Preserved category-based inferences for word learning in school-aged children with Developmental Language Disorder

Estelle Dauvister[a,b]*, B. Jemel[c,d] and C. Maillart[a]

[a]Department of Speech and Language Therapy, Research unit for a life-Course perspective on Health and Education – RUCHE, University of Liège, Liège, Belgium; [b]FRESH fund, F.R.S.-FNRS;[c]Laboratoire de recherche en neurosciences et électrophysiologie cognitive, Hôpital Rivière-des-Prairies, Montréal, Canada; [d]Ecole d'orthophonie et d'audiologie, Faculty of Medecine, Université de Montréal, Montréal, Canada

Correspondence : Rue de l'Aunaie 30, 4000 Liège, Belgium – estelle.dauvister@uliege.be

Preserved category-based inferences for word learning in school-aged children with Developmental Language Disorder

Word learning difficulties are often found in children with Developmental Language Disorder (DLD). Lexical patterns of difficulties appear to be well described in the context of DLD but very little research focuses on their underlying causes. Word learning is known to be an inference-based process, constrained by categorization, which helps the extension of new words to unfamiliar referents and situations. These processes appear integrated in Bayesian models of cognition, which supposes that learning relies on an inductive inference process that recruits prior knowledge and principles of statistical learning (detection of regularities). Taken together, these mechanisms remain underexplored in DLD. Our study aims to define whether children with DLD can draw inductive inferences in a word learning context using categorization. Twenty children with DLD (between 6;0 and 12;6), and 20 language-matched and 16 age-matched controls were exposed to a word learning task where they were given exemplars of objects associated with pseudo-words. The objects belonged to six categories spread across three hierarchical levels. For each item, the children chose which one(s), among a set of test objects from the same categories, could be labelled the same way (word extension). Results showed that school-aged children with DLD could extend new words to broader categories as well as their typically developing (TD) peers. Nevertheless, none of the DLD or TD children showed a specification of-their categorization of familiar instances that referred to more restricted instances. Our study suggests preserved abilities in using conceptual knowledge in order to learn new words, which could be used as a compensative strategy in the context of therapy. Further studies are needed to investigate this ability in more complex learning contexts.

Keywords: Developmental Language Disorder; vocabulary; word learning; word extension; categorization

Introduction

Children with Developmental Language Disorder (DLD, previously known as specific language impairment) face many difficulties concerning language acquisition and language use. According to the current consensus (Bishop et al., 2017), DLD is characterized by language deficits occurring without any sensory loss or neurological damage and with nonverbal cognitive abilities that do not meet intellectual disability criteria (i.e. within the normal to low range but not below the normal range). Varying degrees of difficulty with using/understanding language are noted, and they always affect everyday functioning and persist into adulthood (Conti-Ramsden et al., 2012; Conti-Ramsden et al., 2018).

Children with DLD often present with word learning deficits (Kan & Windsor, 2010). They are reported as making more naming errors compared with typically developing (TD) children (Sheng & McGregor, 2010b). McGregor et al. (2002) reported that the performance of young school-aged children in a naming task was linked to the scores of a drawing task and a definition task, indicating that naming difficulties are related to the richness – or poorness – of semantic representations of words. Furthermore, Charest and Skoczylas (2019) analysed the discourse of children with DLD who scored in the normal range in naming tasks. They described more lexico-semantic errors, which they defined as the inappropriate use of words in context. Their findings revealed that even if they know and can use as many words as their TD peers, children with DLD use these words less accurately when considering how to use them in a particular semantic context. These data tend to indicate that children with DLD have poorer knowledge of concepts underlying words preventing them from a refined use that accurately fits with the meaning of the context. Furthermore, this knowledge seems to be less organized in children with DLD than in TD children of the same age (Sheng & McGregor, 2010a). Indeed, these authors mentioned that semantic representations, which allow the storage of conceptual knowledge and the meaning of words, are less connected with each other.

Children with DLD thus seem to have a vocabulary knowledge limited in breadth, i.e. in terms of the number of known words, and depth, i.e. how well words are known in terms of semantic features (McGregor et al., 2013), which is likely underpinned by specific difficulties in word learning. Kan and Windsor (2010) found that children with DLD learned significantly fewer form-meaning associations than TD children matched for chronological age but not among children matched for language measures. Their meta-analysis also revealed that performance was moderated by variables such as age, nonverbal IQ or word type, and task characteristics. Nevertheless, the word learning difficulties encountered by children with DLD and their underlying causes remain misunderstood.

Learning new words implies identifying and encoding sound sequences, mapping a sequence of sounds to a referent or meaning, encoding the sound-meaning association and elaborating and integrating the meaning of this association into a larger set of semantic knowledge (Alt & Plante, 2006; Kan & Windsor, 2010; Waxman & Lidz, 2006). Early in the course of word learning, children have to organize the meaning they have built into categories. According to the internal lexical organization and the taxonomic assumption (Markman, 1990; Xu & Tenenbaum, 2007), children decide which concept of hierarchical taxonomy the word refers. For example, when learning the word 'flower', children must determine whether it refers to all flowers (basic level), a broader range of green living objects ('plant', superordinate level), or more specific instances such as 'roses' or 'tulips' (subordinate level). They can also consider meanings such as a specific part of the referent (an object in our example), for example the petals or the leaf, or the specific context/person attached to the word occurrence. Children build semantic knowledge connecting word representations in an interconnected network of features and use this knowledge to employ the words appropriately in context.

Among the processes involved in word learning, inference and categorization seem to make word learning easier and faster. Inference can be defined as a cognitive ability allowing the extension and generalization of knowledge to unfamiliar situations (Sloutsky et al., 2015). Some authors report inference as essential in the course of learning (Caplan, 2018; Fisher et al., 2015). Regarding word learning, the ability to draw inferences occurs as early as the initial stage of word learning, when mapping a lexical label with the corresponding referent (i.e. the form-meaning association step) is reached (Diesendruck, 2007). The contexts in which a child encounters a new word are rarely clear and explicit. The child has to go through the difficult task of identifying the word's referent and predicting which one might be best (Xu & Tenenbaum, 2007). Furthermore, in later stages of word learning, i.e. when building meaning, the learner has to extract relevant features to identify which category it refers to, and use this knowledge to extend the new words learned (lexical labels and concepts) to unfamiliar referents (Gentner & Namy, 2006).

Again, in some models of word learning (Caplan, 2018; Xu & Tenenbaum, 2007), inference appears helpful, or even necessary, in word extension[1]. As seen above, the amount of possible meanings makes the task confusing, leading some authors to argue that constraints or biases apply to word learning (Caplan, 2018; Markman, 1990) and guide the inference process.

Some of these constraints are well known. For example, Markman (1990) described the whole-object assumption, which supposes that a word refers to a whole entity rather than a part of it. She also explained the mutual exclusivity assumption, which refers to the learner's tendency to assign a new word to a referent that does not yet have a name. The shape bias, defined as a tendency to take into account an object's shape to form categories and extend words, can be considered as another example of this kind of constraint (Collisson et al., 2015).

---

[1] As suggested in Kan and Windsor (2010), word learning is defined as the whole process of encoding sequences of sounds, linking a word to an object, a concept or a meaning, and encoding and accessing the form-meaning association. By contrast, word extension refers to a later stage of word learning during which a word is generalized, i.e. the ability to apply a word to new instances which present similarities. Word extension contributes to word learning as it helps expand and make the initial learning more abstract.

Forming categories also appears to promote the creation of inferences. Categories may represent one of these constraints, as they limit the number of potential meanings, and allow for the generalization of a new word to referents with which they share common features. Categories have a strong inductive force, and some data suggest that category-based inferences contribute to word learning (Fisher et al., 2015; Waxman & Lidz, 2006). The ability to draw inferences based on category information seems to emerge early in development but it is only around the age of 7 that it is used in preference to perceptual information (Sloutsky et al., 2015).

In a series of experiments, Sloutsky et al. (2015) studied the impact of conceptual factors (conceptual knowledge about a category) and attentional factors (factors that draw attention to category-based features) on the ability to draw category-based inferences. They showed that, contrary to older children (7 year olds) and adults, 4 and 5 year olds do not use category-based information unless their attention is directed to the relevant features, even if conceptually rich information is provided. They preferentially rely on perceptual similarity to perform the inference task. So even in young children, an inference process appears to be involved in word learning. This process can be based on perceptual or conceptual cues, which change during development depending on the kind of information available.

