

A study on audiovisual processing in noise in typically reading adults.

Master thesis

Master scriptie

Date: 20 oktober 2018

Naam: Mattanja Pauw

Student ID: S4828089

Opleiding: Master Taalwetenschap – Taal – en Spraakpathologie

Universiteit/afdeling: Radboud Universiteit– Taalwetenschap

Naam begeleider: dr. Margriet Groen

Naam tweede lezer: dr. Esther Janse

0. Abstract

The main question we tried to answer in the current research was: Do adults without dyslexia identify steps from an audio-visually presented continuum similar in a modulated noise condition and a stationary noise condition? In order to contribute to this investigation, we determined: I) if typical readers show different audiovisual perception of phonetic categories in two different noise conditions, II) if typical readers show different reaction times in the two different noise conditions and III) if in typically readers a correlation is present between speechreading ability and the visual influence on audio-visual perception of phonetic categories. 11 typically reading adults were asked to participate in the study. We used a phonetic categorization task with a /p/-/t/-continuum. Five auditory steps along the continuum were combined with five visual steps along the continuum, both presented unimodally and bimodally combined with different types of noise (speech-shaped steady state noise and speech-shaped modulated noise). We did not find a statistically difference in audiovisual perception of the phonetic categories between the two noise conditions, although the figure suggests there is a difference. It might, therefore, not be statistically significant due to the small sample size, further research is suggested. We also did not find a statistically significant difference in the reaction times between responses in the different noise conditions. Since, again, it can be seen from the figure that the reaction times are (mostly) slower for the modulated, we argue that this might also be due to the small sample size. Another explanation could be that the modulation used in the current study did not result in a difference in reaction time between the modulated noise condition and the stationary noise condition. Also, not correlation was found between the speech-reading abilities and the visual influence on the audio-visual perception of phonetic categories.

Content

0. Abstract	2
Content.....	3
1. Introduction.....	4
<i>1.1 Literature Review</i>	<i>7</i>
1.1.1 Reading development.	7
1.1.2 Dyslexia.....	7
1.1.3 Phonological Awareness.....	8
1.1.4 Phonological.....	9
1.1.5 Phonetic categorization.....	10
1.1.6 Visual cues.....	12
2. Method	19
<i>2.1 Participants.....</i>	<i>19</i>
<i>2.2 Reading and Cognitive Measures</i>	<i>19</i>
2.2.1 Reading.....	19
2.2.2 Speechreading.....	20
2.2.3 Phonological awareness.....	20
2.2.4. Non-verbal cognitive ability.....	21
<i>2.3 Experimental Materials and Procedures</i>	<i>21</i>
2.3.1 Phonetic categorization task.	21
2.3.2 General procedure.	22
2.3.3 Statistical analysis.....	23
3. Results	25
<i>3.1 Cognitive Measures.....</i>	<i>25</i>
<i>3.2 Phonetic Categorization Tasks</i>	<i>25</i>
3.2.1 /P/-responses.....	25
3.2.2 Reaction time.....	29
3.2.3 Correlation between the number of /p/-responses and speechreading abilities.....	31
4. Discussion.....	32
4.1 Hypothesis I	32
4.2 Hypothesis II	34
4.3 Hypothesis III.....	35
4.4 General discussion.....	36
5. Conclusion	37
References.....	38

1. Introduction

Learning to read is one of the early challenges we face in life. Despite showing normal IQ, having acceptable educational opportunities and without any sensory or neurological impairment, approximately 10 – 15% of school age children experience difficulties with learning to read and are diagnosed with developmental dyslexia (Snowling, 2000; Vellutino, Fletcher, Snowling and Scanlon, 2004). Since the literature shows a wide variety of types of dyslexia, it is important to first determine which type of dyslexia will be the focus in the current research. The most frequently discussed manifestations of dyslexia are: acquired dyslexia, which occurs after a (traumatic) head injury, and developmental dyslexia. The current research focuses only on developmental dyslexia because it aims to contribute to the ongoing investigation on the underlying mechanisms in this reading deficit. It has been found that the difficulties people with developmental dyslexia experience continue into adulthood (e.g. Van den Bunt et al., 2017). For the majority of children, who learn to read normally, acquiring the skills of reading depends on the ability to form reliable cross-sensory associations between speech-sounds and letter combinations (Hahn et al., 2014). With practice, the retrieval of letter-sound associations becomes increasingly automatic. Although still unclear what exactly is being integrated, it has been found that letter knowledge, the rapidity of the automatized naming and the phonemic awareness (e.g. Blauw, van Atteveldt, Ekkebus, Goebel and Blomert, 2009) are important predictors of the individual differences in reading development (Lervåg, Bråten, & Hulme, 2009; Hulme and Snowling, 2013).

It has been found that some of the children who experience difficulties with learning to read show impaired phonological processing. And even though multiple theories have been proposed, the phonological deficit hypothesis is currently viewed as the main theory explaining the problem to individuals with dyslexia (e.g. Dehaene, 2009; Blomert, 2010). The phonological deficit theory states that an impairment is present on the level of phonemes, which are the elementary constituents of spoken words (Dehaene, 2009, p. 239). However, considering the different manifestations of dyslexia, a discussion is still going on about whether every individual with dyslexia has a phonological deficit (Castles and Friedmann, 2014). And the question it raises altogether, is if this is indeed the case, what is the underlying cause of this phonological deficit? The possible causes for a phonological deficit in dyslexia that have been proposed so far, will be discussed briefly.

Multiple studies have investigated phonological deficits in developmental dyslexia and its different aspects. A strong interaction between learning to read and spoken language development has been shown (Morais, Cary, Alegria, and Bertelson, 1979; Perfetti, Beck, Bell, and Hughes, 1987). More specifically, it has been suggested that dyslexic individuals show speech perception deficits (Boets, Ghesquière, van Wieringen, and Wouters, 2007). It is suggested that it is specifically difficult for

individuals with dyslexia to use phonetic features for selecting the phonological representations (Manis et al., 1997). Therefore, one of the main topics of interest in relation to the phonological deficit is the role of categorical speech perception in people with dyslexia. Liberman, Harris, Hoffman, and Griffith (1957) suggested that speech perception is categorical, meaning that it is difficult for listeners to discriminate acoustic differences when they belong to the same category, but easy to discriminate comparable differences when the phonemes belong to different categories. The results of a number of studies have indicated that (at least some of) the individuals with dyslexia show a categorical perception deficit (Werker & Tees, 1987; Manis et al., 1997; Adlard and Hazan, 1998; Noordenbos and Serniclaes, 2015). In addition, results have shown a correlation between phonological impairment and categorical perception deficits (e.g. Manis et al., 1997; Boets et al., 2007; Hakvoort et al., 2016). In their meta-analysis, Noordenbos and Serniclaes (2015) conclude that in all studies they used in their study, both comparing individuals with dyslexia to chronological-age controls and comparing individuals with dyslexia to reading-level controls, the categorical perception of the dyslexic group was weaker. The studies that are discussed by Noordenbos and Serniclaes (2015) provide results of both identification tasks and discrimination tasks, which are often used when investigating language and speech processing. In identification tasks, participants are asked to identify an object, word, letter, color etc., out of a number of stimuli presented at the same time. A well-known task used to test identification is the phonetic categorization task. In this task, participants are asked, for example, to judge whether they hear stimulus A or B, whilst presented with steps along a continuum. In a discrimination task, on the other hand, a number of stimuli are presented at the same time (also for example objects, words, letters, colors etc.) and the participant is asked to judge whether the stimuli are the same or not (Taniguchi & Tayama, 2010). Noordenbos and Serniclaes (2015) found that the categorical perception deficit was significantly worse for the individuals with dyslexia on both the identification tasks and the discrimination tasks. The effect-size of the categorical perception deficit comparison, however, was significantly larger for the discrimination task than for the identification task (Noordenbos and Serniclaes, 2015).

Boets et al. (2007) decided to include the factor of noise in their study on speech perception and its relationship with phonological and auditory processing in children with a family risk for dyslexia. Studies had so far found more problems for speech perception in the presence of background noise for individuals with dyslexia, even though the results were received from studies that differed in experimental design or participant inclusion criteria. Boets et al. (2007) found a significant relation between the phonological awareness scores and the results of the speech-in-noise tests. In addition, when comparing the low-risk children with the high-risk children, they found that the high-risk children showed slightly but significantly worse speech-in-noise perception, which was especially present in the most difficult listening condition. Ziegler et al. (2009) also investigated speech perception in noise

conditions in developmental dyslexia. They argued that the results so far had been obtained from experiments in optimal settings (meaning: non-natural environments), which are unnatural listening situations. In everyday life the listener is always confronted with surrounding noises when trying to percept speech. Therefore, a study investigating speech perception in quiet conditions only does not provide accurate results on the speech perception in everyday life. In their study in 2009, Ziegler et al. therefore investigated this matter by comparing the results of speech perception in both quiet and noisy conditions. Their results showed that children with dyslexia showed clear speech perception deficits in the noisy conditions, but not in silence. In addition, they found that the speech-perception in noise predicted the variance in reading, even when they controlled for possible other influencing factors.

A lot of studies so far have tested speech perception (in quiet and noisy conditions) and the relation to phonological deficits by using auditory tasks only (e.g. Füllgrabe, Berthommier, and Lorentzi, 2006; Boets et al., 2007; Ziegler et al., 2009; Dole, Hoen and Meunier, 2012), but the benefits of visual cues for speech perception (especially for speech perception in noise) have been demonstrated in multiple studies (e.g. Grant and Seitz, 2000; Brancazio, 2004; Stacey, Kitterick, Morris and Sumner, 2016). Since reading is an audiovisual process, it could be argued that an audiovisual deficit is underlying the reading deficit. The advantage of using an audiovisual task over an audio only task, which has been the most common approach, is that it gives information about whether individuals with dyslexia use different or similar speech cues as individuals without dyslexia.

Francisco, Jesse, Groen and McQueen (2017) investigated whether an audiovisual deficit is underlying the reading difficulties in dyslexia. They did not find a significant difference in the use of the visual information between the adults with and without dyslexia. They did, however, only test audiovisual speech perception in silent conditions. In some audio-only studies, speech perception of children with dyslexia has been showed to be weaker in the noise conditions, but not in silence (Ziegler et al., 2009). It would therefore be interesting to, again, investigate whether and audiovisual deficit is underlying the reading difficulties in (adults with) dyslexia, but use speech perception in noise. Several approaches have been used to investigate a possible audiovisual deficit in individuals with dyslexia, and the three most important to the current research are: (a) the audiovisual benefit approach, (b) the McGurk-identification task and, (c) a phonetic categorization task. The main aim of this thesis is to further investigate the relationship between reading ability and audiovisual processing, and to contribute to the current knowledge on the relationship between reading ability and audiovisual processing. This will be done by clarifying the role of audiovisual cues on speech intelligibility in noise in adults with dyslexia, in comparison to typically reading (TR) adults. Using a similar experimental design as Francisco et al. (2017), but adding noise conditions based on the findings by Ziegler et al. (2009), the main question we try to answer is: Do individuals with and without dyslexia identify the steps from an audio-visually presented continuum in a similar way?

1.1 Literature Review

1.1.1 Reading development. Although reading is a rather recent cognitive skill in comparison to oral communication, about ninety percent of the people learn to read without difficulties if they are instructed properly (Blomert, 2010). For a long time, reading was thought to be a visual skill, but relevant research over the past years has shown that it is primarily a linguistic skill (Vellutino et al., 2004). In order to become a skilled reader, one has to adequately develop both word identification and language comprehension. Vellutino et al. (2004) give the following definition for word identification and language comprehension: “word identification is a lexical retrieval process that involves visual recognition of a uniquely ordered array of letters as a familiar word and implicit (or explicit) retrieval of the name and meaning of that word from memory. Language comprehension involves integration of the meanings of spoken or written words in ways that facilitate understanding and integration of sentences in spoken or written text in the interest of understanding the broader concepts and ideas represented by those sentences” (p.5). This implies that both the identification and the comprehension of the meaning of the text have to be processed within the limits of the working memory.

Children will be familiar with spoken language when they start learning to read, so the first step in acquiring the skill of reading is learning the letters of the alphabet and matching those orthographic symbols to speech sounds. At first, the child learns to remember the corresponding orthographic symbols and speech sounds, but with practice, the retrieval of letter-sound associations will become automatic. In this stage, the orthographic symbols automatically activate the phonological representation (Hahn et al., 2014). According to Blomert (2011), it only takes children a couple of months to know which speech sounds belong to which letters, but the process of automatically integrating them into newly constructed audiovisual objects takes much longer. Blomert (2011) demonstrated this with the finding that accurate word and letter activations in the fusiform cortex also occur relatively late in reading development.

1.1.2 Dyslexia. Children with developmental dyslexia often experience difficulties with this basic letter-speech-sound mapping, which is thought to be the primary source of their word recognition problems (Swan & Goswami, 1997). Jones, Snowling & Moll (2016) found that even though individuals with dyslexia showed similar lexical processing as individuals without dyslexia, the access to the lexical information necessary for fluent and accurate reading is less automated, which resulted in a delay at the naming phase. This impaired reading fluency (e.g. Wolf & Bowers, 1999) and accuracy is seen in both children and adults with dyslexia (e.g. Vellutino et al., 2004). Since people with dyslexia show intact other domains, the question that remains is why is it so complicated for them to match the orthographic symbols with the phonological representation?

