion perception in the typically developing population (Foley et al., 2012; Furl et al., 2014).

The first aim of **Study 3** was to explore whether such a dorsal-ventral functional connectivity also exists during visual-speech recognition in typically developing individuals. If that were the case, I aimed to examine the behavioral relevance of this functional connectivity, i.e., if its strength is associated with the accuracy of visual-speech recognition. The second aim of **Study 3** was to investigate whether alterations of the dorsal-ventral functional connectivity are related to the difficulties in visual-speech recognition in ASD.

5 Empirical Studies

5.1 Study 1 – Intranasal Oxytocin Modulates Voice-Identity Processing in the Human Brain

5.1.1 Summary Study 1

In **Study 1**, I investigated whether intranasal administration of oxytocin can improve recognition of voice identity and modulate the underlying brain responses in the right pSTS/STG in typically developing individuals and individuals diagnosed with ASD. Such an oxytocin-related enhancement would be particularly beneficial for individuals with ASD, who show behavioral and functional brain deficits during voice-identity processing (Schelinski et al., 2016). However, given the dysfunction of the right pSTS/STG in ASD, it was also likely that the efficacy of the intranasal oxytocin on the right pSTS/STG processing might be impeded or more variable in ASD. In **Study 1**, 18 adults with ASD and 18 typically developing controls (pairwise matched on age, gender, handedness and full-scale IQ) participated in an functional magnetic resonance imaging (fMRI) study that was conducted in a randomized, double-blind, placebo-controlled, within-subject, cross-over design. Participants were invited to two fMRI sessions where they received a nasal spray containing either oxytocin (24 International Units) or placebo. 45 minutes after substance administration, a voice-identity recognition experiment was conducted in the MRI environment. Participants listened to blocks of two-word sentences spoken by four male speakers and performed a speaker task and a speech task on the same stimuli. The sentences were assorted into blocks of 16 phonologically similar sentences. At the beginning of each block, participants were presented with a target and asked to decide whether the identity of speaker matched a target speaker (speaker task), or whether the content of a sentence matched a target sentence (speech task). The contrast "speaker task > speech task" was used to specifically target mechanisms for voice-identity recognition. I chose the right pSTS/STG as a region of interest (ROI), because it has been implicated in the perceptual analysis of voice-identity information in typically developing individuals (Maguinness et al., 2018) and it had reduced responses to the contrast "speaker task > speech task" in ASD (Schelinski et al., 2016).

Oxytocin compared to placebo increased the right pSTS/STG responses to "speaker task > speech task" only in the controls, but not in individuals with ASD. In the ASD group, I found

a positive correlation between the right pSTS/STG responses and speaker recognition accuracy in the oxytocin condition, but not in the placebo condition. There were no significant differences in the behavioral recognition accuracy in the speaker task after oxytocin compared to placebo in any of the groups. Finally, I replicated previous findings in ASD (Schelinski et al., 2016) by showing that ASD compared to typically developing controls had reduced Blood Oxygenation Level Dependent (BOLD) responses to the contrast "speaker task > speech task" in the right pSTS/STG.

The findings provide first evidence that brain responses to voice-identity recognition can be modulated by intranasal oxytocin in the typically developing population. In ASD, however, there were no significant changes in the right pSTS/STG responses. I assume that the dysfunction of the right pSTS/STG in ASD might be at the core of the overall lack of oxytocin-related modulation of the right pSTS/STG responses to voice identity in this population. This indicates that atypical sensory brain mechanisms not only contribute to at least some difficulties in social communication in ASD, but might also influence the efficacy of therapeutic interventions that aim to alleviate the behavioral deficits. The potential of intranasal oxytocin to enhance voice-identity processing is likely not universal, but rather dependent on the functional integrity of the underlying brain region (i.e., the right pSTS/STG).

5.1.2 Publication Study 1

Intranasal oxytocin modulates voice-identity recognition in typically developing individuals, but not in ASD

Kamila Borowiak[1,2,3] & Katharina von Kriegstein[2,3]

1 *Berlin School of Mind and Brain, Humboldt University of Berlin, Luisenstraße 56, 10117, Germany,*
2 *Technische Universität Dresden, Bamberger Straße 7, 01187 Dresden, Germany*
3 *Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany*

Published in

Borowiak, K., & von Kriegstein, K. (2020). Intranasal oxytocin modulates voice-identity recognition in typically developing individuals, but not in ASD. *Translational Psychiatry, 10,* 221.

https://doi.org/10.1038/s41398-020-00903-5

5.2 Study 2 - Recognizing Visual Speech: Reduced Responses in Visual-Movement Regions, but not other Speech Regions in Autism

5.2.1 Summary Study 2

In **Study 2**, I aimed to resolve at what level of processing atypical brain mechanisms contribute to visual-speech recognition difficulties in ASD. In a hypothesis-driven approach, I investigated eight ROIs that have been previously associated with visual-speech recognition (Blank & von Kriegstein, 2013): three dorsal-movement regions (bilateral V5/MT, left TVSA) and five speech-related regions (bilateral PCG, bilateral SMA, left IFG). I tested following hypotheses. First, if the behavioral deficits are due to atypical perception of visual movement in the face, I expected lower responses and/or lower functional connectivity in the dorsal-movement regions in ASD compared to typically developing individuals during recognition of visual speech. Second, if the visual-speech recognition difficulties are due to atypical mechanisms for speech processing, I expected comparable responses and functional connectivity in the dorsal-movement regions in both groups, but lower responses and/or functional connectivity in the speech-related regions in ASD in comparison to controls. 17 adults with ASD and 17 typically developing controls (pairwise matched on age, gender, handedness and full-scale IQ) participated in an fMRI study including a visual-speech recognition experiment and a functional ROI localizer. Concurrent eye tracking was recorded to monitor participants' eye movements to account for their potential influence on brain responses (Jiang et al., 2017). In the visual-speech recognition experiment, participants viewed blocks of visual-only videos of three male speakers articulating syllables and performed a visual-speech task and a face-identity task on the same stimuli. The syllables were assorted into blocks of nine videos considering the German viseme classes (Aschenberner & Weiss, 2005). At the beginning of each block, participants were presented with a target and then asked to decide whether a syllable matched a target syllable (visual-speech task), or whether the identity of the speaker matched the target person (face-identity task). The functional ROI localizer included passive viewing of moving and static faces and objects. The dorsal-movement ROIs were defined functionally using the functional ROI localizer and the speech-related ROIs were defined anatomically. I analyzed both local brain responses in the ROIs and their functional connectivity during recognition of visual speech

compared to recognition of face identity. I included eye movements as covariates of no interest into the analysis.

Behavioral results corroborated previous evidence that ASD individuals have difficulties recognizing visual speech and face identity. Gaze behavior was comparable between the groups. Local brain response analysis revealed that compared to the control group, the ASD group showed decreased BOLD responses to visual speech in contrast to face identity in the right V5/MT and in the left TVSA. Responses in the right V5/MT were positively correlated with visual-speech recognition accuracy in the ASD group, but not in the control group. Functional connectivity analysis revealed that connectivity between the dorsal-movement regions and between the left IFG was lower in the ASD group compared to the control group. Finally, BOLD responses in the speech-related regions and their functional connectivity in ASD were comparable to controls.

Our findings are the first to reveal brain mechanisms that underlie visual-speech recognition difficulties in ASD and point towards a dysfunction already at the stage of visual-movement perception in the dorsal-movement regions. Similar gaze behavior in both groups emphasizes the perceptual nature of the visual-speech recognition deficit in ASD, which is likely not due to atypical attention allocation to the face. Our findings support the view that at least some aspects of socio-communication difficulties in ASD can be attributed to atypical sensory perception (Baum et al., 2015; Robertson & Baron-Cohen, 2017; Thye et al., 2018) and challenge other views that have entirely linked them to non-perceptual mechanisms (Baron-Cohen, 1997; Chevallier et al., 2012).

5.2.2 Publication Study 2

Recognizing visual speech: Reduced responses in visual-movement regions, but not other speech regions in autism

Kamila Borowiak[1,2,3], Stefanie Schelinski[2,3] & Katharina von Kriegstein[2,3]

1 *Berlin School of Mind and Brain, Humboldt University of Berlin, Luisenstraße 56, 10117, Germany,*
2 *Technische Universität Dresden, Bamberger Straße 7, 01187 Dresden, Germany*
3 *Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany*

Published in

Borowiak, K., Schelinski, S., & von Kriegstein, K. (2018). Recognizing visual speech: Reduced responses in visual-movement regions, but not other speech regions in autism. *NeuroImage: Clinical*, *20*, 1078-1091.

https://doi.org/10.1016/j.nicl.2018.09.019

5.3 Study 3 - Dorsal-Movement and Ventral-Form Regions are Functionally Connected during Visual-Speech Recognition

5.3.1 Summary Study 3

In **Study 3,** I first examined whether dorsal-movement regions (V5/MT, pSTS/STG (i.e., TVSA)) and ventral-form regions (OFA, FFA) are functionally connected during visual-speech recognition in the typically developing population, and if such functional connectivity is behaviorally relevant for successful recognition of visual speech. Second, I tested if alterations of this dorsal-ventral functional connectivity contribute to visual-speech recognition deficits in ASD. To do this, I conducted a psycho-physiological interaction (PPI) analysis (Friston et al., 1997) on the same data that had been acquired for the purpose of Study 2, where 17 pairwise matched adults with ASD and 17 typically developing controls performed a visual-speech recognition experiment. The PPI analysis assesses the relationship between response changes in different brain regions that occur due to a specific task. Thus, I investigated if responses in the dorsal-movement regions (seed regions) are more related to responses in the ventral-form regions (target regions) during the visual-speech task compared to the face-identity task. Seed regions were defined in the bilateral V5/MT and in the bilateral TVSA in each individual participant. I first localized the individual V5/MT ROIs and the TVSA ROIs using the functional ROI localizer and then searched for the peak of the contrast "visual-speech task > face-identity task" that was located within the respective individual ROIs. Target regions were defined in the bilateral OFA and in the bilateral FFA within and across the groups. The bilateral OFA was defined based on the probabilistic atlas of face-sensitive brain regions (Engell & McCarthy, 2013). The bilateral FFA was localized using the functional ROI localizer.