Constraints and category-based inferences involved in word learning have been studied in children with DLD. Collisson et al. (2015) showed that young children with DLD encountered difficulties developing the shape bias. Their results indicated that young children with DLD tend to be less consistent when they have to choose between shape, colour or texture in word extension. Krzemien et al. (2021) extended these findings and showed that this shape bias, particularly relevant for solid and animate objects, does develop in older children with DLD. Nevertheless, these authors also highlighted difficulties in word extension for categories defined by other types of features such as texture, functional roles or spatial relations. In their literature review of 30 years of research on shape bias, Kucker et al. (2019) linked this bias with the learner's ability to detect statistical regularities (what we refer to as statistical learning) and suggested that these statistical regularities modulate the way in which the shape bias is mobilized. However, the ability to track the statistical regularities of the environment appears to be impaired in children with DLD in various tasks and contexts (see Obeid et al., 2016 for a meta-analysis). In the context of vocabulary, Haebig et al. (2017) reported that children with DLD perform less well in a word segmentation task than TD children or children with autism spectrum disorder. In line with these ideas, word learning in children with DLD seems to be facilitated when the opportunities to implement the detection of regularities are highlighted. This appeared in the study of Aguilar et al. (2018), which showed better word generalization performances when the learning context involved a high variability of exemplars. This also seems to be the case in Krzemien et al. (2021), who showed that school-aged children with DLD benefited from the comparison of several exemplars of a referent (and thus more opportunities to detect regularities) to extend words for texture- or spatially-based categories. However, category-based inference and statistical learning in the context of lexical and particularly semantic aspects of word learning in children with DLD has rarely been explored until now (Obeid et al., 2016).

Bayesian models of word learning

Bayesian models of cognition adopt a statistical point of view of the child's development, whereby human beings permanently employ probabilistic or inductive inference to make predictions about their environment (Perfors et al., 2011). This mechanism guides learning experiences and is based on the interconnection of three core elements: prior probability, likelihood, and posterior probability. Prior probability refers to the learner's use of prior knowledge, which, in a word learning context, generates hypotheses for the meaning of a word. Combined with the probability of observing data, i.e. examples of a referent in our case, under specific hypotheses (likelihood), a posterior probability is computed for each of the hypotheses in order to determine which one is the most likely to explain these data. Learning takes place by eliminating hypotheses that are less likely to be and by selecting the most plausible hypothesis. The cognitive system selects the optimal explanation of a phenomenon and creates a unit of knowledge that can quickly become abstract and be generalized to new instances (Tenenbaum et al., 2006).

According to this model, a learner who is faced with the task of learning a word will generate hypotheses about the potential meaning of a label when performing form-meaning associations. They will use their prior knowledge of concepts and their organization of semantic representations and meanings to evaluate concepts related to the different hypotheses. They will then attribute greater or lesser probability to each one of these hypotheses. For example, when having to associate the lexical form ‘flower’ with its referent, the learner may think about the possibility that the referent might be, the petals of the flower, this flower specifically, all flowers, or plants in general. Each of these possible meanings represent generated hypotheses, associated with a prior intuition/probability of being true.

If the learner’s intuition about the exact meaning remains uncertain after their first encounter with a flower, they will accumulate evidence through various encounters that reinforces or weakens the intuition that the label ‘flower’ refers to the flower and not the stem or all plants. Furthermore, each time children learn a word, they will reinforce the idea that words usually (with greater probability) refer to whole objects. Each piece of the accumulated evidence will thus reinforce and modulate prior knowledge. In other words, the updating of prior knowledge will progressively reinforce the hypothesis that favours the association observed and weaken other hypotheses, i.e. ‘flower’ word-stem referent association or ‘flower’ word-plants referent association.

Bayesian theories of learning are of particular interest in word learning as they offer an integrated framework, combining inductive inference, a statistical learning mechanism, and prior knowledge, known to drive learning experiences (e.g. Kim & Rehder, 2011). Xu and Tenenbaum (2007) define *priors* as constraints that guide word learning as well as knowledge, beliefs or expectations about how words and the world work. *Likelihood* is considered as depending on statistical regularities and information observed from the data. *Inference* refers to the updating of priors and their associated weight.

To test if their Bayesian model could account for word learning, Xu and Tenenbaum (2007) conducted an experiment with children between 3;0 and 5;0 years old and adults. Their experiment included pictures from three semantic categories familiar to participants, i.e. vehicles, animals and vegetables, organized on three levels of taxonomy. Pictures were divided into example and test items. Participants were given a word learning task that included 1- or 3-example items labelled with pseudo-words. The 3-example items depicted the subordinate, basic, or superordinate level of taxonomy. From a set of 24 test items, participants selected which ones could be given the same label. Percentages of choices were computed for each hierarchical level of each category in order to see if, according to the Bayesian account, participants adapted their choices according to the amount of data provided. Adults and children both demonstrated results in line with the predictions of the model, showing a gradient of choices differing according to the number and range of example items presented. Participants adapted their responses following statistical regularities among example items in a word learning task based on known objects.

To sum up, children with DLD have a vocabulary limited in breadth and depth (McGregor et al., 2013). Research suggests that the vocabulary limitations may be at least partially attributed to difficulties in mapping sound sequences or semantic features to objects (e.g. Alt & Plante, 2006). Children with DLD also demonstrate statistical learning deficits in word learning tasks (Evans et al., 2009; Haebig et al., 2017; Obeid et al., 2016). These deficits have mostly been studied with regard to phonological aspects but have yet to be investigated in the lexico-semantic domain (Obeid et al., 2016).

With respect to semantic development, category-based inference seems to play an essential role. To our knowledge, only a few of the above-mentioned studies investigate the semantic and conceptual aspects of word learning in this population in terms of ability to track regularities in order to draw category-based inferences. Krzemien et al. (2021) recently showed that the shape bias appears to be preserved in school-aged children with DLD, who still encounter difficulties with other types of categories. These authors address the form-meaning mapping stage of word learning but do not concentrate on how children with DLD deal with deeper semantic, conceptual knowledge.

Study objectives

Our study aimed to assess the inductive inference and categorization abilities of children with DLD in the context of word learning. It asked the following question: Are school-aged children with DLD able to extend new words and organise their semantic knowledge in categories by drawing on inductive inference as well as their typically developing peers?

Children with DLD were compared with typically developing (TD) children in a word extension task adapted from Xu and Tenenbaum (2007). As category-based inference appears to change during development (Sloutsky et al., 2015), we created two control groups, one group with children matched according to chronological age and nonverbal cognitive abilities, and one group of children with similar lexical development. If the children with DLD had presented with word extension difficulties in our task, the presence of two control groups might have allowed us to determine if their development appeared atypical or delayed, and thus similar to children with similar lexical development. Our hypotheses were as follows:

(1) children with DLD would not benefit from *more* examples to the same extent as their TD peers when extending new words.

 As predicted by the Bayesian model proposed by Xu and Tenenbaum (2007), TD children should present better extension performances, i.e. a more specific pattern of choices (related to the adequate level of taxonomy, in contrast with diffuse patterns of responses when only one exemplar is available), when three exemplars are given than when only one is presented. The opportunity to detect common features, i.e. statistical regularities, among the three exemplars might help word extension.

 In line with the predictions, compared with showing one exemplar, children should: (1) choose less responses from the basic and superordinate levels when shown three exemplars from the subordinate level; (2) show similar patterns when shown three exemplars from the basic level due to the presumption of a basic-level bias; (3) choose all of the subordinate and basic responses as well as more responses from the superordinate level when shown three exemplars from the superordinate category. Since the 1-exemplar test item should lead to greater ambiguity, we suggested that children with DLD should present better extension when shown three exemplars (less ambiguous) compared with one exemplar, but perform less well than their TD peers (either age- or language-matched) given their difficulties in statistical learning paradigms (Obeid et al., 2016).

(2) Compared with TD children, we expected children with DLD to be unable to take into account the *type* of information provided.

 According to the model, a decreasing gradient should be observed among items from the subordinate level to the superordinate level of taxonomy such as: (1) three subordinate exemplars would lead to a majority of subordinate level responses and a minority of basic and superordinate level responses; (2) three basic exemplars would lead to a majority of subordinate and basic responses; (3) three superordinate exemplars would lead to a greater selection of responses from the superordinate level compared with the two other kinds of items.

We hypothesized that children with DLD should show an intermediate pattern of responses (in terms of distribution of choices between the different levels of taxonomy) between A-matched children and L-matched children, given that they have less lexical knowledge but similar concept exposure than the former and similar lexical knowledge but more concept exposure than the latter.

Methods

This research was validated by the Research Ethics Committee of the University of Liege's Faculty of Psychology, Speech Therapy and Educational Sciences (ref. 1718-54). Consent was obtained from the parents of every child and from the children themselves.

Participants

We recruited three groups of participants. The first group consisted of 20 children (15 boys aged from 6;5 to 13;3 years), with a diagnosis of Developmental Language Disorder (DLD). They attended schools for children with special needs in the French-speaking part of Belgium and all of them were monolingual (French only). It should be mentioned that the fact they attended such schools indicates that their difficulties had a functional impact on their academic achievement. Language difficulties were confirmed by at least two scores at 1.25 standard deviation below the mean on the language tasks described below. Their language profiles reflected mostly severe difficulties in phonology and morphosyntax (production and reception, see Table 1). Their lexical abilities fell within the mean scores but were significantly lower than the scores of their age-matched peers. All of them had a nonverbal IQ above the criteria for intellectual disability (ranging from 80 to 128 in our sample) and a poor prognosis without any associated differentiating or biomedical condition. They did not present any history of sensory loss, neurological impairment or prematurity.