For years the deficit was thought to be caused by some kind of visual (memory) impairment (e.g. Goulandris & Snowling, 1991), but since the late 1970s the main view is that dyslexia is a language deficit, specifically a phonological language deficit (Serniclaes, Sprenger-Charolles, Carré and Demonet, 2001; Ramus, 2001; Vellutino et al., 2004; Blomert et al., 2011; Castles and Friedmann, 2014). Phonological impairment can be roughly divided into three dimensions: phonological awareness (PA), phonological processing (PP) and phonetic categorization (PC). These three dimensions will be briefly discussed below.

1.1.3 Phonological Awareness. Yopp and Yopp (2009) define PA as: “phonological awareness is the ability to attend to and manipulate units of sounds in speech (syllables, onsets and rhymes, and phonemes) independent of meaning” (p. 13). This includes the ability to analyze, match and synthesize spoken sounds, but also understanding the variance in sounds (that they are the same, even when they occur in a different phonetic context) (Bishop and Snowling, 2004). These skills are believed to contribute to learning that letters represent sound values, and to the process of matching those letters and sounds (Vellutino et al., 2014). PA has been suggested to, together with letter-knowledge, provide the basis to develop the ability to decode language. Enough evidence has been found over the past years to suggest that PA, indeed, plays an important role in reading development (e.g. Bradley & Bryant, 1983; Ehri et al., 2001), as it has been shown to be a good predictor for reading skill, (Lervåg et al., 2009) and reading accuracy in particular (Poulsen, Juul and Elbro (2012). Ehri et al. (2001) conducted a meta-analysis on articles that reported results about phonemic awareness instruction and its relation to reading development. Results revealed improved reading after phonemic awareness instruction in various types of children (including normally developing readers as well as at-risk and disabled readers). Also, statistically significant contribution of PA instruction was found for reading acquisition, which indicates that PA-instruction helps the process of reading acquisition. This is supported by findings of Vellutino et al. (2004), who found similar results. Since the relationship between PA and reading development/skills has been established in general, an increased interest developed for the role PA plays in individuals with dyslexia. Evidence has been found for impaired PA in individuals with dyslexia (e.g. Bruck, 1992; Snowling, 2000; Catt, Adlof, Hogan and Weismer, 2005; Hogan, Catts and Little, 2005). Bruck (1992) conducted a longitudinal study in which dyslexic and non-dyslexic children were tested on PA twice. The first time they were all between the age of eight and 16, and they were tested again when they were between the age 19 and 27. The most interesting finding in relation to the current study is that dyslexics did not acquire a similar level of PA compared to controls: both children and adults made more errors on all phonological awareness measures than the control groups did. This was regardless of their reading level or age. This thus implies that the PA deficit was present during childhood and continued into adulthood. Vellutino et al. (2004) compared several studies on PA in dyslexics and found that the scores on PA tasks were consistently lower for poor readers in comparison to the scores of

normal readers. They concluded that people with dyslexia show difficulties acquiring phonological awareness and in alphabetic coding, which results in non-specific phonological representations, meaning that the phonological representation of a word lacks phonemic – and phonetic details to define the word’s specific acoustic structure.

One of the predominant views on the proposed problems in acquiring phonological awareness and alphabetic coding in individuals with dyslexia is the one by Blomert (2011). Blomert (2011) hypothesized that both reading deficits and the notorious lack of reading fluency in individuals with dyslexia are caused by an orthographic-phonological binding deficit. This is supported by neuroimaging evidence, which shows a deficit in the integration of letters and speech sounds into automatized audiovisual objects. The orthographic-binding deficit was not only demonstrated in both children and adults with dyslexia (Blau et al., 2010), but also in children with a familial risk of dyslexia who had not started reading yet (Blomert and Willems, 2010). This, again, suggests impairment in PA in people with dyslexia.

1.1.4 Phonological processing. As elucidated in the paragraph above, in order to process language (written or oral), it is important that phonological representations are well established. If this is not the case, it may cause problems in phonological processing. It is commonly believed that a deficit in phonological processing might also be a frequently occurring impairment in developmental dyslexia (Snowling, 2000; Lyon, Shaywitz and Shaywitz, 2003) and a number of tasks have been used to investigate this matter: e.g. the Rapid Automatized Naming-task (RAN) and the Non-word repetition task (NWR). These will be discussed briefly below.

1.1.4.1 Rapid automatized naming. Phonological processing is referred to as a range of cognitive skills involving speech sounds (Bishop and Snowling, 2004) and PA is said to be one of these cognitive skills. A relationship between phonological awareness and phonological processing was proven by multiple studies, which found that the RAN-reading relationship was partially mediated by phonological awareness and letter-knowledge (e.g. Poulsen et al., 2012). RAN is defined as rapid automatized naming, a task in which participants are asked to name e.g. letters, colors, object or pictures out loud, as quickly as possible. The findings by Poulsen et al. (2015) showed especially contributing effects of phonological awareness to the relationship between RAN-scores and reading accuracy. Bexkens et al. (2014) investigated the cognitive processes involved in RAN, and also found a contributing role for PA on RAN.

Jones, Moll and Snowling (2016) investigated whether lexical processing was impaired in individuals with dyslexia. A version of the RAN-task was used, but they included a Stroop-switch component. It was decided to combine the RAN-task with a Stroop-component, referring to the well-known Stroop-test in which participants are asked to name aloud the color of the word whilst it is presented in different colors. They combined these two tests to measure both automatized lexical access (Stroop-task) and fluency (RAN). In the experiment by Jones, et al. (2016), participants were

asked to name words out loud, which changed font color upon fixation. This demanded of the participant to name the color of the font instead of the letter name and withhold the activation of the lexical code. Their findings suggest that, even though the individuals with dyslexia show similar initiation and time course of lexical recognition compared to the control group, they show problems in either the cognitive control mechanism that is responsible for suppression of a phonological response or in the speed in which the output can be computed. Jones et al. (2016) suggest that the early lexical processes appear to be automatic in individuals with dyslexia, but a delay arises at the output stage.

1.1.4.2 Non-word repetition task. Secondly, the non-word repetition task (NWR) will be discussed. In this task, participants are asked to repeat the non-words presented to them (in most tests they were a mix between one – five syllables long). Impaired non-word repetition has been demonstrated in multiple studies on dyslexia (e.g. Catts et al., 2005; De Bree, Rispens and Gerrits, 2007). Catts et al. (2005) showed results consistent with their prediction that children with dyslexia and children with both dyslexia and SLI scored poorly on measures of PA and NWR. In line with these findings, De Bree, Rispens and Gerrits (2007) found that repeating non-words was more difficult for both preschool and school-going children with (a risk of) dyslexia, than it was for typically developing age-matched controls. They also tested for a so-called word-length effect, which indicates the severity of the phonological processing deficit. The results showed that the children with dyslexia had a significantly lower percentage phonemes correct in comparison to the control groups, but only for five-syllable non-words. Their combined findings suggest that the children with (a risk of) dyslexia were found to be characterized by a phonological processing deficit (De Bree et al., 2007). Above, Blomert's (2011) hypothesis was mentioned, stating that reading deficits in individuals with dyslexia are caused by a phonological-binding deficit, which leads to less automatized access to audiovisual objects. The difference between phonological processing and phonological binding is that phonological processing refers to the ability to process phonemes, whereas phonological-binding refers to the ability to match letters or symbols to sounds.

1.1.5 Phonetic categorization. So far, PA and phonological processing have been discussed as possible underlying affected mechanisms in the phonological deficit underlying dyslexia. A variegated amount of evidence has been found to suggest a poorly specified representation of speech sounds in people with dyslexia. The third proposed theory about the mechanism underlying the phonological deficit is impaired phonetic categorization (Manis et al., 1997; Chiappe and Chiappe, 2001; Serniclaes et al., 2001; Bogliotti Serniclaes, Messaoud-Galusi and Charolles, 2008; Vandermosten et al., 2010). Phonetic categorization requires the ability to recognize to which category a sound belongs, and whether sounds belong to the same or a different category (Ehri et al., 2001). It is widely believed that speech sound discrimination is governed by phonemic categories (Serniclaes et al., 2001).

In 2001, Serniclaes et al. investigated the perceptual discrimination of speech sounds in developmental dyslexia by comparing the results of a number of discrimination tasks to those of non-dyslexic children. They found evidence for the theory that the problem with dyslexia is not in the processing of rapid incoming sensory information, but in the construction of phonemic categories. When the stimuli belonged to the same phoneme category, the dyslexic readers were better at discriminating the acoustic differences than the average readers. This confirms earlier research that suggested a less categorical perception of speech sounds in children with dyslexia, as they better perceive within-category differences. Also, they stated that the deficit is present in both the perception of speech and non-speech. They did, however, suggest that there is no causal relationship between speech perception and non-speech perception, since the categorical boundaries of the sounds are different.

In 2008, Bogliotti et al. also studied categorical perception deficit in children with dyslexia as compared to chronological age and reading level controls. Along a voice onset time continuum, /do-to/-syllables were identified and discriminated by the children of all groups. Their results show complimentary evidence for a deficit in the categorical perception of children with dyslexia, as they performed significantly worse on the task discriminating speech sounds than the control children. In addition, they further investigated the findings discussed by Serniclaes et al. (2001) on allophonic perception in children with dyslexia. The allophonic peak index (the difference between the across category versus within category discrimination) was used to examine the reliability of the allophonic perception differences. They found that the children with dyslexia showed reduced phonemic boundaries when discriminating sounds in comparison to the controls. In addition, their discrimination performance was characterized by a nonphonemic discrimination peak, which was located at -20 ms VOT (this is close to the -30 ms peak that has been found in a previous similar study for children with dyslexia). Therefore, they concluded that children with dyslexia base their speech perception on allophones, rather than on phonemes. Using allophonic representations for speech perception is suggested to be a significant handicap for the establishment of grapheme-phoneme correspondence, according to Bogliotti et al. (2008), since using allophonic representations causes problems in the one-to-one correspondence between graphemes and phonemes. An example given is that when a child perceives the allophones /d/, /p/, and /p^h/ instead of the phonemes /p/ and /b/, it will be difficult to assign the letter “p” to both /p^h/ and /p/.

Phonetic categorization is one of the main skills necessary for speech perception, but how does speech perception work? Speech can be seen as an immense variability of acoustic cues, and even though individual speakers all use these acoustic cues differently, listeners are able to accurately recognize, process and analyze speech (Toscano, McMurray, Dennhardt and Luck, 2010). Examples of cues that are used to identify speech are place of articulation, voicing and temporal information. In

order to process language, the speech sounds have to accurately be mapped onto the mental representations of lexical form. Since speech signals are continuously changing along the duration of the utterance, it is required that the processes involved in lexical access and selection are continuously shifting as well (Warren and Marslen-Wilson, 1988). The question is how do listeners transform all these variable acoustic signals into meaningful categories?

1.1.6 Visual cues. When communicating, information is received from both the face (providing visual cues) as well as the voice (providing auditory cues) (Sumbly & Pollack, 1954). The visual cues are presented through the movement of facial muscles, the lips and the tongue of the speaker. The auditory cues are presented through the acoustic waveform. Both provide information about the utterance. For example, to distinguish whether the speaker said /pa/ or /ta/ both hearing and seeing the speaker provides the necessary information. However, seeing the lips close when the speaker says /pa/ but not when saying /ta/ gives more specific information than just the hearing. This clearly indicates the influence of visual cues on speech perception, which has been proven by various studies (Jesse and Janse, 2012; Yeung and Werker, 2013; Lalonde, 2016).

In natural environments, speech characteristics are always competing with acoustic signals coming from surrounding noises. This noise can either be a steady type of noise (e.g. a continuous background noise of wind or machines at the office) or a more fluctuating noise (e.g. other speakers in the same room). These surrounding noises make it more difficult for the listener to separate and distinguish the acoustic cues belonging to the speech and necessary to process the speech. A numerous amount of studies has argued that loud background noise may function as an amplifier for the dominant features that are necessary for speech perception (Davies, 1968; Hockey, 1973). It is said to induce the concentrated attention upon the task-defined dominant aspects of the stimulus (Broadbent, 1971). Speech redundancy can improve speech perception by providing more cues than necessary to identify the spoken words. The perceptual mechanism is enhanced by the combined information of many acoustic cues to rely on different sources and channels (Nittrouer, 2005). Since most natural listening environments contain background noise, in which often speech redundancy is present, it is important to investigate speech perception in similar environments and include variations of noise in the experimental design.