The PPI analysis revealed that the bilateral dorsal-movement regions were functionally connected to the bilateral ventral-form regions during recognition of visual speech compared to face identity. This was the case in both typically developing controls and individuals with ASD. However, parts of the dorsal-ventral functional connectivity were reduced in ASD compared to controls (i.e., right V5/MT- right OFA, left TVSA – left FFA). The strength of the dorsal-ventral connectivity did not correlate with recognition accuracy in the visual-speech task in any of the groups.

The findings corroborate the theoretical account for dynamic face perception by Bernstein and Yovel (2015) by demonstrating that the dorsal-movement regions and the ventral-form regions are concurrently involved in perceiving information from dynamic faces, such as visual speech. This poses a challenge to the traditional model by Haxby et al. (2000) that attributed visual-speech processing only to the dorsal-movement region in the pSTS/STG (i.e., TVSA). Our findings give new insights into the processing of dynamic faces in ASD suggesting that atypicalities of the dorsal-ventral connectivity might contribute to dynamic face perception deficits known in ASD (Foxe et al., 2015; O'Brien et al., 2014; Sato et al., 2012)

5.3.2 Publication Study 3

Dorsal-movement and ventral form regions are functionally connected during visual-speech recognition.

Kamila Borowiak[1,2,3], Corrina Maguinness[2,3] & Katharina von Kriegstein[2,3]

1 *Berlin School of Mind and Brain, Humboldt University of Berlin, Luisenstraße 56, 10117, Germany,*
2 *Technische Universität Dresden, Bamberger Straße 7, 01187 Dresden, Germany*
3 *Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany*

Published in

Borowiak, K., Maguinness, C., & von Kriegstein, K. (2019). Dorsal-movement and ventral-form regions are functionally connected during visual-speech recognition. *Human Brain Mapping*, *41*(4), 952-972.

https://doi.org/10.1002/hbm.24852

6 General Discussion

The findings from the current dissertation lend new evidence and advance our understanding of the sensory brain mechanisms behind the perception of visual and auditory communication signals. They provide important implications for both the ASD and the typically developing population. First, the findings corroborate the view that atypical brain mechanisms at the level of sensory perception might be linked to difficulties in processing some aspects of audiovisual communication that are known in ASD (i.e., recognition of visual speech and voice identity). Second, they improve our current knowledge about how sensory brain regions process communication signals in typically developing individuals, namely the visual communication signals from the dynamic face. Finally, they give new insights into the mechanisms that might determine the efficacy of therapeutic interventions that aim for improving social perception during face-to-face communication.

6.1 Implications for Autism Spectrum Disorder

Most previous studies have considered socio-communication difficulties in ASD in terms of atypical processing at the level of cognition (for review see Lai et al., 2014; Leekam, 2016). The current dissertation delivers empirical support for the view that atypical sensory perception might play an important role for our understanding socio-communication deficits in ASD (Baum et al., 2015; Robertson & Baron-Cohen, 2017). It provides findings indicating that functional atypicalities of sensory perceptual brain regions may be crucial for understanding the etiology of at least some of the behavioral difficulties that are at the core of ASD. Moreover, it highlights that the efficacy of therapeutic interventions that aim to alleviate the behavioral deficits might be modulated by the functional integrity of these sensory brain regions.

The current findings emphasize that knowledge about functions in primary sensory regions outside of the so-called social brain for social processing (Adolphs, 2009; Brothers, 1990), such as the V5/MT, might be essential to advance our knowledge about how the brain of individuals with ASD processes communication signals. So far, mechanisms in the

STS/STG (i.e., TVSA) and the FFA, which are both part of the social brain, have been consistently related to social difficulties in ASD (for reviews see Bölte et al., 2015; Redcay, 2008; Scherf et al., 2015; Yang et al., 2015). It is therefore of particular interest that functional deficits of a region that is not specifically involved in perception of socially relevant information, i.e., V5/MT, were found to be associated with behavioral difficulties in visual-speech recognition in ASD. I propose that this is the case because more general (i.e., non-social) sensory mechanisms provide a critical foundation for further information processing in more specialized brain regions. Thus, sensory brain mechanisms might have a central role for explaining not only behavioral deficits, but also atypical functioning in brain regions that are typically implicated in social cognition. In the context of this dissertation, the atypical mechanisms in the V5/MT might contribute to less reliable perceptual representations of movement. This is in agreement with findings about impairments in general movement perception in ASD (Robertson et al., 2012; Tsermentseli et al., 2008), that have been linked to atypical V5/MT functioning (Robertson et al., 2014; Brieber et al., 2010). Less reliable movement representations likely impair the construction of more advanced representations of visual speech in other regions (Figure 2). On the one hand, functions of the TVSA in the dorsal-movement pathway might be affected – a region known to be specifically involved in processing the visual-speech signal. On the other hand, functions of the OFA and the FFA along the ventral-form pathway could be disturbed, and thus the perception of form information from the moving face (O'Toole et al., 2002). Such an effect of the V5/MT deficits is likely not specific to visual speech and may have an effect on processing of other socially relevant visual information. For example, one might expect that the V5/MT atypicalities could be related to difficulties in recognizing emotion and identity from the moving face (O'Brien et al., 2016; Sato et al., 2012), as Bernstein and Yovel (2015) proposed that the dorsal-ventral connectivity might function as a universal mechanism for perceiving dynamic faces.

An important contribution of the current work is to emphasize that increasing our knowledge about atypical perceptual brain mechanisms might also be relevant for the development of therapeutic interventions for improving social deficits in ASD. In Study 1, I found no significant effect of a pharmacological intervention with intranasal oxytocin on voice-identity processing in ASD. This might be due to the dysfunction of the sensory right

pSTS/STG during voice-identity recognition in ASD. This suggests that the right pSTS/STG likely not only determined the processing of the voice-identity information, but also modulated its malleability using intranasal administration of oxytocin. Thus, the potential of intranasal oxytocin to boost social perception might be reliant on the functional integrity of the underlying sensory brain regions. Additionally, the pSTS/STG was reported to have a reduced expression of oxytocin receptor genes (OXTR) in ASD (Gregory et al., 2009). This might be crucial because OXTR determine the numbering and functioning of oxytocin receptors to which intranasal oxytocin can bind (Yamasue et al., 2012). In contrast, the pSTS/STG in typically developing individuals is characterized by a high expression of OXTR (Paloyelis et al., 2016; Bethlehem et al., 2017). Such variations of the OXTR expression have been shown to be predictive for efficacy of intranasal oxytocin both in typically developing and in ASD individuals (Feng et al., 2015; Watanabe et al., 2017). Taken together, this suggests that not all therapeutic interventions might be suitable to enhance brain mechanisms that are related to aspects of socio-communication deficits in ASD.

From a more general perspective, there has been an ongoing debate about the validity of previous findings favoring the potential of intranasal administration of oxytocin to enhance behavior and brain processes (Lane et al., 2016; Leng & Ludwig, 2016; Young & Berrett, 2015; Walum et al., 2016). A meta-analysis of previous studies that specifically investigated oxytocin-related modulation of social cognition in ASD reported a non-significant low effect size (Ooi et al., 2017). Numerous clinical trials examining the potential of a long-term administration of intranasal oxytocin to reduce the severity of ASD symptoms also provided rather weak evidence for its efficacy (e.g., Anagnostou et al., 2012; Anagnostou et al., 2014; Guastella et al., 2015; Parker et al., 2017). The present findings support the critical view on intranasal oxytocin administration by indicating that its efficacy to improve social deficits in clinical populations might not be as promising and universal as initially presumed, at least not in ASD.

6.2 Implications for Typically Developing Individuals

The findings of the present dissertation contributed to a better understanding of brain mechanisms for the perception of visual and auditory communication signals also in

typically developing individuals. First, in the visual domain, they lend empirical support for the recent theoretical account for dynamic face perception proposed by Bernstein and Yovel (2015) by demonstrating that visual sensory regions in the dorsal-movement pathway and in the ventral-form pathway are functionally connected during visual-speech recognition. Second, in the auditory domain, they show that intranasal oxytocin has the potential to modulate responses in the sensory right pSTS/STG implicated in voice perception - an effect that in typically developing individuals had to date only been documented for visual sensory regions and face perception (Kanat et al., 2015; Domes et al., 2014; Domes et al., 2010).

