ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Overt Subject Pronoun Use in Switch-Reference Contexts in Child Spanish Developmental Language Disorder: A Discriminant Function Analysis

Kendra V. Dickinson,[a] Pedro Antonio Ortiz-Ramírez,[b] Ana Arrieta-Zamudio,[c] John Grinstead,[b] and Blanca Flores-Ávalos[d]

[a] Department of Spanish and Portuguese, Rutgers University, New Brunswick, NJ [b] Department of Spanish and Portuguese, The Ohio State University, Columbus [c] Universidad Nacional Autónoma de México, Mexico City [d] Instituto Nacional de Rehabilitación, Mexico City, Mexico

ARTICLE INFO

Article History:
Received September 7, 2021
Revision received November 19, 2021
Accepted October 24, 2022

Editor-in-Chief: Stephen M. Camarata
Editor: Mary Alt

https://doi.org/10.1044/2022_JSLHR-21-00483

ABSTRACT

Purpose: Our study analyzes probabilistic constraints on subject expression previously found in adult Spanish in the speech of typically developing (TD) Spanish-speaking children and children with developmental language disorder (DLD). Previous work shows that children with DLD produce fewer overt subjects than typically developing children, and that the latter acquire constraints on subject expression as they age into adolescence. Our study complements these findings and provides further substance to the grammatical profile of children whose morphosyntactic development diverges from that of typically developing children.

Method: Data are drawn from unstructured spontaneous production data from a sample of 19 monolingual Mexican, Spanish-speaking children, collected in 2006–2007. This sample includes 19 children diagnosed with DLD and 19 age-matched, typically developing children. We collected all instances of finite verbs that either did or could have occurred with a subject personal pronoun uttered by the child participants and coded them for several factors including tense–mood–aspect, switch reference, and person and number.

Results: We find that children with DLD produce fewer overt subject pronouns in switch reference contexts than typically developing controls, with a significant interaction of group and switch reference. Furthermore, a discriminant function analysis shows that overt pronoun use in switch reference contexts can form part of a useful diagnostic discriminant function, with high levels of sensitivity and specificity.

Conclusions: Overall, we find important differences between TD Spanish-speaking children and those diagnosed with DLD regarding rates of overt subjects and sensitivity to the probabilistic constraint of switch reference. This finding contributes to our understanding of the morphosyntactic profiles of children with DLD, as well as the utility of factors such as switch reference in the identification of language disorders.

The accurate diagnosis of developmental language disorder (DLD)[1] in young children is critical. Underdiagnosis can result in children and families not getting the help they need to cope with children's struggles with language and, later, other areas of scholastic achievement that depend on language, including mathematics (e.g., Grindal et al. 2019; Morgan, 2021). Overdiagnosis, which is also too common, especially in language-minority populations in the United States (e.g., Baugh, 1995; Pray, 2003),

Correspondence to Kendra V. Dickinson: kendra.dickinson@rutgers.edu. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

[1] We will use the term *DLD* in this article following the CATALISE recommendations (Bishop et al., 2017). We have only included children with greater than 85 nonverbal IQ in this study, due to the poorly understood interaction of nonverbal IQ and discourse pragmatics. We aspire to eventually broaden our study to less-constrained groups of children.

can mean that children get inappropriately tracked into special education classes that divert them from financially lucrative academic paths.

To accurately diagnose monolingual child Spanish speakers with DLD, some aspects of morphosyntax do not seem useful, including plural marking and noun–adjective agreement (Castilla-Earls et al., 2020; J. Grinstead et al., 2008). In contrast, other aspects of morphosyntax do seem useful and these aspects include what could be thought of as "definites" in the semantic sense of having referents that are familiar to both the speaker and the hearer. This is true in the domain of noun constructions for definite noun phrases (e.g., Anderson & Souto, 2005; Restrepo & Gutiérrez-Clellen, 2001) and clitic pronouns (e.g., De la Mora, 2004; Merino, 1983) and in the domain of verb constructions for finite verb tense (e.g., J. Grinstead et al., 2013). Of course, the intended sense of "useful for diagnosis" here refers to grammatical constructions that are particularly non-adultlike in their use by children diagnosed with DLD. As explained in greater detail below, the Interface Deficit Hypothesis (J. Grinstead et al., 2013) attempts to account for the fact that aspects of morphosyntax that require children to be sensitive to what is prominent in conversation are more non-adultlike than are those that do not require such sensitivity.

