Purpose: This study compares online word recognition and prediction in preschoolers with (a suspicion of) a developmental language disorder (DLD) and typically developing (TD) controls. Furthermore, it investigates correlations between these measures and the link between online and off-line language scores in the DLD group.

Method: Using the visual world paradigm, Dutch children ages 3;6 (years;months) with (a suspicion of) DLD ($n = 51$) and TD peers ($n = 31$) listened to utterances such as, “Kijk, een hoed!” (Look, a hat!) in a word recognition task, and sentences such as, “Hé, hij leest gewoon een boek” (literal translation: Hey, he reads just a book) in a word prediction task, while watching a target and distractor picture.

Results: Both groups demonstrated a significant word recognition effect that looked similar directly after target onset. However, the DLD group looked longer at the target than the TD group and shifted slower from the distractor to target pictures. Within the DLD group, word recognition was linked to off-line expressive language scores. For word prediction, the DLD group showed a smaller effect and slower shifts from verb onset compared to the TD group. Interestingly, within the DLD group, prediction behavior varied considerably, and was linked to receptive and expressive language scores. Finally, slower shifts in word recognition were related to smaller prediction effects.

Conclusions: While the groups’ word recognition abilities looked similar, and only differed in processing speed and dwell time, the DLD group showed atypical verb-based prediction behavior. This may be due to limitations in their processing capacity and/or their linguistic knowledge, in particular of verb argument structure.

It is estimated that around 4%–7% of preschool children have a developmental language disorder (henceforth DLD; Law et al., 2000; Lindsay & Strand, 2016). Children with DLD have a core deficit in expressive and/or receptive language without the presence of any neurological or genetic disorders, hearing loss, or intellectual impairments (Leonard, 2014). DLD is typically diagnosed at preschool age (Bishop, 2014). At a younger age, one often speaks of a suspicion of DLD. It is, however, difficult to forecast if children who show severe difficulties in language development at an early age remain experiencing problems later in life. In a longitudinal study, Tomblin et al. (2003) followed children with a suspicion of DLD from kindergarten to fourth grade and found that the disorder persisted for 44% of these children. It would be of significant importance to gain a better understanding of which children are likely to catch up and which children are less likely to do so. There is a considerable variation in the nature of the language areas that are affected in DLD and in the severity of the language problems (Leonard, 2014; van Weerdenburg et al., 2006). The heterogeneity of the group children with DLD, especially at a young age, may be a starting point when looking for predictors for later language development. This raises the question of how to best capture the language abilities of young children with (a suspicion) of DLD and, more specifically, how off-line standardized language measures are associated with better or poorer online language outcomes.

There seems to be a wide consensus on two main linguistic domains that are affected in children with DLD:
phonology and (morpho)syntax (e.g., Leonard et al., 2004; Rice et al., 2009; Sheng & McGregor, 2010). One of the first signals in young children with DLD is a poor and/or delayed vocabulary development (e.g., Leonard & Deevy, 2006; Montgomery, 2006), next to delayed phonological development. Previous work has demonstrated that children with DLD experience special difficulty with verbs—a special set of words in the vocabulary—and their argument structure (e.g., Andreu, Sanz-Torrent, Legaz, & MacWhinney, 2012; Bishop, 1997; Verhoeven & Van Balkom, 2004). Furthermore, they show impaired working memory (e.g., Lum et al., 2011). These two components, namely, verbs and their argument structure and working memory, are crucial for successful sentence processing. The verb is specifically important for establishing meaning, as it selects which types of nouns can function as suitable arguments. In addition, it is the listeners’ task to hold and integrate multiple words in working memory when the sentence unfolds over time.

In clinical settings, expressive and receptive language skills are typically assessed as part of the test battery used to diagnose whether children have DLD, and for planning of and monitoring the effect of treatment. These measures are taken by using standardized off-line language tests, which provide information about the words children know. Such tests give no insight into the online recognition of spoken words, while the real problem may actually be better understood by looking at the course of the recognition process itself. Furthermore, to our knowledge, only one study on young children with DLD has looked at whether results from standardized off-line language tasks relate to results from online tasks (Pijnacker et al., 2017).

In this study, we aim to investigate the real-time processing of words in isolation and in sentences in young children with (a suspicion of) DLD. More specifically, we are interested in the unconscious online mental processes during word recognition and prediction. A wealth of research on typically developing (TD) children has shown that they have a remarkable capacity to recognize language with great ease and rapidity (e.g., Marchman & Fernald, 2008; Swingley et al., 1999). This ability is at least partially the result of generating predictions on the basis of incoming auditory input, which allows them to proactively construct interpretations before the speaker has completed the word or the utterance (e.g., Borovsky et al., 2012; Mani & Huettig, 2012). This so-called predictive or anticipatory processing has been described as a core ingredient of efficient communication (Altmann & Mirković, 2009). It is possible that the activation, selection, and integration of words when auditory input unfolds is delayed for children with (a suspicion of) DLD compared to TD children.

Online word recognition and prediction have predominantly been investigated in electrophysiological (e.g., in TD children: Friedrich & Friederici, 2005; in children with DLD: Pijnacker et al., 2017) and eye-tracking studies (e.g., in TD children: see Sedivy, 2010, and Trueswell, 2008, for an overview; in children with DLD: Andreu et al., 2013). In particular, the so-called visual world paradigm (Tanenhaus et al., 1995) has been an excellent technique for assessing real-time precision during spoken language processing. Typically, participants’ task in this paradigm is to listen to speech while looking at a visual display with two to four objects. Eye movements to those objects are closely time-locked to the speech and are taken to represent underlying activation levels of lexical candidates. The paradigm is specifically suitable for children with DLD, as it does not require them to make overt (button press) responses (Schul et al., 2004). The current study therefore used this paradigm to assess both online word recognition and prediction in children with (a suspicion of) DLD.

Word Recognition in TD and DLD Children

Many studies have investigated how TD children process words (e.g., Fernald et al., 1998, 2001; Swingley et al., 1999). Fernald et al. (1998), for example, demonstrated that the speed and accuracy in spoken word recognition increases significantly over the second year. In that study, 15- to 24-month-olds had to look at pictures of familiar objects while one of two objects was being named. Results showed that the 15-month-olds responded inconsistently and only looked at the correct object after the end of the target word; however, the 24-month-olds showed more reliable responses and already shifted their eye gaze halfway through the target word. This provided first evidence that 2-year-olds interpret spoken language incrementally, recognizing words on the basis of partial phonetic information.

Although there is general consensus on the finding that TD children become increasingly efficient in interpreting speech they hear, contradictory findings have been found for whether there is a relation between speech processing skills and vocabulary size. A couple of studies have found evidence for such a link (e.g., Fernald et al., 2001, 2006; Marchman & Fernald, 2008; Zangl et al., 2005). For example, Fernald et al. (2001) grouped 2-year-old children on the basis of how quickly they responded to words. Correlational analyses showed that children with faster mean response times had larger production vocabularies than those with slower response times. However, no such correlations were found in work by Swingley and Aslin (2000, 2002). These divergent findings have been explained by methodological and age differences between the studies.