One of the approaches to investigate the influence of visual cues on speech perception (both in individuals with and without dyslexia) is the audiovisual benefit approach. In 2016, Stacey et al. investigated the beneficial effect of visual cues on speech intelligibility in noisy situations. Their main question of interest was whether the size of the benefit received from visual speech information depends on the presence of informative temporal fine structure information. In order to answer this question, they systematically investigated the perception of sine-wave vocoded speech at a range of SNRs (sound-to-noise ratio). To create this sine wave speech, the formants of the frequencies in the

utterance are tracked. Then, the center frequency of those formants is used for the synthesized sine waves. The perception on the different SNR-conditions were compared with the performance from the speech conditions in which they retained the informative temporal fine structure cues, as is the case in optimal listening conditions. In the experiment they included both typical hearing adults and adults with a cochlear-implant in their study. Their results are in line with the statement that visual information provides larger benefits when the speech is lacking in informative temporal fine structure (TFS). Stacey et al. (2016) cite Moore (2008) who substantiates the suggestion that visual information provides larger benefits when the speech is lacking in informative temporal fine structure, by saying that the absence of informative TFS can hinder the ability to identify the target talk based on vocal characteristics and at the same time hinder the ability to segregate speech from background noise based on cues such as periodicity (Stacey et al., 2016). Therefore, using redundancy and/or additional information such as visual cues may be more beneficial for listeners with a deficit in the access to these temporal fine structure cues (e.g. as has been observed to be the case in (some) individuals with dyslexia (e.g. Rocheron, Lorenzi, Füllgrabe and Dumont, 2002) (Stacey, Kitterick, Morris, & Sumner, 2016). Sumbly and Pollack examined the contribution of visual factors to speech intelligibility in normal hearing adults, as a function of the speech-to-noise ratio (SNR). Their findings show that the visual contribution is higher when a decrease in the SNR is present. In general, they suggest that the listener will perceive speech more accurately when the speaker can both be seen and heard at the same time. A study done by McGettigan et al. (2012) demonstrated greater assistance from visual speech information when the speech lacked auditory clarity. This implies that the worse the auditory cues, the greater the benefit of the visual stimuli. The impact of removing the informative temporal fine structure cues has been studied thoroughly for audio-only situations, but the impact it has on audio-visual perception of speech with background noise conditions has not received as much attention.

Even though it has been argued that noise enhances speech perception, a convincing amount of evidence has been found for an impeding effect, meaning that it can also hinder speech perception (Füllgrabe et al., 2006; Ziegler et al., 2009). For example, the noise may mask acoustic cues necessary to identify or discriminate speech sounds. Also, it may cause distraction from the relevant cues, which makes it harder to percept the speech. One of those studies is the one by Füllgrabe et al. (2006), who investigated masking release for consonant features in temporally fluctuating background noise. Masking is defined as: “the process ... and the amount by which the threshold of hearing for one sound is raised by the presence of another (masking) sound” (Oxenham, 2014). Masking release can then be explained as the process and amount by which the masking is reduced, due to a manipulation of the masking or the target sound (Oxenham, 2014). Füllgrabe et al. (2006) used vowel-consonant-vowel-stimuli (VCV-stimuli), either unprocessed or spectrally degraded to force listeners to use temporal-envelope cues, to measure consonant identification for normal-hearing listeners. The stimuli were

either embedded in a steady state or fluctuating noise masker. As Füllgrabe et al. (2006) said, the range of the background fluctuation frequencies determines the availability of the acoustic cues necessary to identify voicing, place of articulation or manner. Spectral and fast fine structure cues are helpful for the perception of the place of articulation. These cues only need short dips in the background noise in order to be perceived. Results showed that dips of only 4 ms were enough to extract speech cues. Also, findings showed that the highest reception of place of articulation was found for 32 Hz. Peaks of 16 ms were found to be optimal for speech perception mechanisms using mainly spectral or temporal fine-structure information, which is the case of place of articulation. Based on the findings by Füllgrabe et al. (2006), a difference in speech perception can be expected between speech perception in an environment with steady-state noise and in an environment with fluctuating noise, with the latter suggested as being a more optimal environment. Ziegler et al. (2009) continued investigating speech in noise. They suggest that the differences in results found on speech perception in people with dyslexia are because most studies have tested in optimal settings. This meant that the stimuli were provided in quiet surroundings optimal for receiving all acoustic signals necessary. This is, however, according to Ziegler et al. (2009), not representative of natural speech perception environments. Therefore, Ziegler et al. (2009) suggested that to investigate speech perception in people with dyslexia, it had to be in an environment with background noise. An experiment was designed in which the speech perception in a quiet condition was compared to speech perception in four different noise conditions. The noise conditions were as followed: one condition with speech-shaped stationary noise and three conditions with speech-shaped modulated noise with different frequencies (4 Hz, 32 Hz and 128 Hz). 48-vowel-consonant-vowel audio-stimuli were presented to the participants in all condition. Participants were asked to identify each stimulus. They found that children with dyslexia had more difficulty recognizing speech in noise than children without dyslexia (Ziegler et al., 2009). Their results show that the children with dyslexia show a clear speech perception deficit in all noise conditions but not in silence. A significant effect of group and noise modulation frequency was found, but no interaction effect. The biggest effect size was found for the noise condition with a 4 Hz modulation ($d = 1.44$), but 32 Hz, 128 Hz and the stationary noise conditions also showed a large effect size. The results of Ziegler et al. (2009) show that there might be a problem for speech recognition in noise for children with dyslexia. These results, however, are based on audio only stimuli.

Speech perception in individuals with dyslexia (both children and adults) was already shortly discussed, but one of the questions yet to be answered in relation to this is: is a general auditory processing deficit underlying the speech perception impairments in individuals with dyslexia, or is this deficit specific to language? One way to approach this question is by investigating auditory processing of both speech and non-speech material. Serniclaes et al. (2001) suggested a deficit in both auditory processing of speech and non-speech material, but no causal relationship between the two deficits. In

2003, Ramus et al. also found no difference between perception of speech or non-speech for adults with dyslexia. To approach the question even more specifically, studies started to focus on bimodal processing, instead of unimodal processing. Blomert (2011) was one of these studies, and in this study, it was argued that the letter-speech sound objects are special and that the deficit individuals with dyslexia show on these associations should not be generalized to other audiovisual objects. Widmann, Schröger, Tervaniemi, Pakarinen, and Kujala (2012) did, however, find evidence suggesting a difference in processing non-linguistic audiovisual material when comparing dyslexic readers with typical readers. In an electrophysiological study (2012), Widmann and colleagues asked both children with and without dyslexia to indicate whether auditory – and visual patterns were congruent or not. Event-Related Potential (ERP) is a measure that can be used to identify specific times of electrical activity of the cerebral surface. In their study, Widmann and colleagues looked at two constituents of the ERP, namely the N2b and the P3a. Patel and Azzam (2005) give the following definition for the N2b: *“the N2b is a negativity of central cortical distribution seen only during conscious stimulus attention.”* This N2b occurs as a response to irregular presented stimuli. The P3a differs from the N2b in that it is related to the relevance of a stimulus. The results of the study done by Widmann and colleagues indicated whether the children with dyslexia and the typical readers showed a similar N2b. Their findings showed that children with dyslexia showed a later and smaller N2b than the typically reading children. This indicates that the children with dyslexia are less reliable and later in processing audiovisual congruency. Also, they did not show a P3a or early-induced auditory gamma band response when the symbols and sounds were incongruent (Francisco et al. 2017). The results, therefore, suggest an impaired identification process of audiovisual stimuli. Also, no early-induced auditory gamma band response was found when there was a congruency between the symbols and sounds. When neural activity is synchronized, this can be seen by early-induced auditory gamma band responses. A number of studies have related this early-induced auditory gamma band to the integration of auditory and visual information (Widmann, Gruber, Kujala, Tervaniemi and Schröger, 2007). Based on this theory, Widmann et al. (2012) have argued that no or limited integration of audiovisual information is present in people with dyslexia, since they show no early-induced auditory gamma band response. Altogether, this suggests a more general audiovisual deficit in dyslexia rather than a deficit specifically to letter-speech sound associations or language and studies have shown that this impairment in multisensory integration persists into adulthood, like dyslexia itself (e.g. Elbro, Nielsen, & Petersen, 1994).

The discussed audiovisual processing deficit in people with dyslexia might occur due to a difference of the influence the auditory and visual modalities may have on their perception of speech sounds. Several approaches have been used to investigate this possibility, and the three most important to the current research are: (a) the audiovisual benefit approach, (b) the McGurk-identification task and, (c) a phonetic categorization task. These will each be discussed below. The first approach is the

audiovisual benefit approach, which has been elucidated earlier in the current piece. The second approach is the use of the McGurk-identification task. The task is used in multiple studies to investigate speech perception, in general as well as in dyslexia (Hayes, Tiippana, Nicol, Sams, and Kraus, 2003; Groen and Jesse, 2013; Francisco et al., 2017). Hayes et al. (2003) did not find a difference in performance on the McGurk-task between children with learning disabilities and typically developing children in low – and no-noise conditions. Interesting in relation to the current study is that they did, however, find that the children with a learning disability showed more visually based responses in the high-level noise conditions. In 2013, Groen and Jesse studied audiovisual speech perception in children and adolescents with developmental dyslexia. They did not find a difference in the unimodal auditory or visual perception between either the children and adolescents or the dyslexics in comparison to their age-matched controls. In addition, dyslexics and controls did not show a difference in response patterns to McGurk stimuli. They also did not find any differences in audiovisual speech perception between the people with dyslexia and the controls, but their results did show a difference between the age groups, showing more visual-based /k/-responses in general for the adolescents than for the children. A difference between children and adults is in line with previous studies, which showed a larger influence of visual information on audiovisual speech perception for adults than children (e.g. McGurk and MacDonald, 1976; Boliek, Keintz, Norrix, and Obrzut, 2010; Massaro, 1984; Dupont and Ménard, 2005). Results from multiple studies are evident for unimodal sensory differences (e.g. a difference in processing audio-only stimuli between individuals with and without dyslexia) (Bastien-Toniazzo, Stroumza, & Cavé, 2010), but did not find comparable results on whether children with dyslexia reported the same amount of fusion responses. Other studies did not report the unimodal sensory results, or no difference was found (e.g. Francisco et al., 2017). The problem with using a McGurk-task to study audiovisual speech perception is that possible differences could be the consequence of differences in the performance in unimodal conditions as well as in the audiovisual processing. Another problem is that differences between groups could also result from different processing of the incongruent audiovisual information. In natural speech perception environments, the auditory and visual information are (almost) always congruent. Therefore, using incongruent audiovisual information is not ecologically valid since it does not provide congruent information.

The third approach used on potential differences in speech perception in readers with and without dyslexia is the phonetic categorization task. Francisco et al used this task. (2017) to investigate whether an audiovisual deficit is underlying reading impairment. Participants were presented with steps from an audiovisual continuum between the Dutch words /so:t/ and /so:p/. Participants were then asked to indicate whether the speaker had said soot or soop. They did not find a difference in speech perception abilities between adults with and without dyslexia but suggest this might be due to not-

challenging enough conditions to expose speech perception deficits. Therefore, they suggest a similar experiment; again, using a phonetic categorization task, but present the stimuli with background noise.

In summary, readers with and without dyslexia may differ in speech perception due to a difference in phonological representations, phonetic categorization differences or a difference in processing audiovisual stimuli. In addition, evidence has been shown for impaired audiovisual processing non-specific to language in both children and adults with dyslexia. If, indeed, the deficit in people with dyslexia is caused by a general audiovisual deficit rather than a deficit specific to language, it raises the question whether people with dyslexia receive the same benefits of visual cues on speech perception in noise, as has been proven for people without dyslexia.

The current study is a follow-up on the study done by Francisco et al. (2017), who found no difference in speech perception abilities between adults with and without dyslexia. Based on the findings in Ziegler et al. (2009) it is, however, suggested there is a difference in speech perception between people with and without dyslexia in noise conditions, but not in silence. To test this in adults, we will use the same task used by Francisco et al. (2017), namely the Phonetic Categorization Task, but add noise conditions. In the current study, the initial plan was to test for group differences in phonetic identification of consonants placed in audiovisual nonsense syllables. The main question was: Do individuals with and without dyslexia identify the steps from an audio-visually presented continuum in a similar way? The following hypotheses were formulated to help answer this question.

- I. Individuals with dyslexia rely more on the visual cues than people without dyslexia in noise conditions.
- II. Individuals with and without dyslexia rely more on visual cues in the speech-shaped steady state noise condition than in the speech-shaped modulated noise condition.
- III. Individuals with and without dyslexia show longer reaction times when the stimuli are incongruent, with individuals with dyslexia being showing slower reaction times than the individuals without dyslexia;
- IV. A correlation is present between speechreading ability and the visual influence on audio-visual speech perception.

Due to time constraints, however, we were only able to contribute to this investigation partially since we were not able to test adults with dyslexia. Therefore, we only tested adults without dyslexia and the new question was: Do adults without dyslexia identify steps from an audio-visually presented continuum similar in a modulated noise condition and a stationary noise condition? The following hypotheses were investigated in the current study:

- I. Typical readers show different visual influence on audio-visual perception of phonetic categories in a speech-shaped stationary noise condition than in a speech-shaped modulated noise condition;

- II. Typical readers show different reaction times in the modulated noise condition than in the stationary noise condition
- III. Typical readers show a correlation between speechreading ability and the visual influence on audio-visual speech perception.