The findings about the recruitment of the dorsal-ventral connectivity during visual-speech recognition corroborates the view that the dorsal-movement regions and the ventral-form regions are concurrently involved in the perception of different aspects of dynamic faces, including facial emotion and visual speech (Foley et al., 2012; Furl et al., 2014). Considering that Bernstein and Yovel (2015) proposed that the dorsal-ventral network might act as a universal mechanism for perceiving dynamic faces, it remains to be tested whether dorsal-ventral functional connectivity is also recruited during recognition of face identity. The dorsal-ventral connectivity could provide a suitable mechanism for an integration of movement and form information that is processed in each respective pathway, into one coherent percept of a dynamic face. Individuals with prosopagnosia - a selective deficit in recognizing face identity that has been linked to functional atypicalities along the ventral-form pathway (Gomez et al., 2015; Song et al., 2015) - show behavioral difficulties in integrating movement and form cues during face-identity recognition (Maguinness & Newell, 2015). This suggests that movement-form-integration might involve interactions between the dorsal-movement and the ventral-form regions. However, the theoretical model by Bernstein and Yovel (2015) did not propose how and when movement and form information could be integrated during dynamic face perception. The present findings leave the answer to this question open, because the PPI analysis provides no information about the time course and the directionality of the functional connectivity (Friston et al., 1997). Future studies applying methods which are more sensitive to temporal components e.g., magnetoencephalography (MEG) or to directionality e.g., the Dynamic Causal Modeling (DCM) (Friston et al., 2003) might help elucidate this matter.

The finding about the recruitment of the dorsal-ventral functional connectivity during visual-speech recognition also sheds new light on the role of non-social visual sensory regions such as the V5/MT for the perception of dynamic faces in typically developing individuals. So far, most of the theoretical brain models for dynamic face perception have focused mainly on social sensory regions linked to face perception including the pSTS/STG (i.e., TVSA), the OFA and the FFA (Duchaine & Yovel, 2015; Haxby et al., 2000; but see O'Toole et al., 2002). However, the V5/MT might have a key structural position in the architecture of the dorsal-ventral network. It is structurally well connected not only to the dorsal-movement pSTS/STG (Kim et al., 2006), but also to the OFA and the FFA in the ventral-form pathway (Ethofer et al., 2011; Kim et al., 2006). At the same time, the pSTS/STG and the two ventral-form regions share less structural connections (Ethofer et al., 2011; Gschwind et al., 2012; Pyles et al., 2013). Thus, the V5/MT is a suitable candidate region to facilitate the functional connectivity between the dorsal-movement and the ventral-form regions during dynamic face perception. Such a central position of the V5/MT is in agreement with the finding in Study 2 of this dissertation where reduced right V5/MT responses to visual speech were associated with lower visual-speech recognition accuracy in ASD. Further studies are needed to systematically elucidate the functional role of the V5/MT region in the dorsal-network connectivity and its contribution to recognition of dynamic faces. Application of transcranial magnetic stimulation (TMS) or transcranial direct-current stimulation (tDCS), which can induce virtual lesions in the brain (Gandiga et al., 2006; Pascual-Leone, 1999), could facilitate investigation of the causal involvement of the V5/MT in processing dynamic faces and visual speech in typically developing individuals.

The present dissertation demonstrated that intranasal administration of oxytocin can modulate responses in auditory sensory brain regions such as the pSTS/STG during processing of voice identity in typically developing individuals. So far, previous studies reported an oxytocin-related increase of responses to visual social stimuli in visual sensory brain regions (i.e., emotional faces) in contrast to non-social stimuli or to baseline conditions without any stimulation (e.g., Domes et al., 2010; Domes et al., 2013). The present finding advanced the current state of evidence by showing that intranasal oxytocin can boost brain responses to perception of social information, such as voice identity, when this is presented together with other social information, such as speech. This is important, because in every-

day life we are confronted with multiple aspects of such socially relevant communication signals at the same time. Yet, it is still unclear how or if these increased brain responses may contribute to a better comprehension of communication signals. Specifically, the enhanced right pSTS/STG responses to voice identity in Study 1 were not mirrored in an improved behavioral accuracy of voice-identity recognition under oxytocin. Similar findings have also been reported by other recent neuroimaging studies investigating effects of intranasal oxytocin on face recognition (e.g., Kanat et al., 2015; Domes et al., 2013).

The current dissertation raises important questions for future studies that aim to elucidate if intranasal administration of oxytocin can promote perception and comprehension of communication signals in natural every-day life situations. It is intriguing that significant oxytocin-related modulation of brain responses often has not been related to significant changes in behavior. One potential reason for that might be the experimental designs that were applied. Thus, I propose a few implications for future studies. First, experimental paradigms are needed which include two social conditions requiring a selective attention to one specific social information in the presence of other competing social signals. Second, such experimental paradigms should be validated with regard to their sensitivity to reliably detect oxytocin-related changes in behavior. Finally, the validated experimental paradigms could be brought into the MRI environment to test for how oxytocin-related changes in behavior are related to brain responses to social communication signals in comparison to other socially relevant information. In addition, communication signals during natural social interactions are often multimodal (e.g., audiovisual speech) requiring fast integration of components from different modalities. In contrast, most previous studies focused on information processing in visual or auditory modality separately. Two recent studies in ASD demonstrated that intranasal oxytocin can enhance brain responses to audiovisual emotion recognition, when the face and the voice are presented at the same time (Aoki et al., 2015; Watanabe et al., 2012). Future investigations should further explore if oxytocin can improve integration of visual and auditory components of communication signals. Taken together, I suggest that future studies in the field of oxytocin research should put more emphasis on the ecological validity of experimental designs.

7 Future Applications of the Current Findings

The current dissertation works towards a better understanding of fundamental components of audiovisual human communication: recognition of speech movements from the face and person identity from the voice. The characterization of the neuro-functional mechanisms that underlie these processes and their malleability provide novel incentives for future studies and practical applications. These are particularly relevant for individuals with ASD, but also other clinical populations that show communication difficulties.

Impairments of the visual sensory V5/MT region are likely central to the processing of visual communication signals from the dynamic face, at least for the visual speech. This indicates that interventions stimulating functions of the V5/MT might have the potential to alleviate such perceptual difficulties with visual communication signals that are known in ASD (e.g., O'Brien et al., 2014; Sato et al., 2012). Methods of brain stimulation, such as the TMS or the tDCS are suitable to enhance brain responses (Gandiga et al., 2006; Pascual-Leone, 1999). For example, in typically developing individuals, perception of coherent motion can be enhanced after tDCS stimulation of the V5/MT region (Antal, 2004; Battaglini et al., 2017). In addition, visual perceptual trainings have been shown to boost perception of coherent motion (Ball & Sekuler, 1982; Watanabe et al., 2001). Improved processing of general movement might contribute to its more reliable representations in the human brain, and thus to an improved recognition of visual communication signals that contain movement cues. I suggest that intervention programs combining brain stimulation approaches and behavioral tasks on visual-movement perception might have the potential to boost functioning of the dorsal-movement regions and thus to enhance perception of visual communication signals. Brain stimulation studies targeting the V5/MT could also allow us to investigate if improved perceptual representations of movement in the V5/MT could contribute to more efficient processing of the visual-speech signal in the TVSA or modulate the V5/MT functional connectivity to regions in the ventral-form pathway. Finally, studies testing tDCS stimulation over the left pSTS/STG (i.e., TVSA) in ASD are recommended, since tDCS stimulation of this region has been shown to improve visual-speech recognition in typically developing individuals (Riedel et al., 2015).

Future studies could also test if combinining behavioral trainings and brain stimulation might have any effect on voice-identity recognition in ASD. The lack of oxytocin-related enhancement of voice-identity processing in this population does not imply that other intervention approaches should not be effective. To date and to my knowledge, no previous study has specifically investigated if brain stimulation methods could enhance voice-identity processing. The finding that TMS over the right STS/STG can modulate discrimination between vocal and non-vocal information (Bestelmeyer et al., 2011) suggests that brain stimulation approaches could be effective to improve voice-identity recognition in the human brain. Furthermore, it would be interesting to examine if intranasal oxytocin administration can boost processing of visual speech in ASD, because this behavioral deficit is also underpinned by a dysfunction in portions of the pSTS/STG (i.e., TVSA). This could inform us if the failure of oxytocin to influence voice-identity processing in the pSTS/STG was due to the dysfunction of this region in ASD (in case of no improvement of visual-speech recognition) or due to other factors such as for example the auditory or emotionally neutral nature of the stimuli (i.e., in case of improved visual-speech recognition). Alternatively, one should examine if intranasal oxytocin could benefit visual-speech recognition via mechanisms in the V5/MT. Finally, further investigation of OXTR expression in the brain of individuals with ASD is required to characterize which brain regions, and thus processes, are suitable to benefit from intranasal oxytocin administration.

Developing therapeutic interventions to improve the perception of visual and auditory communication signals could benefit not only individuals with ASD, but also individuals diagnosed with other neurodevelopmental disorders such as schizophrenia spectrum disorders (SSD) and dyslexia (APA, 2013). SSD is of particular interest in the context of ASD symptomatology. ASD and SSD are both associated with social communication and interaction difficulties and there has been an ongoing discussion about a partially shared etiology of the two conditions (for review see Chisholm et al. 2015). In the context of social communication, SSD and dyslexia have been both associated with impaired perception of visual speech (de Gelder et al., 1998; de Gelder et al., 2003; Mohammed et al., 2006; Silverstein, 2016) and voice identity (Alba-Ferrara et al., 2012; Chhabra et al., 2012; Perea et al., 2014; Perrachione et al., 2011). Similar to ASD, the behavioral deficits in SSD and dyslexia might also be linked to atypical perceptual mechanisms such as deficits in pitch perception

(Gold et al., 2012; Ziegler et al., 2012) or general movement perception (Benassi et al., 2010; Silverstein et al., 2016). However, the brain mechanisms underlying these behavioral difficulties are largely unknown (Zhang et al., 2008). Comparative investigation of visual-speech and voice-identity processing in ASD, SSD and dyslexia could be informative to increase our understanding of the mechanisms behind the behavioral social difficulties and to create therapeutic approaches that are tailored for their improvement.