Most of the work examining word recognition in children with DLD has used behavioral tasks in which participants had to make overt responses using a button-box (e.g., Dollahgahan, 1998; Edwards & Lahey, 1996; Montgomery, 2000; Montgomery et al., 1990). Those studies revealed that school-aged children with DLD are typically slower than their TD age-matched peers at recognizing words in sentences, indicating that they experience lexical retrieval problems. To our knowledge, the number of studies looking at the online recognition of spoken language in children with DLD is limited (visual world paradigm: Andreu, Sanz-Torrent, & Guardia-Olmos, 2012; McMurray et al., 2010; electroencephalography: Haebig et al., 2017, 2018; Pijnacker et al., 2017; Sabisch et al., 2006). In a visual world paradigm, McMurray et al. (2010) examined phonological processing
of isolated nouns in 17-year-old adolescents with DLD. The authors found that their initial lexical activation was similar to the control group but that later looks were more atypical and distributed among potential phonological competitors. Andreu, Sanz-Torrent, and Guardia-Olmos (2012) were interested in online noun and, in particular, verb processing, as children with DLD have shown a substantial delay in the use and (off-line) understanding of verbs and functional morphology (e.g., Andreu, Sanz-Torrent, Legaz, & MacWhinney, 2012; Bishop, 1997; Verhoeven & Van Balkom, 2004). The researchers therefore tested children with DLD between ages 5 and 8 years on their online recognition of nouns and verbs. Researchers therefore tested children with DLD between ages 5 and 8 years on their online recognition of nouns and verbs. Their eye gaze behavior was compared to that of TD children who were either matched on age or on their mean length of utterances (MLU). Eye movements were recorded as they searched an array of pictures in response to hearing a noun (e.g., a cake) or a verb (e.g., a girl painting on a canvas). Overall, all groups recognized nouns faster than verbs and one-verb-argument structures faster than two- and three-verb-argument structures. More specifically, children with DLD were slower than age-matched controls but not slower compared to MLU-matched controls.

Only few studies have used electroencephalography to assess online spoken word recognition in children with DLD. Haebig et al. (2018), for example, tested preschoolers with and without DLD on a word processing task in which pictures either matched or mismatched with an auditory label. No group differences were found in accuracy, and mismatched trials elicited an N400 of equal size and peak latency in both groups. However, the TD children demonstrated a more robust late positive component on mismatch trials, indicating that they may have a more advanced profile of postlexical reanalysis and integration. Furthermore, Pijnacker et al. (2017) tested slightly older preschool children with DLD and found that the semantic processing of nouns in sentences was affected in the children with DLD in comparison to a TD group of peers, as reflected by a smaller and delayed N400 effect for the DLD group. The N400 effect was associated with language comprehension, vocabulary, and grammar. That is, smaller N400 effects in the DLD group were associated with lower levels of language skills. The authors suggested that one of the possible explanations for these findings is that the DLD group experiences problems with verb argument structure (cf. Andreu, Sanz-Torrent, Legaz, & MacWhinney, 2012; Sabisch et al., 2006; Thordardottir & Ellis Weismer, 2002).

In summary, studies on word recognition in TD children have shown that, at age 2 years, they are able to understand spoken language incrementally. In addition, some studies on TD children have suggested that the speed of language processing is related to vocabulary development and input. In contrast, school-aged children with DLD have been demonstrated to experience serious difficulties with understanding words and verbs off-line. Online word and verb processing have hardly been investigated in children with DLD, and the results that have been found are inconsistent. These conflicting findings may be due to the difference in stimulus presentation across studies. More specifically, some studies have presented words in isolation or in a short carrier sentence together with a picture (e.g., Haebig et al., 2018), whereas others have offered words in relatively complex sentences without any supporting visual information (e.g., Pijnacker et al., 2017). It seems that children with DLD experience little problems with recognizing words in isolation, but they do find it hard to recognize words in sentences. A possible reason for this could be that the latter task is far more complicated, asking for a higher level of semantic integration. Another explanation for the conflicting results may be that the children that have been investigated had reached ceiling performance due to their relatively old age (except for Haebig et al., 2018). It would therefore be interesting to tap into children’s word recognition abilities at an even younger age, as it remains unclear whether relations between processing words and off-line language measures exist in children with (a suspicion of) DLD.

Word Prediction in TD and DLD Children

TD children cannot only rapidly recognize words as they unfold over time, but they have also shown to be able to make predictions as to which (morpho)syntactic structure (e.g., Brouwer et al., 2017; Deevy et al., 2017; Lew-Williams & Fernald, 2007; Trueswell & Gleitman, 2004) or lexical item comes next (e.g., Borovsky et al., 2012; Brouwer et al., 2019; Fernald et al., 2008; Mani & Huetting, 2012; Nation et al., 2003). This predictive behavior is seen as one of the mechanisms why language processing is so effortless and accurate. In order to proactively construct interpretations, it is important that knowledge of lexical semantics, and especially verbs, is intact. Verbs are complex as they can specify multiple arguments in roles such as agent, theme, and recipient, which asks for activation of both semantic and syntactic specification of the verb. Mani and Huetting (2012), for example, examined whether 2-year-olds could use the semantics of the verb to predict an upcoming noun. Using the visual world paradigm, they presented sentences such as “The boy eats/sees the big cake” while displaying two objects on the screen (e.g., a cake and a tree). Upon hearing the semantically constraining verb “eat,” children predicted that the direct object is likely to be something edible (i.e., the cake). However, upon hearing the neutral verb “see,” no such predictive behavior was observed. This same study also showed that children’s prediction skills were correlated with the size of their expressive vocabulary. Borovsky et al. (2012) demonstrated a positive link between 3- to 10-year-old children’s receptive vocabulary size and their prediction abilities. Note that these children were older than the children in Mani and Huetting’s (2012) study.

Although anticipatory behavior has been well attested in TD children, only three recent studies have examined whether children with DLD are able to move their eyes to an object prior to auditory input. First, Andreu et al. (2013) conducted three eye-tracking experiments with adults, 5-year-old children with DLD, and an age-matched and an
MLU-matched control group. They presented sentences such as “The boy carefully trims the paper” while displaying four objects on the screen (the target “paper” and three unrelated distractors). Results revealed that children with DLD were able to use verb-specific semantic information to predict upcoming referents, although their anticipatory looks to the target were fewer in number compared to age-matched controls and adults. No differences were observed between children with DLD and MLU-matched controls. In two follow-up experiments, it was demonstrated that, for all three child groups, predictive eye movements were also modulated by the semantic fit or typicality of the object. Only a slight difference was observed between children with DLD and age-matched controls with respect to overall anticipatory looks at the target. Secondly, Andreu et al. (2016) investigated whether 5- to 8-year-old children with DLD use verbs to predict arguments (themes, goals, and instruments) and adjuncts (locatives). Similar control groups as in Andreu et al. (2013) were investigated as well. Sentences such as “The man carefully reads a storybook in bed” were presented in the presence of four objects of which one was the target, one was the competitor, and the other two distractors. Results showed that all groups anticipated upcoming arguments and adjuncts rapidly. However, the number of looks to the arguments were overall higher than to the adjuncts for adults and children with and without DLD. And finally, Devey and Leonard (2018) tested whether 5-year-old children with DLD and 3-year-old TD children were sensitive to number information on fronted auxiliaries during online comprehension of questions (e.g., “Are/Is the nice little dog/s running?”). In a looking-while-listening task, children with DLD showed no anticipatory looks to target pictures, while TD children did. These data revealed that children with DLD have a weaker command of tense-agreement forms and, in particular, have difficulty understanding the relation between subject–verb sequences and the preceding information in questions.

Taken together, there is evidence to suggest that both TD children and children with DLD, around preschool age, are able to use verb semantics during online prediction. However, at the same time, it seems that younger children with DLD may not be able to readily extract information from the verb to predict upcoming input. To understand if online prediction skills are impaired in children with DLD, more studies are needed with younger children. Furthermore, previous studies have incorporated measures of on-line word recognition or prediction, but not both at the same time, which made it impossible to establish links between the two measures. In addition, no evidence exists for whether there are relations between predictive behavior and standardized languages measures in children with (a suspicion of) DLD.