2. Method

2.1 Participants

A total of 11 participants were recruited. It was initially planned to test both people with and without dyslexia, but due to time constraints this was not possible. Therefore, only people without dyslexia participated in the experiment. All participants were between 17 – 29 years old ($M = 24,9$ year, $SD = 2,6$; five males and six females) and participated voluntarily. The native language for all participants was Dutch, all were right-handed, and all had normal or corrected-to-normal vision. None of the participants were excluded from the main experiment based on the scores on the cognitive and reading measures.

Hearing sensitivity was measured for all participants to make sure their hearing thresholds were sufficient. A pure-tone standard audiometric screening test was used to make sure participants detected pure tones for a range of frequencies (0.125 to 4 kHz) in each ear at 20 dB HL. Two out of the 11 participants did not complete the audiometric screening test due to software problems. Three of the participants did were not able to detect all pure tones in each ear at 20 dB HL, but they were able to at 30 dB HL. Therefore, and due to the small sample size, it was decided to not exclude any of the participants from the main analysis.

2.2 Reading and Cognitive Measures

2.2.1 Reading. A text reading task from a standardized Dutch reading and writing test-battery for dyslexia diagnosis in adolescents and adults ('Test voor gevorderd Lezen & Schrijven') was used to assess reading accuracy and speed (Test voor gevorderd Lezen en Schrijven; Depessemier & Andries, 2009). The participant was asked to read a 582-word text out loud, which consisted of three paragraphs varying in reading difficulty (easy, medium and difficult). The participants were not allowed to pre-read the text in silence and were instructed to focus on accuracy more than reading speed. The reading was audiotaped for further analyses and scored afterwards. If the participant could not (fully) read a word after 5 seconds, it was prompted by the experimenter and the participant would start reading again starting with the following word. The experimenter measured the time needed to read the entire text and counted the numbers of errors made. These included additions, omissions, inversions and replacements (as instructed in the test manual). When the participants corrected a mistake spontaneously it was also noted. The total number of errors per participants was calculated and the time to complete the task was the total time in seconds taken to read the entire text. The raw scores of the two measures were transformed into percentiles using the norms provided in the test manual. As according to the test instructions: a score under the 5th percentile on either number of errors or time indicates a reading problem. The raw scores for both the number of errors and time were used in a

Kendall's Tau correlation to investigate whether a correlation was present between reading speed and the number of errors on the reading task.

2.2.2 Speechreading. The importance of visual cues (e.g. lip movement, jaw movement) on speech perception has been elucidated in the introduction above. One way to investigate the ability to use visual cues is a speechreading task. Therefore, we included a speechreading task in the current study. Also, it was used to check whether a correlation was present between speechreading ability and the number of /p/-responses on the phonetic categorization task. In order to do so, a Spearman's correlation test was done.

A forced-choice visual-only syllable identification task, taken from Jesse and Janse (2012), was used to assess speechreading. The stimuli used were the same ones as used by Jesse and Janse (2012), and consisted of 10 consonant-vowel syllables. The consonants came from five Dutch viseme classes (bilabial: /p/, /m/; labiodental: /f/, /v/; non-labial front fricatives: /s/, /z/; other non-labial front consonants: /t/, /n/; and other non-labial back consonants: /k/, /x/). The same vowel (/ø/) was used for all syllables. Six blocks were presented, each consisting of 10 silent videos of a speaker's face pronouncing each of the consonant-vowel syllables. They were presented in random order. A set of possible responses was given on the screen after each video. The participants were asked to indicate which consonant (out of the ten options) the speaker had pronounced. They were instructed to press the corresponding key on a computer keyboard. If a response was not given in 5 seconds, the next video was presented. No feedback was given. Overall accuracy (the proportion of correct answers) was computed.

2.2.3 Phonological awareness. The task used to test the phonological awareness was a phonological sub-test 'omkeren' (reverse) of the Dutch test-battery Gletschr. (Test voor gevorderd Lezen en Schrijven; Depessemer & Andries, 2009). This task was taken from the website of the Gletschr and carried out via a PowerPoint. The computer with the PowerPoint faced the experimenter the entire time the test was being assessed. The participant was given headphones and could not see the computer screen. The participant was asked to judge whether the second word (of two) was the first word pronounced and spelled backwards (e.g. 'gak' – 'kag'). They were instructed to say ja (yes) or juist (correct) if they thought they were, and nee (no) or onjuist (incorrect) if they did not match. During the experiment, they were not allowed to write anything down. Six practice items were given before the actual experiment started. Feedback was given during the practice items. Before the main task started the participants were instructed to give as many correct answers as possible as fast as they could. The next item was presented right after the participant gave an answer, both correct or incorrect. No feedback was given during the main task. Both the answers and total time were entered on the score form, and after computed to a total score.

2.2.4. Non-verbal cognitive ability. To assess non-verbal cognitive ability, a subtest (the matrix reasoning-task) of the Dutch adaptation of the Wechsler Adult Intelligence Scale – Fourth Edition was used (Wechsler, 2012). This task was used to check the cognitive ability of the participants, independent of language. The scores on this task might give helpful insights for possible outliers on the main task since it gives information about attention and concentration of the participant(s).

With the matrix reasoning-task, an incomplete matrix of abstract figures was presented to the participant, who was then asked to select the figure that best completed the matrix. Participants were first presented two practice items, which provided the two possible matrixes that were presented in the main part of the task. Participants were presented either a 2x2 matrix with one missing figure, or a row of 5 with one figure missing. In both cases five answer possibilities were provided. Items were presented until the participant made three consecutive errors or three errors on four consecutive items, or until the end of the task was reached. The answers were entered on the score form, and the number of correct responses was used to compute a standardized score ($M = 13.27$, $SD = 2.87$).

2.3 Experimental Materials and Procedures

2.3.1 Phonetic categorization task. A phonetic categorization task was used to investigate the hypotheses. Steps from an audiovisual continuum between the Dutch non-words /so:p/ and /so:t/ were presented to the participants. The stimuli used were the same as in Francisco et al. (2017). Both the audio and visual stimuli were recorded by a male native speaker of Dutch with a Sony DCR-HC1000E camera and two Sennheiser microphones. In the videos, the speaker's head and the top of his shoulders were visible. Videos were digitized as uncompressed 720 x 576 .avi files in PAL-format. The audio sampling rate was 44.1 kHz. Two 21-step continua (one auditory-only and one visual-only) were created (see van der Zande et al., 2013, for details). Based on the pilot carried out by Francisco et al. (2017), the same five audio steps and five visual steps were selected for the main experiment. Each visual step was combined with each audio step, resulting in 25 (congruent and incongruent) videos. The final stimuli list consisted of 5 audio-only stimuli, 5 visual-only stimuli and 25 audiovisual stimuli, which makes a total of 35 stimuli. The audio-only and visual-only stimuli were included as a baseline to compare the scores of the audiovisual stimuli to. Following Ziegler et al. (2009), two conditions were included in the experiment: a speech-shaped stationary noise condition and a speech-shaped modulated noise condition. Gaussian noise was used for the noise-mask in both conditions, with a 10-ms rise/fall (Füllgrabe, 2006). Also, in both conditions the noise masker was added to each stimulus at a 0-dB signal-to-noise ratio (SNR) by setting the intensity level of the noise to the mean intensity of the target speech stimulus. Both Füllgrabe et al. (2006) and Ziegler et al. (2009) used (slightly) different noise tokens on

each stimulus, which we therefore did as well. Different noise was created for different repetitions, with a total of 8 repetitions per stimulus.

The experiment consisted of two times 8 blocks: 8 blocks of the speech-shaped stationary noise condition (audio, visual and audio-visual stimuli intermixed), followed by 8 blocks of the noise condition with the speech-shaped modulated noise (audio, visual and audio visual intermixed). The order of presentation was counterbalanced across participants.

2.3.2 General procedure. Participants were informed with the experimental procedures prior to the experiment and were asked to sign a declaration of consent. The procedures performed in the present study were in accordance with the ethical standards of the Radboud University.

The experiment took place at the laboratory at the Centre for Language Studies (CLS-lab), room 12.19, of the Radboud University in Nijmegen. The phonetic categorization task and the speechreading task were assessed in a soundproof booth that was located in this room to avoid possible distraction during the task. The other (cognitive) tasks were assessed outside the booth. All experimental tasks were to be completed the same day. In order to avoid attention loss, the experimental tasks were intermixed. This also to reduce the possible influence between the tasks. The tasks were presented in the following order: a hearing screening, the phonological awareness task, the reading task, the phonetic categorization task and the speechreading task at the end. Due to the length of the phonetic categorization task, we decided to assess the matrix reasoning-task in between noise condition 1 and noise condition 2 of the phonetic categorization task to switch attention and give the participant a break.

Presentation software was used to make and run the experimental tasks (Version 17.0, www.neurobs.com). A laptop (XPS, with a Mobile PC Display) was used to run the experiment. The screen resolution was set to 1600 X 900. Sennheiser headphones were used to present the audio (The model used was HD 201), which was set at the same volume for every participant. The audio was presented diotically.

The reading task and the phonological awareness task were recorded for further analysis. The audio recorder used was a LS-P1 digital handheld Olympus audio recorder. Only the speechreading task and the phonetic categorization task were presented in Presentation software on the laptop. Both tasks had the same presentation sequence: (a) a 50-ms black screen; (b) a fixation cross, which was presented for 250 ms; (c) a 250-ms black screen; and (d) the stimulus presentation. All videos lasted 2 seconds, were always played completely and presented in the center of the screen. After the stimulus was shown, the response options were presented on the screen. The participants were instructed to respond by pressing the one of the response buttons. The buttons used in the experiment were the left and right shift-key (possible responses to the stimuli) and the ENTER (to start the experiment/continue the experiment after a break). The next trial was presented if a response was not given within 5 seconds.

To familiarize the participants with the procedure, practice blocks always preceded the experimental blocks. These practice blocks consisted of five practice trials. Feedback was not given in either the practice blocks nor the experimental blocks.

2.3.3 Statistical analysis. For the statistical analysis of the data received in the current experiment, IBM SPSS statistics, version 24 was used. Each hypothesis was analyzed and the specific approach per hypothesis is briefly discussed below. Even though both unimodal and bimodal stimuli were used in the phonetic categorization task, only the bimodal stimuli were used for the statistical analysis. Initially, it was planned to use the unimodal stimuli as a baseline to compare the responses of the bimodal responses to. But, since this information was not necessary to be able to discuss hypothesis (a), (b) and/or (c), it was decided to not include this in the statistical analysis.

First, hypothesis I: Typically reading adults show different visual influence on audio-visual perception of phonetic categories in a speech-shaped stationary noise condition than in a speech-shaped modulated noise condition. To test this hypothesis, the data received from the phonetic categorization task was used. The dependent variable was the number of /p/-responses on each stimulus. A repeated measure design was used to analyze the data relevant to this hypothesis. Three within-subject factors were created: Noise Type, which included two levels (one level representing modulated noise and the other level representing stationary noise), Visual Step, which include five levels (the five chosen visual steps on the /t-/p/-continuum) and Audio Step, which also included five levels (the five chosen auditory steps on the /t-/p/-continuum). Version, which represented the order of administration of the noise conditions, was admitted as a between-subject factor. Version included two levels, one representing the version where the modulated noise condition was presented first, followed by the stationary noise condition and the other level representing the reverse order presentation.

Füllgrabe et al. (2006) found that fluctuating noise makes it possible for the listener to use these fluctuations to receive acoustic cues about what is being said. In addition, Ziegler et al. (2009) found that speech perception was better in fluctuating noise conditions than in stationary noise conditions for children with dyslexia. These studies have been discussed broadly in the introduction. Based on these findings and the findings by Francisco et al. (2017), it is expected that the overall number of /p/-responses is higher in the stationary noise condition than in the modulated noise condition. This because the listener is expected to rely more on the visual cues when the auditory cues are (even more) masked. Therefore, we expected a significant main effect for the Noise Type, Visual Step and Audio Step. Also, a significant interaction effect is expected for Noise Type, Visual Step and Audio step.

The second hypothesis was: (b) Typically reading adults show different reaction times in the modulated noise condition than in the stationary noise condition. In order to test this hypothesis, we used the reaction time data from the phonetic categorization task as the dependent variable. Again, we

conducted a repeated measure design similar to the one described above, with the same independent variables as were used to test the first hypothesis.

As was mentioned above, we expected more /p/-responses in the stationary noise condition than in the modulated noise condition. Again, based on the findings by Füllgrabe et al. (2006) and Ziegler et al. (2009), it is expected that the modulated noise condition provides more auditory cues to the listener, making speech perception easier. We argue that the incongruent stimuli are therefore more confusing in the modulated noise condition than in the stationary noise condition, resulting in longer reaction times on the phonetic categorization task in the modulated noise condition. Following these expectations, we expected to find a main effect for the Noise Type, Visual Step and Audio Step. Also, a significant interaction effect is expected for Noise Type, Visual Step and Audio step.

The third hypothesis was the following: (c) In typically reading adults, a correlation is present between speechreading ability and the visual influence on audio-visual speech perception. This test was done in order to see if the speechreading ability is related to the scores on the phonetic categorization task. To investigate this hypothesis, a Spearman's correlation test was done. The variables used were two continuous variables, namely: the speechreading scores and the number of /p/-responses given on the phonetic categorization task. Since the sample was small ($n = 11$) we additionally used the Kendall's Tau correlation test to check whether the outcome was similar to the Spearman's correlation test.