The second group consisted of 20 TD children (11 boys aged from 4;10 to 10;9 years) matched in lexical performance (L-matched) and socioeconomic status. They had similar scores (more or less 8 points) for the word designation task.

The third group consisted of 16 typically developing children (6 boys aged from 6;4 to 11;7 years) matched in chronological age (A-matched – more or less 4 months) and nonverbal IQ.

Children from both control groups attended regular schools in the French-speaking part of Belgium and scored in the mean range in language tests. Like the children with DLD, TD children were monolingual (French). They did not have any sensory loss, brain injury, neurological impairment or prematurity, or other biomedical conditions.

As seen in Table 1, groups differed significantly in all language measures: children with DLD achieved poorer scores than their TD peers, except for matching measures.

[Insert Table 1]

Materials and procedure

Language and cognitive measures

All children received a language assessment using standardized tests to confirm diagnosis for children with DLD and exclude language difficulties for TD children. Receptive and expressive language tests were administered for phonological, lexical and morphosyntactic abilities. We used a word repetition task (Répétition de mots from the Evaluation du Langage Oral (E.L.O.) battery, Khomsi, 2001) for expressive phonology. We conducted a word designation task (French adaptation of the Peabody Picture Vocabulary Test-revised (PPVT), Dunn et al., 1993) and a word denomination task (Lexique en production from the Evaluation du Langage Oral (E.L.O.) battery, Khomsi, 2001) for receptive and productive lexical abilities. A sentence designation task (Epreuve de COmpréhension Syntaxico-Sémantique (E.CO.S.SE), Lecocq, 1996) and a sentence completion task (Production d'Enoncés from the Evaluation du Langage Oral (E.L.O.) battery, Khomsi, 2001) were administered to assess receptive and expressive morphosyntax. For these tests, children's performances were compared to a norm established on the basis of similar chronological ages for the French adaptation of the PPVT and the E.CO.S.SE test, and on similar grade levels for the E.L.O. battery. Nonverbal IQ was measured using the Weschler Nonverbal Scale of Ability (Weschler & Naglieri, 2009).

Word learning task

We drew inspiration from Xu and Tenenbaum (2007) and designed a two-part word learning task. Each part was similarly designed as follows.

Materials. For each part, we selected three semantic categories that appear in the early development of vocabulary and are highly familiar to children. This led to a total of six categories for the entire task (vegetables, vehicles and animals for Part 1, fruits, toys and clothes for Part 2). For each semantic category, we selected 15 colour photographs presenting category members distributed across three levels of taxonomy (superordinate, basic, subordinate).

Among the 15 pictures, seven were selected to form test exemplars. For each category, we created one 1-exemplar test item including one picture and three 3-exemplar test items, one for each level of taxonomy including three pictures each, with one constant picture across the items. For example, we designed four test items for the animal category as follows: 1-exemplar test item represented a red bird; 3-exemplar test items represented the same red bird and two other red birds (subordinate level), two birds of different colour (green and yellow, basic level), and two other animals (a whale and a pig, superordinate level). This led to 12 test exemplars in total for each part of the task, four per semantic category (see Figure 1). For each category, the criteria used to determine that an object belonged to the superordinate and basic levels of taxonomy, were similar to the natural ones (e.g. animals, vegetables or vehicles as superordinate members; birds, peppers and trucks as basic levels). Nevertheless, for some of the categories used, the subordinate level of taxonomy appeared more familiar than for others (e.g. regarding animals, Dalmatians refer to a concept more familiar than lamb's lettuce regarding vegetables or tankers and vehicles). In order to avoid the impact of prior knowledge on our task as much as possible, the criteria used for categorization needed to be equally familiar (or unfamiliar) for each child. Therefore, we chose a similar criterion that seemed salient and highly familiar, i.e. colour, as a defining characteristic of belonging to the subordinate level for each category.

The eight remaining photographs of each category (24 pictures) formed a 4*6 grid of response items. For each category, these eight pictures were distributed in such a way that two belonged to the subordinate level (e.g. red birds), two to the basic level (e.g. a blue and a pink bird) and four to the superordinate level (e.g. a bear, a bee, a seal and a cat). Following this distribution, since the more restricted levels of taxonomy are by definition, included in the broader ones, participants could select maximum two pictures from the subordinate level and four pictures from a basic level (two specific to the basic level plus the two pictures specific to the subordinate level). They could choose eight pictures from a larger superordinate category (two from the subordinate level, two specific to the base, and four that were not birds and belonged to the superordinate level). The eight pictures of the other two categories represented non-matching pictures for the target category. The grid of responses was the same for all of the test items.

[Insert Figure 1]

Each test item was associated with a disyllabic non-word built according to a Consonant-Vowel-Consonant-Vowel (CVCV) structure and the phonotactic rules of French. The use of words, or language in general, appears to be linked to categorization (Ferguson et al., 2015). It has been shown that lexical labels influence object categorization in the first few months of life, and may constitute an additional feature on which a learner can rely to form categories (Fulkerson & Waxman, 2007). Furthermore, conceptual categories, those that are labelled with lexical forms, may differ from perceptual ones (Clark, 2017). For 3-exemplar items, one similar non-word was given once for each picture.

Procedure. We programmed and administered the task on a 15-inch laptop computer using OpenSesame software (Mathôt et al., 2012). We told the children that a Martian who spoke a different language needed help to collect sets of objects. They were shown one or three images at the top of the screen and heard a Martian word (pseudo-word in French, e.g. '*baté'* - */bate/*) for the objects depicted. The children were told that they had to pick out all the '*baté'*, but only the '*baté'*, among the pictures below. They had to pay attention to the fact that the Martians did not want any objects that were not '*baté'*. The examiner explained that the number of selected pictures could change depending on the trials and the children were free to select what seemed to be a good answer. When the children had selected all of the '*baté'*, they clicked on a 'next' button to see the next trial. For each part, the test began with the 1-exemplar test items and continued with the 3-exemplar test items, allowing us to observe a change in their response with a larger number of exemplars. In order to avoid the influence of any previous experience that may have suggested good or bad answers, and to examine the children's spontaneous behaviour, no training set was included before the task. After the task was completed, we used a homemade denomination task containing each of the items used. This verification process allowed us to test the knowledge of each child for the items used and to control any effect due to a difference in lexical knowledge concerning the materials. Each child knew at least 88 out of the 90 items. The whole task took approximately 20 minutes to complete. For more details about the pictures - non-word pairs and the entire procedure, see Appendix 1.

Collected data. For each item, we counted the number of pictures chosen by children for each level of taxonomy in each category and the number of pictures chosen outside of the target category. For instance, for the 3-exemplar test item from the superordinate level of the animals category, we counted how many pictures a child had selected among the subordinate level pictures (red birds), the basic level pictures (blue bird and pink bird) and the superordinate level pictures (bear, cat, seal and bee) for the animal category. We also counted how many pictures they had selected among pictures from the two other categories (vegetables and vehicles). By counting the responses separately for each level of taxonomy, it was possible to take into account the precision of the choices a child made according to the test item presented. The specificity of word extension was reflected by differences in the proportions of responses between the three levels of taxonomy according to the quantity and type of exemplars provided.

Results

Statistical analyses

We conducted Bayesian statistical analyses in accordance with our theoretical framework. Bayesian statistics indicate a degree of change in our previous beliefs, reweighting the credit given to each of the *a priori* hypotheses according to the results obtained (Kruschke & Liddell, 2018; Wagenmakers et al., 2018). Whereas classical analyses are based on the acceptance or rejection of the alternative hypothesis according to arbitrary thresholds, Bayesian statistics express an amount of evidence for or against a specific hypothesis/model in comparison with other models. They can lead to evidence against a specific hypothesis, until there is a strong enough amount of evidence against the null hypothesis to make it implausible.

Analyses were conducted on JASP (Love et al., 2019) using default priors. We shall present Bayesian factors (BF) as results of our analyses. BF represent the way the data observed favour one model over another. BF_{10} expresses the ratio of the likelihood of an alternative model over the null model, and vice-versa for BF_{01}. We refer to the classification proposed by Lee and Wagenmakers (2013) to qualify the BF values: $BF \leq 1$ indicates data are insufficient – no evidence is given, $1 \leq BF \leq 3$ reports anecdotal evidence, $3 \leq BF \leq 10$ provides moderate evidence, $10 \leq BF \leq 30$ accounts for strong evidence, $30 \leq BF \leq 100$ corresponds to very strong evidence and $BF \geq 100$ decisive evidence for the tested hypothesis.