This Study

The purpose of this study is threefold. The first aim is to examine online word recognition and prediction in young children with (a suspicion of) DLD. More specifically, in this study, we tested 2- to 4-year-olds, which is a much younger age range than has been tested in previous work. They were recruited from special treatment groups for preschool children with (a suspicion of) DLD. These children were diagnosed with a suspicion of DLD, as it is difficult to distinguish between children with DLD and late talkers at such a young age and DLD is commonly identified only at 4–5 years of age (e.g., Moyle et al., 2011). Most experimental studies strive for testing homogenous groups. However, here, we aimed for testing a clinically relevant group. We included children whose language deficit was prominent in their expressive language only as well as children who had a combination of expressive and receptive language problems (see Method section for more detailed information). This combination is more representative of the actual heterogeneous population, in particular at this young age. Moreover, this may also aid in the second aim, which is to identify the potential bases for individual differences in language processing in children with (a suspicion of) DLD, and to study the link between online and off-line language measures (receptive vocabulary, language comprehension, expressive vocabulary, and expressive grammar). Looking into this may help to understand individual differences in language processing within this heterogeneous clinical group. In addition, any discrepancy between online and off-line measures may highlight the added value of tapping into the process of language recognition (with online measures) rather than only looking at the outcomes of the process (as measured with off-line languages tests). The final aim is to investigate possible relations (for the DLD group and TD group together) between the online measure of recognizing words and the online measure of predicting upcoming information. In particular, the question was whether processing efficiency in recognition and prediction, as measured in proportion of fixations and response latencies, can be linked to each other.

We therefore set up two visual world experiments. In the word recognition experiment, we presented 2- to 4-year-old Dutch children with (a suspicion of) DLD and an age-matched control group with utterances such as “Kijk, een hoed!” (Look, a hat!), while two objects were presented on the screen (the target and a distractor). In the word prediction experiment, sentences such as “Hè, hij leest gewoon een boek” (literally translated: Hey, he reads just a book) were presented to the same two groups. These sentences contained a semantically constraining verb that predicted which of two objects displayed would be named.

Given most of the previous findings, we expected that children with (a suspicion of) DLD at age 2- to 4-year-old may not have difficulties with recognizing words online, as the task is rather simple (i.e., matching a word in a carrier sentence to a picture), and the words are highly familiar. Note, however, that it may be possible that the children with (a suspicion of) DLD show a more distributed late activation pattern compared to that of TD children (McMurray et al., 2010). Furthermore, we expected children with (a suspicion of) DLD to have more difficulties with predicting words on the basis of verb semantics compared to TD children, as
this task requires more advanced linguistic knowledge, such as how to integrate words in a sentence, and more cognitive processing resources. In addition, we expected that any differences we found between the DLD group and the TD group would also be reflected within the DLD group in effects of the off-line language measures. Based on the (to our knowledge) only study that found significant relationships between online and off-line measures in children with DLD (Pijnacker et al., 2017), we decided to include both receptive and expressive language measures. When considering the (mixed) findings in children with a typical development, it was expected that recognition ability and/or speed may be positively associated with vocabulary (e.g., Fernald et al., 2001). Similarly, as previous results on the link between prediction ability and off-line language measures have also demonstrated mixed findings, it was hypothesized that prediction ability and/or speed may be positively associated with expressive vocabulary (Mani & Huettig, 2012) and/or with receptive vocabulary (Borovsky et al., 2012). Concerning the link between processing efficiency in recognition and prediction, we hypothesized that the response times of both tasks would positively correlate with each other. The faster children will recognize words, the faster they will be in predicting upcoming words.

Method

Participants

Two groups of monolingual Dutch children took part in this study: a group of 68 toddlers with a suspicion of DLD and a group of 36 toddlers with TD. For simplicity reasons, we refer to the former as the DLD group and the latter as the TD group. For 11 of the participating children, we failed to obtain eye-tracking data: Two children (in the DLD group) were excluded because they refused to cooperate, two (one in the DLD group and one in the TD group) due to a technical problem, and seven (five in the DLD group and two in the TD group) due to an unsuccessful calibration. Of all the remaining participants, the data of another 11 children (nine in the DLD group and two in the TD group) were excluded from statistical analysis because they did end up with an insufficient number of trials per experiment (see Pretesting Results and Exclusion of Trials section for the different steps for exclusion). The final data set that we used for the statistical analyses consisted of a DLD group of 51 children and a TD group of 31 children.

The average age of the DLD group (41 boys, 10 girls) was 41 months (range: 31–48 months, SD = 4.6), and the average age of the TD group (25 boys, six girls) was 40 months (range: 34–54 months, SD = 5.1). There were no significant differences in the number of boys, \( \chi^2 (1, N = 82) < 0.01, p = .98, \) and average age, \( t (80) = 0.35, p = .73, \) between the groups. All children had normal vision, no persistent hearing loss, no specific neurological or syndromic disorder, and there was no indication of a cognitive disability. All children were full-term infants, except for two children in the DLD group that were born after 35 weeks of pregnancy. The educational level of the mothers\(^1\) in the TD group was significantly higher than of the mothers in the DLD group, DLD: 33% low, 28% middle, 39% high; TD: 3% low, 10% middle, 87% high; \( \chi^2 (2, N = 82) = 18.61, p < .001, \) Cramér’s \( V = 0.48. \)

Children with (a suspicion of) DLD were recruited from special treatment groups for toddlers with DLD within the Royal Dutch Kentalis. They were all diagnosed by a speech-language pathologist as having severe language difficulties in the context of nonverbal intelligence within the normal range. In order to rule out the possibility that the language problems were caused by any hearing loss, hearing was assessed by an audiologist as part of the diagnostic procedure. Children completed a battery of standardized language tests: Nonverbal cognitive skills were estimated by means of a Dutch nonverbal intelligence test (Snijders-Oomen Niet-Verbale Intelligietest [SON-R 2.5–7]; Tellegen et al., 1998), receptive language skills were evaluated with the Peabody Picture Vocabulary Test-Dutch Version (Schlichting, 2005), and the Schlichting Test for Language Comprehension (Schlichting & Lutje Spellberg, 2010a) and expressive language skills (expressive vocabulary and grammar) were measured by the Schlichting Test for Language Production (Schlichting & Lutje Spellberg, 2010b). All children with (a suspicion of) DLD in our study had at least one expressive score of 1 SD or more below the mean. They had a combination of expressive and receptive language problems or expressive language problems only. The mean standardized (quotient) scores per test are given in Table 1. TD children were recruited from the subject pool of the Baby & Child Research Center of Radboud University in Nijmegen. There was no indication of any developmental disorder in these children. Parents signed an informed consent prior to participation, and children received a little present for their participation.

Stimuli

The study comprised two different experiments: the word recognition experiment and the word prediction experiment. For the word recognition experiment, 20 (unique) simple Dutch phrases were constructed ending with a target noun (e.g., “Kijk, een hoed”; Look, a hat!), recorded by a female native speaker of Dutch. Each target noun was only used once in the experiment. Three different interjection words were used to at the start of the sentences (kijk, hé, oh; look, hey, oh). The duration of the initial part of the phrase up to the start of the target noun was always 2,000 ms, the duration of the target noun varied and was on average 644 ms (SD = 106 ms). Measurements were made using the Praat software (Boersma & Weenink, 2006).

For the word prediction experiment, 16 (unique) Dutch sentences were devised containing a verb that predicted

---

\(^1\)The Dutch school system is different from that in the United States with differentiation in education levels after primary school. Low educational level is comparable to less than high school or the lowest level of high school. Middle educational level is comparable to high school (higher than lowest level) or community college, and high educational level is comparable to college or university.
which of the two objects displayed on the screen would be named. Each predicting verb and target noun was only used once in the experiment. Two different sentence constructions were used in which the target noun that was predicted by the verb was either the subject or the object of the sentence (e.g., prediction of the subject “Kijk, daar loopt een dikke koe”; literally: “Hey, he reads just a book”). These sentences were recorded in the same recording session by the same female native speaker of Dutch. To make sure that there would be enough time after the predictive verb to measure anticipatory looks to the target picture before the actual acoustic start of the target word, there was always an adjective or adverb between the verb and the noun. The words between the verb and the noun (adjective or adverb and indefinite article) were neutral, that is, they did not favor the target or the distractor picture. The duration of the initial part of the sentence up to the start of the predictive verb was always 2,000 ms; the duration of the rest of the sentence including the verb and the target noun varied (mean durations: onset verb until onset noun 1,290 ms, \( SD = 95 \) ms; onset verb until end of the sentence 1,873 ms, \( SD = 126 \) ms; offset verb until the end of the sentence 1,346 ms, \( SD = 119 \) ms; verb 527 ms, \( SD = 57 \) ms; noun 583 ms, \( SD = 100 \) ms).