3. Results

3.1 Cognitive Measures

A summary of the performance on the reading task and the other cognitive tasks is shown in Table 1. Kendall's Tau correlation was used to investigate whether a correlation was present between reading speed and the number of errors on the reading task. Results showed a medium positive correlation between the reading speed and errors made ($r_t = .486$, $p = .041$).

Table 1:
Mean scores and standard deviations on the cognitive measures and reading tasks.

Measure	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Phonological awareness (total score)	49.46	7.99	63.12	36.11
Reading accuracy (errors)	9.73	5.24	22	0
Reading speed (seconds)	255 s	28.81 s	313 s	229 s
Non-verbal cognitive ability – Matrix Reasoning (<i>standardized score</i>)	13.27	2.87	18	8
Speechreading accuracy (number of items correct)	1.67	1.49	3.83	0.5

3.2 Phonetic Categorization Tasks

3.2.1 /P/-responses. To investigate the hypothesis whether typically reading adults show different visual influence on audio-visual perception of phonetic categories in a speech-shaped stationary noise condition and in a speech-shaped modulated noise condition, we used the number of /p/-responses on phonetic categorization task. The mean percentage was calculated for the percentage of /p/-responses on each auditory step on the /t-/p/-continuum combined with the visual steps on the /t-/p/-continuum. The overview of the group means in both the stationary noise condition and the modulated noise condition is given in Figure 1 for each bimodal stimulus. The auditory stimuli are placed on the horizontal axes, *a* standing for auditory. The number after *a* stands for the step on the continuum, *a*1 being the least /p/-like auditory step and *a*21 being the most /p/-like auditory step. All auditory steps are combined with both stationary noise and modulated noise. The abbreviation used for the stationary noise condition is *ss* and *mod* for the modulated noise condition. The dots represent the visual steps with modulated noise and the triangles represent the visual steps with stationary noise. The colors indicate the visual step on the /t-/p/-continuum, with the lightest color representing the least /p/-like visual step (*v*0) and the darkest color the most /p/-like visual step (*v*100).

A repeated measure design was used to analyze the data. The order of presentation was included in the design as the between-subject factor with two levels (level one representing the version in which the participants were first presented with the modulated noise condition and then the stationary noise condition and level two representing the version in which participants were first presented with the stationary noise condition, then the modulated noise condition). Results showed no

significant effect for the order of presentation ($p = .305$). Therefore, order was not included in the further analysis.

Three within-subject factors were created: Type (the type of noise), including two levels (modulated noise and stationary noise), Visual Step, including five levels (the five visual steps on the /t-/p/-continuum and Audio Step, also including five levels (the five auditory steps on the /t-/p/-continuum. The assumption of Mauchly's Sphericity was not met for Type, Visual, Audio, Visual*Audio or Type*Visual*Audio, therefore a Greenhouse-Geisser correction was applied. A main effect was found for Type ($F(1,9)=10.44$, $p = .01$, Visual ($F(4,36)=50.62$, $p < .001$) and Audio ($F(4,36)=10.75$, $p < .005$). The significant main effect for Type means that the total number of /p/-responses are significantly different for the noise conditions. The significant main effect for Visual tells us that the total number of /p/-responses is significantly different for the visual steps, and the significant main effect for Audio means that the total number of /p/-responses is also significantly different for the auditory steps. The fact that Type is significant suggests that there is, as we hypothesized, a difference in the number of /p/-responses in the modulated noise condition and the stationary noise condition. To check whether this difference was still present when in combination with the auditory and visual steps, we looked at the interaction effect. No three-way interaction effect was found between the type of noise and the visual and auditory steps. This means that there is no significant difference in the number of /p/-

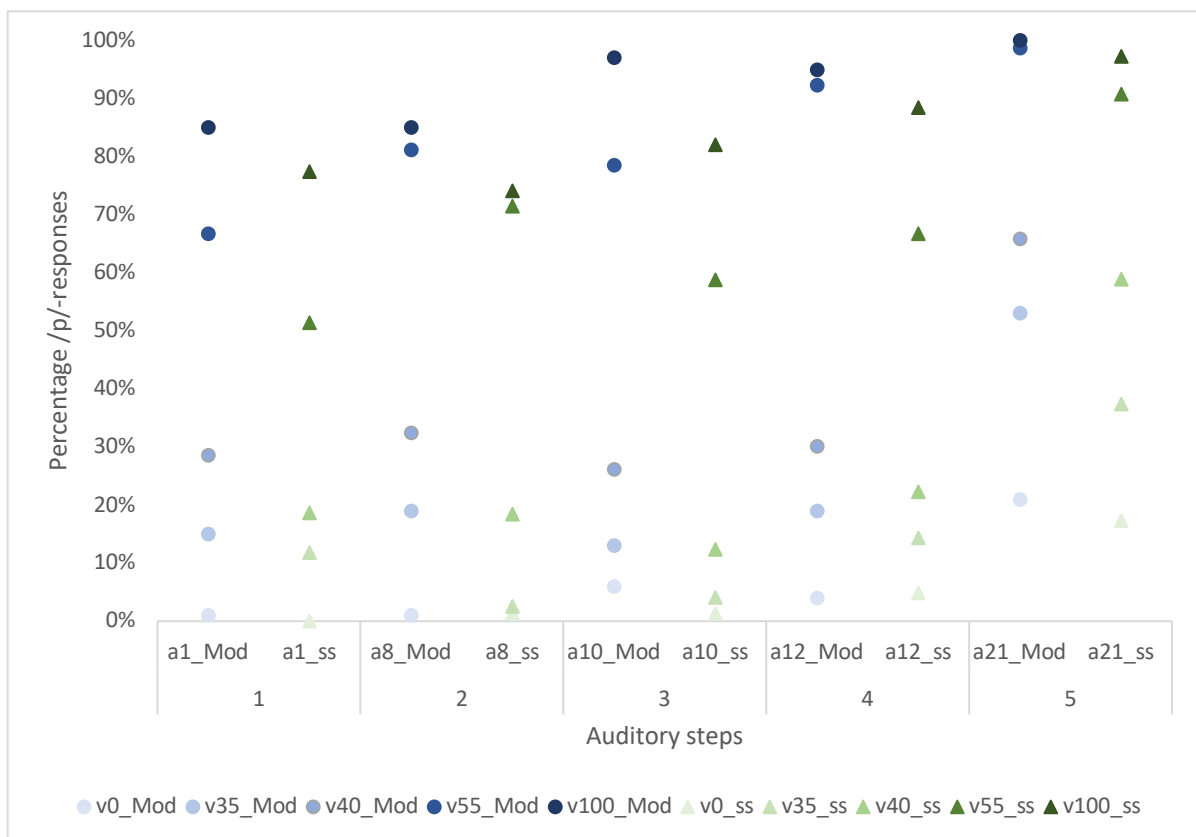


Figure 1: Mean percentage of /p/-responses on the combined steps of the visual (the lightest shade = least /p/-like and the darkest shade is most /p/-like) and the auditory continuum (1 = least /p/-like to 5 = most /p/-like).

responses for each combination of the type of noise, the visual steps and the auditory steps, even though a difference in the total number of /p/-responses between the two noise conditions was seen.

There was, however, a significant interaction effect found between the /p/-responses on the visual and auditory steps ($F(16,144)=3.39$, $p < .001$). Since the small sample size could be the cause of the non-significance of the three-way interaction effect, we were interested in checking whether there was a significant interaction effect between Audio and Visual. This could tell us if there, even though there was not when combined with Type, was a difference between these combinations on the number of /p/-responses. If indeed there was, bigger sample sizes might influence the significance level in a (to our research question) favorable way, which can be helpful for further research. Therefore, we created a new variable based on these findings. The new variable, named *combined variable* (C_v# for the visual steps variable and C_a# for the auditory steps variable) was created by averaging the number of /p/-responses to the modulated stimuli with the number of /p/-responses to the stationary noise. Again, a repeated-measure design was conducted with the new variable, leaving out the type of noise. The results again showed a significant main effect for both Visual ($F(4,40)=52.32$, $p < .001$) and Audio ($F(4,40)=10.37$, $p = .005$), meaning that both have a significant effect on the number of /p/-responses. Also, a significant interaction effect was again found between the /p/-responses on the visual and auditory steps ($F(16,160)=3.56$, $p = .008$), meaning that there was a significant difference between the number of /p/-responses on the combined visual and auditory steps. To follow up this interaction effect and in order to understand which specific stimuli differ from each other, pairwise comparisons were

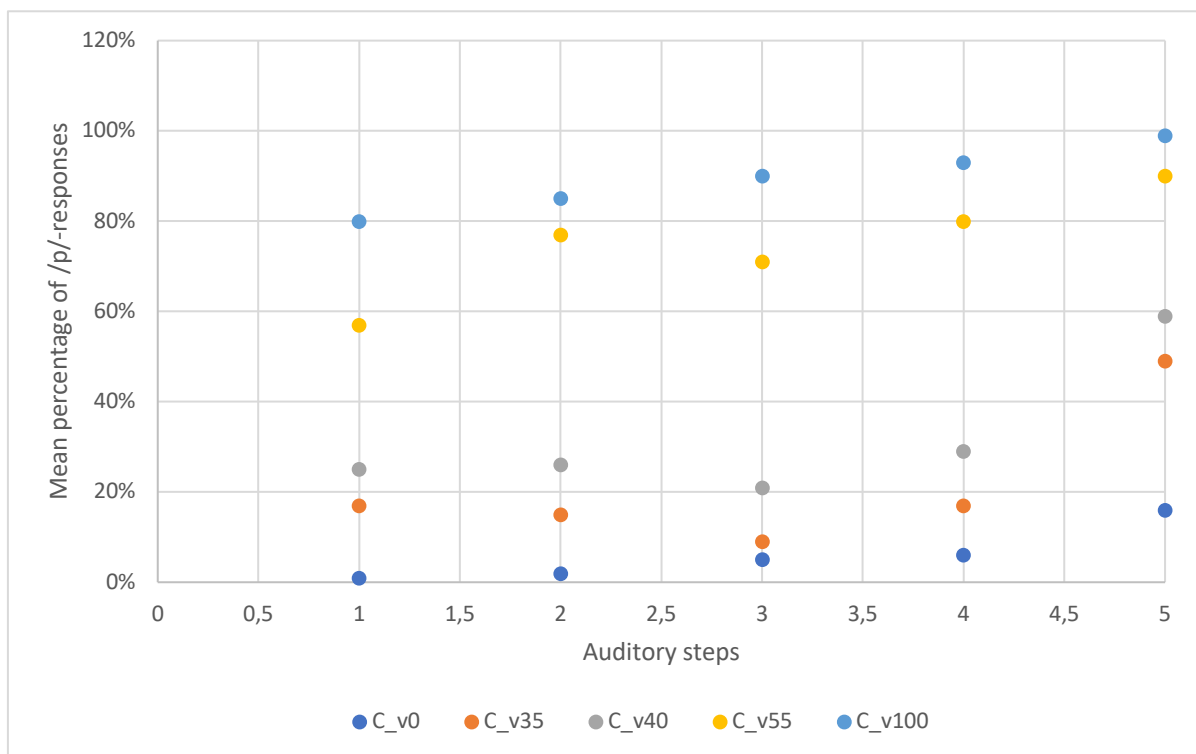


Figure 2: The mean percentage of /p/-responses on the combined steps of the visual and auditory continuum for the stationary noise responses and modulated noise responses combined into a new variable, namely the combined variable (labelled in the graph as C_v#).

done for each visual step on the auditory steps separately. These results are shown in Figure 2. The significant p-values are displayed in table 2. Significant difference is shown between the three least /p/-like visual steps and the two most /p/-like visual steps, as can be seen by the significant values presented in Figure 2. The most /p/-like visual step differs significantly from the three least /p/-like visual steps on each auditory step, and also from the second most /p/-like visual step on the first auditory step. The latter finding indicates a significantly weaker visual influence for the second most /p/-like visual step combined with the least /p/-like auditory step.

Table 2: The significant values of the pairwise comparison analysis are presented in this table. The color of the value indicated the corresponding visual step. The new variable representing the combined visual steps are presented as C_v#, C_v0 being the least /p/-like visual step and C_v100 being the most /p/-like visual step. The new variable representing the combined auditory steps are presented as C_a#, C_a1 being the least /p/-like auditory step and C_a5 being the most /p/-like auditory step.

