CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Morphosyntactic weaknesses in Developmental Language Disorder: the role of structure and agreement configurations

Vincenzo MOSCATI[1]*, Luigi RIZZI[1,2], Ilenia VOTTARI[1], Anna Maria CHILOSI[3], Renata SALVADORINI[3], and Maria Teresa GUASTI[4]

[1]University of Siena, Italy, [2]University of Geneva, Switzerland, [3]IRCCS Stella Maris, Italy, and [4]University of Milano-Bicocca, Italy
*Corresponding author. DISPOC – Department of Social, Political and Cognitive Science, Complesso S. Niccolò, Via Roma 56–53100 Siena (Italy). E-mail: moscati.v@gmail.com

Abstract

Agreement is a morphosyntactic dependency which is sensitive to the hierarchical structure of the clause and is constrained by the structural distance that separates the elements involved in this relation. In this paper we present two experiments, providing new evidence that Italian-speaking children with Developmental Language Disorder (DLD), as well as Typically Developing (TD) children, are sensitive to the same hierarchical and locality factors that characterise agreement in adult grammars. This sensitivity holds even though DLD children show accrued difficulties in more complex agreement configurations. In the first experiment, a forced-choice task was used to establish whether children are more affected in the computation of S-V agreement when an element intervenes hierarchically or linearly in the agreement relation: DLD children are more subject to attraction errors when the attractor intervenes hierarchically, indicating that DLD children discriminate between hierarchical and linear configurations. The second experiment, also conducted through a forced-choice task, shows that the computation of agreement in DLD children is more 'fragile' than in TD children (and also in children with a primary impairment in the phonological domain), in that it is more sensitive to the factors of complexity identified in Moscati and Rizzi's (2014) typology of agreement configurations. To capture the agreement pattern found in DLD children, we put forth a novel hypothesis: the FRAGILE COMPUTATION OF AGREEMENT HYPOTHESIS. Its main tenet is that DLD children make use of the same grammatical operations employed by their peers, as demonstrated in Experiment 1, but difficulties increase as a function of the complexity of the agreement configuration.

CrossMark

Introduction

Developmental Language Disorder (DLD) is defined as a clinical condition that identifies a group of children whose poor language abilities create obstacles to communication and learning in everyday life (see Bishop, Snowling, Thompson, Greenhalgh, & the CATALISE-2 Consortium, 2017, on the choice of this term rather than Specific Language Impairment). These problems are unlikely to resolve spontaneously and are not associated with other known biomedical conditions. Although several researchers have shown that children with DLD have problems in non-linguistic areas like attention (Cohen, Vallance, & Barwick, 2000), memory (Archibald & Gathercole, 2006), and motor activities (Hill, 2001), language still stands out as the major and more affected capacity. The primary domain of impairment may reside in various areas of language, and an accurate linguistic description of the disorder is fundamental for diagnostic accuracy, in the absence of any reliable neurobiological signature.

In this paper, we will mainly focus on Italian, a language whose rich morphology could be exploited to shed light on DLD children's morphosyntactic deficit in general. A look at previous studies on Italian shows that verbal morphology, determiners, and clitic pronouns have all been found to be problematic. In particular, inflectional morphemes participating in agreement relations can be specifically affected. Similar problems have been attested in DLD children speaking other languages as well, suggesting that the agreement relation may be a locus of specific difficulty. Among the different types of agreement relations, problems with verbal agreement morphology are perhaps the best documented, showing a certain variability across languages. At one extreme, we find English, where tense and/or agreement are highly problematic for DLD children, and their incorrect expression is considered to be a hallmark of the language disorder (a.o. Clahsen, Bartke, & Göllner, 1997; Conti-Ramsden, Botting, & Faragher, 2001; Hoover, Storkel, & Rice, 2012; Oetting & Horohov, 1997; Rice & Wexler, 1996). Similar problems have also been well documented in other Germanic languages, with differences across persons of the verbal paradigms (for Dutch: Blom, Vasić, & de Jong, 2014; de Jong, 1999; Duinmeijer, 2016; for Swedish: Hansson, Nettelbladt, & Leonard, 2000; for German: Ott & Höhle 2013; Rice, Noll, & Grimm, 1997). Agreement impairments have also been documented in Hebrew-speaking children with DLD (Dromi, Leonard, Adam, & Zadunaisky-Ehrlich, 1999) and for third person singular morphemes in Arabic (Abdalla & Crago, 2008). In Romance languages, most of the documented problems of DLD children concern auxiliary omission (for Italian: Leonard, Bortolini, Caselli, McGregor, & Sabbadini, 1992; for French: Paradis & Crago, 2001). Substitution errors have also been reported (see Bortolini, Caselli, & Leonard, 1997, for the third person plural morpheme in the present indicative). Difficulties with subject–verb agreement have also been reported using grammaticality judgements in Dutch- and Italian-speaking children with DLD (Cantiani, Lorusso, Perego, Molteni, & Guasti, 2015; Rispens & Been, 2007). This brief review shows that in many languages, to a variable extent, verbal agreement seems to be one of the weaknesses of DLD children.

Agreement in the nominal domain has received less attention in the literature, but substitution errors have occasionally been reported. In Dutch, DLD children seem to have problems with neuter nouns in establishing the correct gender agreement with the determiner (Duinmeijer, 2016; Orgassa & Weerman, 2008).

In French, Roulet-Amiot and Jakubowicz (2006) found that DLD children committed gender errors on determiners and adjectives. Although substitutions have been occasionally observed (see also special cases of allomorphy, e.g., *la/el* substitutions in Spanish **la agua* instead of *el agua* (Restrepo & Gutiérrez-Clellen, 2001), or *lo/il* substitutions in Italian **il zaino* instead of *lo zaino* (Caselli, Leonard, Volterra, & Campagnoli, 1993)), omissions are the most common error type found in DLD children across languages. This has been documented in several studies on English (McGregor & Leonard, 1994; Polite, Leonard, & Roberts, 2011; Rice & Wexler, 1996), German (Clahsen, 1991; Eisenbeiss, Bartke, & Clahsen, 2005), Italian (Bortolini *et al.*, 1997; Bottari, Cipriani, Chilosi, & Pfanner, 2001), Spanish (Bedore & Leonard, 2005; Restrepo & Gutiérrez-Clellan, 2001), French (Roulet 2007; Royle & Stine, 2013), and Greek (Chondrogianni, Marinis, Edwards, & Blom, 2015; Tsimpli & Stavrakaki, 1999).

Finally, turning to the agreement on past participles triggered by a direct object clitic, no specific study exists. Attention has been devoted to the ingredients of this relation separately: either past participles or clitics. Past participles in the DLD population have been studied in German, establishing a lack of impairment in the use of suffixes and prefixes (e.g., *ge-brauch-t*; Rothweiler & Clahsen, 1994). However, it is important to point out that German participles do not display any agreement morphology. Several studies have looked at clitics in Romance languages, some of which have compared the omission of clitics with the omission of articles since these forms are often homophonous. A sharp difference emerges, with clitic omissions being attested to a greater extent and for a longer period than determiner omissions. Italian children with DLD still omit clitics at the age of five years (Bortolini, Arfé, Caselli, Degasperi, Deevy, & Leonard, 2006; Guasti *et al.*, 2016), replacing them with the corresponding noun phrases at the age of seven years (Arosio, Branchini, Barbieri, & Guasti, 2014; Guasti *et al.*, 2016). Moreover, sometimes, five-year-old children with DLD also use an incorrect gender form of the clitic (Leonard & Dispaldro, 2013). Similar problems have also been documented in Spanish (Bedore & Leonard, 2001), Greek (Stavrakaki & van der Lely, 2010; Tsimpli, 2001) and French (Grüter, 2005; Hamann *et al.*, 2003; Jakubowicz, Nash, Rigaut, & Gérard, 1998; Tuller, Delage, Monjauze, Piller, & Barthez, 2011). Concerning French, it is worth mentioning that, when production of clitics is compared with production of the homophonous articles, direct object clitics are considerably more impaired than articles (Jakubowicz *et al.*, 1998).

This overview shows that the morphosyntactic abilities of children with DLD may be weak with determiners, verbal inflections, and clitic pronouns. Going back to the specific case of Italian, all these three elements enter into agreement relations, and they could pose problems with a different degree of severity for DLD children. Despite the efforts that have been made to capture DLD children's morphosyntactic difficulties, when it comes to agreement many important issues are still left untouched. In our view, at least three issues need to be considered. The first is whether children who present a morphosyntactic impairment compute agreement relying on the same grammatical operations and structural configurations adopted by TD children, i.e., whether they compute agreement on a hierarchical representation. The second issue concerns the fact that most of the existing studies focused on single agreement relations, without a close comparison between different configurations in the same population. The

third issue is that the available evidence mostly comes from production data, making it hard to disentangle phonological production deficits from (morpho) syntactic ones.

To at least partially overcome these limitations, we would like to provide new evidence on the source of agreement errors. We will first look at DLD children's hierarchical computation of Subject–Verb agreement. Then, we will compare three agreement relations: Subject–Verb (S-V), Determiner–Noun (Det-N), and Clitic–Past-Participle (Cl-PPart) agreement.

The paper is organised as follows. Initially, we will discuss agreement dependencies in structural terms considering existing psycholinguistic evidence mostly coming from the study of S-V ATTRACTION ERRORS in adults. Then, we will extend the discussion to agreement configurations other than S-V, namely Det-N and Cl-PPart agreement, presenting the ranked typology of agreement configurations proposed in Moscati and Rizzi (2014) and extending it to formulate a new hypothesis, the FRAGILE COMPUTATION OF AGREEMENT HYPOTHESIS (FCAH). Finally, we will present the results of two new experiments in which we compare TD and DLD children's performance using a forced-choice task.

The hierarchical computation of agreement and attraction errors

Grammatical agreement in natural languages is computed on local configurations, e.g., a verb typically agrees with a nominal in the same simple clause, not with an element in a subordinate or superordinate clause. Moreover, locality is computed in hierarchical, not in linear terms. In a structure as in (1), the verb (here the copula) agrees with the structurally local head noun of the subject DP (*picture*), not with the linearly adjacent but structurally more distant noun *trees*.

(1) The picture of the trees is on the table.

Elicited production experiments have shown that the production system may sometimes erroneously perform agreement with the linearly closer noun, i.e., this noun determines an 'attraction error' (see Bock & Cutting, 1992; Franck, Vigliocco, & Nicol, 2002; Pearlmutter, Garnsey, & Bock, 1999; Wagers, Lau, & Phillips, 2009, among many other references) as in the following sentence:
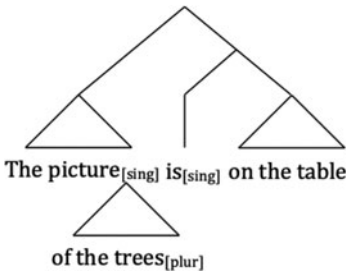
(2) *The picture of the trees are on the table.

However, the detailed study of attraction errors in such environments has shown that the adult production system is strongly sensitive to hierarchical factors. First, the overwhelming majority of the elicited sentences shows correct agreement with the structurally closest element, attraction errors being a minute fraction of the produced utterances (roughly 5% in this configuration, according to Franck, Lassi, Frauenfelder, & Rizzi, 2004). Second, production errors are more frequent in configurations of structural rather than linear intervention, as in (2). For instance, Franck, Lassi, Frauenfelder, and Rizzi (2006) have shown that object clitics in French, structurally intervening between the subject and the inflected verb (3), determine about twice as many errors as purely linear intervening elements such as the adnominal complement in the French equivalent of (2) (see also Franck, Soare, Frauenfelder, & Rizzi, 2010):
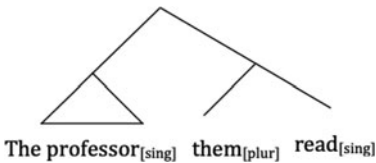
(3) *Le professeur$_{SING}$ les$_{PLUR}$ lisent$_{PLUR}$
    (instead of: le professeur$_{SING}$ les$_{PLUR}$ lit$_{SING}$)
    'The professor them read'

The conclusion that can be drawn from previous studies on adults' production errors is that a STRUCTURAL TYPE OF INTERVENTION, exemplified by sentences like (3) and corresponding to the representation in (3′) (English morphemes are used for simplicity), triggers more attraction errors than a LINEAR TYPE OF INTERVENTION as in (2) and represented in (2′).

(2′) Linear intervention

The picture$_{[sing]}$ is$_{[sing]}$ on the table

of the trees$_{[plur]}$

(3′) Structural intervention

The professor$_{[sing]}$ them$_{[plur]}$ read$_{[sing]}$

This asymmetry in the rate of attraction errors can thus be considered a hallmark of the hierarchical computation of agreement relations. Based on these insights, we propose to exploit attraction errors to investigate the nature of the operation used to realize agreement by children with DLD (and TD controls). Franck *et al.* (2004) attempted to elicit attraction errors in children with and without language impairment in a linear intervention configuration. They found that typically developing children are sensitive to the feature of the intervening element, while children with DLD were not. However, Franck *et al.* used an elicited production task that could have shadowed intervention effects in those DLD children with general production problems. Therefore, we decided to test attraction effects with a receptive task, rather than a productive task. If DLD children compute (or try to compute) agreement employing the same hierarchical operation as adults, they will show a higher rate of attraction errors in configurations like (3′) compared to those like (2′). If instead, they resort to a qualitatively different operation, e.g., relying on the mere linear proximity between the subject and its verb, no difference between the error rate in the linear and structural configurations should be found. In fact, in this case, one may even expect that linear configurations should be more prone to attraction errors since a longer chunk separates the verb and its subject (i.e., *of the trees* in (2′)) than in cases of structural configurations (where a clitic pronoun is intervening, as in (3′)).