In addition to the target words in the word experiment and prediction experiment, respectively 16 and 20 simple (to be expected familiar) nouns were selected as distractors that matched the target nouns in the number of syllables. There was no initial phonemic overlap between the target and distractor words. The visual stimuli were digitized photographs of objects on a gray background. All target and distractor nouns and verbs that we used in this experiment were words that were expected to be familiar to children in this age range, based on a vocabulary index with commonly known words for preschool children. For the children with (a suspicion of) DLD, given their language problems, it could occur that some of the words we used were unfamiliar to them. This was assessed by means of a pretest (see Procedure section for more details).

**Table 1. Language and nonverbal intelligence measures for the children with (a suspicion of) DLD.**

| Standardized measures (quotient scores) | Children in the DLD group |
|----------------------------------------|---------------------------|
|                                        | N  | M (SD) | Range       |
| Receptive Vocabularya                   | 50 | 94 (16) | (60–133)    |
| Language Comprehensionb                 | 51 | 84 (14) | (55–112)    |
| Expressive Vocabularyc                  | 50 | 69 (15) | (55–110)    |
| Grammard                               | 47 | 73 (8)  | (55–92)     |
| Nonverbal IQd                          | 50 | 101 (13) | (73–137)    |

*Note. DLD = developmental language disorder. aPeabody Picture Vocabulary Test. bSchlichting Test for Language Comprehension. cSchlichting Test for Language Production. dSnijders-Oomen Niet-Verbale Intelligientetest 2.5–7.*
to draw the child’s attention to the screen. The two pictures (target and distractor) appeared simultaneously on the screen, centered in the left and right half-split of the screen. Each picture was only displayed once during the experiment, and the target appeared randomly on the left or right side of the screen, but never more than 3 times on the same side in succession. After 700 ms, the auditory stimuli began, while the pictures remained on the screen. The critical word (the noun in the word experiment and the verb in the prediction experiment) always started 1,300 ms after the onset of the sentence (thus 2,000 ms after the onset of the pictures). After the offset of the sentence (which differed between sentences as the duration of the sentences varied) the pictures remained on the screen for another 1,000 ms. Between trials, there was a blank screen for 1,000 ms. The eye-tracking task lasted about 8 min. At the end of the full test session, children received a small compensation for their participation. The trial structure of both experiments is depicted in Figure 1.

**Pretesting Results and Exclusion of Trials**

Of the nouns used in the word recognition experiment (tested in the pretest with the picture book), the children with DLD knew, on average, 96.5% and the TD children 99.6%. Of the verbs and target nouns used in the word prediction experiment (as indicated by parents), the children with DLD knew, on average, 89.2% of the verbs and 91.0% of the nouns, in comparison to, respectively, 92.2% and 94.1% for the TD children. Consequently, for each participant, a trial was deleted from the final eye-tracking data set if any of the relevant words (nouns and/or verbs) was unfamiliar to the child, according to the outcome of the pretest. Subsequently, for each participant, we calculated the number of remaining trials per experiment (as a result of missing data in the eye-tracking and deletion of trials due to unfamiliar words). A criterion of 50% remaining trials was set (i.e., at least 10 trials in the word recognition experiment and eight trials in the word prediction experiment) for inclusion of the data of that particular participant in the final data set. On average, the number of remaining trials in the word experiment was 17.4 ($SD = 2.5$) for the DLD group and 19.2 ($SD = 1.2$) for the TD group; in the word prediction experiment, the mean number of remaining trials was 12.5 ($SD = 2.4$) and 14.8 ($SD = 1.3$), respectively.

**Data Coding Procedure**

The eye-tracking data for both experiments were automatically coded as fixations, saccades, or “unclassified” using the fixation filter IVT-filter of Tobii Studio (Olsen, 2012). The minimal fixation duration was set to 60 ms. The timing of the fixations was established relative to the acoustic onset of the critical word (i.e., the target noun in the word recognition experiment and the verb in the word prediction experiment). Gaze position was categorized by the left and right half of the screen in which the picture appeared. Fixations were coded as directed to the target or the distractor.

**Dependent Variables and Analysis**

In both experiments, we were interested in two dependent variables: fixations to the target compared to the distractor (before and after the onset of the critical word) and the reaction times (RTs) of the shifts from distractor to target (after the onset of the critical word). For the statistical analyses, we selected a time window before the onset of the critical word and a time window after the onset of the critical word. As it takes, on average, 200 ms to plan and launch an eye movement (Matin et al., 1993), these windows were aligned to the onset of the critical word +200 ms. The duration of the prewindow was 1,300 ms (referred to as the pre-noun window in the trials of the word recognition experiment and as the pre–verb window in the word prediction experiment). In both cases, the onset of these prewindows was aligned to the onset of the sentences. The duration of the time windows after the onset of the critical words differed for the two experiments: In the word recognition experiment, the noun window was 1,300 ms, and in the word prediction experiment, the verb window was 1,100 ms. The duration of the noun window was selected on the basis of visual inspection of the eye-tracking data (combined for both groups of children): After 1,300 ms, there was a sudden decrease in the proportion of fixations on the screen (i.e., the children started to lose interest in the pictures). The duration of the verb window in the word prediction experiment was determined by the earliest acoustic onset of the target nouns. Because in this experiment, we were interested in anticipatory looks based on the verb, we wanted to make sure that, during the verb window, participants could not have heard the onset of the target nouns.

For both experiments, the statistical analyses consisted of four different steps, followed by one last analysis examining the relationship between the outcomes of both experiments. For the word recognition experiment, we first studied the main effect: the recognition effect. We examined whether children looked more at the target picture upon hearing the target word than before hearing the target, and we looked at whether there was a difference between groups in the size of the word recognition effect. To test this, we conducted a generalized mixed-effects logistic regression analysis (Jaeger, 2008) in R (Version 3.4.2; R Core Team, 2013) using the lme4 package (Bates et al., 2015). In each model, group (TD vs. DLD) and window (pre-noun window vs. noun window) were entered as categorical fixed effects, which were coded as numeric contrasts (Group: TD children as −0.5 and DLD children as +0.5; Window: pre-noun window as −0.5 and noun window as +0.5; see Barr, 2008). Main effects of Group, Window, and its interaction were directly entered into the model. The package emmeans (Length, 2019; Searle et al., 1980) was used to interpret significant interactions in these models. The maximal random effects structure that converged in the model was implemented, which included random intercepts for participants and items, as well as random slopes for Group × Participants and Group × Items.
Secondly, we focused on the time course of the recognition process, that is, how rapidly the children recognized the target word, and tested whether the time course differed between the two groups. We did this by examining how fixation behavior unfolded over time in the noun window. For this analysis, time (in 10-ms bins) was entered in the model as a continuous fixed effect. This continuous factor was centered (Baguley, 2012, pp. 590–621) and divided by 1,000 to prevent model convergence issues. We examined main effects of group and time and its interaction. The maximal random effects structure included random intercepts for participants and items, as well as random slopes for Group × Items.

Thirdly, to examine whether there was a difference in the speed of word recognition between the DLD group and TD group, an independent-samples t test was performed on the RTs of the shifts from the distractor to the target (see also Fernald et al., 2006). For this purpose, only those trials were selected for which the participant happened to look at the distractor just before the onset of the critical noun. Trials were labeled as target-initial or distractor-initial by determining the location of the last fixation in the 200-ms time window after the critical word, and only the distractor-initial trials were included in the analysis. This resulted in an average of 7.6 (SD = 2.2) trials in the DLD group and 8.3 (SD = 2.6) trials in the TD group.
In addition, we examined the predictive relationship between target noun recognition (1 = target, 0 = distractor) and the off-line language abilities of the children with (a suspicion of) DLD across the two recognition time windows by performing a generalized mixed-effects logistic regression analysis (Jaeger, 2008) in R (Version 3.4.2; R Core Team, 2013) using the lme4 package (Bates et al., 2015). In each model, the categorical fixed effect Window was entered as numeric contrast (pre-noun window was coded as −0.5 and noun window as +0.5; see Barr, 2008) in interaction with one of the four off-line language measures (Receptive Vocabulary, Language Comprehension, Expressive Vocabulary, and Grammar). These continuous measures were centered (Baguley, 2012, pp. 590–621) and divided by 10 to prevent model convergence issues. We examined the main effects of window and each off-line language measure and their interaction. The maximal random effects structure that converged in the model was implemented, which included random intercepts for participants and items.