Visual step	C_a1	C_a2	C_a3	C_a4	C_a5
C_v0	$p < .001$ $p = .002$	$p < .001$ $p < .001$	$p < .001$ $p = .002$	$p < .001$ $p < .001$	$p < .001$ $p = .001$ $p = .004$ $p = .023$
C_v35	$p = .003$ $p = .031$	$p < .001$ $p < .001$	$p < .001$ $p < .001$	$p < .001$ $p < .001$	$p = .006$ $p = .023$
C_v40	$p = .004$	$p = .002$ $p = .003$	$p < .001$ $p = .001$	$p < .001$ $p = .003$	$p = .030$ $p = .004$
C_v55	$p = .002$ $p = .031$ $p = .016$	$p < .001$ $p < .001$ $p = .003$	$p = .002$ $p < .001$ $p = .001$	$p < .001$ $p < .001$ $p = .003$	$p = .001$
C_v100	$p < .001$ $p = .003$ $p = .004$ $p = .016$	$p < .001$ $p < .001$ $p = .002$	$p < .001$ $p < .001$ $p < .001$	$p < .001$ $p < .001$ $p < .001$	$p < .001$ $p = .006$ $p = .030$

Francisco et al. (2017) calculated the difference between the mean percentage of /p/-responses to the most /p/-like visual step minus the mean percentage of /p/-responses to the least /p/-like visual step, to further illustrate their findings by showing the visual benefit for each auditory step. A standardized score was then computed to account for a possible improvement due to a ceiling-effect. This score was computed by taking the difference as was calculated above and divide this number by 100 minus the mean percentage of /p/-responses to the most /p/-like visual step. However, in this study each score was divided by 101 instead of 100 to avoid calculation errors due to having to divide by 0. These results are shown in Figure 3. Again, the auditory steps are displayed as a, a1 being the least /p/-like auditory step and a21 being the most /p/-like auditory step. The abbreviation used for the modulated noise condition is *mod* and for the stationary noise condition *ss*.

As can be seen from Figure 3, the mean change in percentage of /p/-responses on the two least /p/-like auditory steps is around 5% for both the modulated noise condition and the stationary noise condition. The difference gets bigger on the two most /p/-like auditory steps for the stationary noise condition, with the biggest change of /p/-responses on the most /p/-like auditory step. In the modulated noise condition, the three most /p/-like auditory steps show the biggest change in the mean percentage of /p/-responses. The biggest change, however (around 78%), can be seen in the combined most /p/-like visual and auditory step. Looking at the graph, these data suggest greater influence of the visual stimuli on the auditory categorization in the modulated noise condition in comparison to the stationary noise condition. The fact that no significant difference is found might be due to the sample size, which is relatively small.

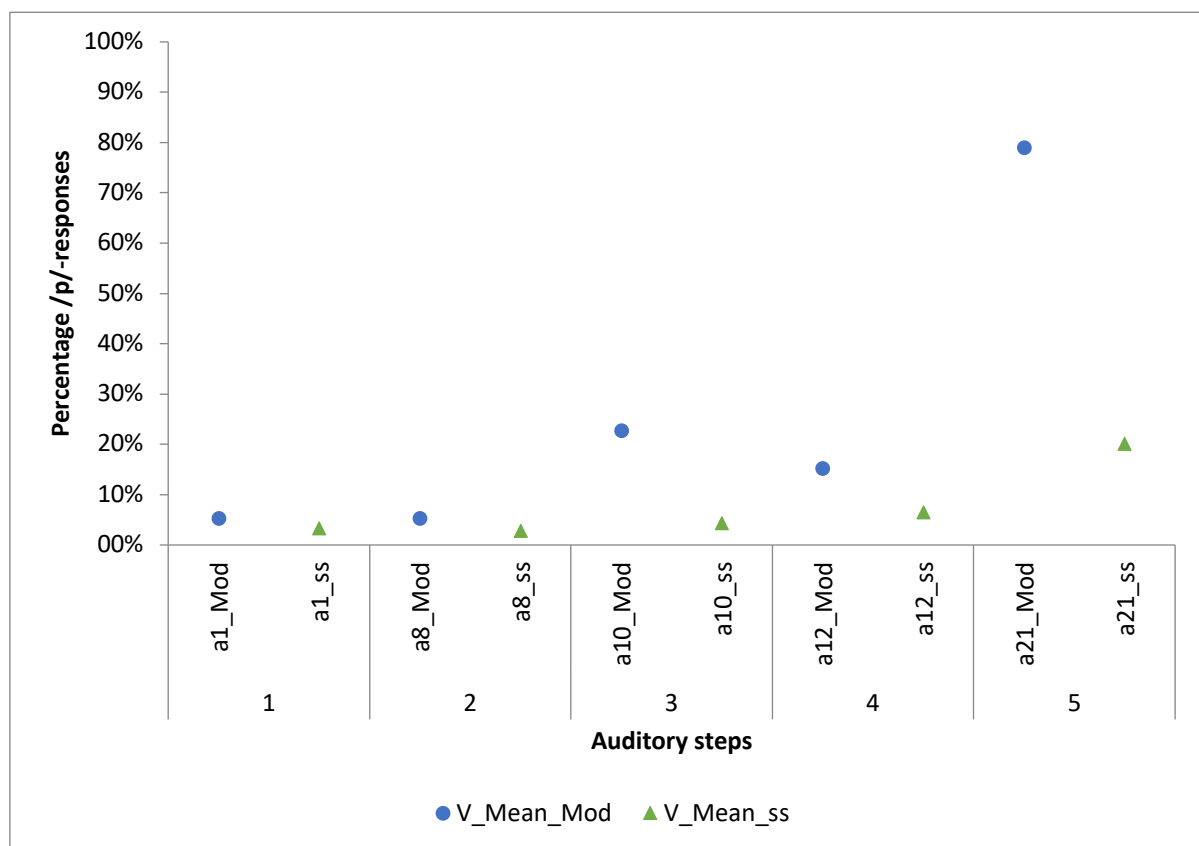


Figure 3: Mean change in the percentage of /p/-responses on the combined steps of the visual (dots = modulated noise condition (mod), triangles = stationary noise condition (ss)) and auditory continuum (1 = least /p/-like to 5 = most /p/-like).

3.2.2 Reaction time. The reaction times for the responses on the phonetic categorization task were used to investigate the second hypothesis: typically reading adults show different reaction times in the modulated noise condition than in the stationary noise condition. Figure 4 shows the mean of the reaction times for each auditory step on the /t-/p/-continuum combined to the visual steps on the /t-/p/-continuum. The auditory stimuli are placed on the horizontal axes, a standing for auditory. The number after a stands for the step on the continuum, a1 being the least /p/-like auditory step and a21

being the most /p/-like auditory step. All auditory steps are combined with both stationary noise and modulated noise. The abbreviation used for the stationary noise condition is ss and mod for the modulated noise condition. The dots represent the visual steps with modulated noise and the triangles represent the visual steps with stationary noise. The colors indicate the visual step on the /t-/p/-continuum, with the lightest color representing the least /p/-like visual step (v0) and the darkest color the most /p/-like visual step (v100).

Again, a repeated measure design was used to analyze the data. The order in which the noise stimuli were presented was analyzed using a repeated measure design with a between-subject factor added. The order of presentation was included in the design as a between-subject factor with two levels (1 = first the modulated noise condition, then the stationary noise condition and 2 = first the stationary noise condition, then the modulated noise condition). Results showed no significant effect for the order of presentation ($p = .489$).

Reaction time was used as the dependent variable and three within-subject factors were created: Type, including two levels (modulated noise and stationary noise), Visual, including five levels (the five visual steps on the /t-/p/-continuum and Audio, also including five levels (the five auditory steps on the /t-/p/-continuum. The output showed no significant main effect for either Type, Visual or Audio, meaning that none of these factors caused a significant difference in the reaction times of the participants on the phonetic categorization task. Also, no interaction effects were found, meaning that there was no significant difference in the reaction times overall. Since the assumption of equal variances was not met, a log transformation was used to create new data, and the same repeated measure design was conducted. Also, the same variables were used. Yet, again no significant differences were found for either Type, Visual or audio. Again, also no interaction effects were found. Therefore, we were not able to confirm the hypothesis, and we were not able to conclude any difference in reaction times between the two noise conditions on the phonetic categorization task.

When looking at the graph, it can be seen that the reaction times for the modulated noise condition are slightly longer. The mean reaction time is the longest on the least /p/-like auditory step, in combination with the most /p/-like visual step, suggesting that more time is needed to respond when the stimuli are extremes. On the two least /p/-like auditory steps, the most /p/-like visual steps show the slowest reaction times, but on the three most /p/-like auditory steps, the least /p/-like visual steps show the slowest reaction times. In all cases, the modulated noise stimuli show longer reaction times than the stationary noise condition stimuli, but not necessarily for the same visual stimulus on the

continuum. The longer reaction times in the modulated noise condition, even though not statistically significant, are promising for further research with a greater sample size.



Figure 2: Reaction times on the combined steps of the visual (the lightest shade = least /p/-like and the darkest shade is most /p/-like) and the auditory continuum (1 = least /p/-like to 5 = most /p/-like).

3.2.3 Correlation between the number of /p/-responses and speechreading abilities.

To investigate if the overall number of /p/-responses showed a correlation with the speechreading ability, a Spearman's correlation test was done. No significant correlation was found ($p = .773$). Since the sample was small ($n = 11$) we used the Kendall's Tau additional to check whether the outcome was different from the Spearman's correlation test. Again, results showed no significant correlation between participants' total number of /p/-responses on the phonetic categorization task and the speechreading task ($p = .694$). Therefore, we cannot confirm the third hypothesis.

4. Discussion

The aim of the current study was to contribute to this ongoing investigation by exploring the question whether an audio-visual deficit is present in adults with dyslexia. Due to time constraints, however, we were not able to test adults with dyslexia and we were only able to contribute to the above said investigation partially. We looked into the hypotheses from a typically reading adults perspective only and therefore adjusted the hypotheses according to this. Thus, after slightly changing the direction of the current study, we determined if adult typical readers:

- I. showed different visual influence on audio-visual perception of phonetic categories in a speech-shaped stationary noise condition than in a speech-shaped modulated noise condition;
- II. showed different reaction times in the modulated noise condition than in the stationary noise condition
- III. showed a correlation between speechreading ability and the visual influence on audio-visual speech perception.

4.1 Hypothesis I

So, instead of comparing individuals with dyslexia and without dyslexia, as was initially planned, we compared the visual influence on speech perception in two different noise-conditions, but in typically reading adults only. The first hypothesis was investigated by using the number of /p/-responses on the phonetic categorization task in a repeated measure design. The statistical analysis showed no significant difference of visual influence on speech perception in typical reading adults between the speech-shaped stationary noise condition and the speech-shaped modulated noise condition. Specifically, the visual influence was not bigger in one condition over the other. This is not in line with our expectations. The expectations were (based on earlier findings by Füllgrabe et al. (2006) and Ziegler et al. (2009)), that the typically reading adults would depend more on the visual stimulus in the speech-shaped stationary noise condition than in the speech-shaped modulated noise condition. Findings from multiple studies have shown that this was the case, and a variation of explanations have been given on why visual cues would be used more in stationary noise situations than in modulated noise situations. And even though the results in the current study were not significant, Figure 1, does suggest a difference in the mean number of /p/-responses between the two noise conditions. In Figure 1 it can be seen that the modulated noise conditions have more /p/-responses than the stationary conditions. The small sample size, however, might be the cause for the non-significance.

As has been discussed in the introduction, speech perception depends on the availability of acoustic cues. In natural speech, there is always noise that is masking at least some of these acoustic cues, which makes it more difficult to perceive what is being said. Moore (2008) argued that the absence

of information that is normally provided by the temporal fine structure (which is often/partly the case in a surrounding with noise), may hinder identification of the target. Also, it causes problems in the segregation of speech from the background noise based on cues. Füllgrabe et al. (2006) found that fluctuating background noises give the listener the chance to take advantage of these minima of different durations to receive cues about the consonant features. The range of these fluctuations determines the available acoustic cues needed to identify voicing, place of articulation or manner. This would, as we expected, suggest that the speech-shaped modulated noise conditions used in the current study would provide more of those “gaps” that can be used to receive acoustic cues about the /p/ or /t/, than would the speech-shaped stationary noise condition. The results of the current study may be explained by the fact that the cues for the recognition of stop-consonants, and the /p/ and /t/ in particular, might not have been interpretable within the minima of the gaps provided by the modulated noise conditions we used. Füllgrabe et al. (2006) found, that the minima of about 16 ms are optimal for speech perception mechanisms that mainly use spectral or temporal fine structure information, which is the case for place of articulation. Also, highest reception of the place of articulation was found to be at a modulation frequency of 32 Hz. However, using the same noise for each stimulus would give the same spectral or temporal fine structure each time. To prevent for this bias and availability, a slightly different noise modulation was used for each stimulus in the current study. It may have been the case that some stimuli provided the necessary cues for identification of either the /p/ or /t/, but the results show that the typically reading adults could not use the information provided by the varying “gaps” well enough to perceive the necessary cues for the identification of the /p/ or /t/.

The results from the pairwise comparison for Visual and Audio did show significant results for some of the auditory-visual combinations. Visual influence was significantly bigger for some of the auditory steps. As can be seen from Figure 2, especially the two most /p/-like visual steps have great influence on the response, independent of the auditory stimulus that it was combined with. Even on the two least /p/-like auditory steps, the number of -/p/responses is relatively high (more than 75%) when combined to the two most /p/like visual steps. This indicates that the typically reading adults rely fairly strongly on the visual cues when they are (strongly) incongruent with the auditory cues, and noise is present. This is in line with the knowledge so far on general speech perception in noise, which suggests a bigger influence of the visual cues when the auditory cues are not as accessible, as is the case in environments with (a lot of) noise. If the integration of speech sounds and letters is impaired in people with dyslexia, as has been proposed, it is expected that they do not show as much benefit of visual cues as the typically reading adults. It is suggested to investigate this expectation in future research.