How do overt versus null subjects in Spanish fit into this picture? Can they be helpful in identifying children with DLD? In what follows, we explore the potential diagnostic relevance of variationist morphosyntactic analyses of overt subject pronoun use in monolingual Spanish-speaking children. Specifically, we pursue the approach of identifying not the grammatical versus ungrammatical use of morphemes, but rather the probabilistic tendencies of certain linguistic and discourse-pragmatic factors to predict overt subject pronoun use, and we consider the utility of these probabilistic constraints for identifying monolingual Spanish-speaking children with language disorders, following the variationist approach to morphosyntax, rooted in the work of William Labov (Labov, 1973). This work is concerned with the social characteristics of speakers, but also with elements in the linguistic context that may positively or negatively predict the use of linguistic structures. In considering the interaction of variationist morphosyntax and research in language disorders, this work follows Sossaman (2017), Oetting (2019), Oetting et al. (2019), and Oetting et al. (2021) with African American and other varieties of English.

Overt subjects are not grammatically obligatory in finite Spanish sentences. Consider Sentence (2) below, where the verb *tengo* ("I have") occurs with the overt subject pronoun *yo* ("I"), whereas the verb phrase *quieres comer* ("You want to eat") does not. This is worth noting because the situation could have been the reverse for each verb token while, nonetheless, maintaining the same truth conditions and still conforming to the morphosyntactic norms of Spanish.[2]

(1) Speaker 1: ¿Qué Ø hacemos ahora?
what pro.1PL do.1PL.PRES now
"What shall we do now?"

(2) Speaker 2: Yo tengo hambre. ¿Quieres comer Ø?
pro.1SG have.1SG.PRES hunger
want.2SG.PRES eat.INF
"I am hungry. Do you want to eat?"

This variation, namely, the presence/absence of subject personal pronouns (SPPs) with finite verbs, has been extensively studied among adult speakers of Spanish, both in terms of overall rates of usage as well as probabilistic constraints on expression (see Otheguy & Zentella, 2012, for a review). Some research has also been done regarding the acquisition of the SPP constraints and usage patterns in child Spanish, notably Shin and Cairns (2012), Shin and Erker (2015), Shin (2016), and Forsythe (2018). However, absent from this work is the consideration of how children diagnosed with DLD compare with typically developing (TD) children in their acquisition of these constraints.

In this study, we analyze SPP expression among Spanish-speaking TD children and children diagnosed with DLD in Mexico City. We analyze rates of expression, types of overt subjects (full determiner phrases, pronouns, demonstratives, and so forth), as well as grammatical and pragmatic constraints previously found to be significantly predictive of overt pronoun use in studies of adult language. In so doing, we provide further substance to the grammatical profile of children whose morphosyntactic development diverges from that of TD children.

Spanish DLD—Interface Deficit

What is the nature of the DLD? It seems likely that there is more than one locus, with affected children showing significant differences in an array of cognitive and linguistic abilities, including at least working memory, inhibition, lexical development, mathematical ability, and literacy outcomes (see Leonard, 2014, for a review). One line of research has for some time attempted to develop a theory of at least one important aspect of the disorder that seems to prevent syntactic and morphosyntactic constructions from interacting with discourse pragmatics in adultlike ways. This phenomenon has been referred to in TD children as *interface delay* (J. Grinstead, 2004; J. A. R. Grinstead, 1998) and in children with DLD as *interface deficit* (J. Grinstead et al., 2013, 2018).

[2]The pragmatics of the reverse situation would have put the focus on Speaker 1, instead of on Speaker 2, as in the current discourse.