For the word prediction experiment, the same steps were followed, but now the critical word was the verb (rather than the noun) to examine the verb-based prediction of the upcoming target noun. Therefore, for the analysis of the prediction effect, the two levels of the factor window referred to the pre–verb window and the verb window. For the analysis of the time course with time as continuous fixed effect, we focused on the fixations in the verb window only. For both models (Group × Window and Group × Time), the maximal random effects structure included random intercepts for participants and items, as well as random slopes for Group × Items. The analysis of the RTs of the shifts from distractor to target was based on 6.4 distractor-initial trials ($SD = 2.2$) in the DLD group and 7.1 ($SD = 2.1$) trials in the TD group. The regression analysis within the DLD group (to examine the predictive link between verb-based target noun prediction and the off-line language measures) was performed across the pre–verb window (coded as −0.5) and verb window (coded as +0.5).

For both the word recognition as well as the prediction effect, we ran a control analysis to investigate whether the fixed effect educational level (high as the reference group) interacted with group in each model. Note that it was not possible to look at the three-way interaction with window/time due to model convergence issues. None of these control analyses showed a significant interaction between group and educational level (all $p \text{'s} > .1$), except for the analysis on the recognition data across both time windows. Contrary to expectations, the results for the DLD group demonstrated significantly higher target fixations for the children of a mother with a low educational level than for the children with a mother with a high educational level (estimate = .24, $SE = .08$, $z$ value = 2.92, $p = .003$).

Finally, to examine the relationship between the word recognition process and the verb-based prediction of both groups of children together, we performed Pearson correlations between the four dependent variables reflecting the size of the effect (word recognition effect and the verb-based prediction effect), and the RTs of the shifts from distractor to target (in both experiments). The two effects were defined as the difference between the proportion of fixations to the target in the noun/verb window and the proportion of fixations to the target in the pre–noun/pre–verb window. We corrected for the number of comparisons by applying a Bonferroni method, which resulted in a rejection probability of .0125 (four comparisons).

**Results**

**Word Recognition Experiment**

Figure 2 plots the proportion of fixations to the target picture and to the distractor picture during the pre–noun and noun windows, separately for the DLD and the TD group. As can be seen in Figure 2, shortly after the onset of the noun, the proportion of fixations to the target picture increases steadily for both the children with (a suspicion of) DLD and the TD children and reaches a similar maximum around 1,000 ms after the onset of the target word, demonstrating that the children of both groups recognized the words that were named. In line with this, the outcome of the mixed-effects logistic regression analysis revealed a significant effect of window (estimate = .66, $SE = .01$, $z$ value = 55.22, $p < .001$) and no significant effect of group (estimate = −.07, $SE = .08$, $z$ value = −.96, $p = .34$). However, we did find a significant interaction of Window × Group, (estimate = .08, $SE = .02$, $z$ value = 5.45, $p < .001$). Unpacking this interaction and looking at the proportion of fixations in Figure 2 revealed two differences. First, the children in the DLD group started off with fewer target fixations on the target than the TD children in the pre–noun window. Both groups initially preferred the pictures of the distractors over the pictures of the targets, but this preference was stronger for the children in the DLD group. Second, in the noun window, the children in the DLD group showed a slightly higher proportion of target fixations than the children in the TD group at the end of the trial (starting around 1,000 ms after the onset of the target word). It seemed that, when the TD children were already losing interest after having recognized the target word, the children with (a suspicion of) DLD were still looking at the target picture.

When we zoomed in to the noun window only to examine the time course of the recognition effect, we again found no significant effect of Group (estimate = −.14, $SE = .12$, $z$ value = −1.16, $p = .25$), but we did find a significant effect of time (estimate = .74, $SE = .02$, $z$ value = 32.31, $p < .001$) and a significant interaction between group and time (estimate = .28, $SE = .03$, $z$ value = 9.15, $p < .001$). The main effect of time is not a surprise as it indicates that the fixations to the target change significantly over time while the target word unfolds. The significant interaction indicates that the time course of target recognition differs between the two groups. This, again in combination with visual inspection of Figure 2, confirms our previous observation that the children in the DLD group remain looking at the target picture for a longer time than the TD children.
The mean RTs of the shifts from the distractor picture to the target picture (measured from the acoustic onset of the target noun) were 526 ms for the DLD group ($SD = 106$) and 477 ms for the TD group ($SD = 113$). An independent-samples $t$ test showed that this was a medium-sized RT effect, $t(80) = 2.01$, $p = .048$, Cohen’s $d = 0.45$, indicating that children with (a suspicion of) DLD are slower than TD children in shifting from the distractor to the target picture upon hearing the target word.

Next, we examined the predictive relation between target noun recognition and the off-line language performance of the children with (a suspicion of) DLD across the two recognition windows. As expected, the effect of window was significant in all models ($all\ p < .001$). More importantly, a main effect was found for expressive vocabulary (estimate = $-0.06$, $SE = .03$, $z$ value = $-2.22$, $p = .03$), which also significantly interacted with window (estimate = $-0.03$, $SE = .006$, $z$ value = $-5.18$, $p < .001$). The interaction indicates that children with (a suspicion of) DLD with a larger expressive vocabulary looked primarily less at the target in the noun window than the children with smaller vocabulary. In addition, a significant interaction between grammar and window was found (estimate = $0.04$, $SE = .01$, $z$ value = $2.99$, $p = .002$), demonstrating that children with (a suspicion of) DLD with higher grammar skills looked primarily less at the target in the pre-noun window than the children with lower grammar abilities. In the noun window, the grammar skills had no influence on the gazing pattern. Furthermore, the results showed no significant main effects for receptive vocabulary (estimate = $-0.01$, $SE = .02$, $z$ value = $-0.45$, $p = .66$) or language comprehension (estimate = $-0.01$, $SE = .03$, $z$ value = $-0.37$, $p = .71$), and these effects did also not interact with window (receptive vocabulary: estimate = $0.02$, $SE = .006$, $z$ value = $0.39$, $p = .69$; language comprehension: estimate = $-0.01$ $SE = .007$, $z$ value = $-1.45$, $p = .15$) on target noun recognition in the DLD group.

**Word Prediction Experiment**

The proportion of fixations to the target picture and to the distractor picture during the pre–verb and verb windows for the items in the prediction experiment are plotted in Figure 3, again separately for the two experimental groups. As in the previous experiment, there is a difference between fixations to the target and distractor picture in the window before the onset of the word that favors the target...
picture (in this experiment the pre-verb window). Again, the initial picture preference is stronger in the DLD group than in the TD group (but now both groups of children prefer the target pictures over the distractor pictures, possibly because the target pictures were always animate). Interestingly, in contrast to the previous experiment, the pattern clearly looks different for the children in the DLD group than for the children in the TD group; the children with (a suspicion of) DLD show fewer anticipatory looks to the target noun than the TD children. In line with this, the mixed-effects logistic regression analysis showed a significant effect of window (estimate = .74, SE = .01, z-value = 52.75, p < .001), and crucially, a significant interaction of Window by Group (estimate = −.42, SE = .02, z-value = −22.45, p < .001), indicating that the children in the DLD group showed a smaller increase in looks to the target picture across windows (from the pre-verb window to the verb window) than the TD children. There was no main effect of group (estimate = .11, SE = .10, z-value = 1.10, p = .27).