Ziegler et al. (2009) found clear speech perception impairment for children with dyslexia in all noise conditions. They used both a speech-shaped steady-state noise condition and speech-shaped modulated noise conditions, using different modulation frequencies. They did not, however, find a

difference in speech perception between the modulation frequencies (4 Hz versus 32 Hz versus 128 Hz). Place of articulation showed the largest differences of the three consonantal features (place of articulation, voicing and manner) across groups, indicating that the children with dyslexia had the most difficulties perceiving the place of articulation. Also, this difference between groups was found for the average noise condition, the stationary noise condition and the 4 Hz modulated noise condition. As we used a /p/-/t/-continuum in the current study, which are two consonants that only differ in the place of articulation, the finding by Ziegler et al. (2009) is interesting for a study with a similar design comparing adults with and without dyslexia. Based on these findings it would be interesting to investigate whether the difference in perception of the place of articulation is as big between adults with and without dyslexia as it is in children with and without dyslexia, as has been shown by the results of Ziegler et al. (2009). This would (a) indicate whether the difficulty in perceiving the place of articulation in noise is still present in adults with dyslexia, and (b) whether adults, both with and/or without dyslexia, also show the most difficulty in perceiving place of articulation out of the three consonantal features place of articulation, voicing and manner.

In summary, the findings by Füllgrabe et al. (2006) show best reception of place of articulation at a modulation frequency of 32 Hz and Ziegler et al. (2009) found the biggest difference across groups between the stationary noise condition and the 4 Hz modulated noise condition. Lorenzi, Dumont and Füllgrabe (2000) tested children with and without dyslexia and adults on speech perception. They found that the dyslexic children showed poorer reception of voicing for processed speech, and poorer reception of manner, voicing, and place of articulation for unprocessed speech. This is all in favor of repeating the current study but with adults with dyslexia and typically reading adults as the control group. Therefore, it is suggested to (a) repeated the current experiment, but include adults with dyslexia, and (b) use different frequencies for the modulated noise condition to receive more information about the consonantal features and which acoustic cues provide for the necessary information to identify those features.

4.2 Hypothesis II

The second question we addressed in the current study was whether the typically reading adults showed different reaction times in the speech-shaped stationary noise condition than in the speech-shaped modulated noise condition. We expected to find a significant difference, based on the fact that modulated noise makes speech perception more challenging. The results of the current study, however, do not show the expected significant difference in reaction times between the two noise conditions. Looking at the figure, on the other hand, raises the question whether another factor might cause the fact that the difference is not statistically significant. It is clearly seen that there is a difference between the number of /p/-responses on the two noise conditions. A factor that could play a role is the small

sample size. A bigger sample size might make the result that is now only suggested by the figure to actually be statistically significant.

It has been found that children and adults with dyslexia show impairments in speech perception and specifically the integration of speech sound – to – letter – mappings. Therefore, the original expectations were that the people with dyslexia would show slower reaction times in both noise conditions, but especially in the speech-shaped modulated noise condition. This was hypothesized based on the findings that people with dyslexia show impairments in segregating speech sounds, which is the process of separating speech signals from interfering noise(s) (Manis et al., 1997; Boets, et al. 2007). This process is a lot more complex for the listener when the surrounding noises are not steady. In addition, people with dyslexia have shown an impairment in the use of temporal and spectral information for speech perception, both in quiet and in noise conditions (Lorenzi, et al. (2000). In order to investigate this, adults with dyslexia will have to be tested in the future. Since we were not able to do so, we investigated whether typically reading adults showed a difference in reaction times between the two noise conditions. Studies have shown that speech perception in fluctuating noise is more difficult than in steady-state noise, since it requires greater cognitive skill (e.g. Dole, Hoen and Meunier, 2012). Whilst steady-state noise requires attention for continuous input, the modulated noise conditions require more flexible switching skills from the listener. The reception of the auditory cues depends on the ability to segregate the speech sounds from the fluctuations, and we do so by using multiple acoustic cues and a variable amount of processing strategies (Dole et al., 2010). Since this is cognitively more difficult, it was expected for the typically reading adults in the current study that more time is needed to respond in the speech-shaped modulated noise condition than in the speech-shaped stationary noise condition. We found, in contrast to our hypothesis, that there is no significant difference in reaction times across the noise conditions. There are a number of possible explanations for the contradicting results. One is that we still have to consider the small sample size, and therefore the results may not be accurate. A second explanation is that the variability in the modulations was too small in the modulation condition and therefore these stimuli were not associated as more complex than the stationary noise condition stimuli by the listener.

4.3 Hypothesis III

The third question we addressed in the current study is whether the scores on the speechreading task correlated with the visual influence in the phonetic categorization task. A correlation was expected, since multiple studies have proven a correlation between the ability to lipread and audio-visual perception (e.g. Erber, 1974; Macleod and Summerfield, 1986). It was expected that the individuals who showed high(er) scores on the speechreading task would be better at interpreting and therefore relying on the visual cues. The results of the Kendall's Tau correlation test, however, showed no significant

correlation between the scores on the speechreading task and the visual influence in the phonetic categorization task. Since the sample size was relatively small we had to use a statistical test that was robust for the violated assumption of normal distribution. Therefore, we used the Kendall's Tau correlation test. Since studies have proven that the lipreading skills enhances speech perception, a correlation was expected to be found. Again, the small sample size may have contributed to the fact that no correlation was found.

4.4 General discussion

The fact that we were not able to include adults with dyslexia in the current study had the following consequences: First, we were not able to compare group means due to the fact that we were not able to test adults with dyslexia. Earlier studies did not find a difference between both children and adults with and without dyslexia, on speech perception in silence. Ziegler et al., 2009 found no difference in speech perception in an audio-only study between children with and without dyslexia, and Francisco et al., (2017) found the same results testing adults with and without dyslexia audio-visually. But Ziegler et al. (2009) showed promising results for an audio-visual investigation in adults with and without dyslexia, using noise conditions, since they did find a difference between the children with and without dyslexia in the noise conditions. It is, thus, suggested to repeat this study but include adults both with and without dyslexia to compare the group means and investigate the visual influence on audio-visual speech perception in noise.

The sample used in the current study was relatively small due to time constraints. A total of 11 participants were included in the main experiment. Two out of the 11 participants did not complete a hearing screening due to software problems. In addition, three out of the nine other participants did not meet the initial hearing criteria of 20 dB HL. Since the sample was very small, we decided to compensate and change the hearing criteria to 30 dB HL, so we could include all participants. Yet, this has to be considered when interpreting the results. A smaller sample group often results in violation of the assumption of normal distribution. This was also the case in the current study. Both for the number of /p/-responses and the reaction times, the assumption of normal distribution was violated. Therefore, we computed new data using a log-transformation. We then again conducted the repeated-measure design with this new data, using the same within subject factors and between subject factor, and checked whether the log-transformed data showed a different outcome. This was not the case. Yet, this also has to be considered when using this data for interpretation(s) or references in other studies.

5. Conclusion

We tested for the following hypotheses: (a) the typically reading adults show different visual influence on audio-visual perception of phonetic categories in a stationary noise condition as in a modulated noise condition, (b) the typically reading adults show longer reaction times in the speech-shaped modulated noise condition than in the speech-shaped stated noise condition, and (c) there is a correlation between speechreading ability and the visual influence on audio-visual speech perception. Based on the results we conclude that both hypotheses (a) and (b) are true, and we reject hypothesis (c). In summary, we found the following results: no significant difference in visual influence was found between the two noise conditions. When leaving out the factor type of noise, however, there was a significant difference in visual influence between the auditory steps. Especially the most /p/-like visual steps show significant influence, independent of the auditory step. Reaction times did not differ across the two noise conditions, even though a difference was expected based on earlier studies. Also, no correlation was found between the scores on the speechreading task and the scores on the phonetic categorization task. Since we were not able to execute the experiment as we intended to, due to time constraints, we conclude that the findings might not be representative. Further research using a similar design is suggested, especially following the initial experimental design.

References

- Adlard, A., & Hazan, V. (1998). Speech perception in children with specific reading difficulties (dyslexia). *The Quarterly Journal of Experimental Psychology Section A*, 51(1), 153–177. <https://doi.org/10.1080/713755750>
- Bastien-Toniazzo, M., Stroumza, A., & Cavé, C. (2010). Audio-visual perception and integration in developmental dyslexia: An exploratory study using the McGurk effect. *Current Psychology Letters. Behaviour, Brain & Cognition*, 25(3), 1–17. Received from: <https://journals.openedition.org/cpl/4928#authors>
- Bexkens, A., Van den Wildenberg, W. P. M., & Tijms, J. (2014). Rapid automatized naming in children with dyslexia: Is inhibitory control involved? *Dyslexia*, 21(3), 212–234. <https://doi.org/10.1002/dys.1487>
- Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin*, 130(6), 858–886. <https://doi.org/10.1037/0033-2909.130.6.858>
- Blau, V., Van Atteveldt, N., Ekkebus, M., Goebel, R., & Blomert, L. (2009). Reduced neural integration of letters and speech sounds links phonological and reading deficits in adult dyslexia. *Current Biology*, 19(6), 503–508. <https://doi.org/10.1016/j.cub.2009.01.065>
- Blau, V., Reithler, J., Van Atteveldt, N., Seitz, J., Gerretsen, P., Goebel, R., & Blomert, L. (2010). Deviant processing of letters and speech sounds as proximate cause of reading failure: A functional magnetic resonance imaging study of dyslexic children. *Brain*, 133(3), 868–879. <https://doi.org/10.1093/brain/awp308>
- Blomert, L. (2011). The neural signature of orthographic–phonological binding in successful and failing reading development. *NeuroImage*, 57(3), 695–703. <https://doi.org/10.1016/j.neuroimage.2010.11.003>
- Blomert, L., & Froyen, D. (2010). Multi-sensory learning and learning to read. *International Journal of Psychophysiology*, 77(3), 195–204. <https://doi.org/10.1016/j.ijpsycho.2010.06.025>
- Blomert, L., & Willems, G. (2010). Is there a causal link from a phonological awareness deficit to reading failure in children at familial risk for dyslexia? *Dyslexia*, 16, 300–317. <http://dx.doi.org/10.1002/dys.405>
- Boets, B., Ghesquière, P., Van Wieringen, A., & Wouters, J. (2007). Speech perception in preschoolers at family risk for dyslexia: Relations with low-level auditory processing and phonological ability. *Brain and Language*, 101(1), 19–30. <https://doi.org/10.1016/j.bandl.2006.06.009>
- Bogliotti, C., Serniclaes, W., Messaoud-Galusi, S., & Sprenger-Charolles, L. (2008). Discrimination of speech sounds by children with dyslexia: Comparisons with chronological age and reading level controls. *Journal of Experimental Child Psychology*, 101(2), 137–155. <https://doi.org/10.1016/j.jecp.2008.03.006>
- Boliek, C., Keintz, C., Norrix, L., & Obrzut, J. (2010). Auditory-visual perception of speech in children with learning disabilities: The McGurk effect. *Canadian Journal of Speech-Language Pathology and Audiology*, 34(2), 124–131. Received from: https://cjslpa.ca/files/2010_CJSLPA_Vol_34/No_02_81-152/Boliek_keintz_norrix_obrzut_CJSLPA_2010.pdf
- Bradley, L., & Bryant, P.E. (1983). Categorizing sounds and learning to read—a causal connection. *Nature*, 301, 419–421. <https://www.nature.com/articles/301419a0>
- Brancazio, L. (2004). Lexical influences in audiovisual speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, 30(3), 445–463. <https://doi.org/10.1037/0096-1523.30.3.445>
- Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Developmental Psychology*, 28(5), 874–886.
- Castles, A., & Friedmann, N. (2014). Developmental dyslexia and the phonological deficit hypothesis. *Mind & Language*, 29(3), 270–285. <https://doi.org/10.1111/mila.12050>