References

Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annual Review of Psychology*, *60*, 693-716.

Alba-Ferrara, L., Weis, S., Damjanovic, L., Rowett, M., & Hausmann, M. (2012). Voice identity recognition failure in patients with schizophrenia. *The Journal of Nervous and Mental Disease*, *200*, 784-790.

Alcántara, J. I., Weisblatt, E. J., Moore, B. C., & Bolton, P. F. (2004). Speech-in-noise perception in high-functioning individuals with autism or Asperger's syndrome. *Journal of Child Psychology and Psychiatry*, *45*, 1107-1114.

Anagnostou, E., Soorya, L., Brian, J., Dupuis, A., Mankad, D., Smile, S., & Jacob, S. (2014). Intranasal oxytocin in the treatment of autism spectrum disorders: a review of literature and early safety and efficacy data in youth. *Brain Research*, *1580*, 188-198.

Anagnostou, E., Soorya, L., Chaplin, W., Bartz, J., Halpern, D., Wasserman, S., Wang, T., Pepa, L., Tanel, N., Kushki, A., & Hollander, E. (2012). Intranasal oxytocin versus placebo in the treatment of adults with autism spectrum disorders: a randomized controlled trial. *Molecular Autism*, *3*, 16.

Andari, E., Richard, N., Leboyer, M., & Sirigu, A. (2016). Adaptive coding of the value of social cues with oxytocin, an fMRI study in autism spectrum disorder. *Cortex*, 76, 79-88.

Andics, A., McQueen, J. M., Petersson, K. M., Gál, V., Rudas, G., & Vidnyánszky, Z. (2010). Neural mechanisms for voice recognition. *Neuroimage*, *52*, 1528-1540.

Andrews, T. J., & Ewbank, M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *Neuroimage*, *23*, 905-913.

Antal, A., Varga, E. T., Nitsche, M. A., Chadaide, Z., Paulus, W., Kovács, G., & Vidnyánszky, Z. (2004). Direct current stimulation over MT+/V5 modulates motion aftereffect in humans. *Neuroreport*, *15*, 2491-2494.

Anzellotti, S., & Caramazza, A. (2017). Multimodal representations of person identity individuated with fMRI. *Cortex*, *89*, 85-97.

Arnold, P., & Hill, F. (2001). Bisensory augmentation: A speechreading advantage when speech is clearly audible and intact. *British Journal of Psychology*, *92*, 339-355.

Aschenberner, B., & Weiss, C. (2005). Phoneme-viseme mapping for german video-realistic audio-visual-speech-synthesis.

Asperger, H. (1944). Die „Autistischen Psychopathen" im Kindesalter. *European Archives of Psychiatry and Clinical Neuroscience*, *117*, 76-136.

Auer, E. T., & Bernstein, L. E. (2007). Enhanced visual speech perception in individuals with early-onset hearing impairment. *Journal of Speech, Language, and Hearing Research, 50,* 1157-1165.

Ball, K., & Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science*, *18*, 697-698.

Baron-Cohen, S. (1989). The autistic child's theory of mind: A case of specific developmental delay. *Journal of Child Psychology and Psychiatry*, *30*, 285-297.

Baron-Cohen, S. (2008). *Autism and Asperger syndrome*. Oxford University Press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"?. *Cognition*, *21*, 37-46.

Baron-Cohen, S., Lombardo, M. V., Auyeung, B., Ashwin, E., Chakrabarti, B., & Knickmeyer, R. (2011). Why are autism spectrum conditions more prevalent in males?. *PLoS Biology*, *9*, e1001081.

Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., & Brayne, C. (2009). Prevalence of autism-spectrum conditions: UK school-based population study. *The British Journal of Psychiatry*, *194*, 500-509.

Bartz, J. A., Zaki, J., Bolger, N., & Ochsner, K. N. (2011). Social effects of oxytocin in humans: context and person matter. *Trends in Cognitive Sciences*, *15*, 301-309.

Bate, S., Cook, S. J., Duchaine, B., Tree, J. J., Burns, E. J., & Hodgson, T. L. (2014). Intranasal inhalation of oxytocin improves face processing in developmental prosopagnosia. *Cortex, 50*, 55-63.

Battaglini, L., Noventa, S., & Casco, C. (2017). Anodal and cathodal electrical stimulation over V5 improves motion perception by signal enhancement and noise reduction. *Brain Stimulation*, *10*, 773-779.

Baum, S. H., Stevenson, R. A., & Wallace, M. T. (2015). Behavioral, perceptual, and neural alterations in sensory and multisensory function in autism spectrum disorder. *Progress in Neurobiology*, *134*, 140-160.

Belin, P., Fecteau, S., & Bedard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences*, *8*, 129-135.

Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Cognitive Brain Research*, *13*, 17-26.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*, 309.

Benassi, M., Simonelli, L., Giovagnoli, S., & Bolzani, R. (2010). Coherence motion perception in developmental dyslexia: a meta-analysis of behavioral studies. *Dyslexia*, 16, 341-357.

Bernstein, L. E., Demorest, M. E., and Tucker, P. E. (2000). Speech perception without hearing. *Perception & Psychophysics*, 62, 233–252.

Bernstein, M., Erez, Y., Blank, I., & Yovel, G. (2018). An Integrated Neural Framework for Dynamic and Static Face Processing. *Scientific Reports*, *8*, 7036.

Bernstein, L. E., Jiang, J., Pantazis, D., Lu, Z. L., & Joshi, A. (2011). Visual phonetic processing localized using speech and nonspeech face gestures in video and point-light displays. *Human Brain Mapping*, *32*, 1660-1676.

Bernstein, L. E., & Liebenthal, E. (2014). Neural pathways for visual speech perception. *Frontiers in Neuroscience*, *8*, 386.

Bernstein, L. E., Tucker, P. E., & Demorest, M. E. (2000). Speech perception without hearing. *Perception & Psychophysics*, *62*, 233-252.

Bernstein, M., & Yovel, G. (2015). Two neural pathways of face processing: a critical evaluation of current models. *Neuroscience & Biobehavioral Reviews*, *55*, 536-546.

Bestelmeyer, P. E., Belin, P., & Grosbras, M. H. (2011). Right temporal TMS impairs voice detection. *Current Biology*, *21*, 838-839.

Bethlehem, R. A. I., Lombardo, M. V., Lai, M. C., Auyeung, B., Crockford, S. K., Deakin, J., … & Baron-Cohen, S. (2017). Intranasal oxytocin enhances intrinsic corticostriatal functional connectivity in women. *Translational Psychiatry*, *7*, e1099.

Black, M. H., Chen, N. T., Iyer, K. K., Lipp, O. V., Bölte, S., … & Girdler, S. (2017). Mechanisms of facial emotion recognition in autism spectrum disorders: Insights from eye tracking and electroencephalography. *Neuroscience & Biobehavioral Reviews*, *80*, 488-515.

Blank, H., Anwander, A., & von Kriegstein, K. (2011). Direct structural connections between voice-and face-recognition areas. *Journal of Neuroscience*, *31*, 12906-12915.

Blank, H., & von Kriegstein, K. (2013). Mechanisms of enhancing visual–speech recognition by prior auditory information. *NeuroImage*, *65*, 109-118.

Bölte, S., Ciaramidaro, A., Schlitt, S., Hainz, D., Kliemann, D., Beyer, A., …. & Walter, H. (2015). Training-induced plasticity of the social brain in autism spectrum disorder. *The British Journal of Psychiatry, 207*, 149-157.

Born, J., Lange, T., Kern, W., McGregor, G. P., Bickel, U., & Fehm, H. L. (2002). Sniffing neuropeptides: a transnasal approach to the human brain. *Nature Neuroscience*, *5*, 514.

Boucher, J., Lewis, V., & Collis, G. (1998). Familiar face and voice matching and recognition in children with autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*, 171-181.

Boyd, B. A., Baranek, G. T., Sideris, J., Poe, M. D., Watson, L. R., Patten, E., & Miller, H. (2010). Sensory features and repetitive behaviors in children with autism and developmental delays. *Autism Research*, *3*, 78-87.

Brieber, S., Herpertz-Dahlmann, B., Fink, G. R., Kamp-Becker, I., Remschmidt, H., & Konrad, K. (2010). Coherent motion processing in autism spectrum disorder (ASD): an fMRI study. *Neuropsychologia*, *48*, 1644-1651.

Brooks, R., & Meltzoff, A. N. (2002). The importance of eyes: how infants interpret adult looking behavior. *Developmental Psychology*, *38*, 958.

Brothers, L. (1990). The social brain: a project for integrating primate behaviour and neurophysiology in a new domain. *Concepts Neuroscience*, *1*, 27-51.

Brugha, T. S., McManus, S., Bankart, J., Scott, F., Purdon, S., Smith, J., … & Meltzer, H. (2011). Epidemiology of autism spectrum disorders in adults in the community in England. *Archives of General Psychiatry*, *68*, 459-465.

Callan, D. E., Jones, J. A., & Callan, A. (2014). Multisensory and modality specific processing of visual speech in different regions of the premotor cortex. *Frontiers in Psychology*, *5*, 389.

Calvert, G. A., & Campbell, R. (2003). Reading speech from still and moving faces: the neural substrates of visible speech. *Journal of Cognitive Neuroscience*, *15*, 57-70.

Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J., … & David, A. S. (2001). Cortical substrates for the perception of face actions: an fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). *Cognitive Brain Research*, *12*, 233-243.

Campbell, R., Zihl, J., Massaro, D., Munhall, K., & Cohen, M. M. (1997). Speechreading in the akinetopsic patient, LM. *Brain: a Journal of Neurology*, *120*, 1793-1803.