We now turn to the second issue, which also relates to the structural computation of agreement. Once agreement is seen as the result of a computation based on a hierarchical structure, a metric of syntactic complexity that allows us to compare agreement in various configurations can be proposed. In the next section, we will introduce a typology of agreement configurations and a connected complexity metric.

Locality on agreement

Locality on agreement can be characterised in terms of at least two different factors: the first is the structural distance that separates the elements with matching features, the second is the number of the elementary operations which are required. The most local form of agreement is the one in which the two agreeing elements are 'merged' together (i.e., combine in a syntactic structure) and end up sharing the relevant morphological features. In such cases of extreme locality, agreement may be just a reflex of the merge operation. One example is Det-N agreement in number and gender in Italian (as well as other kinds of DP-internal agreement), as in (4):

(4) le   ragazz-e   (Det-N)
     DET.PL.F   girl-PL.F
     'the girls'

The second type of agreement is S-V agreement in person and number, illustrated in (5). The creation of the agreement configuration, in this case, involves the movement of the subject from its thematic position to the Specifier of a functional projection hosting the inflected verb.[1]

(5) le   ragazz-e   mangiano   (S-V)
     DET.PL.F   girl-PL.F eat.PRS.3.PL
     'the girls eat'

The third case is the Cl–PPart agreement in (6):

(6) Il   ragazz-o   le   ha   mangiate   (Cl-PPart)
     DET.S.M boy-S.M   OBJ.PL.F   AUX3.SG   eat.PTCP.PL.F
     'the boy has eaten them'

In this configuration, agreement is checked 'in passing' (Belletti, 2006; Kayne, 1989). The clitic moves from the object position to the specifier of a position hosting the inflected past participle. There, it triggers agreement on the past participle and proceeds to its final destination, the clitic position in the functional structure of the clause. Agreement in (5) and (6) occurs in a Spec-Head configuration and requires

---

[1]Guasti and Rizzi (2002), revising the standard minimalist analysis of agreement (Chomsky, 2000), propose that Subject–Verb agreement involves two operations: the first is called Agree and connects the functional head hosting the verb endowed with the relevant morphosyntactic features and the subject in its vP-internal thematic position; the second is Spec-Head agreement, obtaining after the subject has moved to the Spec of the functional head hosting the verb. Guasti and Rizzi (2002) provided evidence that the Spec-Head configuration plays a crucial role in the computation of agreement. In what follows, in order to simplify the discussion, we will focus on the Spec-Head agreement step.

movement, but at the surface level the two cases crucially differ. While the Subject–Verb configuration remains local in the surface representation, the Clitic–Past Participle configuration is disrupted by a successive derivational step (movement of the object pronoun to the clitic position adjacent to the inflected verb), so that the two agreeing elements are obligatorily separated (at least, by the auxiliary). Based on these considerations, Moscati and Rizzi (2014) have proposed a metric of complexity that can distinguish among the three configurations in terms of the operations involved:

(7) A metric of local complexity, based on Italian
    I.   Det – N   (merge, agreement)
    II.   S – V   (merge, movement, agreement)
    III. Cl – PPart   (merge, movement, agreement, further movement)

The crucial role attributed to movement in determining a scale of complexity in agreement configurations echoes the classical Derivational Theory of Complexity (see Fodor, Bever, & Garrett, 1974, for critical discussion), and converges with the complexity metric introduced for wh-movement in Jakubowicz (2005, 2011).

    A previous study on TD children, reported in Moscati and Rizzi (2014), has shown that this metric of complexity accounts for the fact that the three different configurations are fully mastered at different ages. The more local configuration in (7 I) is the easiest, while the one in (7 III) is the most difficult. If children with DLD follow the same path of development as TD children, but are delayed (Rice & Wexler, 1996), we expect this metric to characterise their language competence.

A new hypothesis: the FRAGILE COMPUTATION OF AGREEMENT

Grammatical complexity and previous accounts on agreement deficits in Developmental Language Disorder

One of the first accounts explicitly casting DLD children's difficulties in the syntactic feature-checking mechanism is Rice and Wexler's (1996) EXTENDED UNIQUE CHECKING CONSTRAINT. It was originally proposed to account for children's omissions of tense or agreement morphemes in English, resulting in the production of bare stems (e.g., *he play <– he plays, he played*). However, this account cannot be straightforwardly extended to other languages in which a dissociation between Tense and Agreement morphology is evident in the speech of DLD children. Looking at Greek, Tsimpli (2001) showed that while DLD children had problems with agreement, tense morphology appeared to be unaffected. Capitalising on this difference, Tsimpli proposed a fundamental distinction based on FEATURE INTERPRETABILITY, where only uninterpretable features in the sense of Chomsky (1995) pose a real challenge for DLD children. Without entering into the details, for which we refer to the original reference, Tsimpli's (2001) observation is particularly relevant since uninterpretable features (e.g., the expression of number on verbs in Subject-V or Object-V agreement) are exactly the ones that must enter into agreement relations.

    The idea that only a subset of grammatical features, the uninterpretable ones, is impaired in DLD children has also been put forth in Clahsen's AGREEMENT DEFICIT ACCOUNT, first proposed in Clahsen *et al.* (1997) and later reformulated in Clahsen (2008). In our understanding, the original formulation is essentially 'procedural' in the sense that children's difficulties lie in their poor capacity to establish a specific

grammatical relation. In later work, this vision has been substantially modified, and in a more recent paper (Clahsen, 2008) the source of the impairment is considered to lie in the children's poor mastery of the full agreement paradigms. This second formulation is a 'lexical' one, in which grammatical features of individual forms may be left underspecified, leading to substitution errors. This change of perspective from a PROCEDURAL to a LEXICAL type of impairment was necessary to account for the selective nature of children's errors, confined to some verbal morphemes but not found with others (see Clahsen & Dalalakis, 1999; Tsimpli & Stavrakaki, 1999, for Greek; Leonard *et al.*, 1992, for Italian; de Jong, 1999; Duinmeijer, 2016, for Dutch).

Tsimpli's and Clahsen's proposals introduced the pivotal difference between morphemes that are an expression of a dependency between two elements, and morphemes that are not. We believe that this distinction is fundamental, but also that at least one further dimension of analysis needs to be considered, that is, the syntactic distance that separates the two elements. This idea is not new, and some theories already incorporate it, but in relation to another set of phenomena related to movement-derived sentences, such as passives (Marinis & Saddy, 2013; van der Lely, 1996), relative clauses (Friedmann & Novogrodsky, 2011), and questions (Deevy & Leonard, 2004; Friedmann & Novogrodsky, 2007; Hamann, 2006; Prévost, Tuller, Barthez, Malvy, & Bonnet-Brilhault, 2017; Stavrakaki, 2006).

One such theory is the REPRESENTATIONAL DEFICIT FOR DEPENDENT RELATIONSHIPS (RDDR; van der Lely, 1998), more recently reformulated in the COMPUTATIONAL GRAMMATICAL COMPLEXITY hypothesis (van der Lely, 2005). A second one (see Marinis, 2011, for a comparison), in the same vein and formulated with an explicit definition of complexity, is the COMPUTATIONAL COMPLEXITY HYPOTHESIS (CCH; Jakubowicz, 2011; see also Prevost *et al.*, 2017), which offers a clear metric based on the number of movement operations. Originally formulated to account for DLD children's difficulties with wh-questions, the CCH shares with Moscati and Rizzi (2014) the idea that movement increases sentence complexity. A novel proposal along these lines will be presented in the next section. Finally, we wish to conclude this review by considering Leonard's (2014) SURFACE ACCOUNT, which was formulated to capture children's difficulties with verbal paradigms. According to Leonard, DLD children's incorrect realisation of agreement morphemes depends on their phonological status: children with DLD are expected to have trouble only with those agreement morphemes that are non-syllabic or phonologically weak. More specifically, children with DLD are expected to have difficulties with unaccented and unfooted syllables. This difficulty is observed in the case of the Italian third person plural morpheme in the present tense, which is replaced by the third person singular. Consider the third person verb *dormono* (sleep-3PL) and its metrical structure formed by a strong–weak–weak syllable sequence. Children are unable to produce the second weak unfooted syllable (-*no*), and they revert to the third person singular form of the verb *dorme* (sleep-3SG), with the metrical structure strong–weak, a trochaic foot. Thus, in this approach, children have difficulties with the production of weak syllables that cannot form a trochaic foot with the preceding strong syllable.

The FRAGILE COMPUTATION OF AGREEMENT HYPOTHESIS

With Clahsen *et al.* (1997) and Tsimpli (2001), we share the idea that one source of the problem lies in the difficulties that some DLD children may have with the grammatical mechanism responsible for sharing a set of grammatical features. However, our

hypothesis aims at covering different agreement configurations other than S-V, placing them along a scale of derivational complexity. We will look at the typology of configurations introduced in (7), which ranks the agreement configurations in terms of the more or less stringent locality conditions that they meet, thus providing a natural gradation of complexity. On the basis of these assumptions, we propose that the computation of agreement is progressively more fragile as the complexity of the configuration increases. Globally, our approach can be spelt out as follows:

(8) i.   DLD children compute agreement hierarchically, much as TD children do;
     ii. DLD children's mastery of agreement is modulated by the locality of the agreement configuration, with less local configurations being more fragile.

The formulation in (8) casts agreement in DLD children in the same mould as in TD children and adults: DLD children compute agreement in terms of a hierarchical structure and not in terms of a purely linear configuration: (8i). This is a point that previous accounts have assumed but never demonstrated. We will address it by asking the following questions: Do DLD children that show a morphosyntactic type of impairment uses a purely linear operation to compute agreement or are they sensitive to the hierarchical configuration expressed in terms of c-command? To answer these questions, we exploit a paradigm which, to our knowledge, was never used with DLD children, in order to establish whether DLD children are also sensitive to well-documented ATTRACTION ERRORS. We aim to determine whether their sensitivity to attraction errors is modulated by the configurations of intervention – structural vs. linear – as documented in the adult literature (Franck *et al.*, 2002; Franck *et al.*, 2006; Franck *et al.*, 2010).

As for (8ii), we expect that, if tightly compared, the three agreement configurations Det-N, S-V, and Cl-PPart will impact differently on the linguistic performance of DLD children. To make point (8ii) operative, we assume the same working notion of locality-based complexity proposed in Moscati and Rizzi (2014) and reported in (7). From (7), the prediction can be derived that S-V and Cl-PPast are more fragile and prone to errors than Det-N; moreover, the Cl-PPart is expected to be the most fragile configuration among the three, as it involves an extra movement step.

A new set of predictions

The set of predictions of the FRAGILE COMPUTATION OF AGREEMENT HYPOTHESIS differ in important respects from the ones of previous accounts. We will briefly consider them in relation to Clahsen's AGREEMENT DEFICIT ACCOUNT and Leonard's SURFACE ACCOUNT.

Let us first consider the AGREEMENT DEFICIT ACCOUNT in its 'procedural view'. If not supplemented by auxiliary hypotheses, it predicts the same performance across conditions, since the checking of uninterpretable features is involved in all three configurations. The 'lexical' view, instead, capitalises on the paradigms' size. In Italian, the S-V agreement paradigm is more articulated than the past participle or the determiner agreement paradigm (see Caprin & Guasti, 2009). If the paradigm's size matters, we would expect Det-N and Cl-PPart (I and III) to pattern together, and differ[2] from S-V agreement (II).

As for the SURFACE ACCOUNT, S-V agreement should be problematic for children with DLD only when the verb is plural, because only the final syllable of plural verbs is

---

[2]We thank an anonymous reviewer for commenting on this point.

unfooted. Det-N agreement should also be problematic, given that determiners are weak elements that precede the noun and remain unfooted. Similarly, Cl-PPart agreement should also be challenging as clitics are weak elements that precede the verb and remain unfooted as articles do. However, within this framework, it is difficult to establish whether there are different degrees of complexity among the three configurations.

The predictions of the three accounts are summarised in (9).

(9) a. FRAGILE COMPUTATION OF AGREEMENT HYPOTHESIS
     I > II > III
   b. AGREEMENT DEFICIT ACCOUNT
     Procedural view: I = II = III
     Lexical view: I = III ≠ II
   c. SURFACE ACCOUNT
     I = III; II difficult with third person plural verb

We now turn to our study. First, we tried to establish whether DLD children compute agreement hierarchically comparing their sensitivity to the presence of a structural vs. a linear intervening element. Then, in the second experiment, we examined the performance of children with DLD in the three different agreement configurations.