While children with DLD appear to not have difficulties using aspects of morphosyntax that are relatively insensitive to discourse-pragmatics, such as plural marking on nouns (J. Grinstead et al., 2008 in Spanish; Oetting, 1992; Rice & Oetting, 1993 in English) or noun–adjective agreement in Spanish (J. Grinstead et al., 2008), it has been shown, in contrast, that what has been referred to as *discourse-sensitive* morphosyntactic constructions are difficult for them. *Discourse-sensitive* in the intended sense refers to the requirement that the speaker considers the interlocutors' perspective on the prominence of antecedents in the Conversational Common Ground, in the sense of Stalnaker (1974). In the domain of noun-related constituents, the set of such constructions that have been studied in monolingual child Spanish DLD includes direct object clitics (De la Mora, 2004; Jackson-Maldonado & Maldonado, 2017; Merino, 1983; Morgan et al., 2009), definite articles (Anderson & Souto, 2005; Restrepo & Gutiérrez-Clellan, 2001), and overt subjects (J. Grinstead et al., 2018). In the verbal domain, it has been argued that finite verbs also constitute a type of "definite" in the sense that verb tense is fundamentally anaphoric, requiring the speaker to assume that the interlocutor can access a prominent speech time-event time antecedent in discourse (e.g., J. Grinstead et al., 2013; J. Grinstead, De la Mora, Vega-Mendoza, & Flores, 2009). The claim of interface deficit is that, in the absence of adultlike abilities to manage the morphosyntax-discourse pragmatic interface, children overassume familiarity of these antecedents, which produces an "egocentric" character in the speech of children, in Piaget's terms. This means using direct object clitic pronouns, definite noun phrases, and null subjects for antecedents that interlocutors are not in fact familiar with.

To be complete, Interface Deficit does not attribute the non-adultlike use of discourse-sensitive morphosyntax to the immaturity of theory of mind development, though children with DLD definitely have lower theory of mind scores than do controls (see Nilsson & de López, 2016, for a relevant meta-analysis). Rather, we note that subcomponents of theory of mind, including belief-tracking (Onishi & Baillargeon, 2005) and intention-tracking (Woodward et al., 2009), appear to be available to infants in the 12- to 15-month-old age range. However, discourse-sensitive aspects of morphosyntax may take much longer than this to develop (e.g., child English definite articles, in the work of Maratsos, 1974, 1976, are still developing at 4 and 5 years of age). It is precisely for this reason that we believe that it is neither discourse-pragmatics itself nor developing theory of mind that is developing to allow the interface between morphosyntax and discourse-pragmatics, but rather the interface itself.

As noted elsewhere in the literature, nonfinite verbs are also used when interlocutors are not familiar with the temporal specification intended by the child (Castilla-Earls et al., 2018, 2020, 2021; J. Grinstead et al., 2013; J. Grinstead, De la Mora, Vega-Mendoza, & Flores, 2009). Note in the following discourse, for example, that even adults many times do not overtly mark tense on verbs, but rather infer it from discourse.

(3) A: ¿Qué quieres hacer?
what want-PRES.2SG do.INF
"What do you want to do?"

B: Comer.
eat.INF
"Eat."

(4) A: ¿Qué querías hacer?
what want-IPFV.2SG do.INF
"What were you wanting to do?"

B: Comer.
eat.INF
"Eat."

(5) A: ¿Qué vas a querer hacer?
what go-PRES.2SG to want.INF do.INF
"What will you want to do?"

B: Comer.
eat.INF
"Eat."

Across these three exchanges, the temporal interpretation of the of the unrealized event in the answer "Eat." is either past, present, or future with respect to the moment of speaking (speech time, in the terms of Reichenbach, 1947), though the verb "comer" does not change morphologically. Similarly, children may assume that their interlocutors know what speech time–event time relation they intend, although they have not specified it morphologically. This is illustrated in the following bare stem example produced by a monolingual child Spanish-speaker diagnosed with DLD, from J. Grinstead et al. (2014, p. 54, Example 4). These are the most common nonfinite verb forms in typical and atypical child Spanish, though usually there is not an overt subject pronoun that occurs with them to make the non-adultlike nature of the verb unambiguous.