Effect of the number of exemplars on word extension in children with DLD

The objectives of the study were twofold. The first objective was to determine whether children with DLD benefit as much as their TD peers from an increased number of exemplars to extend new words, and the second was to determine whether children with DLD benefit as much as their TD peers from the type of information provided in the exemplars.

As no differences appeared between categories, responses for the different categories were computed together according to the type of example given (one exemplar, three exemplars from the subordinate level, three from the basic level and three from the superordinate level). Scores reflect the number of responses given for each level of taxonomy for each item. Results are shown in Figure 2 (see also Table 2), representing patterns of choices made for each kind of item.

[Insert Table 2]

[Insert Figure 2]

In order to answer our first question, Bayesian repeated measures analyses of variance (ANOVA) were conducted on these scores with Type of item (4) x Level of taxonomy (3) as within-subject factors and group (3) as a between-subject factor. Results revealed moderate evidence against a group effect ($BF_{excl} = 6.524$). This suggests that whatever the group, the patterns of choices are quite similar, indicating that children with DLD make similar choices to their TD peers (either A- or L-matched). Furthermore, decisive evidence was found for an effect of the type of items ($BF_{incl} = 4.823e+8$), the level of taxonomy ($BF_{incl} = 6.145e+167$)

and an interaction effect between the type of items and the level of taxonomy (BF_{incl} = 8.686e+20, see Table 3).

[Insert Table 3]

Post-hoc tests revealed strong evidence for a difference of choices between the 1-exemplar condition and the 3-subordinate exemplars condition (BF_{10} = 22.427) and extreme evidence for a difference of choices between the 1-exemplar condition and the 3-basic and superordinate exemplars conditions (respectively BF_{10} = 122663.907 and BF_{10} = 3.818e+8). Patterns on the graphs (Figures 2 a, b, c) show a decreasing gradient of choices from the subordinate level to the superordinate level. However, this gradient differs according to the type of exemplars given: children tend to choose slightly more responses from the subordinate and basic levels when shown 3-subordinate exemplars (for DLD and L-matched children) compared with one exemplar. When shown three exemplars from the basic level, they tend to choose more responses from the basic level. L-matched children also choose more responses from the superordinate level. When shown 3-superordinate exemplars, children choose more responses from the basic and superordinate levels, indicating that children adjust their choices when additional information extends to a broader level (basic or superordinate). However, they demonstrate less specific extension when the information concerns a more restricted level (subordinate).

Effect of the type of exemplars on word extension in children with DLD

We then conducted a second ANOVA on the 3-exemplars items only to address our second objective regarding adaptation to the type of information provided. Given the evidence against group differences, the results of the three groups were compiled together for further analyses.

A Bayesian repeated measures ANOVA with Type of items (3 subordinate, 3 basic and 3 superordinate exemplars) and Level of taxonomy (3) as within-subject factors, revealed decisive evidence for an effect of type of exemplar and level of taxonomy, and for an interaction between both factors (see Table 4). Post-hoc tests indicated strong to decisive evidence for each comparison, i.e. the patterns of choices for each type of item differed from each other: 3-subordinate vs 3-base (BF_{10} = 15.008), 3-base vs 3-superordinate (BF_{10} = 294.891) and 3-subordinate vs 3-superordinate (BF_{10} = 69251.769). As predicted by our hypotheses, and following the model proposed by Xu and Tenenbaum (2007), the type of information seems to guide word extension.

[Insert Table 4]

Discussion

Our study aimed to define whether word learning difficulties encountered by children with DLD could be explained by a deficit in category-based inductive inference. Results revealed that contrary to our hypotheses, children with DLD seem able to extend new words as well as their A- and L-matched TD peers. These results are in line with recent findings suggesting that school-aged children with DLD perform similarly to A-matched and L-matched children regarding solid and animate categories of nouns in a word extension task (Krzemien et al., in press). In our study, children with DLD appeared to benefit as much as TD children from both the quantity and the type of information they received during word extension. Regarding quantity, children extended new words to closer concepts when three exemplars of a similar concept were provided whereas their responses were rather more diffuse when shown only one exemplar of the same concept. Furthermore, children with and without DLD adapted their choices depending on the level of taxonomy to which the three exemplars referred. When shown three red birds, children tended to choose fewer responses from the basic level (birds of different colours) and barely any responses from the superordinate level of taxonomy (a cat, a bear or a bee). They also chose nearly every response from the subordinate and basic levels (red birds and different coloured birds) when shown three exemplars from the basic level (a yellow, a green and a red bird). Despite more responses from the superordinate level than expected, children also chose few responses from the superordinate level in this same case. In our example, they chose fewer responses such as cat, bear, bee or seal when shown three different coloured birds than when shown a red bird, a pig and a whale. Finally, they chose a large amount of exemplars from the subordinate and basic levels, but also from the superordinate level, when shown three exemplars from the superordinate level.

These results run counter to the hypothesis of a deficit of statistical learning in DLD children. As suggested so far, children with DLD encounter difficulties in tasks evaluating statistical learning in diverse modalities and at different ages (Obeid et al., 2016; Haebig et al., 2017). Our findings suggested that they were quite good at extending new words based on the detection of regularities. This could reveal a preserved ability to draw category-based inferences. This could also be explained by at least two interconnected factors: the age of the participants and the type of stimuli. The semantic categories used in our task referred to categories that children learn early in their development (MacRoy-Higgins et al., 2016), defined by salient perceptive features that are acquired earlier and more easily than relational features. This allowed us to check for a difference in semantic knowledge between DLD children and their peers. In contrast, the use of relatively simply defined, early-acquired categories might have led the children with DLD to acquire strong knowledge to perform the task, in line with the idea of a benefit from age and experience as suggested by Krzemien et al. (2021). Our verification procedure did not allow us to control the semantic and conceptual knowledge associated with our items. However, the fact that each child had lexical knowledge for every item included in our task, as well as scores ranging in the mean on standardized vocabulary tests, seemed to confirm this hypothesis. Furthermore, our categories could have been compared to the solid and animate categories defined by the same authors as mastered by school-aged children, either with or without language disabilities. Given that these tests are often composed of concrete, solid and animate categories, these elements could also explain why a large number of children with DLD obtain scores in the mean of TD children at vocabulary tests, even though they often score significantly lower than their peers (Charest &

Skoczylas, 2019). In our study, it could be that a real inference process was not needed to implement category-based induction when the distributions of categories across different levels of taxonomy were known. In line with the model developed by Xu and Tenenbaum (2007), children with DLD may be able to draw inductive inferences when prior knowledge is available but not when this knowledge is unavailable or with more complex categories. Further studies are needed to explore word extension abilities of children with DLD in contexts in which they have less strong prior knowledge.

Our results also confirmed the existence of a suspicious coincidence effect, i.e. it would be suspicious if multiple occurrences of a concept with the same label did not refer to a similar entity, or a similar level of categorization. As explained by these authors, when shown three red birds called '*baté'*, it would have been a coincidence if *'baté'* had referred to a broader level such as birds or animals. Following our results, word learning and word extension in children appear to depend on how many exemplars of a concept they have been exposed to, further specifying their choices according to the features shared by these exemplars. The extent of these effects on generalization behaviour nevertheless differs from what Xu and Tenenbaum (2007) observed in younger children and adults. Our participants adapted their extension in the case of basic or superordinate levels of taxonomy (e.g. extending words adequately for concepts of birds and animals, but to a lesser extent for the concept of red birds). In contrast, the pattern of choices for narrow extension (three subordinate exemplars) was not as specific as expected, revealing an extension failure for three potential reasons.

Firstly, the preference for the basic level could reveal knowledge about the fact that labelling objects at a basic level is informative enough to avoid labelling at a narrower level (Emberson et al., 2019). From a communicative perspective, labelling objects at the basic level often appears functional.

Regarding the impact of prior knowledge about the categories, the observed trend is consistent with the findings revealed by Jenkins et al. (2015) showing that children with stronger category knowledge demonstrate an attenuated suspicious coincidence effect. Children might have learnt that items from narrower categories are usually labelled with multiple morphemes, such as a noun and an adjective, and not only one as presented in our experiment. This might have been especially true in our case, as the characteristics that defined the subordinate level were a combination including colour, such as a red bird or a green pepper. However, the extent to which children chose basic level responses remained higher.