Experiment 1: testing ATTRACTION ERRORS in DLD children

In the first experiment, we tried to assess if DLD children with difficulties in the morphosyntactic domain compute agreement by taking into account the structural architecture of the sentence. We tested Italian DLD children with two different types of sentences, instantiating two different types of intervention:

(10) Linear intervention:
    La   nonn-a   *delle*   *bambin-e* ha/*hanno cucinato   la   pasta
    DET.S.F grandmother-F.S POSS.F.PL girl-F.PL   AUX.3.S/*P cook.PTCP.S.M DET.S.F pasta
    'The grandmother of the girls has/*have cooked the pasta.'

(11) Structural intervention:
    La   nonna   *le* ha/*hanno   abbracciate
    DET.S.F grandmother-F.S   OBJ.F.PL AUX.3.S/*P hug-PTCP.F.P
    'The grandmother has/*have hugged them.'

(10) is a case of linear intervention. The subject is a complex DP including a modifier of the head noun *nonna* 'grandmother', i.e., the plural noun *bambine* 'girls', which is embedded under a prepositional phrase (PP). The nominal *bambine* linearly intervenes between the head noun of the subject DP *nonna* and the verb. However, it stays lower in the constituent structure, and it does not structurally intervene (it does not c-command the inflected auxiliary verb), as in (10′). In contrast, (11) is a case of structural intervention. Here the object is realised by a plural object clitic pronoun, which appears in a preverbal position. The plural object clitic hierarchically intervenes between the subject and the inflected verb, as it c-commands the verb,

and does not c-command the subject (11′.)[3] Hence, the clitic pronoun structurally intervenes in the Spec-Head relation between the DP *nonna* and the verb.

(10′)

The grandmother   has   cooked   the noodles

of the girls

(11′)

The grandmother   them   has   hugged
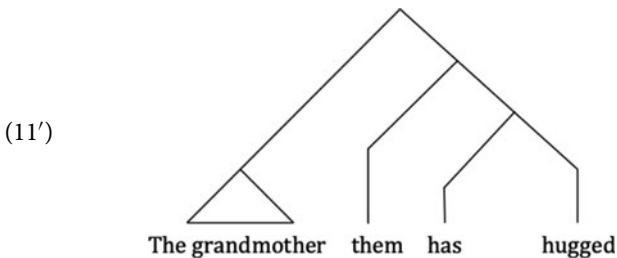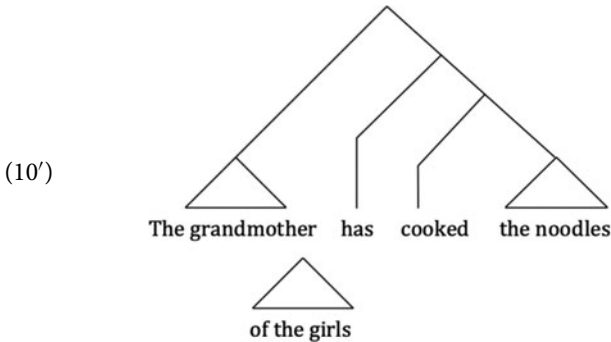
We presented children with a simple event using pictures. The event was later described by two characters: one produced a grammatical sentence and the other an ungrammatical one. Children had to choose who said it better (forced-choice task). Our hypothesis predicts a preference for the structures with agreement governed by the subject noun phrase over the incorrect agreement with the closer nominals (*le bambine* in (10) and the clitic *le* in (11)); moreover, we expect some attraction errors in both configurations, but more errors in the structural intervention configuration (11) than in the purely linear configuration (10). Thus, if DLD children compute agreement hierarchically, much as TD children and adult speakers do, we expect that

1. DLD children will systematically prefer agreement with the structurally correct nominal, the head noun of the subject DP, not with the intervening nominal;
2. DLD children will manifest some attraction errors;
3. DLD children will be sensitive to the distinction between linear vs. hierarchical intervention, along the lines shown by adult elicited production data (Franck *et al.*, 2006; Franck *et al.*, 2010).

___

[3]We adopt the definition of c-command in Chomsky (1995). The definition of hierarchical intervention in terms of c-command is borrowed from Rizzi (1990). Here we do not discuss the exact derived structure of cliticisation. Notice that whether the clitic is adjoined to the inflected verb, as in most standard analyses, or is attached to a higher position, as in (11′), the configurational property is not affected: in both configurations the clitic c-commands the inflected verb and does not c-command the subject, hence it hierarchically intervenes between the subject and the inflected verb.

Method

Participants

Two groups of Italian-speaking children took part in the experiment. Nineteen children diagnosed with DLD aged between 4;3 and 6;0 were included in the first group, while the second consisted of the same number of TD children within a similar age range, between 4;3 and 5;8.

The DLD children were recruited from the Centro Dedalo of Siena, the rehabilitative centre Giovanni XXIII in Lessona (Biella), and a private speech-therapist centre in Biella. All DLD children who we initially selected were diagnosed by clinicians as having difficulties in receptive grammatical tasks but no indication of more general cognitive disorders (as established by clinicians, WPPSI-III, or Raven Coloured Matrices). Since we were primarily interested in the comparison between TD children and children with grammatical comprehension problems, we further screened all DLD children for inclusion in our experimental group. Their morphosyntactic difficulties were assessed through a standardised test (Syntactic Structure Comprehension Test: Marini, Marotta, Bulgheroni, & Fabbro, 2015, Batteria per la Valutazione del Linguaggio in Bambini dai 4 ai 12 anni, henceforth BVL 4-12). All children in the DLD group scored below –1 standard deviation in the Syntactic Structure Comprehension Test, confirming their below-average capacities at comprehending various types of morphosyntactic structures (including relative clauses, reflexive and clitic pronouns, simple S-V agreement structures). Our screening was in line with the diagnosis initially made by the speech-therapists. We will descriptively refer to these subjects as Syntactic-DLD (S-DLD), following the labelling in Novogrodsky and Friedmann (2006), to simply indicate the attested manifestation of a morphosyntactic type of impairment in all the DLD children who took part in Experiment 1.

The TD group of children was recruited from two kindergartens, one in Siena (the Comprehensive Institute of San Benedetto) and one in the province of Biella (the Comprehensive Institute of Candelo). None of them had any known language or cognitive problem. A two-samples t-test shows a non-significant difference ($t(36)$ = –1.5, d = –0.51, p = .12) in the mean age of the two groups. As with the DLD children, we further screened them using the same grammatical comprehension test. All but one[4] scored between 0 and + 2 standard deviations in the Syntactic Structure

Table 1. Number of participants (N), age (mean and standard deviations) and raw score in the BVL (mean and standard deviation) for Typically Developing children (TD) and Syntactic-DLD children (S-DLD)

Group	N	Mean age (SD)	BVL raw scores (SD)
S-DLD	19	5;1(0;6)	19,5 (3,1)
TD	19	4;9(0;6)	34,5 (5,5)

[4]One child in the TD group scored below –1SD (raw score = 23). A later inspection of his performance, however, revealed that in the experimental task this child had no particular trouble, with a global average proportion of correct choices at 83%, at ceiling in the Linear condition and at 75% in the Structural Condition. We thus decided to keep this child in the dataset. This choice had no effect on the general results.
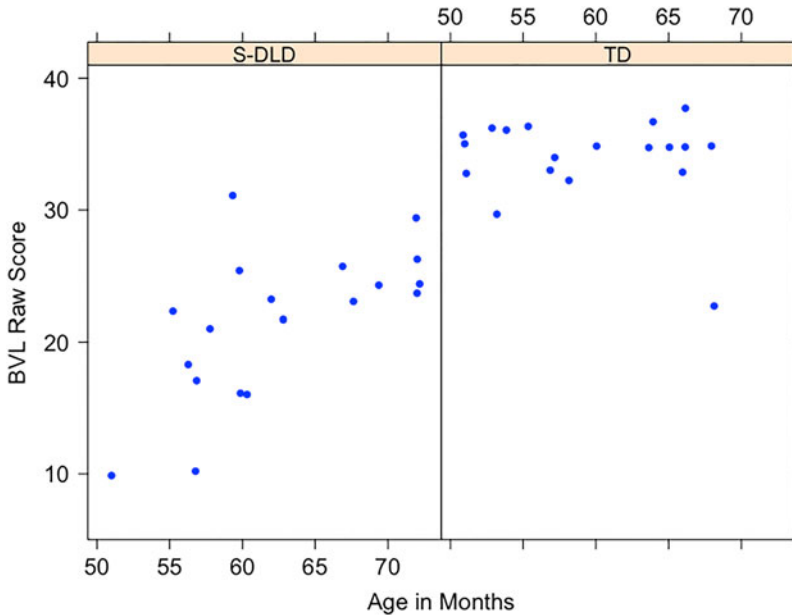
Figure 1. Distribution of raw scores (*y*-axis) at the BVL test (Marini *et al.,* 2014) assessing grammatical comprehension for S-DLD participants and TD as a function of Age (*x*-axis).

Comprehension Test battery. For the BVL score, a two-samples *t*-test confirmed that the score of the TD group was higher than that of the DLD group ($t(36) = 8.4$, $d = 2.72$, $p < .01$). Table 1 summarises the group demographics as well as the mean score in the linguistic test obtained by participants in each group. Individual data are plotted in Figure 1, where raw scores are reported. The study was approved by the Ethics committee of the University of Milano-Bicocca according to the standards of the Helsinki Declaration (1964) (prot. 20974/13).

Procedure and materials

In order to test the attraction effect of a feature mismatch between the head noun in the subject DP and an intervening element, we used a Forced-Choice of Grammatical Form (FCGF) task (Moscati & Rizzi, 2014). In this task, a child-friendly context was created. Each child was told that she was going to see some pictures of people doing something. The child was also told that she had to pay very careful attention to what happened, as at the end of the story two characters (two piglets) would appear on the screen and say what happened. It was explained to the children that the characters liked to play silly games, and only one of them would be 'right'. Thus, the children had to indicate who said it 'right'. A typical experimental trial is illustrated in Figure 2.

   The first image set the stage and presented the characters involved. Then, a second image presented the action; in this case, a grandmother hugging her granddaughters. At this point, the first character appeared and uttered the first sentence. Then, the second one appeared and uttered the second sentence:
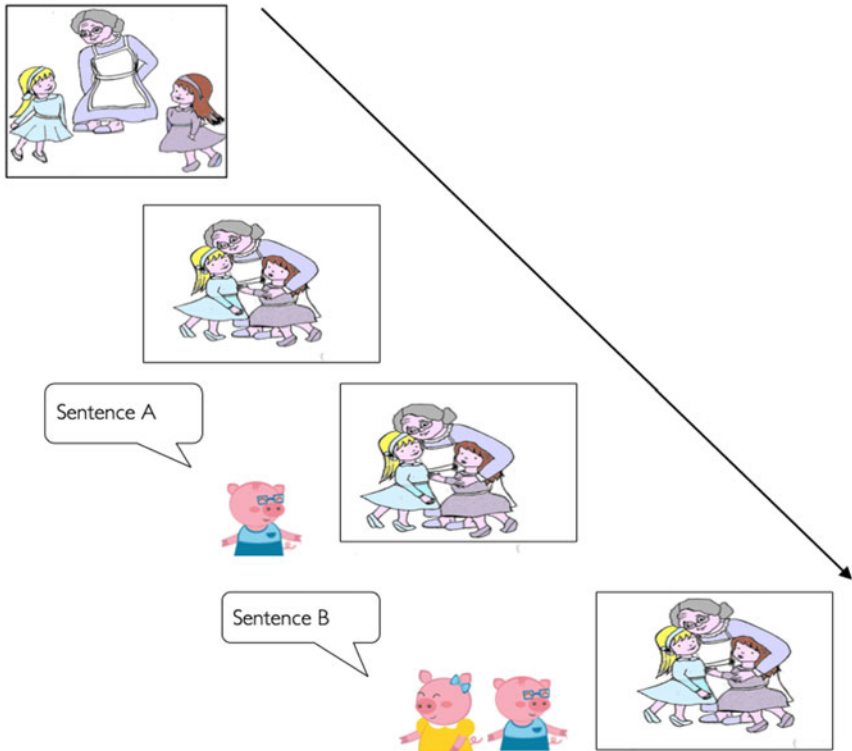
Figure 2. Example of the experimental procedure.

(12)

    a. La   nonn-a   le   ha   abbracciate
       DET.S.F   grandmother-F.S   OBJ.F.PL AUX.3.S hug.PTCP-F.P
       'The grandmother has hugged them'
    b. *La   nonn-a   le   *hanno abbracciate
       DET.S.F   grandmother-F.S   OBJ.F.PL AUX.*3.P hug.PTCP-F.P

Stimuli were counterbalanced to vary the order of presentation of the correct sentence and the character. In the experiment, we varied the type of intervention (structural, linear) and the feature mismatch (plural, singular intervening feature). Thus, we had the four experimental conditions illustrated in Table 2.

    The rationale for varying the number of targets and the intervening element's feature is the following. Some studies (a.o., Bock & Cutting, 1992) have shown that a plural intervening element is more likely to trigger an attraction error, in case of number mismatch, than a singular one. This is plausibly due to the marked status of plural with respect to singular. We would then expect a higher number of attraction errors when the intervening element is plural.