Zooming in to the verb window to examine the time course of the verb-prediction effect, we found no significant effect of group (estimate = −.01, SE = .18, z value = −.05, p = .96), but we did find a significant effect of time (estimate = 2.03, SE = .04, z value = 53.87, p < .001) and a significant interaction between group and time (estimate = −.61, SE = .05, z value = −12.64, p < .001). This interaction combined with the clearly visible difference in slope of the curve in Figure 3 demonstrates that the children with (a suspicion of) DLD are slower in anticipating the target upon hearing the predicting verb than the TD children and do not reach the same proportion of fixations to targets at the end of the verb window.

Given our interest in the individual differences within the DLD group in comparison to the group of TD children, we visualized the distribution of the prediction effect
separately for the two groups of children in a box plot with a bee swarm. Each data point shows the mean prediction effect of each individual child, defined as the difference between the proportion of anticipatory fixations to the target after hearing the predicting verb (verb window) and the proportion of fixations to the target before the verb (pre-verb window). A positive value indicates that the proportion of fixations to the target upon hearing the verb increased. Looking at Figure 4, the first distinctive observation is that the variability of the prediction effect is greater in the DLD group than in the TD group. Furthermore, we see that most children in the TD group have positive prediction values. In contrast, a small but considerable number of children in the DLD group show a negative value, indicating that they did not make use of the verb to predict the upcoming target, at least not consistently. Remarkably, a considerable number of children with (a suspicion of) DLD show prediction behavior that closely resembles that of the TD children.

As for the analysis of the RTs of the shifts from distractor picture to target picture upon hearing the onset of the predictive verb, a significant difference was found (independent t test: t(80) = 2.85, p = .006, Cohen’s d = 0.68): The children in the DLD-group appeared to be 89 ms slower than the TD children in shifting from distractor to target on average 518 ms (SD = 157) for the DLD group and 429 ms (SD = 98) for the TD group.

The regression analyses we performed to examine the predictive relationship between target noun prediction and the offline language performance of the children with (a suspicion of) DLD measured with standardized language tests across the two prediction windows showed several significant effects. As expected, the effect of window was significant in all models (all ps < .001). More importantly, significant interaction effects were found between window and each of the four offline language measures (Receptive Vocabulary: estimate = .10, SE = .007, z value = -13.33, p < .001; Language Comprehension: estimate = .13, SE = .009, z-value = 13.91, p < .001; Expressive Vocabulary: estimate = .16, SE = .008, z-value = 19.48, p < .001; Grammar: estimate = .19, SE = .02, z value = 11.66, p < .001). These interactions demonstrate that children with (a suspicion of) DLD with higher off-line language abilities look more at the target (i.e., are better in prediction) than those with lower abilities in the verb window.

Correlations Between Word Recognition and Prediction

Finally, let us consider the outcomes of the Pearson correlation analysis between the four dependent variables of both experiments (excluding the time course). We found a positive correlation between the RTs of the two experiments, r(82) = 0.30, p = .01; that is, faster shifts from distractor to target in the word recognition experiment are related to faster anticipatory shifts from distractor to target in the verb-based prediction experiment. More interestingly, both RTs are negatively related to the prediction effect. That is, not only do we see that slower anticipatory shifts from distractor to target are associated with a smaller verb-based prediction effect (both measures within the same experiment, r(82) = -0.55, p < .001), we also see that slower shifts to the target in the word recognition experiment are related to a smaller verb-based prediction effect, r(82) = -0.27, p = .01.

Discussion

The purpose of the current study was threefold. First, the process of online word recognition and verb-based prediction in young children with (a suspicion of) DLD was examined in comparison to a group of TD peers. Secondly, the relation between the measures of these two experiments was investigated. And finally, the relation between the two main online measures and the outcomes of offline standardized language measures within the DLD group was examined. This study contributes to the relatively small body of online language processing research with children with DLD and was the first to conduct such research with a clinically relevant group of children with (a suspicion of) DLD at a very young age. We discuss our specific main findings below.

First, as expected, the results for the online word recognition experiment showed that the group of children with (a suspicion of) DLD have little difficulty in recognizing familiar words in a short carrier phrase such as “Kijk, een hoed!” (Look, a hat!). Up until 1,000 ms after the onset of the target word, their eye fixation trajectory does not differ from that of their TD peers. This finding is in line with previous online language processing research with children with DLD (e.g., Andreu, Sanz-Torrent, & Guardia-Olmos, 2012;
Haebig et al., 2018). It suggests that the lexical representations of nouns are sufficiently specified for children with (a suspicion of) DLD. Furthermore, the access to nouns seems to be intact for this group, at least in studies in which a visual aid, in the form of pictures on a display, was present. It needs to be further investigated whether a similar pattern emerges if this group cannot make use of any pre-activation of the concepts involved.

Moreover, the results confirm our hypothesis that children with (a suspicion of) DLD might experience problems when the task is more complex (e.g., unfamiliar words, complex sentences, absence of pictures) and thus requires a higher level of semantic integration processing. For example, it is possible that Pijnacker et al. (2017) demonstrated online word recognition differences between children with DLD and TD children because they presented more complicated sentences without any visual cues, such as pictures. To verify this, follow-up research should investigate whether recognizing words in more complex sentences elicit differences between children with DLD and TD children. Similarly, group differences may show up when less familiar words are used.

Next to the overall similarity in online word recognition, we detected three minor differences in eye gaze patterns between the groups. The first significant difference was that, on average, the speed of the shifts from the distractor picture to the target picture was slightly slower for the DLD group in comparison to the TD group. A possible explanation is that children with DLD have a smaller vocabulary size (and possibly lower intrinsic activation of the words) than TD peers causing a delay in word recognition (e.g., Leonard & Deevy, 2006; Montgomery, 2006). Secondly, we found that the beginning of the online word recognition process was identical for both groups but that the eye gaze patterns diverged around 1,000 ms after the onset of the target word (see Figure 2). More specifically, after having recognized the target word, children in the DLD group remained looking at the target picture a bit longer, while the TD children were already losing interest. A similar pattern has been found in other eye-tracking studies: When gaze patterns of younger and older children are compared, the window of analysis is often shorter for older children or adults, as their interest wanes after the goal (finding the picture) has been reached (Cree, 2012; Paquette-Smith et al., 2016; Weighall et al., 2016). Third, the children in the DLD group looked more to the more attractive picture in the noun window than the TD children, suggesting that their linguistic processing may be less mature than their TD peers. This is in line with research on word learning by Hollich et al. (2000), who studied the role of perceptual salience in word learning in 12-, 19-, and 24-month-olds. When presented one interesting and one boring novel object, the older children changed their attention from the interesting to the boring one when the experimenter pointed to the latter, but the younger infants did not learn the boring object under the same conditions.

While we found that children with (a suspicion of) DLD were rather similar to their TD peers in a simple online word recognition task, and only showed minor timing differences, they clearly differed in the process of predicting upcoming nouns on the basis of a verb. More specifically, prediction was more difficult for children in the DLD group than for their TD peers. Children with (a suspicion of) DLD showed less anticipatory looks to the picture of the target word upon hearing the verb in a sentence such as “Hè, hij leest gewoon een boek” (Hey, he reads just a book); from the onset of the predicting verb, the curve was less steep, and the proportion of fixations to the target never reached the proportion of fixations of the TD children.

This second main finding was in line with our expectations, as previous work has shown that children with DLD have specific problems with verbs, argument structure, and functional morphology (e.g., Andreu, Sanz-Torrent, Legaz, & MacWhinney, 2012; Bishop, 1997; Deevy & Leonard, 2018; Verhoeven & Van Balkom, 2004). Note that, at the same time, children with DLD have also demonstrated to be fully capable to use verb semantics during online prediction (e.g., Andreu et al., 2016, 2013), but these children were older than our group.

In this study, the children with (a suspicion of) DLD, as a group, predicted less and were also slower compared to their TD peers. It is possible that this is due to less robust representations of the verbs (e.g., Capone & McGregor, 2005). In other words, it might be that the lexical representations of the verbs are not (fully) specified yet. Children with DLD might, for example, know the general meaning of the verb “reading,” but they have not stored (or cannot retrieve) the information that the verb requires an object, while the lexical representations of the nouns are sufficiently specified to match the words to the pictures (i.e., in the word recognition experiment). Previous research has shown that verbs are overall acquired later than nouns (Gleitman et al., 2005). Possible reasons for the difficulty of acquiring verbs are that they are less concrete and appear in many different forms due to agreement rules.