Capek, C. M., MacSweeney, M., Woll, B., Waters, D., McGuire, P. K., David, A. S., … & Campbell, R. (2008). Cortical circuits for silent speechreading in deaf and hearing people. *Neuropsychologia*, *46*, 1233-1241.

Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, *125*(8), 1839-1849.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*, 231-239.

Chhabra, S., Badcock, J. C., Maybery, M. T., & Leung, D. (2012). Voice identity discrimination in schizophrenia. *Neuropsychologia*, *50*, 2730-2735.

Chisholm, K., Lin, A., Abu-Akel, A., & Wood, S. J. (2015). The association between autism and schizophrenia spectrum disorders: A review of eight alternate models of co-occurrence. *Neuroscience & Biobehavioral Reviews*, *55*, 173-183.

Chu, Y. H., Lin, F. H., Chou, Y. J., Tsai, K. W. K., Kuo, W. J., & Jääskeläinen, I. P. (2013). Effective cerebral connectivity during silent speech reading revealed by functional magnetic resonance imaging. *PLoS One*, *8*, e80265.

Churchland, P. S., & Winkielman, P. (2012). Modulating social behavior with oxytocin: how does it work? What does it mean?. *Hormones and Behavior*, *61*, 392-399.

Coskun, M. A., Varghese, L., Reddoch, S., Castillo, E. M., Pearson, D. A., Loveland, K. A., … & Sheth, B. R. (2009). Increased response variability in autistic brains?. *Neuroreport*, *20*, 1543-1548.

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Developmental Psychology*, *40*, 271.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208*, 1174-1176.

de Gelder, B., & Vroomen, J. (1998). Impaired speech perception in poor readers: Evidence from hearing and speech reading. *Brain and Language*, *64*, 269-281.

de Gelder, B., Vroomen, J., Annen, L., Masthof, E., & Hodiamont, P. (2003). Audio-visual integration in schizophrenia. *Schizophrenia Research*, *59*, 211-218.

DeRamus, T. P., & Kana, R. K. (2015). Anatomical likelihood estimation meta-analysis of grey and white matter anomalies in autism spectrum disorders. *NeuroImage: Clinical*, *7*, 525-536.

Dobs, K., Schultz, J., Bülthoff, I., & Gardner, J. L. (2018). Task-dependent enhancement of facial expression and identity representations in human cortex. *Neuroimage, 172*, 689-702.

Dodd, B. (1979). Lip reading in infants: Attention to speech presented in-and out-of-synchrony. *Cognitive Psychology*, *11*, 478-484.

Domes, G., Heinrichs, M., Kumbier, E., Grossmann, A., Hauenstein, K., & Herpertz, S. C. (2013). Effects of intranasal oxytocin on the neural basis of face processing in autism spectrum disorder. *Biological Psychiatry*, *74*, 164-171.

Domes, G., Kumbier, E., Heinrichs, M., & Herpertz, S. C. (2014). Oxytocin promotes facial emotion recognition and amygdala reactivity in adults with asperger syndrome. *Neuropsychopharmacology,* 39, 698.

Domes, G., Lischke, A., Berger, C., Grossmann, A., Hauenstein, K., Heinrichs, M., & Herpertz, S. C. (2010). Effects of intranasal oxytocin on emotional face processing in women. *Psychoneuroendocrinology*, *35*, 83-93.

Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science*, *1*, 393-416.

Engell, A. D., & McCarthy, G. (2013). Probabilistic atlases for face and biological motion perception: an analysis of their reliability and overlap. *Neuroimage, 74,* 140-151.

Estes, A., Zwaigenbaum, L., Gu, H., John, T. S., Paterson, S., Elison, J. T., ... & Kostopoulos, P. (2015). Behavioral, cognitive, and adaptive development in infants with autism spectrum disorder in the first 2 years of life. *Journal of Neurodevelopmental Disorders*, *7*, 24.

Ethofer, T., Gschwind, M., & Vuilleumier, P. (2011). Processing social aspects of human gaze: a combined fMRI-DTI study. *Neuroimage*, *55*, 411-419.

Fairhall, S. L., & Ishai, A. (2006). Effective connectivity within the distributed cortical network for face perception. *Cerebral Cortex, 17,* 2400-2406.

Feng, C., Lori, A., Waldman, I. D., Binder, E. B., Haroon, E., & Rilling, J. K. (2015). A common oxytocin receptor gene (OXTR) polymorphism modulates intranasal oxytocin effects on the neural response to social cooperation in humans. *Genes, Brain and Behavior*, *14*, 516-525.

Fine, D. R. (2012). A life with prosopagnosia. *Cognitive Neuropsychology*, *29*, 354-359.

Fisch, G. S. (2012). Nosology and epidemiology in autism: classification counts. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, *160*, 91-103.

Foley, E., Rippon, G., Thai, N. J., Longe, O., & Senior, C. (2012). Dynamic facial expressions evoke distinct activation in the face perception network: a connectivity analysis study. *Journal of Cognitive Neuroscience*, *24*, 507-520.

Fombonne, E., Quirke, S., & Hagen, A (2011). *Epidemiology of pervasive developmental disorders.* In: Amaral, D. G., Dawson, G., & Geschwind, D. H., eds. Autism Spectrum Disorders. New York, NY: Oxford University Press.

Fox, C. J., Hanif, H. M., Iaria, G., Duchaine, B. C., & Barton, J. J. (2011). Perceptual and anatomic patterns of selective deficits in facial identity and expression processing. *Neuropsychologia*, *49*, 3188-3200.

Foxe, J. J., Molholm, S., Del Bene, V. A., Frey, H. P., Russo, N. N., Blanco, D., … & Ross, L. A. (2013). Severe multisensory speech integration deficits in high-functioning school-aged children with autism spectrum disorder (ASD) and their resolution during early adolescence. *Cerebral Cortex*, *25*, 298-312.

Freitag, C. M., Konrad, C., Häberlen, M., Kleser, C., von Gontard, A., Reith, W., … & Krick, C. (2008). Perception of biological motion in autism spectrum disorders. *Neuropsychologia*, *46*, 1480-1494.

Fridenson-Hayo, S., Berggren, S., Lassalle, A., Tal, S., Pigat, D., Bölte, S., … & Golan, O. (2016). Basic and complex emotion recognition in children with autism: cross-cultural findings. *Molecular Autism*, 7, 52.

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological Reviews*, *91*, 1357-1392.

Friederici, A. D. (2015). White-matter pathways for speech and language processing. In *Handbook of Clinical Neurology* (Vol. 129, pp. 177-186). Elsevier.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage, 6*, 218-229.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *Neuroimage*, *19*, 1273-1302.

Frith, U. (1989). *Autism: Explaining the enigma.* Oxford: Blackwell.

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind". *Cognition*, *50*, 115-132.
Furl, N., Henson, R. N., Friston, K. J., & Calder, A. J. (2014). Network interactions explain sensitivity to dynamic faces in the superior temporal sulcus. *Cerebral Cortex*, *25*, 2876-2882.
Gamer, M., Zurowski, B., & Büchel, C. (2010). Different amygdala subregions mediate valence-related and attentional effects of oxytocin in humans. *Proceedings of the National Academy of Sciences*, *107*, 9400-9405.
Gandiga, P. C., Hummel, F. C., & Cohen, L. G. (2006). Transcranial DC stimulation (tDCS): a tool for double-blind sham-controlled clinical studies in brain stimulation. *Clinical Neurophysiology*, *117*, 845-850.
Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, *12*, 495-504.
Giese, M. A., & Poggio, T. (2003). Cognitive neuroscience: neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience*, *4*, 179.
Gillberg, C. (1998). Asperger syndrome and high-functioning autism. *The British Journal of Psychiatry*, *72*, 200-209.
Gimpl, G., & Fahrenholz, F. (2001). The oxytocin receptor system: structure, function, and regulation. *Physiological Reviews*, *81*, 629-683.
Giraud, A. L., Price, C. J., Graham, J. M., Truy, E., & Frackowiak, R. S. (2001). Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron*, *30*, 657-664.
Gold, R., Butler, P., Revheim, N., Leitman, D. I., Hansen, J. A., Gur, R. C., … & Javitt, D. C. (2012). Auditory emotion recognition impairments in schizophrenia: relationship to acoustic features and cognition. *American Journal of Psychiatry*, *169*, 424-432.
Gomez, J., Pestilli, F., Witthoft, N., Golarai, G., Liberman, A., Poltoratski, S., … & Grill-Spector, K. (2015). Functionally defined white matter reveals segregated pathways in human ventral temporal cortex associated with category-specific processing. *Neuron*, *85*, 216-227.
Gordon, I., Jack, A., Pretzsch, C. M., Vander Wyk, B., Leckman, J. F., Feldman, R., & Pelphrey, K. A. (2016). Intranasal oxytocin enhances connectivity in the neural circuitry supporting social motivation and social perception in children with autism. *Scientific Reports*, 6, 35054.
Grabrucker, A. M. (2013). Environmental factors in autism. *Frontiers in Psychiatry*, *3*, 118.

Grace, S. A., Rossell, S. L., Heinrichs, M., Kordsachia, C., & Labuschagne, I. (2018). Oxytocin and brain activity in humans: A systematic review and coordinate-based meta-analysis of functional MRI studies. *Psychoneuroendocrinology*, *96*, 6-24.