    For each condition, we had six items. We also added six extra pairs of SVO control sentences, in which the correct sentence was compared with a sequence of words having

Table 2. Experimental conditions

Type of intervention	Intervening element's number	Examples
Structural	Plural	La nonna *le* ha/*hanno abbracciate
		'the grandmother has/*have hugged them'
Linear	Plural	La nonna *delle bambine* ha/*hanno cucinato la pasta
		'the grandmother of the girls has/*have cooked the pasta'
Structural	Singular	Le principesse *la* hanno/*ha salutata
		'the princesses have/*has said hello to her'
Linear	Singular	Le infermiere *della signora* hanno/*ha controllato la lista
		'the nurses of the lady have/*has checked the list'

the same lexical material but arranged in a random order, for a total of 30 sequences of pictures. Moreover, we had a short warm-up in which participants were familiarised with the task. During the familiarisation, some objects appeared on the screen and were referred to differently by two characters: one using the correct noun and the other one using the wrong one. This was intended to familiarise participants with the procedure and to ensure that they were paying sufficient attention to the task.

Results

All children completed the task, showing no general attention problem. DLD children were able to retain in memory the sentence pair and their performance with the four SVO control sentences was very good (above 95% accuracy). For what concerns agreement, a first preliminary question was whether S-DLD children generally preferred agreement with the correct nominal – the head noun of the subject DP – or were systematically misled by the closer nominal. The overall proportion of correct choices in the S-DLD group was 64.8%. Although S-DLD children performed less well than children in the TD control group, who provided the correct answer in 87.7% of cases, their global performance was above chance (one-sample *t*-test: $t(18) = 5.23$, $p < .001$). Looking at individual data (Figure 3), the majority of S-DLD children (12/19[5]) answered above chance (at least 15 correct answers out of 24, cumulative binomial distribution with a cut-off point below $p < .01$). This finding indicates that S-DLD children were sensitive to the hierarchical configuration of agreement: they were not systematically misled by linear proximity and by the global complexity of the structure.

Now we turn to their performance in the experimental conditions. Aggregate descriptive statistics are plotted in Figures 4 and 5. Data were analysed in R (v. 3.4.3;

---

[5]Among the other seven children, four of them gave 13 or 14 correct answers, while only two were exactly at chance (12 correct answers) and only one below (11 correct answers).
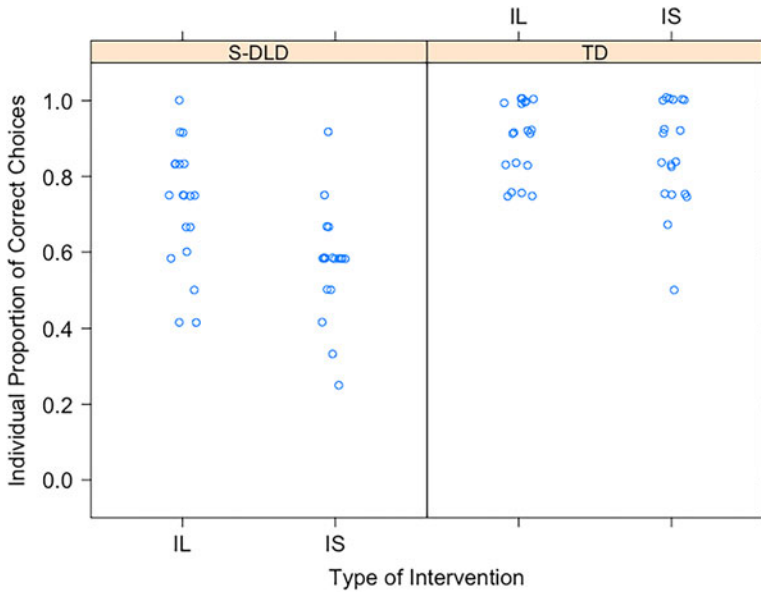
Figure 3. Proportion of correct choices (*y*-axis) for each participant in the two groups of children (S-DLD and TD) as a function of structural (SI) and linear (LI) intervention.
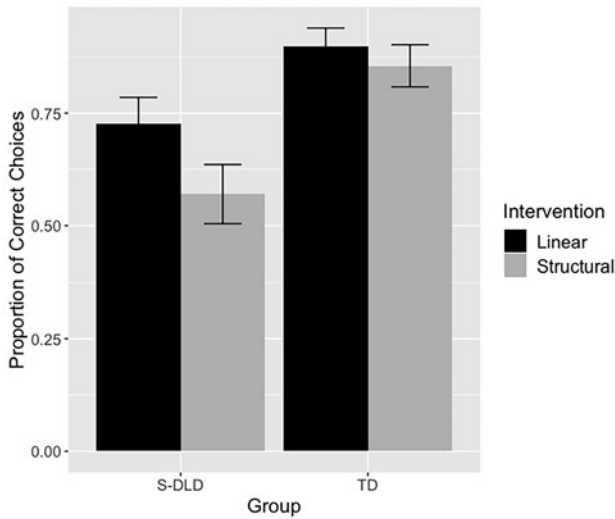
Figure 4. Proportions of correct choices (*y*-axis) in the two conditions of intervention (linear and structural) by the two groups of children (Syntactic-DLD and TD, *x*-axis). Error bars, 95% Confidence Interval.

R Core Team, 2019) using the *lme* function of the *lme4* package v.1.1-21 (Bates, Maechler, Bolker, & Walker, 2015; CRAN.R-project.org/package = lme4). We ran different generalised mixed models with Group, Type of Intervention, and Number
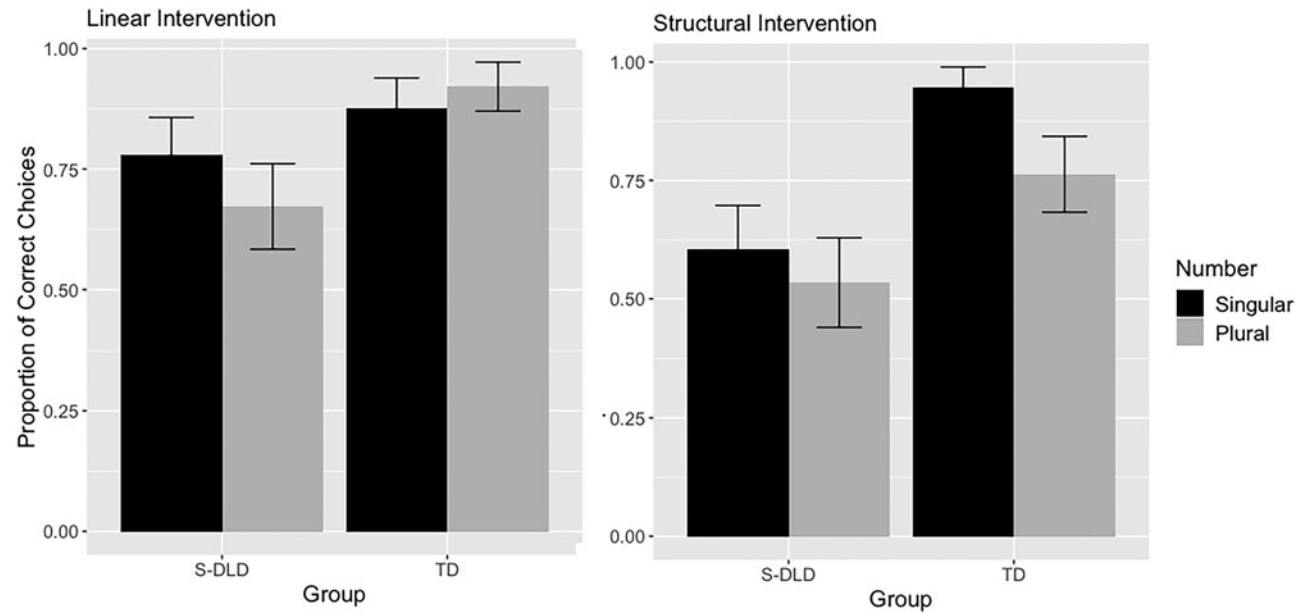
Figure 5. Proportion of correct choices (*y*-axis) in the linear condition (left) and in the structural condition (right) as a function of the Number feature (singular or plural) of the intervening nominal element in the two groups of children (S-DLD and TD, *x*-axis). Error bars, 95% Confidence Interval.

as predictors, or fixed effects, and Subject and Item as random effects. Models were compared through a likelihood ratio test (Baayen, Davidson, & Bates, 2008) using the *lrtest* function (*lmtest* package v.0.9-37, CRAN.R-project.org/package = lmtest). The simpler best-fitting model was the one including the full three-way interaction between our predictors.[6] Details of the model are reported in Appendix C, Table 1. It revealed a marginally significant main effect of Type of Intervention and Group, plus significant 2-way interactions between Group x Type of intervention, Group x Number, and Type of Intervention x Number. We also found a significant three-way interaction. To interpret the interactions, we performed a further analysis keeping the two groups separate. We will refer to the different models reported in Appendix C when presenting the descriptive statistics plotted in the figures.

Comparing the children's performance in the Linear and Structural conditions, Figure 4 shows a clear difference between the two groups, and a greater difference between Linear and Structural intervention in the S-DLD group (resulting in a two-way interaction between Group x Type of Intervention; see Table 1, Appendix C). In the S-DLD group, a significant main effect of Type of Intervention was found (Appendix C, Table 2), confirming that S-DLD children have more trouble in the Structural than the Linear condition.

Turning to the effect of number, we plot it in Figure 5, keeping separate the Linear (left) and the Structural (right) conditions. Considering the TD group first, a comparison between the two panels in Figure 5 shows a greater difficulty with a plural intervening element in the Structural Condition only. This is confirmed by the interaction between Number and Type of Intervention in the TD group (Appendix C, Table 3), due to the effect of number in the structural (Appendix C, Table 4) but not in the Linear condition (Appendix C, Table 5). Concerning S-DLD children, the same comparison between the two panels in Figure 5 shows that the effect of number in the two conditions is low and it is only marginal (Appendix C, Table 2).

Taken together, the results show a clear difference between the two groups, with S-DLD children having more troubles than TD in the Structural configuration, with a marginal main effect of plurality. As for TD children, the most interesting result is the interaction between Number and Type of Intervention, with the plural intervening feature being harder in the Structural Condition. This can be straightforwardly accounted for by the fact that their performance is at near ceiling in all the experimental conditions, except for the one that is also the most challenging for adults (a.o., Franck *et al.*, 2002).

Intermediate discussion

A critical finding of the first experiment is that S-DLD children show a good global performance, preferring the verbal form that agrees with the head noun of the subject. If considered in general, their choices are not random: the noun linearly closer to the verb is dispreferred over the structurally closer subject.

Our first experiment confirmed that attraction errors are modulated by the type of intervening element in the S-DLD group. S-DLD children find it harder to ignore elements that carry different grammatical traits from the subject in the condition of

---

[6]The model with the three-way interactions and random intercepts presented a better fit over models with two-way interactions and over the simplest model without interactions. A further likelihood ratio test showed that adding random slopes to the model only marginally ($p = .057$) increased the model's fit.

Structural Intervention. This configuration is the one particularly challenging for S-DLD, regardless of the number specification of the intervening element. We take this result as showing that the presence of the clitic has a significant impact by increasing the complexity of the computation of agreement. This also demonstrates that S-DLD children are sensitive to the presence of an element that introduces a feature mismatch in a position that is structurally in between the subject and the verb. This fact supports the idea that the S-DLD children are also capable of building the right structural representation of the sentence, and (at least try) to perform the feature-checking operation.

S-DLD children show an even more evident effect of structure when compared to the control TD group since, in TD children, the structural effect is modulated by the Number feature. The more challenging configuration is, in fact, the one with a plural structural intervening element (i.e., a plural clitic). This result indicates that TD children at age five already behave like adults (see Bock & Cutting, 1992; Franck et al., 2002) and find a plural intervening feature more 'attractive' when it is structurally inserted between the subject and the verb. To the best of our knowledge, this is the first study on preschoolers that has replicated the adults' results previously obtained with elicited production techniques.

We also wish to point out that in both groups, a plural subject poses no additional challenge. On the contrary, more difficulties are found when children have to choose a singular verb that agrees with a singular subject across a plural intervening element. Therefore, nothing suggests that plural forms of the verb add complexity in our forced-choice task, contra what one would have expected under Leonard's SURFACE HYPOTHESIS.

In sum, the results of Experiment 1 support the view that S-DLD children compute agreement in the same way as TD children, i.e., adopting a hierarchical representation where structural intervention is more disruptive than linear intervention.

Experiment 2: a comparison between three agreement configurations

In the second experiment, we compared agreement in three different configurations: Det-N, S-V, and Cl-PPart. According to the FRAGILE COMPUTATION OF AGREEMENT HYPOTHESIS, we predict that these three configurations display different degrees of increasing syntactic complexity, as a function of the movement-based complexity metric introduced in Moscati and Rizzi (2014). The prediction is that DLD children with a morphosyntactic impairment might have trouble with computing agreement relations, particularly in the more complex configurations (i.e., S-V and Cl-PPart), while the simpler one may be spared (i.e., Det-N). In order to compare the three configurations, we built minimal pairs, testing the same sentence structure, but varying the type of agreement violation.