Considering the familiar nature of the categories, children may have learnt that colour is usually an irrelevant feature to classify these instances (Kucker et al., 2019), leading them to ignore it which is consistent with this acquired knowledge. Prior knowledge may also modulate the impact of frequency (Harris et al., 2008). A learner equipped with prior knowledge about relevant features of a category is inclined to ignore frequent information inconsistent with their knowledge. Additionally, the shape bias is a knowledge acquired by both DLD and TD children (Krzemien et al., 2021) that guides their choices in word extension tasks and may have been mobilized massively. Prior knowledge can modulate the mechanism that detects regularities and integrates them with the knowledge already available (Harris et al., 2008) such that, even if colour appeared as a frequent feature in our stimuli, the child may have ignored it because of its inconsistency with their prior knowledge about categories. From a developmental perspective, as prior knowledge changes and develops with age and experience, younger children, such as those tested by Xu and Tenenbaum (2007), may be less influenced in their choices by factors such as shape bias and knowledge about colour. In contrast, school-aged children, either with or without DLD, may rely heavily on this prior knowledge to make their choices in word extension tasks. This could explain differences in results with Xu and Tenenbaum (2007). Contrary to our results, in their study, 3- and 4-year-old children showed a strong restriction in their choices in favour of items from the subordinate level when seeing three exemplars from this level. Finally, compared to school-aged children, adults may master their knowledge or have developed more cognitive control to reconsider statistical regularities beyond the knowledge they acquired. Their prior knowledge could also be richer and more nuanced, leading them to consider more characteristics in their choices.

Thus, the influence of prior knowledge might partially explain our results. Furthermore, the use of prior knowledge could be linked to the age factor mentioned earlier. As Sloutsky et al. (2015) mentioned, attentional factors seem to drive inferences until the age of 7, whereas conceptual factors are preferentially used from 7 years old. Variations in the use of prior knowledge might be associated with this shift from attentional to conceptual factors. In parallel with its increase with age, prior knowledge could help the shift, or become extensive enough to combine with conceptual factors and reinforce their use.

Secondly, it seems that a higher quantity of information can lead to broader extensions but cannot explain narrower extensions, and that atypical exemplars of a category lead to a more specific pattern of choices than typical exemplars (Emberson et al., 2019). Atypical exemplars usually possess salient features that distinguish them from typical exemplars and lead to the grouping of these instances in a subcategory. In our case, colour could have been a feature that was is not salient enough to drive the child’s attention and affect their choices when extending new words.

Thirdly, both variability and predictability seem to play a role in language processing. These concepts could be inversely related: a highly variable context leads to easier categorization but prevents the learner from predicting future words, and vice versa (Cassani et al., 2018). Variability seems to facilitate word learning in children with DLD and help category formation by making relevant properties salient (Aguilar et al., 2018). On the contrary, too much variability leads to higher attentional demands. A balance may be necessary to provide enough variability to facilitate category abstraction but enough predictability to allow for category-based inference (Fisher et al., 2015) and a reduction in attentional demands. It could be argued that three very similar exemplars were insufficient in our study to allow for feature abstraction and category formation at a narrower level. Further studies are needed to determine how both factors vary and the mutual influence between them.

Limitations

Our study has some methodological issues that need to be addressed. Firstly, the stimuli and the pictures used to represent them might have influenced the results. As already mentioned, the categories were chosen to be very familiar to children. Moreover, the choice of colour as a relevant feature to identify the subordinate level is not completely consistent with natural categories. This criterion was chosen to ensure consistency across all categories and their subordinate level, and to avoid an impact of differences in prior knowledge for more precise terms or concepts as much as possible. For one part of the experiment, most of the examples and stimuli were similar to the original task of Xu and Tenenbaum (2007). The basic level for the animal category was adapted to make a colour subordinate less dependent on a breed of dogs, which is often associated with this criterion of colour. We assumed that these categories were equally familiar to children with and without DLD. However, as Kucker et al. (2019) suggested, the colour criterion might not have been relevant for categorization at this age, and therefore the choices made by our participants must be read in the light of this idea. Furthermore, children with DLD are delayed in their semantic organization compared with TD peers and behave like younger children when tested for word associations (Sandgren et al., 2021). It would be useful to test whether children with and without DLD had similar underlying knowledge about selected pictures and stimuli.

Secondly, while we included pseudo-words in our study, we did not investigate the phonological aspects of word learning, nor did we check for form-meaning mapping after delay. The inclusion of non-words in association with pictures of existing objects (and thus existing words) addresses methodological issues. The present study aimed at highlighting the behaviour of children when they have to learn concepts associated with words they do not know, but not words themselves. This implies adding words to pictures in order to see how children behave in this particular case, but does not necessitate that they encode phonological aspects of the words presented. However, further studies could examine how children with DLD encode word forms in this type of learning situation.

We did not conduct statistical power analysis for two main reasons. First, our population of interest manifest a disorder, and is therefore a small population. In this case, additional recruitment barriers may be encountered, making it difficult to form groups that are large enough to meet the power criteria. However, Bayesian statistical approaches remain a valid way of treating the results (Kruschke, 2010). The second reason is related to the Bayesian nature of the analyses that have been conducted. Power analyses using Bayesian methods involve knowing at least the direction in which the data will go, and defining priors as close as possible to the expected results (Kennedy, 2015; Kruschke, 2010). Given the exploratory nature of our experiment, and in the absence of previous results or studies on similar topics or questions, it is difficult to estimate effects or generate fictive data that could fit the actual results and data collected. Nevertheless, when estimating statistical power with the G*Power software (Faul et al., 2007; Kennedy, 2015), dedicated to power analysis of classical statistical tests, a total sample size of 24, i.e. eight subjects per group, was proposed. Thus, it appeared that our sample sizes were sufficient, but it might be interesting to replicate these findings with larger groups to ensure greater power.

Finally, the ages of our participants covered a wide range. Regarding lexical development, several authors have suggested a reorganization of lexico-semantic knowledge around the ages of 7 or 8 (Sandgren et al., 2021; Sloutsky et al., 2015). Due to the way our groups are constituted, we were not able to distinguish age groups and investigate the differences between children under 6 or 7 years old and those over 8 years old. From a developmental perspective, it would have been interesting to contrast these groups. Although we did not find significant differences between our groups, including the younger L-matched group, and even found evidence against a group effect, this distinction opens area for further research to clarify the developmental changes that might be occurring. Especially, it would be interesting to investigate whether such reorganization can also been observed in children with DLD, and at what age.

Conclusions

In conclusion, both children with and without DLD adjusted their extension behaviour according to the number and type of exemplars in word learning. Our results suggested for the first time that children with DLD can mobilize processes such as word extension given the presence of a background in terms of conceptual knowledge and use this acquired knowledge in new contexts of learning. The reason why children with DLD performed as well as their TD peers remains unclear but could reveal an ability to use prior knowledge or to draw statistical inferences that could be used as a basis to overcome word learning difficulties and facilitate word learning. This may lead to evaluating word learning over word knowledge in DLD children in a dynamic way, in order to see if they can already mobilize these processes in word learning. Therapy could either rely on or reinforce these abilities as a basis to compensate word learning difficulties. Research is needed to understand more precisely the influence of prior knowledge on word learning and extension in a developmental perspective. Finally, the decreasing gradient of responses along with the type of items shown did not appear as marked as supposed. We reviewed several potential factors that may have led to different results, offering perspectives for future studies and highlighting the importance of considering the word learning abilities of DLD children in therapy.

Acknowledgements. We thank the parents, children, and teachers who participated in the project. This work was supported by the F.R.S.-FNRS FRESH fund under Grant 6200320F.

Disclosure of interest. The authors report no conflict of interest

References:

Aguilar, J. M., Plante, E., & Sandoval, M. (2018). Exemplar Variability Facilitates Retention of Word Learning by Children With Specific Language Impairment. *Language, Speech, and Hearing Services in Schools*, *49*(1), 72-84. https://doi.org/10.1044/2017_LSHSS-17-0031

Alt, M., & Plante, E. (2006). Factors That Influence Lexical and Semantic Fast Mapping of Young Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *49*(5), 941-954. https://doi.org/10.1044/1092-4388(2006/068)

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & and the CATALISE-2 consortium. (2017). Phase 2 of CATALISE : A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, *58*(10), 1068-1080. https://doi.org/10.1111/jcpp.12721

Caplan, S. (2018). Word Learning as Category Formation. *Proceedings of the society for computation in linguistics, 1*, Article 19. https://doi.org/10.7275/R54X560G

Cassani, G., Grimm, R., Daelemans, W., & Gillis, S. (2018). Lexical category acquisition is facilitated by uncertainty in distributional co-occurrences. *PLOS ONE, 13*(12), e0209449. https://doi.org/10.1371/journal.pone.0209449

Charest, M., & Skoczylas, M. J. (2019). Lexical Diversity Versus Lexical Error in the Language Transcripts of Children With Developmental Language Disorder : Different Conclusions About Lexical Ability. *American Journal of Speech-Language Pathology*, *28*(3), 1275-1282. https://doi.org/10.1044/2019_AJSLP-18-0143

Clark, E. V. (2017). Semantic categories in acquisition. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (2nd ed., pp. 397-421). Elsevier.