At the same time, the smaller and slower prediction effect for the children in the DLD group could also be due to limitations in processing (e.g., Kail, 1994; Miller et al., 2001). It may be difficult for children with DLD to recognize verbs online because of a lack of processing capacity. In this task, children have to look ahead in time and they seem not to be able to do this as quickly as their TD peers. Another reason for the differences between the two groups on the prediction task might be that the task asks for advanced linguistic knowledge, which might be impaired in children with (a suspicion of) DLD. As put forward by Pijnacker et al. (2017), DLD children may experience problems with semantic integration of multiple words in a sentence, as is the case in the current prediction experiment. Relatedly, it is also possible that working memory deficits influence prediction skills in children with DLD (e.g., Lum et al., 2011).

The current set of findings is not able to disentangle these different explanations, but this was also not the aim of the current study. We were mainly interested in the variation in performance within the DLD group. We found that the size of the prediction effect varied considerably within
the group of children with (a suspicion of) DLD (see Figure 4), indicating that some of these children showed similar prediction processes as their TD peers, while others showed a more deviant pattern. This suggests that not all children with (a suspicion of) DLD have difficulty with online receptive language processing. This is an important result as an understanding of the range of individual differences could have implication for the diagnosis and treatment of children with (a suspicion of) DLD. Furthermore, identifying which components of online language processing vary between individuals could aid us in tackling new theoretical questions. For example, why are some children with (a suspicion of) DLD good predictors and others not? Part of the answer to this question may be found in our third main finding.

The third main finding comes from our investigation of predictive relationship in which we examined to what extent online word recognition and prediction performance in children with (a suspicion of) DLD are associated with their off-line language abilities. For online word recognition, we found that both expressive language measures (tapping into expressive vocabulary and grammar) had an effect on target recognition in the noun window. Within the DLD group, the children with lower expressive language scores looked longer at the target than children with higher expressive scores. This is reversed to what we initially hypothesized based on correlations between online word recognition and off-line language measures in young children with a typical language development (e.g., Fernald et al., 2001). Yet, these findings are in accordance to the difference in eye gaze patterns we found between the two groups of children, which showed that children in the DLD group remained looking at the target picture a bit longer than the TD children did. In line with this, within the DLD group, children with lower expressive language skills looked longer at the target picture in comparison to children with higher language skills. For online word prediction, we found that all four off-line language measures (reflecting expressive and receptive language skills at both the word level and sentence level) had a positive effect on word prediction. Again, these findings within the DLD groups are in accordance with the differences we found between the DLD group and the TD group. Children within the DLD group with lower language scores have more difficulty with verb-based word prediction than children within the DLD group with higher language scores.

The final main finding is that the analysis of the relation between the online measures of the prediction experiment and the word recognition experiment showed that the size of the verb-based prediction experiment is related to the speed of the word recognition effect. TD children and children with (a suspicion of) DLD who are slower in the recognition of familiar words in a simple sentence seem to demonstrate a smaller verb-based prediction effect. This provides the first evidence that the skills necessary for these online language processing tasks are related to each other.

The clinical implications of this study mainly come from the combination of off-line and online measures. It is important to recognize that off-line measures, which are commonly used in the clinical practice to assess the language development of children, do not give the full picture. While off-line measures provide information on the outcome of a language process, online measures can capture any difficulties in the underlying process. This study showed, in particular, that although children with (a suspicion of) DLD have little difficulty recognizing familiar words, the timing of the recognition process did differ. In comparison to TD peers, they shift slower from distractor picture to target picture upon hearing the onset of the target word and they remain looking longer at the target picture after having recognized the word. This (combined with the correlation between the RTs on the word recognition task and the RTs on the word prediction task) suggests that processing issues are at stake for children with (a suspicion of) DLD. When recognizing words in short sentences, this may not lead to any problems, but as soon as sentences start to become more complex, any slight decrease in the processing speed of the individual words may hamper language comprehension. Noticeably, within the group of children with (a suspicion of) DLD, we found that the off-line expressive language measures, and not the receptive language measures, had an effect on the (late) looking behavior in this comprehension task. Furthermore, our finding that there are differences within the DLD group, in particular in the ability of verb-based prediction, triggers a question that many parents and clinicians of young children with a suspicion of DLD asks themselves: Can we predict which of these children will catch up as they grow older, and which children will not? Future studies should investigate whether online tasks are indeed able to predict later language acquisition skills.

There are a number of limitations of this study. First, the analyses on the RTs of the shifts are based on a smaller number of trials, because the data are restricted by only including distractor-initial looks. Secondly, inherent to the visual world paradigm, visual information was provided to the participants. This could have made the linguistic situation easier for the groups under investigation than when no visual cues were present. Thirdly, due to the young age of the children, it is not (yet) certain if a DLD is the cause of the language problems of the children in the DLD group or if the children are simply late talkers. At the same time, as mentioned before, this corresponds to the clinical reality that makes this study clinically only more relevant. Future research should reveal which children will continue having language problems and are still diagnosed with DLD at a later age, and which children will catch up and show a normal language development. A final limitation is that the experimental groups differed in one aspect: maternal education. Although this factor only had a significant effect on the recognition results across time windows when it was added as a covariate, the effect goes into the opposite direction as expected. Ideally, in future studies, maternal educational level should not differ across groups.

In conclusion, this study elaborated on previous studies by testing a clinically relevant sample of children with DLD at a young age on both online word recognition and prediction. Such data aid in a better understanding of early problems in language processing, which may help
to develop useful diagnostic tools and treatments for children with DLD. Our results revealed that our groups demonstrated different online language processing trajectories. Most importantly, we found that, although children with DLD had similar online word recognition skills, they seem to have different online prediction abilities. These contrasting results may be due to task demands as word prediction is cognitively more challenging. Moreover, these results are also related to both expressive and receptive off-line skills. In addition, children with DLD may fail to fully grasp verb argument structure, which affects processing times. The outcomes of this study prompt further research on the different aspects of language comprehension in young children with DLD and the variability we see within this group of unique children.

Acknowledgments
This research was funded by Royal Dutch Kentalis. We thank Judith Pijnacker, Syrthe Niessen, and Constance Visser for their help in designing and setting up the experiments. A warm thank you is extended to all of the children and their parents who took part in this study, and to the developmental language disorder treatment groups of Kentalis and the Baby & Child Research Center of Radboud University for assisting in the recruitment of participants and hosting of this research. Finally, we thank Mirjam Otters and Iris van de Laar for assisting in testing the children, Bob Rosbag for his help with the preprocessing and analyses of the eye-tracking data, and the clinical members of the sounding board for their helpful feedback during the entire project.