Method

Participants

A total of 64 children took part in the second experiment. DLD Children were recruited at the IRCCS Fondazione Stella Maris in Calambrone (Pisa), the Centro Dedalo of Siena, the rehabilitative centre Giovanni XXIII in Lessona (Biella), and a private speech-therapy studio in Biella. Children in the control group were recruited from a kindergarten of the Comprehensive Institute of Candelo (Biella). The first group of

22 DLD children was selected using the same criteria as in Experiment 1: following the diagnostic indication of clinicians, we initially selected children with reported difficulties in grammatical comprehension. Successively, children in this group were further screened using the Syntactic Structure Comprehension Test (BVL 4-12). All of them, except one, showed low performance in grammatical comprehension (between 0 and –2 SD). For this group, we maintained the descriptive label of S-DLD employed previously.

In Experiment 2, we also added a second group of 12 DLD children who were following individual language therapy at the time, but without any reported problem in grammatical comprehension. Clinicians individuated for all of them a primary impairment in the phonological domain. We label children in this second group as non-Syntactic DLD (nS-DLD), simply to differentiate them from the first group. We administered the Syntactic Structure Comprehension Test (BVL 4-12). The majority, 9 out of 12 children in the nS-DLD group, had a grammatical score in the BVL above 1 SD, confirming the speech-therapists' indication that they present no evidence of a grammatical comprehension impairment. Three of them instead scored below the average (between 0 and –1 SD). For these three children, based on their comprehension score, we cannot exclude a grammatical deficit akin to one of the children in the S-DLD group. For this reason, we excluded them from the nS-DLD group. In this way, we obtained a principled distinction between S-DLD and nS-DLD based not only on the clinician's indications but also on their BVL score.

We added this group because the kind of agreement morphemes examined in this study may be singled out for two reasons: they express syntactic dependencies between positions, and they typically are phonologically weak syllables. Potentially, these properties make such morphemes selectively vulnerable in pathology. On the phonological side, Bortolini *et al.* (2006) have proposed that phonologically weak syllables, such as clitics and articles, are optionally omitted by children with DLD in production. A similar fate is reserved for the third person plural morpheme of the verbal paradigm. In this case, children with DLD do not produce it, and use a third person singular morpheme instead. These properties have been observed in production, but it is not known whether children with such a phonological impairment have difficulties with the comprehension of these morphemes. Although the BVL includes sentences with clitics, articles, and third person plural verbs, it does not tap specifically into them, as there are only a few items per type, and there are many types of sentences (e.g., passives, relative clauses) in which other features of the structure can help children in understanding sentences. We thus feel that it is important to test, employing minimal pairs, the abstract grammatical capacity to deal with agreement in children diagnosed with a production deficit. A potentially important outcome of this comparison would be to investigate the possibility that the same morphemes may be affected for different reasons in different groups: their special phonological status may count for one group of children with DLD and the morphosyntactic status for another group. This distinction may be important because the same morphemes could be the object of distinct treatments for the two groups of children.

We also had a third group of 31 age-matched TD children. All children in the TD group had a score on the BVL 4-12 above 1 SD. Individual scores on grammatical comprehension for children in each group are reported in Figure 6. Stars indicate the four children (1 S-DLD with a high BVL score, three nS-DLD with a low BVL score) who were excluded on the basis of the second screening.
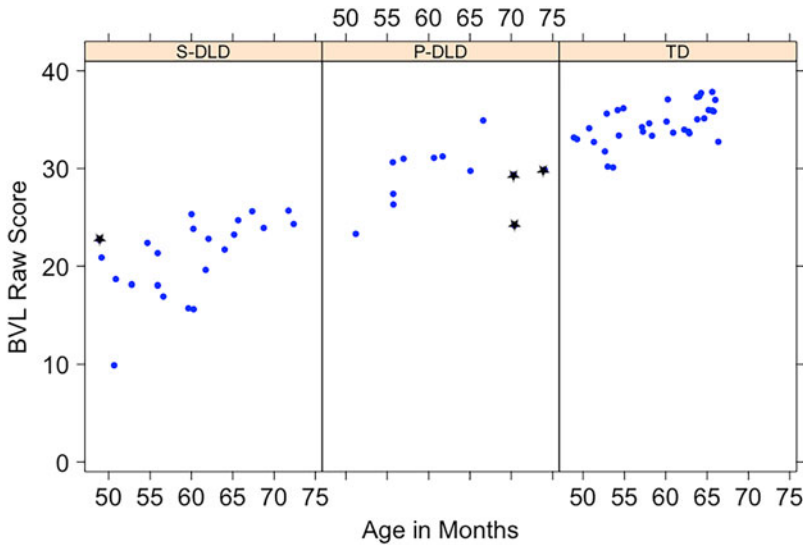
Figure 6. Distribution of raw scores (*y*-axis) at the BVL test (Marini *et al.*, 2014) assessing grammatical comprehension for S-DLD, nS-DLD, and TD participants as a function of Age (*x*-axis). Stars indicate the scores obtained by children that were excluded from the S-DLD and the nS-DLD group.

Table 3. Number of participants (N), means (standard deviations) of raw score in the BVL, and age (standard deviations) for each group (Typically developing children, TD, nS-DLD children and S-DLD children)

Group	N	Mean age	Mean raw score BVL
TD	31	4;8 (5.6)	34.6 (2.0)
nS-DLD	9	5;1 (6.8)	29.4 (3.4)
S-DLD	21	4;9 (6.8)	21.0 (3.8)

Aggregated scores obtained in the BVL and age means per each group are given in Table 3. A two-samples *t*-test revealed a significant difference ($t(51) = 16.30$, $d = 4.58$, $p > .001$) in the BVL score between the S-DLD group and the TD group This difference was also significant ($t(28) = 5.44$, $d = 2.16$, $p < .001$) between the S-DLD and the nS-DLD group.

Procedure and materials

The testing procedure was similar to the one already illustrated for Experiment 1. We used an FCGF task in which children had to choose the grammatical alternative between two sentences that minimally differed in a single morpheme: a correct sentence and an alternative one that introduced a violation of agreement. As in the first experiment, participants sat in front of a laptop computer and saw sequences of two pictures illustrating simple events. Then a pair of sentences was presented, each one uttered by a different character (a male or a female piglet) that appeared on the

Table 4. Experimental conditions in Experiment 2. Underling signals the elements entering in the correct agreement. In brackets, the referent of the pronoun, as shown in the picture set, not pronounced in the target sentence.

Conditions	Example: '(the cake) the girl has eaten it'					
D-N	(la torta)	la/*le	bambina	la	ha	mangiata
	(the cake)	DET.F.S/*F.PL	girl-F.S	Obj.F.S	AUX.3.S.	eat.PTCP.F.S
S-V	(la torta)	la	bambina	la	ha/*hanno	mangiata
	(the cake)	DET.F.S	girl-F.S	Obj.F.S	AUX.3.S./3.PL	eat.PTCP.F.S
Cl. Part	(la torta)	la	bambina	la	ha	mangiata/*ate
	(the cake)	DET.F.S	girl-F.S	Obj.F.S	AUX.3.S.	eat.PTCP.F.S/*F.P

Table 5. Percent of correct choices in each group (SD)

Group	S-DLD	nS-DLD	TD
Mean percentage of correct choices	77.4 (2.1)	94.4 (1.8)	89.9 (1.2)

screen. At this point, children had to choose who said it right, either repeating the correct sentence or indicating the character's picture.

We used six different sets of pictures, and for each of them children heard pairs of sentences with the three types of agreement violations for a total of 18 minimal pairs plus 12 fillers. For example, a picture sequence with a girl eating a cake was seen three times throughout the session, each time pairing the same correct sentences with an alternative introducing one of the three violations (assessing Det-N, S-V, and Cl-PPart agreement). Stimuli were counterbalanced to vary the presentation order of the correct sentence and items were pseudo-randomised to avoid the presentation of the same picture set twice in a row. The experimental conditions are summarised in Table 4. To keep the experiment duration to a minimum, only the singular was used for the correct answers. In this way, the experimental design was within-subject limited to three conditions per group (3 × 3).

Results

Table 5 reports the overall proportion of correct choices for each group. When the overall performance was considered, the S-DLD children obtained the lowest performance, with 77.4% of correct choices. Although being the lowest, the S-DLD group as a whole did not answer at chance (one-sample t-test: $t(20) = 9.36$, $p < .001$), and individual scores, based on the binomial distribution with a cut-off point below $p < .01$, indicate that 17/21 children in the S-DLD were above chance with at least 12 correct answers out of 18 trials. This fact shows once again that they have no general difficulty with the task and that S-DLD children choose the correct sentence in the majority of the cases. The other two groups, the nS-DLD and the TD group, presented a generally higher performance with 94.4% and 89.9% correct responses; all children performed above chance in the nS-DLD group and all but one in the TD group.
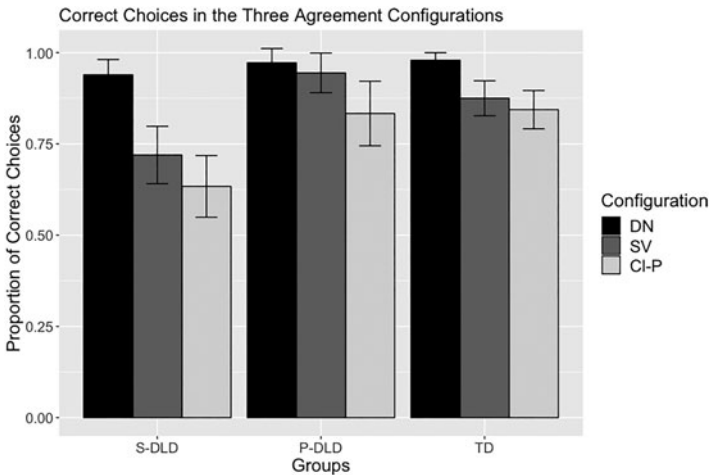
Figure 7. Proportion of correct choices (*y*-axis) for each agreement configuration (Determiner–Noun, DN; Subject–Verb, SV; Clitic–Past participle, Cl-P) by S-DLD, nS-DLD, and TD children (*x*-axis). Error bars, 95% Confidence Interval.

We now turn to the performance of children in the three agreement configurations. Figure 7 shows that the three groups all obtained a score well above 90% in the Det-N agreement condition. In the S-V agreement condition, the difference between groups became instead more evident, with S-DLD children providing correct answers only 73.0% of the time. This proportion is lower than 94.4% of nS-DLD and 87.7% of TD children. In the Cl-PPart condition, S-DLD children provided an even lower proportion of correct answers: they made the right choice in only 64.8% of the cases, while nS-DLD and TD children made the correct one in 88.8% and 84.3% of the cases.

As in Experiment 1, we analyse the data in R by using linear mixed-effects models, using Group and Agreement Configuration as fixed effects and Subject and Item as random effects. This time, the likelihood ratio test revealed that adopting the simpler model with main effects only and random intercepts was justified since it did not differ from the alternative model that included the interaction between Group and Agreement Configuration. Including random slopes for subject and item did not increase the model's fit. The complete outcome of the model is reported in Appendix D (Table 1). It revealed a main effect of Group and Agreement Configuration, confirming that the probability of choosing the correct alternative was lower in the S-DLD group than in the other two groups.

For what concerns the main effect of Configuration, we initially set its reference level to Cl-PPart. This was done using the *levels* function in R. In this way, the model provided an estimate of the probability of having a correct answer in the Cl-PPart condition compared to the other two. The model showed that this probability was significantly lower in the Cl-PPart configuration than in the Det-N configuration and that it was the same in the Cl-PPart and the S-V configurations.

Since we were also interested in comparing S-V and Det-N, we ran the same model again, but this time using S-V (and not Cl-PPart) as the reference level (Appendix D, Table 2). In this way, we had the desired comparison between S-V and Det-N, with the
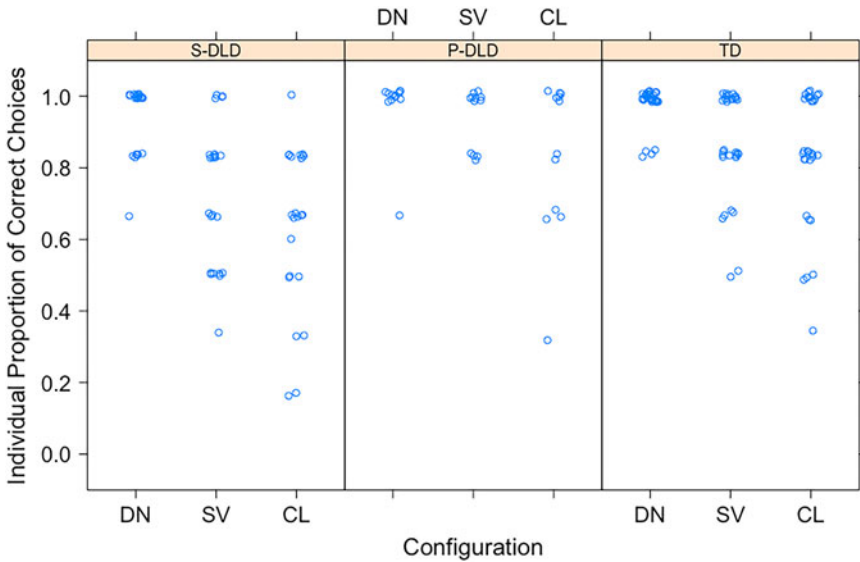
Figure 8. Proportion of correct choices (*y*-axis) for each participant in the three groups of children (S-DLD, nS-DLD, and TD) as a function of the agreement configuration (Determiner–Noun, DN; Subject–Verb, SV; Clitic–Past participle, Cl-P; *x*-axis).

difference between S-V and Det-N also being significant: the probability of making the right choice in the S-V was lower than in the Det-N Configuration.