(6) *The investigator talks to the child about what the child has asked Santa Claus for.*

Investigator: Le hacen una carta a Santa Claus.
You write a letter to Santa Claus.

Child: Y yo pide un carro de *Batman Inicia*.
And I ask for (root + theme vowel)
a car from Batman Begins

Investigator: ¿Pediste un carro de *Batman Inicia* o lo vas a pedir apenas?
You asked for a Batman Begins car or you are about to ask for it?

If this is correct, then child Spanish speakers' production of nonfinite verbs falls into a semantic natural class of overassumed definites, with the nominal definite direct object clitic pronouns, null subjects, and definite noun phrases marked with definite articles. In this project, we pay special attention not to clitic pronouns, but rather to tonic personal pronouns used in subject position, yet another semantic definite that could be overassumed to be familiar to interlocutors by children with DLD, on the interface deficit hypothesis.

Pronominal Subjects in Adult Spanish

To compare the development of adultlike constraints on SPPs among children, it is necessary to establish what is known about constraints on adult subject expression in null subject languages. Subject expression in adult Spanish has a long-standing research tradition, with a particular focus on the presence versus absence of SPPs, examining both rates of and constraints on expression. Regarding rates of expression, a summary by Orozco and Guy (2008) of previous work on the topic notes wide-ranging regional variation, though we will be focusing on the Spanish of Mexico City for our study. For example, Cameron (1993) found rates as high as 44.7% for a group of Puerto Rican speakers, while Otheguy et al. (2007) found rates as low as 19% among Mexican speakers in New York City. Because the children in this study are from Mexico City, it is also critical to note previous findings on overall pronoun rates among adults of the same origin. In their study of an oral corpus of Spanish of Mexico City, Lastra and Martín Butragueño (2015) found that overt SPPs were used 21.7% of the time. A general observation that follows from these and other studies is that Caribbean speakers make greater use of SPPs than do speakers from the Latin American Mainland, but null subjects are still more common cross-dialectally than are overt SPPs.

Despite differences in overall rates, a striking degree of homogeneity has been observed with regard to the linguistic factors that shape SPP expression in Spanish. The factors that have consistently been found to be significant predictors of patterns of pronoun usage include person/number subject–verb agreement and tense–mood–aspect (TMA) marking of the associated verb, discourse continuity as measured by switch/maintenance of referent, and reflexivity of the verb, among others. This pattern of variation has been used by researchers to situate the linguistic behavior of different groups of speakers within the context of established trends in subject pronoun use rates as well as to model linguistic behavior through the investigation of language internal predictors (Cameron, 1993; Erker & Otheguy, 2016; Flores-Ferrán, 2002; Otheguy & Zentella, 2012; inter alia).

Verbal person and number agreement have been found to be important predictors of SPP expression across the Spanish language. For example, Cameron (1993) found that speakers in San Juan, Puerto Rico, used overt SPPs 50% of the time in the singular and 26% in the plural, and that speakers in Madrid, Spain, used overt pronouns 19% and 7% of the time in the same contexts. This illustrates that while overall rates of SPP expression are different between the two speech communities, both groups are sensitive to person and number subject–verb agreement, employing fewer overt subjects in the plural. Similar patterns have been found across other dialects of Spanish (cf. Otheguy et al., 2007; Otheguy & Zentella, 2012; Silva-Corvalán, 1994; inter alia).

In Mexico City in particular, Lastra and Martín Butragueño (2015) found that verbal person and number agreement was the strongest predictor of subject usage patterns in their study of adult oral Spanish from the city. More specifically, their study found that third-person singular verbs were the most likely to occur with overt SPPs (27.1%), followed by first-person singular (24.7%), and then second-person singular (16.6%). Overt SPPs were employed with plurals less often than with singulars, 11.9% and 10.6% of the time with third- and first-person plurals, respectively.