- Catts, H. W., Adlof, S. M., Hogan, T. P., & Weismer, S. E. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech Language and Hearing Research*, 48(6), 1378–1396. [https://doi.org/10.1044/1092-4388\(2005/096\)](https://doi.org/10.1044/1092-4388(2005/096))
- Chiappe, P., Chiappe, D. L., & Siegel, L. S. (2001). Speech perception, lexicality, and reading skill. *Journal of Experimental Child Psychology*, 80(1), 58–74. <https://doi.org/10.1006/jecp.2000.2624>
- Davies, D. (1968). Physiological and psychological effects of exposure to high intensity noise. *Applied Acoustics*, 1(3), 215–233. [https://doi.org/10.1016/0003-682x\(68\)90024-8](https://doi.org/10.1016/0003-682x(68)90024-8)
- De Bree, E., Rispens, J., & Gerrits, E. (2007). Non-word repetition in Dutch children with (a risk of) dyslexia and SLI. *Clinical Linguistics & Phonetics*, 21(11-12), 935–944. <https://doi.org/10.1080/02699200701576892>
- Dehaene, S. (2009). *Reading in the brain: the science and evolution of a human invention*. New York, USA: Viking Adult.
- Depessemier, P., & Andries, C. (2009). *Test voor Gevorderd Lezen en Schrijven [Test for Advanced Reading and Writing]*. Antwerpen, Belgium–Apeldoorn, the Netherlands: Garant.
- Dole, M., Hoen, M., & Meunier, F. (2012). Speech-in-noise perception deficit in adults with dyslexia: Effects of background type and listening configuration. *Neuropsychologia*, 50(7), 1543–1552. <https://doi.org/10.1016/j.neuropsychologia.2012.03.007>
- Dupont, S., & Menard, L. (2005). Vowel production and perception in French blind and sighted adults. *The Journal of the Acoustical Society of America*, 117(4), 2622. <https://doi.org/10.1121/1.4778327>
- Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the national reading panel's meta-analysis. *Reading Research Quarterly*, 36(3), 250–287. <https://doi.org/10.1598/rrq.36.3.2>
- Elbro, C., Nielsen, I., & Petersen, D. K. (1994). Dyslexia in adults: Evidence for deficits in non-word reading and in the phonological representation of lexical items. *Annals of Dyslexia*, 44(1), 203–226. Received from: <https://link.springer.com/article/10.1007/BF02648162>
- Erber, N. P. (1974). Visual perception of speech by deaf children: Recent developments and continuing needs. *Journal of Speech and Hearing Disorders*, 39(2), 178–185. <https://doi.org/10.1044/jshd.3902.178>
- Francisco, A. A., Jesse, A., Groen, M. A., & McQueen, J. M. (2017). A general audiovisual temporal processing deficit in adult readers with dyslexia. *Journal of Speech Language and Hearing Research*, 60(1), 144–158. https://doi.org/10.1044/2016_jslhr-h-15-0375
- Füllgrabe, C., Berthommier, F., & Lorenzi, C. (2006). Masking release for consonant features in temporally fluctuating background noise. *Hearing Research*, 211(1-2), 74–84. <https://doi.org/10.1016/j.heares.2005.09.001>
- Goulandris, N. K., & Snowling, M. (1991). Visual memory deficits: A plausible cause of developmental dyslexia? Evidence from a single case study. *Cognitive Neuropsychology*, 8(2), 127–154. <https://doi.org/10.1080/02643299108253369>
- Grant, K. W., & Seitz, P. (2000). The use of visible speech cues for improving auditory detection of spoken sentences. *The Journal of the Acoustical Society of America*, 108(3), 1197. <https://doi.org/10.1121/1.1288668>
- Groen, M. A., & Jesse, A. (2013). Audiovisual speech perception in children and adolescents with developmental dyslexia: No deficit with McGurk stimuli.
- Hahn, N., Foxe, J. J., & Molholm, S. (2014). Impairments of multisensory integration and cross-sensory learning as pathways to dyslexia. *Neuroscience & Biobehavioral Reviews*, 47, 384–392. <https://doi.org/10.1016/j.neubiorev.2014.09.007>
- Hayes, E. A., Tiippana, K., Nicol, T. G., Sams, M., & Kraus, N. (2003). Integration of heard and seen speech: A factor in learning disabilities in children. *Neuroscience Letters*, 351(1), 46–50. [https://doi.org/10.1016/s0304-3940\(03\)00971-6](https://doi.org/10.1016/s0304-3940(03)00971-6)

- Hockey, R. (1973). Changes in information-selection patterns in multisource monitoring as a function of induced arousal shifts. *Journal of Experimental Psychology*, 101(1), 35–42. <https://doi.org/10.1037/h0035467>
- Hogan, T. P., Catts, H. W., & Little, T. D. (2005). The relationship between phonological awareness and reading. *Language Speech and Hearing Services in Schools*, 36(4), 285–293. [https://doi.org/10.1044/0161-1461\(2005/029\)](https://doi.org/10.1044/0161-1461(2005/029))
- Hulme, C., & Snowling, M. J. (2013). The interface between spoken and written language: Developmental disorders. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1634), 1–8. <https://doi.org/10.1098/rstb.2012.0395>
- Jesse, A., & Janse, E. (2012). Audiovisual benefit for recognition of speech presented with single-talker noise in older listeners. *Language and Cognitive Processes*, 27(7-8), 1167–1191. <https://doi.org/10.1080/01690965.2011.620335>
- Jones, M. W., Snowling, M. J., & Moll, K. (2016). What automaticity deficit? Activation of lexical information by readers with dyslexia in a rapid automatized naming Stroop-switch task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(3), 465–474. <https://doi.org/10.1037/xlm0000186>
- Lalonde, K., & Holt, R. F. (2016). Audiovisual speech perception development at varying levels of perceptual processing. *The Journal of the Acoustical Society of America*, 139(4), 1713–1723. <https://doi.org/10.1121/1.4945590>
- Lervåg, A., Bråten, I., & Hulme, C. (2009). The cognitive and linguistic foundations of early reading development: A Norwegian latent variable longitudinal study. *Developmental Psychology*, 45(3), 764–781. <https://doi.org/10.1037/a0014132>
- Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54(5), 358–368. <https://doi.org/10.1037/h0044417>
- Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, 53(1), 1–14. <https://doi.org/10.1007/s11881-003-0001-9>
- Manis, F. R., Mcbride-Chang, C., Seidenberg, M. S., Keating, P., Doi, L. M., Munson, B., & Petersen, A. (1997). Are speech perception deficits associated with developmental dyslexia? *Journal of Experimental Child Psychology*, 66(2), 211–235. <https://doi.org/10.1006/jecp.1997.2383>
- Massaro, D. W. (1984). Children's perception of visual and auditory speech. *Child Development*, 55(5), 1777–1788. <https://doi.org/10.2307/1129925>
- McGettigan, C., Faulkner, A., Altarelli, I., Obleser, J., Baverstock, H., & Scott, S. K. (2012). Speech comprehension aided by multiple modalities: Behavioural and neural interactions. *Neuropsychologia*, 50(5), 762–776. <https://doi.org/10.1016/j.neuropsychologia.2012.01.010>
- McGurk, H., & McDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746–748. Received from: <http://www.nature.com.ru.idm.oclc.org/articles/264746a0.pdf>
- Moore, B. C. J. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *Journal of the Association for Research in Otolaryngology*, 9(4), 399–406. <https://doi.org/10.1007/s10162-008-0143-x>
- Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, 7(4), 323–331. [https://doi.org/10.1016/0010-0277\(79\)90020-9](https://doi.org/10.1016/0010-0277(79)90020-9)
- Nittrouer, S. (2005). Age-related differences in weighting and masking of two cues to word-final stop voicing in noise. *The Journal of the Acoustical Society of America*, 118(2), 1072–1088. <https://doi.org/10.1121/1.1940508>
- Noordenbos, M. W., & Serniclaes, W. (2015). The categorical perception deficit in dyslexia: a meta-analysis. *Scientific Studies of Reading*, 19(5), 340–359. <https://doi.org/10.1080/10888438.2015.1052455>
- Oxenham, A. J. (2014). Masking and masking release. *Encyclopedia of Computational Neuroscience*, 1–4. https://doi.org/10.1007/978-1-4614-7320-6_432-1

- Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly*, 33(3), 283–319. <http://psycnet.apa.org/record/1987-33995-001>
- Poulsen, M., Juul, H., & Elbro, C. (2012). Multiple mediation analysis of the relationship between rapid naming and reading. *Journal of Research in Reading*, 38(2), 124–140. <https://doi.org/10.1111/j.1467-9817.2012.01547.x>
- Ramus, F. (2001). Outstanding questions about phonological processing in dyslexia. *Dyslexia*, 7(4), 197–216. <https://doi.org/10.1002/dys.205>
- Rocheron, I., Lorenzi, C., Füllgrabe, C., & Dumont, A. (2002). Temporal envelope perception in dyslexic children. *NeuroReport*, 13(13), 1683–1687. Received from: https://journals.lww.com/neuroreport/Abstract/2002/09160/Temporal_envelope_perception_in_dyslexic_children.23.aspx
- Serniclaes, W., Sprenger-Charolles, L., Carre', R., & Demonet, J. (2001). Perceptual discrimination of speech sounds in developmental dyslexia. *Journal of Speech Language and Hearing Research*, 44(2), 384–399. [https://doi.org/10.1044/1092-4388\(2001/032\)](https://doi.org/10.1044/1092-4388(2001/032))
- Snowling, M. J. (2000). *Dyslexia* (2nd ed.). Oxford: Blackwell.
- Stacey, P. C., Kitterick, P. T., Morris, S. D., & Sumner, C. J. (2016). The contribution of visual information to the perception of speech in noise with and without informative temporal fine structure. *Hearing Research*, 336, 17–28. <https://doi.org/10.1016/j.heares.2016.04.002>
- Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America*, 26(2), 212–215. <https://doi.org/10.1121/1.1907309>
- Swan, D., & Goswami, U. (1997). Phonological awareness deficits in developmental dyslexia and the phonological representations hypothesis. *Journal of Experimental Child Psychology*, 66(1), 18–41. <https://doi.org/10.1006/jecp.1997.2375>
- Taniguchi, K., & Tayama, T. (2010). Comparison between discrimination and identification processes using line-drawing stimuli. *Journal of Vision*, 10(7), 991. <https://doi.org/10.1167/10.7.991>
- Toscano, J. C., McMurray, B., Dennhardt, J., & Luck, S. J. (2010). Continuous perception and graded categorization. *Psychological Science*, 21(10), 1532–1540. <https://doi.org/10.1177/0956797610384142>
- Van den Bunt, M. R., Groen, M. A., Ito, T., Francisco, A. A., Gracco, V. L., Pugh, K. R., & Verhoeven, L. (2017). Increased response to altered auditory feedback in dyslexia: a weaker sensorimotor magnet implied in the phonological deficit. *Journal of Speech Language and Hearing Research*, 60(3), 654–667. https://doi.org/10.1044/2016_jslhr-l-16-0201
- Vandermosten, M., Boets, B., Luts, H., Poelmans, H., Golestani, N., Wouters, J., & Ghesquiere, P. (2010). Adults with dyslexia are impaired in categorizing speech and nonspeech sounds on the basis of temporal cues. *Proceedings of the National Academy of Sciences*, 107(23), 10389–10394. <https://doi.org/10.1073/pnas.0912858107>
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, 45(1), 2–40. <https://doi.org/10.1046/j.0021-9630.2003.00305.x>
- Warren, P., & Marslen-Wilson, W. (1988). Cues to lexical choice: discriminating place and voice. *Perception & Psychophysics*, 43, 21–30.
- Wechsler, D. (2012). *Wechsler Adult Intelligence Scale—Fourth Edition—NL, afname en scoringshandleiding [administration and scoring manual]*. Amsterdam, the Netherlands: Pearson Assessment and Information BV.
- Werker, J., & Tees, R. (1987). Speech perception in severely disabled and average reading children. *Canadian Journal of Psychology*, 41(1), 48–61. Received from: <https://ovidsp-tx-ovid-com.ru.idm.oclc.org/sp-3.31.1b/ovidweb.cgi?QS2=434f4e1a73d37e8c64b5117be1bf78b109ce55b44252390724b97b0234e3af3df0eb4e8d8dfe0c4d9ed448e8b32b4500887eeca9bdd5d1747d2386bc549daa4b08b325c78ba379058e9c9bc01fd138c53eb8cd171264ce9484112027f9bafa3719ef48dc3c68d3029d201a478dd2cee10788053c50eace12bd15dc01f19616ef9cbdd3ea24cc97e0cedfc8bea9>

1ed4eec14a8a7b2266468067b3f819990966a5209a071eb2429356231e16ea3ea2b85f8d851f
dabd3226259b13abcb4d1925c133cd2e9952c153d40f171a3c71af19ec6e1c898a8e5181d731c
78178debb3179495427208906ed8c22fb3e7513061c23865ae1118b133eefdd9b5d7e2deb28
5bd230b873ea0650

- Widmann, A., Gruber, T., Kujala, T., Tervaniemi, M., & Schroger, E. (2007). Binding symbols and sounds: Evidence from event-related oscillatory gamma-band activity. *Cerebral Cortex*, 17(11), 2696–2702. <https://doi.org/10.1093/cercor/bhl178>
- Widmann, A., Schröger, E., Tervaniemi, M., Pakarinen, S., & Kujala, T. (2012). Mapping symbols to sounds: Electrophysiological correlates of the impaired reading process in dyslexia. *Frontiers in Psychology*, 3(60), 1–12. <https://doi.org/10.3389/fpsyg.2012.00060>
- Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, 91(3), 415–438. <https://doi.org/10.1037/0022-0663.91.3.415>
- Yeung, H. H., & Werker, J. F. (2013). Lip movements affect infants' audiovisual speech perception. *Psychological Science*, 24(5), 603–612. <https://doi.org/10.1177/0956797612458802>
- Yopp, H. K., & Yopp, R. H. (2009, January). Phonological awareness is child's play! *Young children*, 2009(January), 1–9. Retrieved from http://teachingcommons.cdl.edu/tk/modules_teachers/documents/PhonologicalAwarenessIsChildsPlay.pdf
- Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science*, 12(5), 732–745. <https://doi.org/10.1111/j.1467-7687.2009.00817.x>