In sum, the analysis revealed that the effect of Configuration was due to the difference between the two more complex configurations, Cl-PPart and S-V, on the one side, and the Det-N Configuration on the other. Cl-PPart and S-V substantially patterned together and were more difficult than the Det-N configuration. Thus, although the proportion of correct answers was descriptively higher in the S-V than in the Cl-PPart condition (Figure 7), this difference did not reach statistical significance.

To inspect the individual performance, we also plotted, in Figure 8, the proportion of correct choices of each child in the three conditions. Figure 8 shows that, in the S-DLD group, the individual performance becomes more scattered towards lower values in the S-V condition and even more so in the Cl-PPart condition. A pattern that is indicatively found, to a lesser extent, also in the TD and in the nS-DLD group.

Zooming in on the performance of S-DLD children, we observe that the overall majority of children (16) mirrored at the individual level the overall aggregated pattern, with the Det-N condition being the one with the higher proportion of correct choices, followed by S-V and Cl-PPart.[7]

---

[7]Five children departed from the general pattern. Four subjects showed a lower performance in the S-V condition, compared to the CL-PPart condition. However, also for these four subjects, the D-N condition was the less problematic one. Only one child made more errors in the Det-N condition than in the S-V one. However, the performance of this child was one of the highest in general: he made only one error in both the Cl-PPart and the Det-N condition, and none in the S-V condition.

Discussion of Experiment 2

Experiment 2 highlights a significantly lower performance of S-DLD group in comparison with the other two groups. Moreover, the forced-choice task revealed a clear difference between Det-N and Cl-PPart agreement in all three groups: whereas the D-N configuration is unproblematic for all the three experimental groups, the Cl-PPart configuration was the most challenging in general. These two configurations are at the two extremes of the scale of locality proposed in Moscati and Rizzi (2014). The asymmetry between more local (Det-N) and less local (Cl-PPart) configurations is predicted by the FRAGILE COMPUTATION OF AGREEMENT HYPOTHESIS, and the results of Experiment 2 thus support our idea that various agreement configurations pose different challenges to children and that they must be differentiated in accordance with a finer-grained syntactic typology.

Beside the Det-N and the Cl-PPart agreement configurations, which represent the clearest cases, in Experiment 2 we also tested the S-V agreement configuration. Most of the elicited production literature on Italian language acquisition suggests that, at earlier stages of development Subject–Verb agreement (Caprin & Guasti, 2009; Guasti 1993/94; Pizzuto & Caselli, 1992) is indeed correctly produced, noticeably earlier that Cl-PPart Agreement (Moscati & Tedeschi, 2009; Schaeffer, 2000). However, the studies above also indicate that, at the end of the fourth year, both types of agreement are fully mastered by typically developing children, in the light of the virtual absence of substitution errors. In our experiment, we descriptively observed a small, although not significant, difference between S-V and Cl-PPart at around age five, thus at a developmental stage where elicited production does not show any difference any more. In this respect, the forced-choice task seems to be more sensitive to reveal certain subtle syntactic differences, if minimal pairs can be built on them.[8]

General discussion and conclusion

In this paper, we have explored DLD children's difficulties in establishing a morphosyntactic relation between two elements that agree in grammatical features. Our main goal was to draw a precise comparison between different syntactic configurations in which this feature-sharing operation takes place, observing DLD children's performance with various types of agreement. To allow a finer-grained comparison, we extended the forced-choice methodology previously adopted with TD children (Moscati & Rizzi, 2014) to DLD children acquiring Italian. The advantage of using a forced-choice procedure is that it is fairly simple and sensitive,

---

[8]We tested S-V agreement using sentences containing a clitic pronoun carrying the same gender and number features as the subject, so as to avoid intervention effects of the kind found in Experiment 1. The rationale for this choice was to allow a tight comparison between the three agreement configurations, keeping full structural parallelism between the three conditions: the ungrammatical alternative was always built on the very same sentence. A reviewer pointed out that the simple presence of a clitic pronoun may have added an element of difficulty in the evaluation of S-V agreement. This cannot be excluded, but a comparison between S-V agreement in previous works employing the same forced-choice task shows the same graded pattern regardless of the presence (Moscati & Rizzi, 2014) or the absence (Moscati & Rizzi, 2013) of a clitic, as long as no intervention effect is induced by a feature-mismatch. Moreover, Arosio and Giustolisi (2019) have shown that children at the age of five years produce clitics at a high rate when the feature of the subject and of the clitic are the same as in our experiment.

compared, e.g., to elicited production. Besides, it also permits fine-grained morphosyntactic experimental manipulations, to allow a more detailed assessment of the DLD linguistic competence.

The first question we wanted to address was whether DLD children with manifested morphosyntactic difficulties compute agreement in the same way as adults and TD children do, that is, adopting hierarchical representations of agreement dependencies. In Experiment 1, we tested the fundamental hypothesis that agreement is computed hierarchically by S-DLD and TD children. The crucial finding was that both groups discriminate the correct hierarchical subject–verb agreement from an incorrect linear form of agreement with the linearly closest nominal. The main difference between the TD and S-DLD groups resided in the size of the effect, with S-DLD children finding it harder to ignore the feature mismatch introduced by the structurally intervening element, an object clitic pronoun. So, the structural effect was even sharper in S-DLD children. An important corollary of this finding was that Italian S-DLD children do not ignore the clitic, trying to process its syntactic features actively. Our results show that, although being phonologically weak, a clitic has a strong disruptive effect on S-V agreement.

For their very general nature, we expect that the pattern of results that emerged from attraction errors goes beyond the particular case of Italian. Not only should other Romance languages with clitics align to Italian, but we also expect attraction errors to emerge across the board whenever they can be tested by manipulating the level of embedding or the linear distance.

Besides the structural/linear distinction, a further point we wish to make concerns the role of number. The plural/singular distinction offers crucial insights about previous hypotheses about Italian DLD children's difficulties with verbal morphology. In particular, Leonard's Surface Account predicts that plural verbal morphology should be harder than singular because it is associated with an additional weak syllable. Results from Experiment 1 did not confirm this expectation: no statistically significant difference was found between singular and plural verbal forms, and the descriptive trend was the opposite, with singular forms of the verb being harder (if preceded by a plural intervening element) to handle than plural ones.

One novel aspect of our research, which deserves to be underscored here, is connected to the nature of the experimental task. Whereas much of the literature on agreement in DLD is based on production (natural production corpus analysis or experiments of elicited production), our task of forced-choice of a grammatical structure relies on the reception of syntactic configurations. As such, our results are novel in that they highlight difficulties in the DLD population in a receptive task. Perhaps this point can be pushed a bit further. The forced-choice task arguably is less taxing in terms of computational resources than production: on the one hand, it makes the stimulus to be evaluated immediately available to the experimental subject, as an alternative to a deviant one; on the other hand, it does not require the activation of the articulatory system. The fact that difficulties with agreement continue to emerge in a less taxing task seems to suggest that such difficulties in the DLD population may be linked to a specific impairment in the abstract grammatical computation, which underlies both production and comprehension. In contrast, the hypothesis of a domain-general limitation in computational resources in the DLD population would be consistent with very different manifestations of the difficulty in production and comprehension, linked to the more or less taxing nature of the experimental task.

Having assessed the asymmetry between linear versus structural configurations, in Experiment 2 we moved one step further and directly tested the predictions of the FRAGILE COMPUTATION OF AGREEMENT HYPOTHESIS, based on an explicit scale of syntactic locality. This hypothesis predicts that the Det-N agreement configuration will be the easiest, as it is the most local one. The results of Experiment 2 confirm that the Det-N configuration stands out for being the one where fewer agreement mistakes are found. This result is expected under the FCAH but cannot be straightforwardly captured by other accounts that do not differentiate between different structural configurations, such as Clahsen's Agreement Deficit or Leonard's Surface accounts (see (9)). In Experiment 2, we found the same graded pattern for all the three experimental groups, with the Det-N configuration being easier than the other two, not only for TD children (a result that replicates the previous one in Moscati & Rizzi, 2014), but also for nS-DLD and S-DLD. In this last group, however, the performance was generally much lower than in the other two. Thus, together with the finding in Experiment 1, the difference between nS-DLD and S-DLD holds the promise of offering a basis for reliably discriminating children with a morphosyntactic deficit.

Our scale of complexity incorporates the idea that movement increases the cognitive load, in a way similar to the Derivational Complexity Metric presented in Jakubowicz (2005) about the graded difficulties posed by interrogative sentences. Jakubowicz (2011) and Prevost *et al.* (2017) investigated French DLD children's ability to cope with Wh-sentences, showing that their difficulties are proportional to the number of movement operations. The easiest configurations are the ones where no movement is involved, and Wh-pronouns appear *in situ* (13), followed by questions where the Wh-element is fronted (14), and last by questions that require multiple movements (15):

(13) Tu pousses qui?   Wh-in situ (no movement)
     'You push who(m)?'
(14) Qui$_i$ tu pousses t$_i$?   Wh-fronting (1 movement)
     'Who(m) you push?'
(15) Qui$_i$ c'est t$_i$ que tu pousses t$_i$?   Wh-fronting + intermediate movement
                                (2 movements)
     'Who(m) it is that you push?'

Our proposal capitalises on the insight that movement increases complexity, and extends it by considering the interplay between movement and agreement: an agreement configuration not involving movement is simpler than agreement configurations necessarily associated with movement. We leave for future work the important issue of whether different computational procedures like agreement and movement impose comparable computational demands on developing systems, and an attempt to devise a unified complexity metric based on performed operations.

The inventory of agreement configurations presented here does not mean to be exhaustive, and other configurations could be included, such as subject–verb agreement with preverbal and postverbal subjects, agreement between a nominal expression and an adjectival predicate (on which see Moscati & Rizzi, 2014) etc. Nevertheless, in this preliminary exploration, we decided to limit our study only to a subset of the possible agreement configurations, to avoid several potentially confounding factors. Another limitation of our analysis is that it is entirely based on

Italian. Considering other configurations in other languages should be instrumental in refining the typology of agreement configurations, thus making it possible to obtain a finer-grained characterisation of morphosyntactic deficits in DLD. We are convinced that our typology could be applied in a comparative perspective, but caution is needed to accommodate cross-linguistic variation. Consider, for example, Det-N agreement, where Case could add another dimension of complexity to the otherwise simple agreement configuration in languages like German (see Clahsen, 1991; Eisenbeiss, Bartke, & Clahsen, 2005), or the distinction between the predicative and attributive function of the adjective, that introduce different agreement paradigms, as in the case of Dutch (see Duinmeijer, 2016). There is one important outcome of our Italian-based research, which we expect to be generalisable across languages and agreement systems, namely that DLD children acquiring Italian are sensitive to the configurational structure of the sentence in computing agreement, i.e., they compute agreement through the same formal ingredients (c-command, locality) that are used by typically developing children and adults. No qualitative forms of deviance (such as exclusive sensitivity to purely linear locality) are observed in DLD children, but only an accrued fragility in dealing with complex agreement configurations. We believe this sensitivity reveals a deep property of DLD children's syntactic capacities, and we expect our results to be generalisable in a cross-linguistic perspective.

References

Abdalla, F., & Crago M. (2008). Verb morphology deficit in Arabic-speaking children with specific language impairment. *Applied Psycholinguistics*, *29*, 315–40.

Archibald L. M. D., & Gathercole, S. E. (2006). Short-term and working memory in SLI. *International Journal of Language and Communication Disorders*, *41*, 675–93.

Arosio, F., Branchini, C., Barbieri, L., & Guasti, M. T. (2014). Persistency of direct object clitic omission in Italian school-age children with SLI. *Clinical Linguistics and Phonetics*, *28*, 1–25.

Arosio, F., & Giustolisi, B. (2019). Agreement and interference in direct object clitic production in Italian monolingual children. *Frontiers in Communication*, *3*, e00066.

Baayen, H. R., Davidson, D. J., & Bates D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bedore L., & Leonard, L. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *44*, 905–24.

Belletti, A. (2006). (Past) Participle agreement. In M. Everaert & H. van Riemsdijk (Eds.), *The Blackwell companion to syntax, Vol. III* (pp. 491–521). Oxford: Blackwell.