Switch reference, also termed *discourse continuity*, has also been found to play an important role in SPP expression in Spanish. A token would be considered a switch in reference when the subject of the preceding finite verb is different from the subject of the token in question. Conversely, a token would be considered to have the same referent when the subject of the preceding finite verb is the same as the subject of the verb in question (cf. Otheguy & Zentella, 2012; Silva-Corvalán, 1982; inter alia). Across varieties of Spanish, it has been found that speakers tend to employ higher rates of overt SPPs in cases of switch reference than in cases in which the referent is maintained. For example, Silva-Corvalán (1982) found that speakers of Mexican Spanish in Los Angeles used overt SPPs 25% of the time in cases of continuity of reference but used overt SPPs 53% of the time in the cases of a switch in reference. With regard to Mexico City in particular, Lastra and Martín Butragueño (2015) found that speakers employed overt SPPs in cases of switch reference 32.7% of the time and did so only 16.3% in cases of reference continuity.

Verbal TMA have also been found to play a role in subject expression in adult Spanish. More specifically, it has been found that verbal forms with ambiguous morphology (i.e., paradigms in which the same verbal morphology is used for more than one person and number combination) such as the imperfect indicative and the

conditional, which neutralize agreement for first- and third-person singular forms, promote overt SPP expression. Conversely, paradigms with unambiguous verbal morphology, such as the preterit (simple past perfective), promote the use of null subjects (Erker & Guy, 2012; Otheguy et al., 2007; inter alia). In Mexico City, Lastra and Martín Butragueño (2015) found that overt subjects were used 31.4% of the time with verbs in the imperfect, where the first- and third-person singulars share identical verbal morphology, but only 15.5% of the time with preterit verbs, where each person and number combination has its own unique verbal morphology.

Finally, the presence of additional linguistic mechanisms that encode subject reference, such as a clitic or reflexive pronoun that is coreferential with the subject, play a role in subject expression in Spanish. Specifically, verbs occurring without a reflexive clitic pronoun promote pronoun expression more than verbs with reflexive clitic pronoun (Carvalho & Child, 2011; Otheguy & Zentella, 2012; inter alia). For example, in his study of Yucatán Spanish, Michnowicz (2015) found that SPPs were used 14% of the time in the presence of a coreferential reflexive clitic (e.g., *Yo me senté.* "I sat down."), but 20.5% of the time without one (e.g. *Canté mi canción favorita.* "I sang my favorite song."). It has been posited that this is because the presence of a coreferential clitic pronoun can serve the same function of disambiguating reference as an SPP does (see Silva-Corvalán, 1982, p. 109).

Other factors have been found to influence subject expression in Spanish, including syntactic priming (Torres Cacoullos & Travis, 2011), verbal lexical frequency (Erker & Guy, 2012), speech genre (Travis, 2007), verb semantics (Erker & Guy, 2012), and clause type (Otheguy & Zentella, 2012; Shin & Montes-Alcalá, 2014), among others. However, the factors detailed in this article, namely, verb person and number and TMA, switch reference, and reflexivity, have been found cross-dialectally to be some of the strongest and most common predictors of overt SPP use in Spanish.

Pronominal Subjects in Child Spanish

Now, let us turn to subject expression in child Spanish, the focus of this study. First, it is critical to note that studies have found that TD Spanish-speaking children use far fewer overt subjects than do adults (J. Grinstead, 2004; J. A. R. Grinstead, 1998; J. Grinstead & Spinner, 2009; Shin & Cairns, 2012). In a study of a related language, Italian, Serratrice (2005) found that accessibility and unique identifiability of a referent in discourse (whether anaphorically or deictically), as well as the shared knowledge of interlocutors, impact children's expression of subjects. Her findings show that uninformative and discourse-activated referents are more likely to be expressed as null subjects, suggesting that children are sensitive to information structure. Furthermore, the author found that as child mean length of utterance in words (MLUw) increased, so did expression of all types of overt subjects, including demonstrative pronouns, proper names, full noun phrases (determiner phrases), bare noun phrases, indefinite pronouns, and quantifiers. However, the greatest increase was attributed to subject expression among all overt subject options. Her findings show that as children begin to express more complex grammar, rates of null subjects decrease, and those subjects that are expressed are increasingly instantiated as SPPs, as in Valian and Eisenberg (1996) for Portuguese.