Collisson, B. A., Grela, B., Spaulding, T., Rueckl, J. G., & Magnuson, J. S. (2015). Individual differences in the shape bias in preschool children with specific language impairment and typical language development : Theoretical and clinical implications. *Developmental Science*, *18*(3), 373-388. https://doi.org/10.1111/desc.12219

Conti-Ramsden, G., Durkin, K., Toseeb, U., Botting, N., & Pickles, A. (2018). Education and employment outcomes of young adults with a history of developmental language disorder. *International Journal of Language & Communication Disorders*, *53*(2), 237-255. https://doi.org/10.1111/1460-6984.12338

Conti-Ramsden, G., St Clair, M. C., Pickles, A., & Durkin, K. (2012). Developmental Trajectories of Verbal and Nonverbal Skills in Individuals With a History of Specific Language Impairment : From Childhood to Adolescence. *Journal of Speech, Language, and Hearing Research*, *55*(6), 1716-1735. https://doi.org/10.1044/1092-4388(2012/10-0182)

Diesendruck, G. (2007). Mechanisms of word learning. In E. hoff & M. Shatz, *Blackwell Handbook of Language Development* (pp.257-276). Blackwell Publishing. https://doi.org/10.1002/9780470757833.ch13

Dunn, L. M., Theriault-Whalen, C. M., & Dunn, L. M. (1993). *Echelle de vocabulaire en images peabody*. Toronto: Psychan.

Emberson, L. L., Loncar, N., Mazzei, C., Treves, I., & Goldberg, A. E. (2019). The blowfish effect : Children and adults use atypical exemplars to infer more narrow categories during word learning. *Journal of Child Language, 46*(05), 938-954. https://doi.org/10.1017/S0305000919000266

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 52*(2), 321-335. https://doi.org/10.1044/1092-4388(2009/07-0189)

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3 : A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. https://doi.org/10.3758/BF03193146

Ferguson, B., Havy, M., & Waxman, S. R. (2015). The precision of 12-month-old infants' link between language and categorization predicts vocabulary size at 12 and 18 months. *Frontiers in Psychology*, *6*, Article 1319. https://doi.org/10.3389/fpsyg.2015.01319

Fisher, A. V., Godwin, K. E., Matlen, B. J., & Unger, L. (2015). Development of Category-Based Induction and Semantic Knowledge. *Child Development*, *86*(1), 48-62. https://doi.org/10.1111/cdev.12277

Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not Tones) facilitate object categorization : Evidence from 6- and 12-month-olds. *Cognition*, *105*(1), 218-228. https://doi.org/10.1016/j.cognition.2006.09.005

Gentner, D., & Namy, L. L. (2006). Analogical Processes in Language Learning. *Current Directions in Psychological Science, 15*(6), 297-301. https://doi.org/10.1111/j.1467-8721.2006.00456.x

Haebig, E., Saffran, J. R., & Ellis Weismer, S. (2017). Statistical word learning in children with autism spectrum disorder and specific language impairment. *Journal of Child Psychology and Psychiatry*, *58*(11), 1251-1263. https://doi.org/10.1111/jcpp.12734

Harris, H. D., Murphy, G. L., & Rehder, B. (2008). Prior knowledge and exemplar frequency. *Memory & Cognition*, *36*(7), 1335-1350. https://doi.org/10.3758/MC.36.7.1335

Jenkins, G. W., Samuelson, L. K., Smith, J. R., & Spencer, J. P. (2015). Non-Bayesian Noun Generalization in 3- to 5-Year-Old Children : Probing the Role of Prior Knowledge in the Suspicious Coincidence Effect. *Cognitive Science*, *39*(2), 268-306. https://doi.org/10.1111/cogs.12135

Kan, P. F., & Windsor, J. (2010). Word Learning in Children With Primary Language Impairment : A Meta-Analysis. *Journal of Speech Language and Hearing Research*, *53*(3), 739. https://doi.org/10.1044/1092-4388(2009/08-0248)

Kennedy, J. E. (2015). Beware of Inferential Errors and Low Power With Bayesian Analyses : Power Analysis is Needed for Confirmatory Research. *The Journal of Parapsychology*, 79, 53-64.

Kim, S., & Rehder, B. (2011). How prior knowledge affects selective attention during category learning : An eyetracking study. *Memory & Cognition*, *39*(4), 649-665. https://doi.org/10.3758/s13421-010-0050-3

Khomsi, A. (2001). *ELO: Evaluation du langage oral.* Paris: ECPS.

Kruschke, J. K. (2010). *Doing Bayesian Data Analysis : A Tutorial with R and BUGS* (Academic Press). Academic Press.

Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian New Statistics : Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review*, *25*(1), 178-206. https://doi.org/10.3758/s13423-016-1221-4

Krzemien, M., Thibaut, J.-P., Jemel, B., Levaux, E., & Maillart, C. (2021). How do children with developmental language disorder extend novel nouns? *Journal of Experimental Child Psychology*, *202*, 105010. https://doi.org/10.1016/j.jecp.2020.105010

Kucker, S. C., Samuelson, L. K., Perry, L. K., Yoshida, H., Colunga, E., Lorenz, M. G., & Smith, L. B. (2019). Reproducibility and a unifying explanation : Lessons from the shape bias. *Infant Behavior and Development*, *54*, 156-165. https://doi.org/10.1016/j.infbeh.2018.09.011

Lecocq, P. (1996). *L'ECOSSE: Une épreuve de compréhension syntaxico-sémantique.* Villeneuve-d'Ascq: Presses Universitaires du Septentrion.

Lee, M. D. & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q. F., Smíra, M., Epskamp, S., Matzke, D., Wild, A., Knight, P., Rouder, J. N., Morey, R. D., & Wagenmakers, E.-J. (2019). **JASP** : Graphical Statistical Software for Common Statistical Designs. *Journal of Statistical Software*, *88*(2). https://doi.org/10.18637/jss.v088.i02

MacRoy-Higgins, M., Shafer, V. L., Fahey, K. J., & Kaden, E. R. (2016). Vocabulary of Toddlers Who Are Late Talkers. *Journal of Early Intervention*, *38*(2), 118-129. https://doi.org/10.1177/1053815116637620

Markman, E. M. (1990). Constraints Children Place on Word Meanings. *Cognitive Science*, *14*(1), 57-77. https://doi.org/10.1207/s15516709cog1401_4

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame : An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314-324. https://doi.org/10.3758/s13428-011-0168-7

McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic Representation and Naming in Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *45*(5), 998-1014. https://doi.org/10.1044/1092-4388(2002/081)

McGregor, K. K., Oleson, J., Bahnsen, A., & Duff, D. (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth : *Vocabulary deficits*. *International Journal of Language & Communication Disorders*, *48*(3), 307-319. https://doi.org/10.1111/1460-6984.12008

Obeid, R., Brooks, P. J., Powers, K. L., Gillespie-Lynch, K., & Lum, J. A. G. (2016). Statistical Learning in Specific Language Impairment and Autism Spectrum Disorder : A Meta-Analysis. *Frontiers in Psychology*, *7*, Article 1245. https://doi.org/10.3389/fpsyg.2016.01245

Perfors, A., Tenenbaum, J. B., Griffiths, T. L., & Xu, F. (2011). A tutorial introduction to Bayesian models of cognitive development. *Cognition*, *120*(3), 302-321. https://doi.org/10.1016/j.cognition.2010.11.015

Sandgren, O., Salameh, E.-K., Nettelbladt, U., Dahlgren-Sandberg, A., & Andersson, K. (2021). Using a word association task to investigate semantic depth in swedish-speaking children with developmental language disorder. *Logopedics Phoniatrics Vocology*, *46*(3), 134-140. https://doi.org/10.1080/14015439.2020.1785001

Sheng, L., & McGregor, K. K. (2010a). Lexical–Semantic Organization in Children With Specific Language Impairment. *Journal of Speech Language and Hearing Research*, *53*(1), 146. https://doi.org/10.1044/1092-4388(2009/08-0160)

Sheng, L., & McGregor, K. K. (2010b). Object and Action Naming in Children With Specific Language Impairment. *Journal of Speech Language and Hearing Research*, *53*(6), 1704. https://doi.org/10.1044/1092-4388(2010/09-0180)

Sloutsky, V. M., Deng, W., Fisher, A. V., & Kloos, H. (2015). Conceptual influences on induction : A case for a late onset. *Cognitive Psychology*, *82*, 1-31. https://doi.org/10.1016/j.cogpsych.2015.08.005

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, *10*(7), 309-318. https://doi.org/10.1016/j.tics.2006.05.009

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I : Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35-57. https://doi.org/10.3758/s13423-017-1343-3

Waxman, S. R. & Lidz J L. (2006). Early word learning. In D. Kuhn, R. S. Siegler, W. Damon, & R. M. Lerner (Eds.), *Handbook of Child Psychology* (6th ed., pp. 299-335). John Wiley & Sons Inc.