Bishop, V. M. D., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & 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*, 1068–80.

Blom, E., Vasić, N., & de Jong, J. (2014). Production and processing subject–verb agreement in monolingual Dutch children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *57*, 952–65.

Bock, J. K., & Cutting, J. C. (1992). Regulating mental energy: performance units in language production. *Journal of Memory and Language*, *31*, 99–127.

Bortolini, U., Arfé, B., Caselli, M. C., Degasperi, L., Deevy, P., & Leonard L. B. (2006). Clinical marker for specific language impairment in Italian: the contribution of clitics and non-word repetition. *International Journal of Language and Communication Disorders*, *41*, 695–712.

Bortolini, U., Caselli, M. C., & Leonard, L. (1997). Grammatical deficits in Italian-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *40*, 809–20.

Bottari, P., Cipriani, P., Chilosi, A. M., & Pfanner, L. (2001). The Italian determiner system in normal acquisition, specific language impairment, and childhood aphasia. *Brain and Language*, 77, 283–93.

Cantiani, C., Lorusso, M. L., Perego, P., Molteni, M., & Guasti, M. T. (2015). Developmental Dyslexia with and without language impairment: ERPs reveal qualitative differences in morphosyntactic processing. *Developmental Neuropsychology*, 40, 291–312.

Caprin, C., & Guasti, M. T. (2009). The acquisition of morphosyntax in Italian: a cross-sectional study. *Applied Psycholinguistics*, 30(1), 23–52.

Caselli, C., Leonard, L., Volterra, V., & Campagnoli, M. G. (1993). Toward mastery of Italian morphology: a cross-sectional study. *Journal of Child Language*, 20, 377–93.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalist inquiries: the framework. In R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by step* (pp. 89–155). Cambridge, MA: MIT Press.

Chondrogianni, V., Marinis, T., Edwards, S., & Blom, E. (2015). Production and on-line comprehension of definite articles and clitic pronouns by Greek sequential bilingual children and monolingual children with specific language impairment. *Applied Psycholinguistics*, 36(5), 1155–91.

Clahsen, H. (1991). *Child language and developmental dysphasia*. Amsterdam: John Benjamins.

Clahsen, H. (2008). Chomskyan syntactic theory and language disorders. In M. J. Ball, M. Perkins, N. Mueller, & S. Howard (Eds.), *The handbook of clinical linguistics* (pp. 165–83). Oxford: Blackwell.

Clahsen, H., Bartke, S., & Göllner, S. (1997). Formal features in impaired grammars: a comparison of English and German DLI children. *Journal of Neurolinguistics*, 10, 151–71.

Clahsen, H., & Dalalakis, J. (1999). Tense and agreement in Greek DLI: a case study. *Essex Research Reports in Linguistics*, 24, 1–2.

Cohen, N. J., Vallance, D. D., Barwick, M., Im, N., Menna, R., Horodezky, N. B., & Isaacson, L. (2000). The interface between ADHD and language impairment: an examination of language, achievement, and cognitive processing. *Journal of Child Psychology and Psychiatry*, 41, 353–62.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment. *Journal of Child Psychology and Psychiatry*, 42, 741–8.

Deevy, P., & Leonard L. B. (2004). The comprehension of wh-questions in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 47(4), 802–15.

de Jong, J. (1999). *Specific language impairment in Dutch: inflectional morphology and argument structure*. (Doctoral dissertation, University of Groningen).

Dromi, E., Leonard, L., Adam, G., & Zadunaisky-Ehrlich, S. (1999). Verb agreement morphology in Hebrew-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 42, 1414–31.

Duinmeijer, I. (2016). *Persistent grammatical differences in specific language impairment. Rule learning or rule implementation?* Utrecht, LOT.

Eisenbeiss, S., Bartke, S., & Clahsen, H. (2005). Structural and lexical case in child German: evidence from language-impaired and typically developing children. *Language Acquisition*, 13, 3–32.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language: an introduction to psycholinguistics and generative grammar*. New York: McGraw-Hill.

Franck, J., Cronel-Ohayon, S., Chillier, L., Frauenfelder, U. H., Hamann, C., Rizzi, L., & Zesiger, P. (2004). Normal and pathological development of subject–verb agreement in speech production: a study on French children. *Journal of Neurolinguistics*, 17(2), 147–80.

Franck, J., Lassi, G., Frauenfelder, U. H., & Rizzi, L. (2006). Agreement and movement: a syntactic analysis of attraction. *Cognition*, 101, 173–216.

Franck, J., Soare, G., Frauenfelder, U. H., & Rizzi, L. (2010). Object-interference in subject–verb agreement: the role of intermediate traces of movement. *Journal of Memory and Language*, 62, 166–82.

Franck, J., Vigliocco, G., & Nicol, J. L. (2002). Subject–verb agreement errors in French and English: the role of syntactic hierarchy. *Journal of Language and Cognitive Processes*, 17(4), 371–404.

Friedmann, N., & Novogrodsky, R. (2007). Is the movement deficit in syntactic SLI related to traces or to thematic role transfer? *Brain and Language*, 101(1), 50–63.

Friedmann, N., & Novogrodsky, R. (2011). Which questions are most difficult to understand? The comprehension of Wh questions in three subtypes of DLI. *Lingua*, 121, 367–82.

Grüter, T. (2005). Comprehension and production of French object clitics by child second language learners and children with specific language impairment. *Applied Psycholinguistics*, 26, 363–91.

Guasti, M.-T. (1993/94). Verb syntax in Italian child grammar: finite and non-finite verbs. *Language Acquisition*, 3(1), 1–40.

Guasti, M. T., Palma, S., Genovese, E., Stagi, P., Saladini, G., & Arosio, F. (2016). The production of direct object clitics in pre-school and primary school-aged children with specific language impairments. *Clinical Linguistics and Phonetics*, 30, 663–78.

Guasti, M. T., & Rizzi, L. (2002). Agreement and tense as distinctive syntactic positions. evidence from acquisition. In G. Cinque (Ed), *Functional structure in DP and IP: The Cartography of Syntactic Structures* (Vol. 1, pp. 167–94). New York: Oxford University Press.

Hamann, C. (2006). Speculations about early syntax: the production of wh-questions by normally developing French children and French children with SLI. *Catalan Journal of Linguistics*, 5, 143–89.

Hamann, C., Ohayon, S., Dubè, S., Frauenfelder, U., Rizzi, L., Starke, M., & Zesiger, P. (2003). Aspects of grammatical development in young French children with SLI. *Developmental Science*, 6, 151–8.

Hansson, K., Nettelbladt, U., & Leonard, L. (2000). Specific language impairment in Swedish: the status of verb morphology and word order. *Journal of Speech, Language, and Hearing Research*, 43, 848–64.

Hill, E. (2001). Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders*, 36(2), 149–71.

Hoover, J., Storkel, H., & Rice, M. (2012). The interface between neighbourhood density and optional infinitives: normal development and specific language impairment. *Journal of Child Language*, 39, 835–62.

Jakubowicz, C. (2005). *The Language Faculty: (ab)normal development and interface constraints*. Paper presented at GALA 2005, Siena.

Jakubowicz, C. (2011). Measuring derivational complexity: new evidence from typically-developing and SLI learners of L1-French. *Lingua*, 121, 339–51.

Jakubowicz, C., Nash, L., Rigaut, C., & Gérard, C.-L. (1998). Determiners and clitic pronouns in French-speaking children with SLI. *Language Acquisition*, 7, 116–60.

Kayne, R. (1989). Facets of Romance past participle agreement. In P. Benincà (Ed), *Dialect variation and the theory of grammar* (pp. 85–103). Dordrecht: Foris.

Leonard, B. L. (2014). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Leonard, B. L., Bortolini, U., Caselli, M. C., McGregor, K. K., & Sabbadini, L. (1992). Morphological deficits in children with specific language impairment: the status of features in the underlying grammar. *Language Acquisition*, 2, 151–79.

Leonard, B. L., & Dispaldro, M. (2013) The effects of production demands on grammatical weaknesses in specific language impairment: the case of clitic pronouns in Italian. *Journal of Speech, Language, and Hearing Research*, 56(4), 1272–86.

Marini, A., Marotta, L., Bulgheroni, S., & Fabbro, F. (2014). *Batteria per la Valutazione del Linguaggio in Bambini dai 4 ai 12 anni (BVL_4-12)*. Firenze: Giunti O.S.

Marinis, T. (2011). On the nature and cause of Specific Language Impairment: a view from sentence processing and infant research. *Lingua*, 121, 463–75.

Marinis, T., & Saddy, D. (2013). Parsing the passive: comparing children with specific language impairment to sequential bilingual children. *Language Acquisition*, 20(2), 155–79.

McGregor, K., & Leonard, L. (1994). Subject pronoun and article omissions in the speech of children with specific language impairment: a phonological interpretation. *Journal of Speech and Hearing Research*, 37, 171–81.

Moscati, V., & Rizzi, L. (2013). A typology of agreement processes and its implications for language development. In S. Baauw, F. A.C. Drijkoningen, L. Meroni, & M. Pinto (Eds.), *Romance languages and linguistic theory 2011: selected papers from 'Going Romance' Utrecht 2011* (pp. 143–56). Amsterdam: John Benjamins.

Moscati, V., & Rizzi, L. (2014). Agreement configurations in language development: a movement-based complexity metric, *Lingua*, 140, 67–82.

Moscati, V., & Tedeschi, R. (2009). The delay of Past Participle agreement. *Proceedings of the XXXIII Boston University Conference on Language Development* (Vol. 2, pp. 379–90). Somerville, MA: Cascadilla Press.

Novogrodsy, R., & Friedmann, N. (2006). The production of relative clauses in syntactic SLI: a window to the nature of the impairment. *Advances in Speech Language Pathology*, 8(4), 364–75.

Oetting, J., & Horohov, J. (1997). Past tense marking by children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research*, *40*, 62–74.

Orgassa, A., & Weerman, F. (2008). Dutch gender in specific language impairment and second language acquisition. *Second Language Research*, *24*(3), 333–64.

Ott, S., & Höhle, B. (2013). Verb inflection in German-learning children with typical and atypical language acquisition: the impact of subsyllabic frequencies. *Journal of Child Language*, *40*, 169–92.

Paradis, J., & Crago, M. (2001). The morphosyntax of specific language impairment in French: evidence for an extended optional default account. *Language Acquisition*, *9*, 269–300.

Pearlmutter, N. J., Garnsey, S. M., & Bock, J. K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, *41*, 427–56.

Pizzuto, E., & Caselli, M. C. (1992). The acquisition of Italian morphology: implications for models of language development. *Journal of Child Language*, *19*, 491–557.

Polite, E., Leonard, L., & Roberts, F. (2011). The use of definite and indefinite articles by children with specific language impairment. *International Journal of Speech-Language Pathology*, *13*, 291–300.

Prévost, P., Tuller, L., Barthez, M. A., Malvy, J., & Bonnet-Brilhault, F. (2017). Production and comprehension of French wh-questions by children with ASD. *Applied Psycholinguistics*, *38*(5), 1095–131.

R Core Team (2019). R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Online <www.R-project.org>.

Restrepo, M. A., & Gutiérrez-Clellan, V. (2001). Article use in Spanish-speaking children with SLI. *Journal of Child Language*, *28*, 433–52.

Rice, M. L., Noll, K. R., & Grimm, H. (1997). An extended optional infinitive stage in German-speaking children with specific language impairment. *Language Acquisition*, *6*(4), 255–95.

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research*, *39*, 1239–57.

Rispens, J. E., & Been, P. (2007). Subject–verb agreement and phonological processing in developmental dyslexia and specific language impairment (SLI): a closer look. *International Journal of Language and Communication Disorders*, *42*(3), 293–305.

Rizzi, L. (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.

Rothweiler, M., & Clahsen, H. (1994). Dissociations in SLI children's inflectional systems: a study of participle inflection and subject–verb agreement. *Journal of Logopedics and Phoniatrics*, *18*, 169–79.

Roulet, L. (2007). L'accord grammatical de genre dans la dysphasie de de ´veloppement [Grammatical gender agreement in children with SLI]. *Psychologie francaise*, *52*, 55–70.

Roulet-Amiot, L., & Jakubowicz, C. (2006). Production and perception of gender agreement in French SLI. *Advances in Speech-Language Pathology*, *8*(4), 335–46.

Royle, P., & Stine, I. (2013). The French noun phrase in preschool children with SLI: morphosyntactic and error analyses. *Journal of Child Language*, *40*(5), 945–70.

Schaeffer, J. (2000). *The acquisition of direct object scrambling and clitic placement*. Amsterdam: John Benjamins.

Stavrakaki, S. (2006). Developmental perspectives on Specific Language Impairment: evidence from the production of wh-questions by Greek SLI children over time. *Advances in Speech-Language Pathology*, *8*, 384–96.

Stavrakaki, S., & van der Lely, H. (2010). Production and comprehension of pronouns by Greek children with specific language impairment. *British Journal of Developmental Psychology*, *28*, 189–216.

Tsimpli, I. M. (2001). LF-interpretability and language development: a study of verbal and nominal features in Greek normally developing and SLI children. *Brain and Language*, *77*, 432–48.