Further studies have focused on SPPs and the comparison between adult and child usage patterns. In one such study, Shin and Erker (2015) compare the use of SPPs among 24 monolingual Spanish-speaking children in Oaxaca, Mexico, to Mexican-born adults living in New York City, finding both similarities and differences between the two groups. First, the authors find that children (ages 6–8 years) use far fewer (9%) overt SPPs overall than the adults in the study (21%). In addition, they also find that female children used more overt SPPs than the male children did, following previous studies that have shown that female language development outpaces male acquisition (see Lange & Zaretsky, 2021, for a review). Furthermore, they find that children were sensitive to the person–number combination of the verb, using an overt SPP with first-person singular (1SG) verbs and in switch reference contexts most often. Several other studies (Austin et al., 1998; Shin & Cairns, 2012) also find the highest rates of SPP expression in children with 1SG verbs, noting similarity to adult Spanish regarding the constraints on, if not the proportion of, SPP expression.

This points to a finding that has been put forth from several other studies, namely, that child Spanish acquirers tend to exhibit conservative behavior with regard to SPP expression when compared with adults (Austin et al., 1998; Shin, 2016; Shin & Cairns, 2012; Shin & Erker, 2015; inter alia). For example, in addition to lower rates of SPP expression overall, Shin and Cairns (2012) found particularly low rates (6.1% presence) with third-person singular (3SG) referents among Mexican children.

She suggests that children in the age range of her sample (6–8 years old) tend to alternate between full lexical noun phrases (NPs) and null subjects for 3SG referents, to the exclusion of overt SPP subjects. Shin suggests that the motivations for this observation are multiple. On one hand, she suggests that children may not yet be fully sensitive to discourse-pragmatic constraints on 3SG SPP expression. On the other hand, she notes that in many cases, it is unclear whether children are simply omitting overt 3SG SPPs or if they are replacing overt SPP with full lexical NPs. This is an important distinction, because it raises the question of what is occurring in the underlying system for subject expression that Spanish-acquiring

children are working with. They are overusing lexical NPs and they omit SPPs. Which of these behaviors, if any, is compensatory for the other? To that end, they urge that "...all types of subjects with third person referents need to be included in future studies to thoroughly answer this question" (Shin & Cairns, 2012, p.138). In a related language, Serratrice (2005) suggests that 3SG subjects are most often omitted in child Italian because 3SG verbs are the most frequent in their language. These various analyses suggest not only that there is a combination of contributing factors to the observed patterns in child subject expression, but that future research is needed to investigate their underlying motivations.

Beyond purely grammatical analyses of child subject expression, several studies have also endeavored to investigate the roles of verbal morphology and pragmatic versus syntactic antecedents in the licensing of null subjects in child Spanish. For example, Austin et al. (1998) remark that in general, Spanish-speaking children master verbal morphology earlier than English-speaking children, but that because pro-drop in Spanish is not a purely syntactic phenomenon, acquirers take time to develop adultlike usage patterns. In the same vein, the authors also claim that Spanish-acquiring children understand the global possibility for pro-drop quickly but need to acquire its discourse-pragmatic constraints.

This claim is corroborated to some degree by Shin and Erker (2015), who find that children's use of overt SPPs increases with age. Relatedly, Shin (2016) found that though 6- to 7-year-olds did not show preference for overt subjects in switch reference contexts, 8- to 9-year-olds did, but not at adult levels, and finally that children ages 14–15 years reached adult levels of overt SPP expression in these contexts, suggesting that children acquire these patterns over time. Several studies have also found that TD monolingual Mexican children begin to become sensitive to certain discourse pragmatic constraints such as switch reference and person and number by approximately age 8 years (Shin & Cairns, 2012; Shin & Erker, 2015; Shin, 2016), but acquire others, such as verb semantics, clause type, and TMA later, and do not appear adultlike until adolescence (Shin, 2016).