Gregory, S. G., Connelly, J. J., Towers, A. J., Johnson, J., Biscocho, D., Markunas, C. A., ... & Langford, C. F. (2009). Genomic and epigenetic evidence for oxytocin receptor deficiency in autism. *BMC Medicine*, 7, 62.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, *7*(5), 555.

Grosbras, M. H., Beaton, S., & Eickhoff, S. B. (2012). Brain regions involved in human movement perception: A quantitative voxel-based meta-analysis. *Human Brain Mapping*, *33*, 431-454.

Grossman, E. D., Battelli, L., & Pascual-Leone, A. (2005). Repetitive TMS over posterior STS disrupts perception of biological motion. *Vision Research*, *45*, 2847-2853.

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience*, *12*, 711-720.

Gschwind, M., Pourtois, G., Schwartz, S., Van De Ville, D., & Vuilleumier, P. (2012). White-matter connectivity between face-responsive regions in the human brain. *Cerebral Cortex, 22*, 1564-1576.

Guastella, A. J., Einfeld, S. L., Gray, K. M., Rinehart, N. J., Tonge, B. J., Lambert, T. J., & Hickie, I. B. (2010). Intranasal oxytocin improves emotion recognition for youth with autism spectrum disorders. *Biological Psychiatry*, 67, 692-694.

Guastella, A. J., Gray, K. M., Rinehart, N. J., Alvares, G. A., Tonge, B. J., Hickie, I. B., … & Einfeld, S. L. (2015). The effects of a course of intranasal oxytocin on social behaviors in youth diagnosed with autism spectrum disorders: a randomized controlled trial. *Journal of Child Psychology and Psychiatry*, *56*, 444-452.

Guillon, Q., Hadjikhani, N., Baduel, S., & Rogé, B. (2014). Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neuroscience & Biobehavioral Reviews*, *42*, 279-297.

Hall, D. A., Fussell, C., & Summerfield, A. Q. (2005). Reading fluent speech from talking faces: typical brain networks and individual differences. *Journal of Cognitive Neuroscience*, *17*, 939-953.

Haigh, S. M., Heeger, D. J., Dinstein, I., Minshew, N., & Behrmann, M. (2014). Cortical variability in the sensory-evoked response in autism. *Journal of Autism and Developmental Disorders*, *45*, 1176-1190.

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of autism and Developmental disorders*, *24*, 129-154.

Happé, F. (1999). Autism: cognitive deficit or cognitive style?. *Trends in Cognitive Sciences*, *3*, 216-222.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223-233.

Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron*, *22*, 189-199.

Herrington, J. D., Baron-Cohen, S., Wheelwright, S. J., Singh, K. D., Bullmore, E. T., Brammer, M., & Williams, S. C. (2007). The role of MT+/V5 during biological motion perception in Asperger Syndrome: An fMRI study. *Research in Autism Spectrum Disorders*, *1,* 14-27.

Hickok, G., & Poeppel, D. (2015). Neural basis of speech perception. In *Neurobiology of Language* (pp. 299-310). Academic Press.

Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental Review*, *24*, 189-233.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, *3*, 80.

Hollander, E., Bartz, J., Chaplin, W., Phillips, A., Sumner, J., Soorya, L., … & Wasserman, S. (2007). Oxytocin increases retention of social cognition in autism. B*iological Psychiatry*, *61*, 498-503.

Hunnius, S., & Geuze, R. H. (2004). Developmental changes in visual scanning of dynamic faces and abstract stimuli in infants: A longitudinal study. *Infancy*, *6*, 231-255.

Hyde, D. C., Flom, R., & Porter, C. L. (2016). Behavioral and neural foundations of multisensory face-voice perception in infancy. *Developmental Neuropsychology*, *41*, 273-292.
Iarocci, G., Rombough, A., Yager, J., Weeks, D. J., & Chua, R. (2010). Visual influences on speech perception in children with autism. *Autism*, *14*, 305-320.
Idring, S., Rai, D., Dal, H., Dalman, C., Sturm, H., Zander, E., … & Magnusson, C. (2012). Autism spectrum disorders in the Stockholm Youth Cohort: design, prevalence and validity. *PLoS One*, *7*, e41280.
Irwin, J. R., & Brancazio, L. (2014). Seeing to hear? Patterns of gaze to speaking faces in children with autism spectrum disorders. *Frontiers in Psychology*, *5*, 397.
Irwin, J. R., Tornatore, L. A., Brancazio, L., & Whalen, D. H. (2011). Can children with autism spectrum disorders "hear" a speaking face?. *Child Development*, *82*, 1397-1403.
Jiang, J., Borowiak, K., Tudge, L., Otto, C., & von Kriegstein, K. (2017). Neural mechanisms of eye contact when listening to another person talking. *Social Cognitive and Affective Neuroscience*, *12*, 319-328.
Kanat, M., Heinrichs, M., & Domes, G. (2014). Oxytocin and the social brain: neural mechanisms and perspectives in human research. *Brain Research*, *1580*, 160-171.
Kanat, M., Heinrichs, M., Schwarzwald, R., & Domes, G. (2015). Oxytocin attenuates neural reactivity to masked threat cues from the eyes. *Neuropsychopharmacology,* 40, 287.
Kanat, M., Spenthof, I., Riedel, A., van Elst, L. T., Heinrichs, M., & Domes, G. (2017). Restoring effects of oxytocin on the attentional preference for faces in autism. *Translational Psychiatry*, *7*, e1097.
Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, *2*, 217-250.
Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302-4311.
Kendrick, K. M., Guastella, A. J., & Becker, B. (2017). Overview of human oxytocin research. In *Behavioral Pharmacology of Neuropeptides: Oxytocin* (pp. 321-348). Springer, Cham.
Kim, Y. S., Leventhal, B. L., Koh, Y. J., Fombonne, E., Laska, E., Lim, E. C., ... & Song, D. H. (2011). Prevalence of autism spectrum disorders in a total population sample. *American Journal of Psychiatry*, *168*, 904-912.

Kisilevsky, B. S., Hains, S. M., Lee, K., Xie, X., Huang, H., Ye, H. H., … & Wang, Z. (2003). Effects of experience on fetal voice recognition. *Psychological Science*, *14*, 220-224.

Klin, A. (1991). Young autistic children's listening preferences in regard to speech: a possible characterization of the symptom of social withdrawal. *Journal of Autism and Developmental Disorders*, *21*, 29-42.

Koldewyn, K., Whitney, D., & Rivera, S. M. (2011). Neural correlates of coherent and biological motion perception in autism. *Developmental Science*, *14*, 1075-1088.

Kreiman, J., & Sidtis, D. (2011). Voices and listeners: Toward a model of voice perception. *Acoustics Today*, *7*, 7-15.

Kuhl, P. K., & Meltzoff, A. N. (1984). The intermodal representation of speech in infants. *Infant Behavior and Development, 7*, 361-381.

LaBar, K. S., Crupain, M. J., Voyvodic, J. T., & McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cerebral Cortex*, *13,* 1023-1033.

Lai, M. C., Lombardo, M. V., & Baron-Cohen, S. (2014). Autism. *Lancet, 383*, 896-910

Landgraf, R., & Neumann, I. D. (2004). Vasopressin and oxytocin release within the brain: a dynamic concept of multiple and variable modes of neuropeptide communication. *Frontiers in Neuroendocrinology*, *25*, 150-176.

Lane, A., Luminet, O., Nave, G., & Mikolajczak, M. (2016). Is there a publication bias in behavioural intranasal oxytocin research on humans? Opening the file drawer of one laboratory. *Journal of Neuroendocrinology*, *28*.

Lartseva, A., Dijkstra, T., & Buitelaar, J. K. (2015). Emotional language processing in autism spectrum disorders: a systematic review. *Frontiers in Human Neuroscience*, *8*, 991.

Lee, G. Y., & Kisilevsky, B. S. (2014). Fetuses respond to father's voice but prefer mother's voice after birth. *Developmental Psychobiology*, *56*, 1-11.

Lee, H. J., Truy, E., Mamou, G., Sappey-Marinier, D., & Giraud, A. L. (2007). Visual speech circuits in profound acquired deafness: a possible role for latent multimodal connectivity. *Brain*, *130*, 2929-2941.

Leekam, S. (2016). Social cognitive impairment and autism: what are we trying to explain?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*, 20150082.

Leekam, S. R., Nieto, C., Libby, S. J., Wing, L., & Gould, J. (2007). Describing the sensory abnormalities of children and adults with autism. J*ournal of Autism and Developmental Disorders*, *37*, 894-910.

Leng, G., & Ludwig, M. (2016). Intranasal oxytocin: myths and delusions. *Biological Psychiatry*, *79*, 243-250.

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences*, *109*, 1431-1436.

Lin, I. F., Yamada, T., Komine, Y., Kato, N., Kato, M., & Kashino, M. (2015). Vocal identity recognition in autism spectrum disorder. *PLoS One*, *10*, e0129451.

Lischke, A., Berger, C., Prehn, K., Heinrichs, M., Herpertz, S. C., & Domes, G. (2012). Intranasal oxytocin enhances emotion recognition from dynamic facial expressions and leaves eye-gaze unaffected. *Psychoneuroendocrinology*, 37, 475-481.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: an MEG study. *Nature Neuroscience*, *5*, 910.

Liu, J., Harris, A., & Kanwisher, N. (2010). Perception of face parts and face configurations: an fMRI study. *Journal of Cognitive Neuroscience*, *22*, 203-211.

Lombardo, M. V., Chakrabarti, B., Bullmore, E. T., Baron-Cohen, S., & MRC AIMS Consortium. (2011). Specialization of right temporo-parietal junction for mentalizing and its relation to social impairments in autism. *Neuroimage*, *56*, 1832-1838.