Tsimpli, I. M., & Stavrakaki, S. (1999). The Effects of a morpho-syntactic deficit in the determiner system: the case of a Greek DLI child. *Lingua*, *108*, 31–85.

Tuller, L., Delage, H., Monjauze, C., Piller, A., & Barthez, M. (2011). Clitic pronoun production as a measure of atypical language development in French. *Lingua*, *121*, 423–41.

van der Lely, H. K. J. (1996). Specifically language impaired and normally developing children: verbal passive vs adjectival passive sentence interpretation. *Lingua*, *98*, 243–72.

van der Lely, H. K. J. (1998). SLI in children: movement, economy and deficits in the computational syntactic system. *Language Acquisition*, *72*, 161–92.

van der Lely, H. K. J. (2005). Domain-specific cognitive systems: insight from grammatical specific language impairment. *Trends in Cognitive Sciences*, 9, 53–9.

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: representations and processes. *Journal of Memory and Language*, 61, 206–37.

Appendix A: test sentences Experiment 1

Condition 1: structural intervention, plural intervening feature

(1)   la nonna le *ha/*hanno* abbracciate
       the grandmother CL$_{F,P}$ has/*have hugged
       'the grandmother has hugged them'

(2)   la bambina le *ha/*hanno* spente
       the little girl CL$_{F,P}$ has/*have switched off
       'the little girl has switched them off'

(3)   la maestra le *ha/*hanno* sgridate
       the teacher CL$_{F,P}$ has/*have scolded
       'the teacher has scolded them'

(4)   la signora le *ha/*hanno* lavate
       the lady CL$_{F,P}$ has/*have washed
       'the lady has washed them'

(5)   la dottoressa le *ha/*hanno* controllate
       the doctor CL$_{F,P}$ has/*have checked
       'the doctor has checked them'

(6)   la rana le *ha/*hanno* mangiate
       the frog CL$_{F,P}$ has/*have eaten them
       'the frog has eaten them'

Condition 2: structural intervention, singular intervening feature

(7)   le cuoche la *hanno/*ha* preparata
       the cooks CL$_{F,S}$ have/*has prepared
       'the cooks have prepared it'

(8)   le mamme la *hanno/*ha* guardata
       the mothers CL$_{F,S}$ have/*has looked
       'the mothers have looked at her'

(9)   le giraffe la *hanno/*ha* bagnata
       the giraffes CL$_{F,S}$ have/*has wet
       'the giraffes have wet her'

(10) le principesse la *hanno/*ha* salutata
       the princesses CL$_{F,S}$ have/*has said hello
       'the princesses have said hello to her'

(11) le tartarughe la *hanno/*ha* superata
       the turtles CL$_{F,S}$ have/*has passed
       'the turtles have passed it'

(12) le streghe la *hanno/*ha* preparata
       the witches CL$_{F,S}$ have/*has prepared
       'the witches have prepared it'

Condition 3: linear intervention, plural intervening feature

(13) la ragazza con le collane *ha/*hanno* sollevato la scatola
       the girl with the necklaces has/*have lifted up the box
       'the girl with the necklaces has lifted the box up'

(14) la scimmia con le scarpe *ha/*hanno* mangiato la banana
the monkey with the shoes has/*have eaten the banana
'the monkey with the shoes has eaten the banana'

(15) la nonna delle bambine *ha/*hanno* cucinato la pasta
the grandmother of the little girls has/*have cooked the pasta
'the grandmother of the girls has cooked the pasta'

(16) la bambina con le trecce *ha/*hanno* accarezzato la mucca
the little girl with the braids has/*have touched the cow
'the girl with the braids has touched the cow'

(17) la postina con le borse *ha/*hanno* consegnato la lettera
the post girl with the bags has/*have delivered the letter
'the post girl with the bags has delivered the letter'

(18) la signora con le valigie *ha/*hanno* salutato la ragazza
the lady with the bags has/*have greeted the girl
'the lady with the bags has greeted the girl'

Condition 4: linear intervention, singular Intervening feature

(19) le cagnoline della signora *hanno/*ha* rincorso la pallina
the dogs of the lady have/*has chased the ball
'the dogs of the lady have chased the ball'

(20) le infermiere della signora *hanno/*ha* controllato la lista
the nurses of the lady have/*has checked the list
'the nurses of the lady have checked the list'

(21) le galline della vicina *hanno/*ha* fatto le uova
the hens of the neighbour have/*has laid the eggs
'the hens of the neighbour have laid the eggs'

(22) le fatine con la bacchetta *hanno/*ha* toccato la bambina
the fairies with the wand have/*has touched the girl
'the fairies with the wand have touched the girl'

(23) le gattine della bambina *hanno/*ha* rotto la bottiglia
the kitties of the little girl have/*has broken the bottle
'the kitties of the little girl have broken the bottle'

(24) le spine della rosa *hanno/*ha* punto la principessa
the thorns of the rose have/*has stung the princess
'the thorns of the rose have stung the princess'

Appendix B: test sentences Experiment 2

Picture set 1

(1a) la/*le mamma la ha raccolta   Condition 1: Det–Noun violation
$det_{F,S/*F,P}$ mother $cl_{F,S}$ $aux_{3P,S}$ $plucked_{F,S}$
'the mum has plucked a rose'

(1b) la mamma la ha/*hanno raccolta   Condition 2: Subj–Verb violation
$det_{F,S}$ mother $cl_{F,S}$ $aux_{3P,S/*3P,P}$ $plucked_{F,S}$

(1c) la mamma la ha raccolta/*e   Condition 3: Cl–participle violation
det mother $cl_{F,S}$ $aux_{3P,S}$ $plucked_{F,S/*F,P}$

Picture set 2

(2a) la/*le bambina la ha mangiata   Condition 1: Det–Noun violation
det$_{F,S/*F,P}$ girl cl$_{F,S}$ aux$_{3P,S}$ eaten$_{F,S}$
'the girl has eaten it'

(2b) la bambina la ha/*hanno mangiata   Condition 2: Subj–Verb violation
det$_{F,S}$ girl cl$_{F,S}$ aux$_{3P,S/*3P,P}$ eaten$_{F,S}$

(2c) la bambina la ha mangiata/*e   Condition 3: Cl–Participle violation
det$_{F,S}$ girl cl$_{F,S}$ aux$_{3P,S}$ eaten$_{F,S/*F,P}$

Picture set 3

(3a) la/*le nonna la ha spenta   Condition 1: Det–Noun violation
det$_{F,S/*F,P}$ grandmother cl$_{F,S}$ aux$_{3P,S}$ blew out$_{F,S}$
'the grandmother has blew it out'

(3b) la nonna la ha/*hanno spenta   Condition 2: Subj–Verb violation
det$_{F,S}$ grandmother cl$_{F,S}$ aux$_{3P,S/*3P,P}$ blew out$_{F,S}$

(3c) la nonna la ha spenta/*e   Condition 3: Cl–Participle violation
det$_{F,S}$ grandmother cl$_{F,S}$ aux$_{3P,S}$ blew out$_{F,S/*F,P}$

Picture set 4

(4a) la/*le principessa la ha aperta   Condition 1: Det–Noun violation
det$_{F,S/*F,P}$ princess cl$_{F,S}$ aux$_{3P,S}$ opened
'the princess has opened it'

(4b) la principessa la ha/*hanno aperta   Condition 2: Subj–Verb violation
det$_{F,S}$ princess cl$_{F,S}$ aux$_{3P,S/*3P,P}$ opened

(4c) la principessa la ha aperta/*e   Condition 3: Cl–Participle violation
det$_{F,S}$ princess cl$_{F,S}$ aux$_{3P,S}$ opened$_{F,S/*F,P}$

Picture set 5

(5a) la/*le signora la ha chiusa   Condition 1: Det–Noun violation
det$_{F,S/*F,P}$ lady cl$_{F,S}$ aux$_{3P,S}$ closed
'the lady has closed it'

(5b) la signora la ha/*hanno chiusa   Condition 2: Subj–Verb violation
det$_{F,S}$ lady cl$_{F,S}$ aux$_{3P,S/*3P,P}$ closed

(5c) la signora la ha chiusa/*e   Condition 3: Cl–Participle violation
det$_{F,S}$ lady cl$_{F,S}$ aux$_{3P,S}$ closed$_{F,S/*F,P}$

Picture set 6

(6a) la/*le ragazza la ha accarezzata   Condition 1: Det–Noun violation
det$_{F,S/*F,P}$ girl cl$_{F,S}$ aux$_{3P,S}$ stroked
'the girl has stroked it'

(6b) la ragazza la ha/*hanno accarezzata   Condition 2: Subj–Verb violation
det$_{F,S}$ girl cl$_{F,S}$ aux$_{3P,S/*3P,P}$ stroked

(6c) la ragazza la ha accarezzata/*e   Condition 3: Cl–Participle violation
det$_{F,S}$ girl cl$_{F,S}$ aux$_{3P,S}$ stroked $_{F,S/*F,P}$

Appendix C: data analysis for Experiment 1

Table 1. Fixed effect of Group, Type of Intervention, and Number from best-fitting logistic regression of probability of correct answers

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.04	0.32	6.31	0.00 ***
S-DLD	−0.70	0.41	−1.69	0.09
Structural	0.97	0.52	1.86	0.06
Plural	0.52	0.46	1.12	0.26
S-DLD:Structural	−1.84	0.60	−3.04	0.00 **
S-DLD:Plural	−1.11	0.56	−1.96	0.04 *
Structural:Plural	−2.27	0.67	−3.36	0.00 ***
S-DLD:Structural:Plural	2.52	0.78	3.19	0.00 **

Notes. Formula in R: Answer ~ Group * Type_of_Int * Number + (1 | Item) + (1 | Subject). AIC = 905.1; BIC = 953.2; LogLik = −442.5, Deviance = 885.1. Reference categories: Group = TD, Type of intervention = Linear, Number = Singular. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table 2. Fixed effect of Type of Intervention and Number for the S-DLD Group from logistic regression of probability of correct answers

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.41	0.29	4.87	0.00***
Structural	−0.9118	0.34	−2.63	0.00**
Plural	−0.62	0.35	−1.73	0.08
Structural:Plural	0.18	0.47	0.37	0.70

Notes. Formula in R: Answer ~ Type_of_Int * Number + (1 | Item) + (1 | Subject). AIC = 579.3; BIC = 604.0; LogLik = −283.7, Deviance = 567.3. Reference categories: Type of intervention = Linear, Number = Singular. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table 3. Fixed effect of Type of Intervention and Number for the TD Group from logistic regression of probability of correct answers

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.13	0.34	6.26	0.00***
Structural	0.95	0.50	1.88	0.059
Plural	0.51	0.45	1.13	0.25
Structural:Plural	−2.31	0.65	−3.51	0.00***

Notes. Formula in R: Answer ~ Type_of_Int * Number + (1 | Item) + (1 | Subject). AIC = 325.5; BIC = 350.3; LogLik = −156.8, Deviance = 313.5. Reference categories: Type of intervention = Linear, Number = Singular. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table 4. Fixed effect of Number for the TD Group in the Structural Intervention Condition from logistic regression of probability of correct answers

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	3.25	0.53	6.11	0.00***
Plural	−1.87	0.49	−3.75	0.00***

Notes. Formula in R: Answer ~ Number + (1 | Item) + (1 | Subject). AIC = 174.4; BIC = 188.1; LogLik = −83.2, Deviance = 166.4. Reference categories: Number = Singular. Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table 5. Fixed effect of Number for the TD Group in the Linear Intervention Condition from logistic regression of probability of correct answers

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.06	0.34	5.91	0.00***
Plural	0.50	0.45	1.10	0.27

Notes. Formula in R: Answer ~ Number + (1 | Item) + (1 | Subject). AIC = 155.4; BIC = 169.1; LogLik = −73.7, Deviance = 147.4. Reference categories: Number = Singular. Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Appendix D: data analysis for Experiment 2

Table 1. Fixed effect of group and condition from best-fitting logistic regression of probability of correct answers

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.69	0.25	2.76	0.00**
nS-DLD	1.86	0.47	3.95	0.00***
TD	1.10	0.28	3.91	0.00**
SV	0.39	0.24	−1.58	0.11
DN	2.39	0.37	6.40	0.00***

Notes. Formula in R: Response ~ Group + Condition + (1 | Item) + (1 | Subject). AIC 774.6, BIC 809.8, logLik −380.3, deviance 760.6. Reference categories: Agreement Configuration = CL, Group = S-DLD. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1/

Table 2. Fixed effect of group and condition. Same as in Table 1 with Reference category = SV.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.08	0.25	4.22	0.00***
nS-DLD	1.86	0.47	3.5	0.00***
TD	1.10	0.28	3.91	0.00**
CL	−0.39	0.24	−1.58	0.11
DN	1.99	0.37	5.03	0.00***

Notes. Formula in R: Response ~ Group * Condition + (1 | Item) + (1 | Subject). AIC 774.6, BIC 809.8, logLik −380.3, deviance 760.6. Reference categories: Agreement Configuration = SV, Group = S-DLD. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.