Wechsler, D., & Naglieri, J. (2009). *Echelle non verbale d'intelligence de Wechsler*. Montreuil: ECPA.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*(2), 245-272. https://doi.org/10.1037/0033-295X.114.2.245

Figure 1. Example of stimuli.

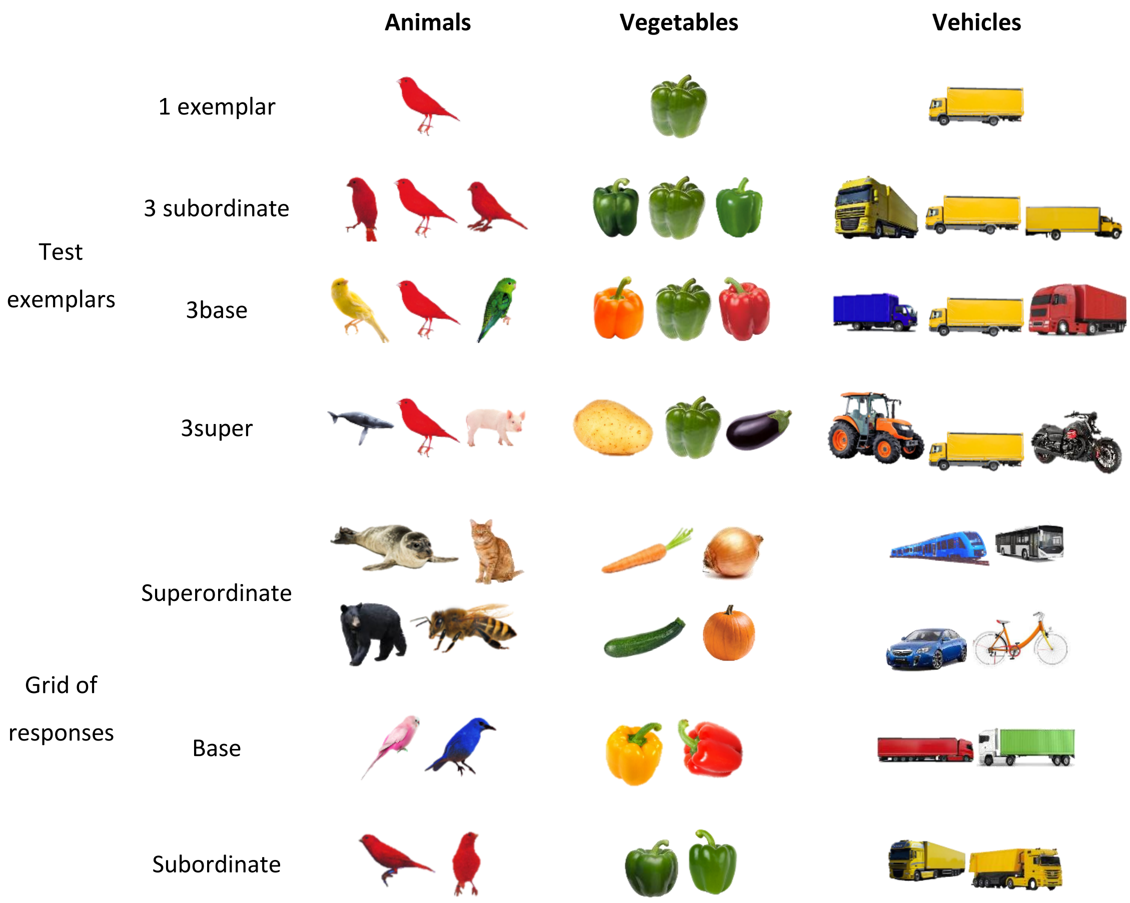
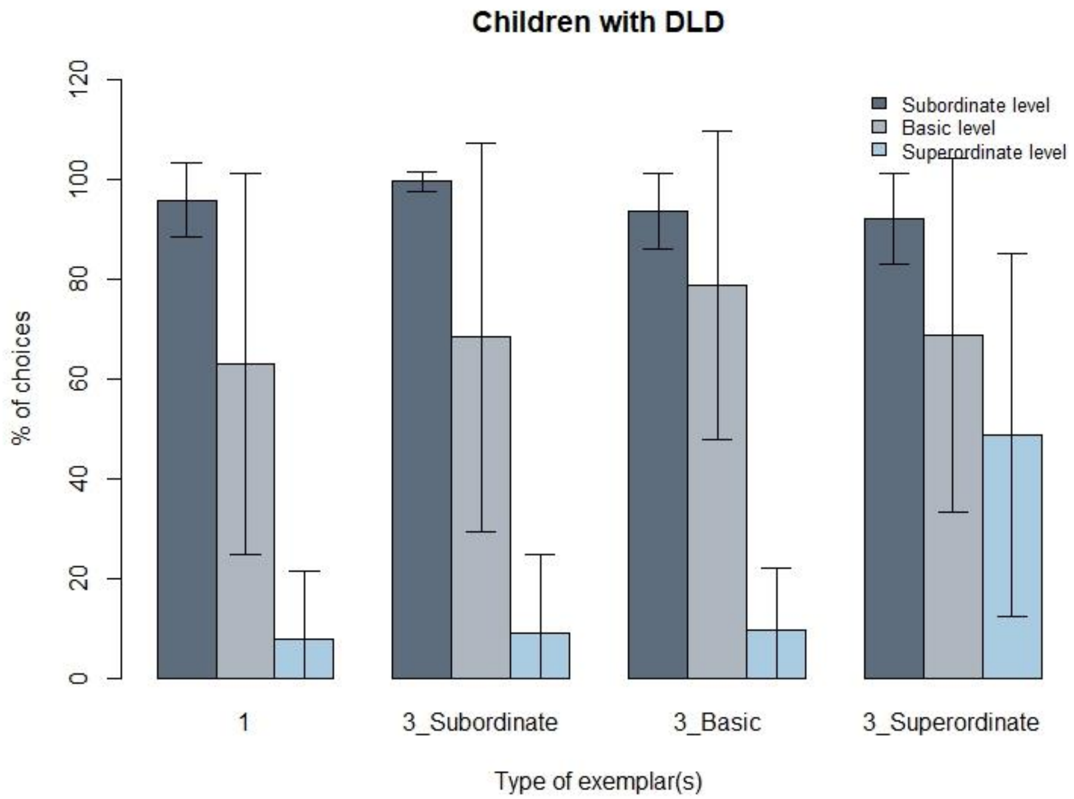
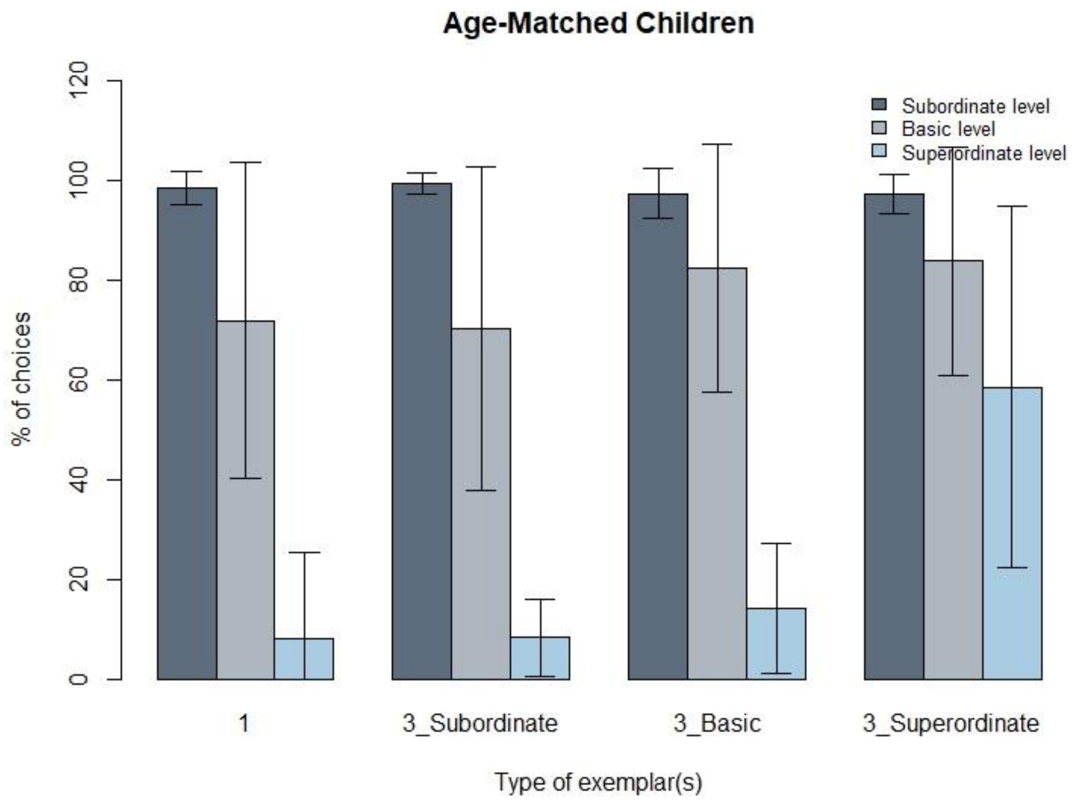


Figure 2. Pattern of responses for each group, each type of exemplar(s) and each level of taxonomy (in %).



a.



b.