Discriminant Function Analyses of Child Spanish DLD

Above, we saw the claim that children diagnosed with DLD are known to have problems with discourse-sensitive morphosyntactic constructions, that overt subject pronouns are such a discourse-sensitive morphosyntactic construction, and that TD children do not acquire adultlike probabilistic constraints on the use of these subject pronouns until well into elementary or even early secondary school years. Based on these facts, overt subject pronouns would appear to be good prima facie candidates for an interface delay-type construction in TD children, and for an interface deficit-type construction in Spanish-speaking children diagnosed with DLD. At the level of basic science, it is of interest to know whether non-adultlike use of overt subject pronoun usage constraints forms part of the linguistic profile of Spanish-speaking children with DLD. It gives greater theoretical weight to the interface deficit construct and opens new avenues for exploring the cognitive systems that underlie this atypical set of phenotypes.

At the level of applied science, it is of interest to know whether Spanish-speaking children's use of overt subject pronouns could aid in the diagnosis and treatment of DLD. In previous research with monolingual Spanish-speaking children with DLD (J. Grinstead et al., 2013, 2018), it has been shown that a combination of different measures of Spanish grammar can provide high levels of sensitivity (ability to accurately identify children with DLD) and specificity (ability to accurately identify typically developing children). However, there has yet to be reported, to our knowledge, a discriminant function analysis with children diagnosed with DLD, using probabilistic language use constraints, such as those that influence overt subject pronoun use, as part of the discriminant function.

What has been reported along these lines includes the role of tense marking, an elicited production measure of which, in J. Grinstead et al. (2013), was shown to provide 89% sensitivity and 89% specificity, when measured using an elicited production task. The same study found slightly higher rates of sensitivity adding in spontaneous production measures of (morpho-) syntax (MLUw, mean length of utterance in morphemes, and the Subordination Index) and a receptive measure of tense, but specificity was somewhat reduced. In J. Grinstead et al. (2018), it was shown that the overall rate of overt subject use (all grammatical overt subject types in all discourse contexts) could also be useful in a discriminant function analysis. In particular, the authors showed that while rate of overt subjects by itself did not produce excellent results (66.7% sensitivity and 92.9% specificity), adding the rate of overt subjects to the two previously mentioned tense measures and MLUw yielded a discriminant function capable of 91.7% sensitivity and 92.9% specificity, which are considered "good" rates of sensitivity and specificity (Plante & Vance, 1994).[3]

In summary, we see that there has been substantial work identifying the constraints that influence overt

[3]Methodologically, it is worth noting that the first of the two mentioned reports was a nonparametric K Nearest Neighbors discriminant function analysis, due to the nonnormal distribution of the tense data, and the second was a quadratic discriminant function analysis, because of the multivariate heteroscedasticity of the subject data, though it was normally distributed.

subject pronoun use in Spanish and that children obey some, but not all, of these constraints. Furthermore, we have seen that linguistic dimensions of morphosyntax, measured in receptive and expressive measures of tense, overall rate of all overt subject use, and mean length of utterance, have been successfully used in discriminant function analyses to identify children with DLD. In what follows, we attempt to determine whether TD monolingual Spanish-speaking children and those with DLD show different behavior with respect to their use of these probabilistic overt subject pronoun use constraints and whether such differences as may exist could be diagnostically useful.

Research Questions

1) Given that children with DLD have been shown to use fewer overt subjects, are the grammatical subject types (pronoun, full NP, and demonstrative) distributed differently between preschool-age children with DLD and age-matched controls?
2) Are Spanish-speaking preschool-age children with DLD sensitive to the same probabilistic constraints on subject pronoun expression as their TD counterparts?
3) If not, in what way are the constraints on their use of overt subject pronouns different from TD aged-matched controls?
4) If there are significant differences in subject overt pronoun use between TD Spanish-speaking children and those with DLD, can such a difference profitably form part of a discriminant function that will allow for high levels of sensitivity and specificity in distinguishing these groups of children?