References
Altmann, G. T. M., & Mirkovic, J. (2009). Incrementality and prediction in human sentence processing. Cognitive Science, 33(4), 583–609. https://doi.org/10.1111/j.1551-6709.2009.01022.x
Andreu, L., Sanz-Torrent, M., & Corella-Olmos, J. (2012). Auditory word recognition of nouns and verbs in children with Specific Language Impairment (SLI). Journal of Communication Disorders, 45(1), 20–34. https://doi.org/10.1016/j.jcomdis.2011.09.003
Andreu, L., Sanz-Torrent, M., Legaz, L. B., & MacWhinney, B. (2012). Effect of verb argument structure on picture naming in children with and without specific language impairment (SLI). International Journal of Language & Communication Disorders, 47(6), 637–653. https://doi.org/10.1111/j.1460-6984.2012.01070.x
Andreu, L., Sanz-Torrent, M., & Rodriguez-Ferreiro, J. (2016). Do children with SLI use verbs to predict arguments and adjuncts: Evidence from eye movements during listening. Frontiers in Psychology, 6, 1–12. https://doi.org/10.3389/fpsyg.2015.01917
Andreu, L., Sanz-Torrent, M., & Trueswell, J. C. (2013). Anticipatory sentence processing in children with specific language impairment: Evidence from eye movements during listening. Applied Psycholinguistics, 34(1), 5–44. https://doi.org/10.1017/S0142716411000592
Baguley, T. S. (2012). Serious stats: A guide to advanced statistics for the behavioral sciences. Macmillan.
Barr, D. J. (2008). Analyzing ‘visual world’ eye-tracking data using multilevel logistic regression. Journal of Memory and Language, 59(4), 457–474. https://doi.org/10.1016/j.jml.2007.09.002
Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1). https://doi.org/10.18637/jss.v067.i01
Bishop, D. (1997). Uncommon understanding. Psychology Press.
Bishop, D. V. (2014). Ten questions about terminology for children with unexplained language problems. International Journal of Language & Communication Disorders, 49(4), 381–415. https://doi.org/10.1111/1460-6984.12101
Boersma, M., & Weenink, D. (2006). Praat: Doing phonetics by computer. University of Amsterdam. http://www.praat.org
Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one’s age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. Journal of Experimental Child Psychology, 112(4), 417–436. https://doi.org/10.1016/j.jecp.2012.01.005
Brouwer, S., Özkan, D., & Kuıntay, A. C. (2019). Verb-based prediction during language processing: The case of Dutch and Turkish. Journal of Child Language, 46(1), 80–97. https://doi.org/10.1017/S0305000918000375
Brouwer, S., Sprenger, S. A., & Unsworth, S. (2017). Processing grammatical gender in Dutch: Evidence from eye movements. Journal of Experimental Child Psychology, 159, 50–65. https://doi.org/10.1016/j.jecp.2017.01.007
Capone, N., & McGregor, K. K. (2005). The effect of semantic representation on toddlers’ word retrieval. Journal of Speech, Language, and Hearing Research, 48(6), 1468–1480. https://doi.org/10.1044/1092-4388(2005/0102)
Creel, S. C. (2012). Preschoolers’ use of talker information in online comprehension. Child Development, 83(6), 2042–2056. https://doi.org/10.1111/j.1467-8624.2012.01816.x
Deevey, P., & Leonard, L. B. (2018). Sensitivity to morphosyntactic information in preschool children with and without developmental language disorder: A follow-up study. Journal of Speech, Language, and Hearing Research, 61(12), 3064–3074. https://doi.org/10.1044/2018_JSLHR-L-18-0038
Deevey, P., Leonard, L. B., & Marchman, V. A. (2017). Sensitivity to morphosyntactic information in 3-year-old children with typical language development: A feasibility study. Journal of Speech, Language, and Hearing Research, 60(3), 668–674. https://doi.org/10.1044/2016_JSLHR-L-15-0153
Dollaghan, C. (1998). Spoken word recognition in children with and without specific language impairment. Applied Psycholinguistics, 19(2), 193–207. https://doi.org/10.1017/S014271640001031
Edwards, J., & Lahey, M. (1996). Auditory explicit decisions of children with specific language impairment. Journal of Speech and Hearing Research, 39(6), 1263–1273. https://doi.org/10.1044/jshr.3906.1263
Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. Developmental Psychology, 42(1), 98–116. https://doi.org/10.1037/0012-1649.42.1.98
Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. Psychological Science, 9(3), 228–231. https://doi.org/10.1111/1467-9280.00044
Fernald, A., Swingley, D., & Pinto, J. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. Child Development, 72(4), 1003–1015. https://doi.org/10.1111/1467-8624.00331
Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In J. A. Sekirina, E. M. Fernandez, & H. Clahsen (Eds.), Developmental psycholinguistics: On-line methods in children’s language processing. John Benjamins. https://doi.org/10.1075/lald.44.06fer

Alphen et al.: Language Processing in Preschoolers With DLD
Schlichting, L. (2005). Peabody Picture Vocabulary Test–Dutch Version (3rd ed.). Pearson.

Schlichting, L., & Lutje Spelberg, H. (2010a). Schlichting Test voor Taalbegrip [Schlichting Test for Language Comprehension]. Bohn Stafleu van Loghum.

Schlichting, L., & Lutje Spelberg, H. (2010b). Schlichting Test voor Taalproductie [Schlichting Test for Language Production]. Bohn Stafleu van Loghum.

Schul, R., Stiles, J., Wulfeck, B., & Townsend, J. (2004). How “generalized” is the “slowed processing” in SLI? The case of visuo-spatial attentional orienting. *Neuropsychologia, 42*(5), 661-671. https://doi.org/10.1016/j.neuropsychologia.2003.10.010

Searle, S. R., Speed, F. M., & Milliken, G. A. (1980). Population marginal means in the linear model: An alternative to least squares means. *The American Statistician, 34*(4), 216-221. https://doi.org/10.1080/00031305.1980.10483031

Sedivy, J. (2010). Using eyetracking in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 115–138). John Benjamins. https://doi.org/10.1075/fllt.27.08sed

Sheng, L., & McGregor, K. K. (2010). Lexical-semantic organization in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*(1), 146–159. https://doi.org/10.1044/1092-4388(2009/08-0160)

Swingley, D., & Aslin, R. (2000). Spoken word recognition and lexical representation in very young children. *Cognition, 76*(2), 147–166. https://doi.org/10.1016/S0010-0277(00)00081-0

Swingley, D., & Aslin, R. (2002). Lexical neighborhoods and the wordform representations of 14-month-olds. *Psychological Science, 13*(5), 480–484. https://doi.org/10.1111/1467-9280.00485

Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71*(2), 73–108. https://doi.org/10.1016/S0010-0277(99)00021-9

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. E. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632–1634. https://doi.org/10.1126.science.7777863

Tellegen, P. J., Winkel, M., Wijnberg-Williams, B. J., & Laros, J. A. (1998). Snijders-Oomen Niet-Verbale Intelligentietest, SON-R 2.5–7 [Snijders-Oomen Nonverbal Intelligence Test, SON-R 2.5–7]. Swets & Zeitlinger.

Thordardottir, E., & Ellis Weismer, S. (2002). Verb argument structure weakness in specific language impairment in relation to age and utterance length. *Clinical Linguistics and Phonetics, 16*(4), 233–250. https://doi.org/10.1080/02699200110116462

Tomblin, J. B., Zhang, X., Buckwalter, P., & O’Brien, M. (2003). The stability of primary language disorder. *Journal of Speech, Language, and Hearing Research, 46*(6), 1283–1296. https://doi.org/10.1044/1092-4388(2003/100)

Trueswell, J., & Gleitman, L. R. (2004). Children’s eye movements during listening: Evidence for a constraint-based theory of parsing and word learning. In J. M. Henderson & F. Ferreira (Eds.), *Interface of language, vision, and action: Eye movements and the visual world*. Psychology Press.

Trueswell, J. C. (2008). Using eye movements as a developmental measure within psycholinguistics. In I. A. Sekerina, E. M. Fernandez, & H. Chalshes (Eds.), *Language processing in children*. John Benjamins. https://doi.org/10.1075/hald.44.05tru

Verhoeven, A. L., & Van Balkom, J. (2004). Classification of developmental language disorders. Erlbaum.

van Weerdenburg, M., Verhoeven, L., & van Balkom, H. (2006). Towards a typology of specific language impairment. *The Journal of Child Psychology and Psychiatry, 47*(2), 176–189. https://doi.org/10.1111/j.1469-7610.2005.01454.x

Weighall, A. R., Henderson, L. M., Barr, D. J., Cairney, S. A., & Gaskell, M. G. (2016). Eye-tracking the time course of novel word learning and lexical competition in adults and children. *Brain and Language, 167*, 13–27. https://doi.org/10.1016/j.bandl.2016.07.010

Zangl, R., Klarman, L., Thal, D., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. *Journal of Cognition and Development, 6*(2), 179–208. https://doi.org/10.1207/s15327647jcd0602_2