LoParo, D., & Waldman, I. D. (2015). The oxytocin receptor gene (OXTR) is associated with autism spectrum disorder: a meta-analysis. *Molecular Psychiatry*, *20*, 640.

MacSweeney, M., Amaro, E., Calvert, G. A., Campbell, R., David, A. S., McGuire, P., … & Brammer, M. J. (2000). Silent speechreading in the absence of scanner noise: an event-related fMRI study. *Neuroreport*, *11*, 1729-1733.

Maguinness, C., & Newell, F. N. (2015). Non-rigid, but not rigid, motion interferes with the processing of structural face information in developmental prosopagnosia. *Neuropsychologia*, *70*, 281-295.

Maguinness, C., Roswandowitz, C., & von Kriegstein, K. (2018). Understanding the mechanisms of familiar voice-identity recognition in the human brain. *Neuropsychologia*, *116*, 179-193.

Marassa, L. K., & Lansing, C. R. (1995). Visual word recognition in two facial motion conditions: Full-face versus lips-plus-mandible. *Journal of Speech, Language, and Hearing Research,* 38, 1387-1394.

Mathias, S. R., & von Kriegstein, K. (2014). How do we recognise who is speaking. *Frontiers in Bioscience (Schol Ed)*, *6*, 92-109.

McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, *9*, 605-610.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746.

McLeod, P. (1996). Preserved and Impaired Detection of Structure From Motion by a'Motion-blind" Patient. *Visual Cognition*, *3*, 363-392.

Meyer-Lindenberg, A., Domes, G., Kirsch, P., & Heinrichs, M. (2011). Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. *Nature Reviews Neuroscience*, *12*, 524.

Miller, M., Bales, K. L., Taylor, S. L., Yoon, J., Hostetler, C. M., Carter, C. S., & Solomon, M. (2013). Oxytocin and vasopressin in children and adolescents with autism spectrum disorders: sex differences and associations with symptoms. *Autism Research*, *6*, 91-102.

Milne, E. (2011). Increased intra-participant variability in children with autistic spectrum disorders: evidence from single-trial analysis of evoked EEG. *Frontiers in Psychology*, *2*, 51.

Modahl, C., Green, L. A., Fein, D., Morris, M., Waterhouse, L., Feinstein, C., & Levin, H. (1998). Plasma oxytocin levels in autistic children. *Biological Psychiatry*, *43*, 270-277.

Mohammed, T., Campbell, R., Macsweeney, M., Barry, F., & Coleman, M. (2006). Speechreading and its association with reading among deaf, hearing and dyslexic individuals. *Clinical Linguistics & Phonetics*, *20*, 621-630.

Neumann, I. D., & Slattery, D. A. (2016). Oxytocin in general anxiety and social fear: a translational approach. *Biological Psychiatry*, *79*, 213-221.

Nickl-Jockschat, T., Habel, U., Maria Michel, T., Manning, J., Laird, A. R., Fox, P. T., … & Eickhoff, S. B. (2012). Brain structure anomalies in autism spectrum disorder-a meta-analysis of VBM studies using anatomic likelihood estimation. *Human Brain Mapping*, *33*, 1470-1489.

Nomi, J. S., & Uddin, L. Q. (2015). Face processing in autism spectrum disorders: from brain regions to brain networks. *Neuropsychologia*, *71*, 201-216.

O'Brien, J., Spencer, J., Girges, C., Johnston, A., & Hill, H. (2014). Impaired perception of facial motion in autism spectrum disorder. *PLoS One*, *9*(7), e102173.

Ockleford, E. M., Vince, M. A., Layton, C., & Reader, M. R. (1988). Responses of neonates to parents' and others' voices. *Early Human Development*, *18*, 27-36.

O'Connor, K. (2012). Auditory processing in autism spectrum disorder: a review. *Neuroscience & Biobehavioral Reviews*, *36*, 836-854.

Okada, K., & Hickok, G. (2009). Two cortical mechanisms support the integration of visual and auditory speech: A hypothesis and preliminary data. *Neuroscience Letters*, *452*, 219-223.

Ooi, Y. P., Weng, S. J., Kossowsky, J., Gerger, H., & Sung, M. (2017). Oxytocin and autism spectrum disorders: a systematic review and meta-analysis of randomized controlled trials. *Pharmacopsychiatry*, *50*, 5-13.

O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. *Trends in Cognitive Sciences*, *6*, 261-266.

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: relationship to theory of mind. Journal of Child Psychology and Psychiatry, and Allied Disciplines, 32, 1081-1105.

Paloyelis, Y., Doyle, O. M., Zelaya, F. O., Maltezos, S., Williams, S. C., Fotopoulou, A., & Howard, M. A. (2016). A spatiotemporal profile of in vivo cerebral blood flow changes following intranasal oxytocin in humans. *Biological Psychiatry*, *79*, 693-705.

Parker, K. J., Garner, J. P., Libove, R. A., Hyde, S. A., Hornbeak, K. B., Carson, D. S., … & Hardan, A. Y. (2014). Plasma oxytocin concentrations and OXTR polymorphisms predict social impairments in children with and without autism spectrum disorder. *Proceedings of the National Academy of Sciences*, *111*, 12258-12263.

Parker, K. J., Oztan, O., Libove, R. A., Sumiyoshi, R. D., Jackson, L. P., Karhson, D. S., ... & Carson, D. S. (2017). Intranasal oxytocin treatment for social deficits and biomarkers of response in children with autism. *Proceedings of the National Academy of Sciences*, *114*, 8119-8124.

Pascual-Leone, A. (1999). Transcranial magnetic stimulation: studying the brain-behaviour relationship by induction of 'virtual lesions'. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *354*, 1229-1238.

Paulesu, E., Perani, D., Blasi, V., Silani, G., Borghese, N. A., De Giovanni, U., … & Fazio, F. (2003). A functional-anatomical model for lipreading. *Journal of Neurophysiology*, *90*, 2005-2013.

Pedersen, C. A., & Prange, A. J. (1979). Induction of maternal behavior in virgin rats after intracerebroventricular administration of oxytocin. *Proceedings of the National Academy of Sciences*, *76*, 6661-6665.

Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, *16*, 504-510.

Perea, M., Jiménez, M., Suárez-Coalla, P., Fernández, N., Viña, C., & Cuetos, F. (2014). Ability for voice recognition is a marker for dyslexia in children. *Experimental Psychology*.

Perrachione, T. K., Del Tufo, S. N., & Gabrieli, J. D. (2011). Human voice recognition depends on language ability. *Science*, *333*, 595-595.

Pitcher, D., Duchaine, B., & Walsh, V. (2014). Combined TMS and fMRI reveal dissociable cortical pathways for dynamic and static face perception. *Current Biology*, *24*, 2066-2070.

Pyles, J. A., Verstynen, T. D., Schneider, W., & Tarr, M. J. (2013). Explicating the face perception network with white matter connectivity. *PLoS One*, *8*, e61611.

Riedel, P., Ragert, P., Schelinski, S., Kiebel, S. J., & von Kriegstein, K. (2015). Visual face-movement sensitive cortex is relevant for auditory-only speech recognition. *Cortex*, *68*, 86-99.

Redcay, E. (2008). The superior temporal sulcus performs a common function for social and speech perception: implications for the emergence of autism. *Neuroscience & Biobehavioral Reviews*, *32*, 123-142.

Robertson, C. E., & Baron-Cohen, S. (2017). Sensory perception in autism. *Nature Reviews Neuroscience*, *18*, 671.

Robertson, C. E., Martin, A., Baker, C. I., & Baron-Cohen, S. (2012). Atypical integration of motion signals in autism spectrum conditions. *PLoS One*, *7*, e48173.

Robertson, C. E., Thomas, C., Kravitz, D. J., Wallace, G. L., Baron-Cohen, S., Martin, A., & Baker, C. I. (2014). Global motion perception deficits in autism are reflected as early as primary visual cortex. *Brain*, *137*, 2588-2599.

Ronald, A., & Hoekstra, R. A. (2011). Autism spectrum disorders and autistic traits: a decade of new twin studies. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *156*, 255-274.

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., & Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex*, *17*, 1147-1153.

Rossion, B., Dricot, L., Goebel, R., & Busigny, T. (2011). Holistic face categorization in higher order visual areas of the normal and prosopagnosic brain: toward a non-hierarchical view of face perception. *Frontiers in Human Neuroscience*, *4*, 225.

Roswandowitz, C., Kappes, C., Obrig, H., & von Kriegstein, K. (2017). Obligatory and facultative brain regions for voice-identity recognition. *Brain*, *141*, 234-247.

Roswandowitz, C**.,** Maguinness, C., & von Kriegstein, K. (2019) *Deficits in Voice-Identity Processing: Acquired and Developmental Phonagnosia.* In: The Oxford Handbook of Voice Perception (Frühholz, S., & Belin, P., ed).

Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., & Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proceedings of the National Academy of Sciences*, *104*, 7295-7300.

Saalasti, S., Kätsyri, J., Tiippana, K., Laine-Hernandez, M., von Wendt, L., & Sams, M. (2012). Audiovisual speech perception and eye gaze behavior of adults with Asperger syndrome. *Journal of Autism and Developmental Disorders*, *42*, 1606-1615.

Saemundsen, E., Magnússon, P., Georgsdóttir, I., Egilsson, E., & Rafnsson, V. (2013). Prevalence of autism spectrum disorders in an Icelandic birth cohort. *BMJ Open*, *3*, e002748.

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. *Cognitive Brain Research*, *20*, 81-91.