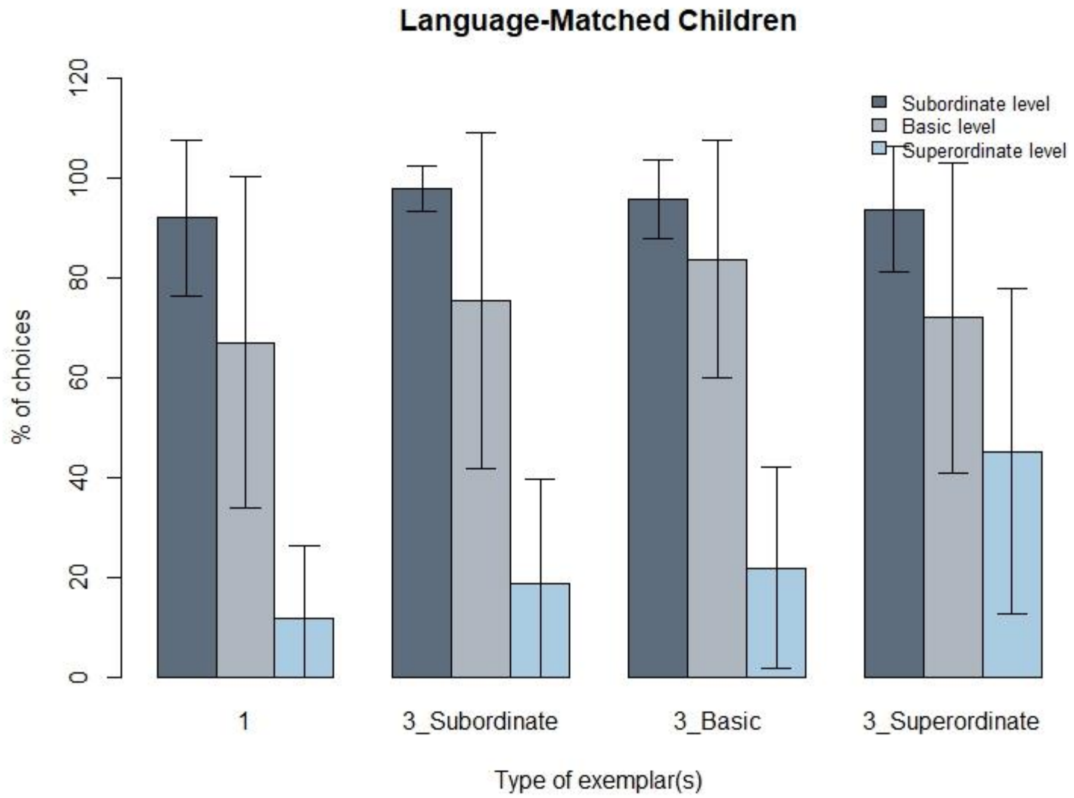

Language-Matched Children
Subordinate level
Basic level
Superordinate level
% of choices
0
20
40
60
80
100
120
1
3_Subordinate
3_Basic
3_Superordinate
Type of exemplar(s)
c.

Table 1. Characteristics of the three groups

	DLD	**Age-matched**		**Language-matched**	
	n=20	n=16		n=20	
Age Year;Month (mean)	9;7 (sd. 22.6 months)	9;3 (sd. 17.8 months)		7;3 (sd. 19.7 months)	
NVIQ Mean(SD)	96.75(13.21)	97.56(10.28)	*t*(34)=0.2	100.3(10.55)	*t*(38)=0.94
Word repetition Raw-score(SD)	21.1(8.23)	31.69(1.01)	*t*(34)=5.1***	29.55(3.07)	*t*(38)=4.3***
Max. score	32	32		32	
Designation (lexical) Raw-score(SD)	82.5(19.33)	115.81(19.25)	*t*(34)=5.15***	83.3(20.86)	*t*(38)=0.13
Norm. score (SD)	86.55(12.17)	118.62(10.47)		111.4(14.36)	
Denomination task Raw-score(SD)	29.45(5.09)	36.63(4.62)	*t*(34)=4.38***	31.11(3.4)	*t*(37)=1.19
Max. score	50	50		50	
Designation (Syntax) Error-score(SD)	17(7.73)	7.06(3.98)	*t*(34)=-4.68***	15(7.35)	*t*(37)=-0.83
Sentence completion Raw-score(SD)	10.85(5.03)	20.81(3.12)	*t*(34)=6.92***	16.53(3.76)	*t*(37)=3.98***
Max. score	25	25		25	

Note. DLD and age-matched participants were evenly distributed across the age ranges; language-matched participants were mainly between 5 and 7 year-old. The means and SDs used for each test depend on the chronological age or grade level of the children (see test manuals for more precisions). Max. score refers to the maximal score for a test and was indicated when relevant. NVIQ = non verbal IQ, SD = standard deviation, Norm.Score = normalized score, Error-score = number of errors, *** p <.001

Table 2. Percentage of choices for each type of item and level of taxonomy (mean (s.d.), in %), in each group.

		DLD	AM	LM	All
		n=20	n=16	n=20	n=56
1	**Subordinate**	95.833 (7.404)	98.44 (3.36)	92.08 (15.64)	95.24 (10.64)
	Base	62.91 (38.19)	71.875 (31.75)	67.08 (33.17)	66.96 (34.23)
	Superordinate	7.71 (13.60)	8.07 (17.38)	11.67 (14.72)	9.23 (14.90)
3_subordinate	**Subordinate**	99.58 (1.86)	99.48 (2.08)	97.92 (4.58)	98.96 (3.20)
	Base	68.33 (38.96)	70.31 (32.49)	75.42 (33.60)	71.43 (34.81)
	Superordinate	8.96 (15.72)	8.33 (7.76)	18.75 (20.88)	12.28 (16.62)
3_base	**Subordinate**	93.75 (7.59)	97.40 (5.02)	95.833 (7.88)	95.54 (7.10)
	Base	78.75 (30.88)	82.29 (24.88)	83.75 (23.80)	81.55 (26.43)
	Superordinate	9.58 (12.40)	14.06 (12.99)	21.875 (20.27)	15.25 (16.41)
3_superordinate	**Subordinate**	92.08 (9.16)	97.40 (3.99)	93.75 (12.65)	94.20 (9.66)
	Base	68.75 (35.55)	83.85 (22.87)	72.08 (31.10)	74.26 (30.87)
	Superordinate	48.75 (36.30)	58.59 (36.24)	45.21 (32.73)	50.30 (34.84)

Note. DLD = Developmental Language Disorder, AM =age-matched, LM = language-matched

Table 3. Analysis of effects of the Bayesian repeated measures ANOVA

Effects	P(incl)	P(incl|data)	BF_{excl}	BF_{incl}
Type	0.26	$1.15e^{-21}$	$2.07e^{-9}$	**$4.82e^{+8}$**
Level	0.26	$1.15e^{-21}$	$1.63e^{-168}$	**$6.15e^{+167}$**
Group	0.26	0.13	**6.52**	0.15
Type ✻ Level	0.26	1.00	$1.15e^{-21}$	**$8.69e^{+20}$**
Group ✻ Type	0.26	0.01	11.65	0.09
Group ✻ Level	0.26	0.00	37.35	0.03
Group ✻ Type ✻ Level	0.05	$2.25e^{-6}$	139.14	0.01

Note. Type = Type of item; Level = Level of taxonomy; $BF_{excl/incl}$ refers to the evidence against/for a main effect of the factor when excluded from/included in the explicative models.

Table 4. Results of the Bayesian repeated measures ANOVA

Models	**P(M)**	**P(M|data)**	**BF_M**	**BF_{10}**	**error %**
Null model (incl. subject)	0.20	$2.70e^{-132}$	$1.08e^{-131}$	1.00	
Type + Level + Type ✻ Level	0.20	1.00	$1.39e^{+19}$	$3.71e^{+131}$	1.25
Type + Level	0.20	$2.88e^{-19}$	$1.15e^{-18}$	$1.07e^{+113}$	1.22
Level	0.20	$1.36e^{-23}$	$5.43e^{-23}$	$5.03e^{+108}$	0.67
Type	0.20	$5.07e^{-132}$	$2.03e^{-131}$	1.88	0.65

Note. Type = Type of item; Level = Level of taxonomy; P(M) = prior probability for each model; P(M|Data) = posterior probability for each model; BF_M = Change from prior to posterior model odds; BF_{10} = Bayes Factor for each model compared to null model; error % = estimation error of the BFs.