Sato, W., Toichi, M., Uono, S., & Kochiyama, T. (2012). Impaired social brain network for processing dynamic facial expressions in autism spectrum disorders. *BMC Neuroscience*, *13*, 99.

Schall, S., Kiebel, S. J., Maess, B., & von Kriegstein, K. (2015). Voice identity recognition: functional division of the right STS and its behavioral relevance. *Journal of Cognitive Neuroscience*, *27*, 280-291.

Schelinski, S., Borowiak, K., & von Kriegstein, K. (2016). Temporal voice areas exist in autism spectrum disorder but are dysfunctional for voice identity recognition. *Social Cognitive and Affective Neuroscience*, *11*, 1812-1822.

Schelinski, S., Riedel, P., & von Kriegstein, K. (2014). Visual abilities are important for auditory-only speech recognition: evidence from autism spectrum disorder. *Neuropsychologia*, *65*, 1-11.

Schelinski, S., Roswandowitz, C., & von Kriegstein, K. (2017). Voice identity processing in autism spectrum disorder. *Autism Research*, *10*, 155-168.

Scherf, K. S., Elbich, D., Minshew, N., & Behrmann, M. (2015). Individual differences in symptom severity and behavior predict neural activation during face processing in adolescents with autism. *NeuroImage: Clinical, 7*, 53-67.

Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., & Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences*, *12*, 106-113.

Schultz, J., & Pilz, K. S. (2009). Natural facial motion enhances cortical responses to faces. *Experimental Brain Research*, 194, 465-475.

Senju, A., & Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends in Cognitive Sciences*, *13*, 127-134.

Sidtis, D., & Kreiman, J. (2012). In the beginning was the familiar voice: personally familiar voices in the evolutionary and contemporary biology of communication. *Integrative Psychological and Behavioral Science*, *46*, 146-159.

Silverstein, S. M. (2016). Visual perception disturbances in schizophrenia: a unified model. In *The Neuropsychopathology of Schizophrenia* (pp. 77-132). Springer, Cham.

Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., & Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision Research*, *49*, 2705-2739.

Skuk, V. G., Palermo, R., Broemer, L., & Schweinberger, S. R. (2017). Autistic traits are linked to individual differences in familiar voice identification. *Journal of Autism and Developmental Disorders*, 1-21.

Smith, E. G., & Bennetto, L. (2007). Audiovisual speech integration and lipreading in autism. *Journal of Child Psychology and Psychiatry*, *48*, 813-821.

Sofroniew, M. V. (1983). Morphology of vasopressin and oxytocin neurones and their central and vascular projections. *Progress in Brain research*, *60*, 101-114.

Song, S., Garrido, L., Nagy, Z., Mohammadi, S., Steel, A., Driver, J., … & Furl, N. (2015). Local but not long-range microstructural differences of the ventral temporal cortex in developmental prosopagnosia. *Neuropsychologia*, *78*, 195-206.

Stevens, F. L., Wiesman, O., Feldman, R., Hurley, R. A., & Taber, K. H. (2013). Oxytocin and behavior: evidence for effects in the brain. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *25*, 96-102.

Streri, A. (2012). Crossmodal interactions in the human newborn. *Multisensory Development*, 88-112.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America*, *26*, 212-215.

Surguladze, S. A., Calvert, G. A., Brammer, M. J., Campbell, R., Bullmore, E. T., Giampietro, V., & David, A. S. (2001). Audio–visual speech perception in schizophrenia: an fMRI study. *Psychiatry Research: Neuroimaging*, *106*, 1-14.

Thomas, S. M., & Jordan, T. R. (2004). Contributions of oral and extraoral facial movement to visual and audiovisual speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 873.

Thye, M. D., Bednarz, H. M., Herringshaw, A. J., Sartin, E. B., & Kana, R. K. (2018). The impact of atypical sensory processing on social impairments in autism spectrum disorder. *Developmental Cognitive Neuroscience*, *29*, 151-167.

Tick, B., Bolton, P., Happé, F., Rutter, M., & Rijsdijk, F. (2016). Heritability of autism spectrum disorders: a meta-analysis of twin studies. *Journal of Child Psychology and Psychiatry*, *57*, 585-595.

Tryfon, A., Foster, N. E., Sharda, M., & Hyde, K. L. (2018). Speech perception in autism spectrum disorder: An activation likelihood estimation meta-analysis. *Behavioural Brain Research*, *338*, 118-127.

Tsermentseli, S., O'Brien, J. M., & Spencer, J. V. (2008) Comparison of form and motion coherence processing in autistic spectrum disorders and dyslexia. *Journal of Autism and Developmental Disorders*, 38, 1201–1210.

Turner-Brown, L. M., Baranek, G. T., Reznick, J. S., Watson, L. R., & Crais, E. R. (2013). The First Year Inventory: a longitudinal follow-up of 12-month-old to 3-year-old children. *Autism*, *17*, 527-540.

Tye-Murray, N., Hale, S., Spehar, B., Myerson, J., & Sommers, M. S. (2014). Lipreading in school-age children: the roles of age, hearing status, and cognitive ability. *Journal of Speech, Language, and Hearing Research*, *57*, 556-565.

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review*, *121*, 649.

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences*, *102*, 1181-1186.

van Zuiden, M., Frijling, J. L., Nawijn, L., Koch, S. B., Goslings, J. C., Luitse, J. S., … & Olff, M. (2017). Intranasal oxytocin to prevent posttraumatic stress disorder symptoms: a randomized controlled trial in emergency department patients. *Biological Psychiatry*, *81*, 1030-1040.

von Kriegstein, K., Eger, E., Kleinschmidt, A., & Giraud, A. L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Cognitive Brain Research*, *17*, 48-55.

von Kriegstein, K., & Giraud, A. L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage*, *22*, 948-955.

von Kriegstein, K., Smith, D. R., Patterson, R. D., Ives, D. T., & Griffiths, T. D. (2007). Neural representation of auditory size in the human voice and in sounds from other resonant sources. *Current Biology*, *17*, 1123-1128.

Walum, H., Waldman, I. D., & Young, L. J. (2016). Statistical and methodological considerations for the interpretation of intranasal oxytocin studies. *Biological Psychiatry*, *79*, 251-257.

Watanabe, T., Náñez, J. E., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature*, *413*, 844.

Watanabe, T., Otowa, T., Abe, O., Kuwabara, H., Aoki, Y., Natsubori, T., … & Yamasue., H. (2017). Oxytocin receptor gene variations predict neural and behavioral response to oxytocin in autism. *Social Cognitive and Affective Neuroscience*, *12*, 496-506.

Weigelt, S., Koldewyn, K., & Kanwisher, N. (2012). Face identity recognition in autism spectrum disorders: a review of behavioral studies. *Neuroscience & Biobehavioral Reviews*, *36*, 1060-1084.

World Health Organization (2004). *International Statistical Classification of Diseases and Related and Related Health Problems (10th ed.).* World Health Organization, Geneva.

Woynaroski, T. G., Kwakye, L. D., Foss-Feig, J. H., Stevenson, R. A., Stone, W. L., & Wallace, M. T. (2013). Multisensory speech perception in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *43*, 2891-2902.

Yamasue, H., & Domes, G. (2017). Oxytocin and autism spectrum disorders. In *Behavioral Pharmacology of Neuropeptides: Oxytocin* (pp. 449-465). Springer, Cham.

Yamasue, H., Yee, J. R., Hurlemann, R., Rilling, J. K., Chen, F. S., Meyer-Lindenberg, A., & Tost, H. (2012). Integrative approaches utilizing oxytocin to enhance prosocial behavior: from animal and human social behavior to autistic social dysfunction. *Journal of Neuroscience*, *32*, 14109-14117a.

Yang, D. Y. J., Rosenblau, G., Keifer, C., & Pelphrey, K. A. (2015). An integrative neural model of social perception, action observation, and theory of mind. *Neuroscience & Biobehavioral Reviews*, *51*, 263-275.

Yardley, L., McDermott, L., Pisarski, S., Duchaine, B., & Nakayama, K. (2008). Psychosocial consequences of developmental prosopagnosia: A problem of recognition. *Journal of Psychosomatic Research*, *65*, 445-451.

Young, L. J., & Barrett, C. E. (2015). Can oxytocin treat autism?. *Science*, *347*, 825-826.

Zäske, R., Hasan, B. A. S., & Belin, P. (2017). It doesn't matter what you say: FMRI correlates of voice learning and recognition independent of speech content. *Cortex*, *94*, 100-112.

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *Journal of Neuroscience*, *11*, 641-649.

Zhang, Z. J., Hao, G. F., Shi, J. B., Mou, X. D., Yao, Z. J., & Chen, N. (2008). Investigation of the neural substrates of voice recognition in Chinese schizophrenic patients with auditory verbal hallucinations: an event-related functional MRI study. *Acta Psychiatrica Scandinavica*, *118*, 272-280.

Ziegler, J. C., Pech-Georgel, C., George, F., & Foxton, J. M. (2012). Global and local pitch perception in children with developmental dyslexia. *Brain and Language*, *120*, 265-270.

Zink, C. F., & Meyer-Lindenberg, A. (2012). Human neuroimaging of oxytocin and vasopressin in social cognition. *Hormones and Behavior*, *61*, 400-409.

Appendix

A List of Figures

1 Introduction

3 Mechanisms for Perception of Visual and Auditory Communication Signals

Selbständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe und dass die aus fremden Quellen direkt oder indirekt übernommenen Gedanken in der Arbeit als solche kenntlich gemacht worden sind.

Kamila Borowiak
Leipzig, den 09.09.2019