Method

Participants

Data are drawn from unstructured spontaneous production data from a sample of 19 monolingual Mexican, Spanish-speaking children, collected in 2006–2007. Informed consent was obtained via a university institutional review board–approved consent form, signed by a parent or guardian, for each participant in the study. This sample includes 19 children diagnosed with DLD (M_{age} = 66.8 months, SD = 7.0) and 19 age-matched, typically developing children (M_{age} = 66.8 months, SD = 6.0), with the mean ages of the two groups not being statistically different from each other ($p > .05$). The sex of the children in the samples was not matched. There were 12 male and seven female children in the DLD group, whereas there were nine male children and 10 female children in the TD group. The children in the DLD group were diagnosed following the conventions in Leonard (1997, 2014). Children passed an American Speech-Language-Hearing Association protocol hearing test at conventional levels, showed no signs of atypical oral structure, and gave no evidence of atypical social interactions nor of frank neurological damage. Furthermore, children were given a list of two-syllable nonce words to imitate, simulating the segmental and phonotactic structures of Spanish verbs. Children had to successfully imitate four of five such nonce words to ensure the absence of articulatory difficulties. Finally, children were given four subtests of the locally normed Batería de Evaluación de Lengua Española (BELE; Rangel et al., 1988), two of which were lexical (expressive and receptive) and two of which were morphosyntactic (expressive and receptive). In order to be included in our DLD sample, children had to have scores of at least 1.25 *SD*s below the mean for their age on this standardized test, on one expressive and one receptive BELE subtest. Several dozen children who appeared to fit the DLD profile were excluded for having what appeared to be severe articulatory difficulties. Our TD children had to score within 1 *SD* of their age mean on all tests.[4] Children in our sample come from a speech and hearing clinic and from a day care/preschool that serve a broad socioeconomic spectrum of Mexican society. We collected linguistic background on children to assure that they were monolingual, but did not collect further racial or ethnic information. At the time, the local members of the research team concluded that the children with DLD and the age controls were of similar enough SES and no detailed SES data were collected. All of the children diagnosed with DLD were receiving speech-language services at the same institution. Participant characteristics are summarized in Table 1.

Procedure

Language Sample Transcription and Reliability

A roughly 20-min unstructured spontaneous production language sample from each child was taken by members of the research team, who were strangers to the children and were native speakers of Mexico City Spanish. Children's language was elicited by asking them about what they do with their friends, their families, or about their favorite movie. We note that while it may be possible to use a more focused, elicited production-type task to produce a language sample, we have here followed the convention in variationist sociolinguistics to work with the most ecological and least contrived conversational context. In this way, our results could be maximally comparable to other work with both adult and child Spanish-speaking populations. For example, the sociolinguistic interviews

[4]The BELE was the only locally normed, developmental test of Spanish in 2006–2007, when these data were collected, though sensitivity and specificity rates for diagnosis of DLD were not given when it was published.

Table 1. Participant characteristics.

Variable	DLD	TD
Age range	56–76	57–77
Age in months (*SD*)	66.8 (7.0)	66.8 (6.0)
BELE 4 subtest composite (*SD*)	18.9 (4.06)	39.8 (10.41)

Note. DLD = developmental language disorder; TD = typically developing; BELE = Batería de Evaluación de la Lengua Española.

used in Shin and Erker (2015) and Shin (2016) uses prompts such as "Tell me about the best day of your life." or "Tell me about a birthday party." These types of prompts tend to elicit narrative discourse, as did our prompts. Though we do not call them sociolinguistic interviews, we believe that the product that was elicited was comparable.

The transcribers and coders for this project were native-speaker undergraduate students in linguistics and neuropsychology at Mexican universities, from Mexico City. For both transcription and coding, research assistants were normed on an initial, common set of recordings. During norming, a recording was transcribed or coded and agreement was checked. Norming was complete when transcribers or coders reached at least 90% agreement. Then, each transcript was checked by a second transcriber and half of all transcriptions had 10% of the utterance retranscribed by a second transcriber. This 10% of utterances was compared between the first and the second transcripts and showed agreement between 90% and 99% at the word level. Mean agreement was 95.4%. Finally, for each transcriber's number of words per utterance, a Krippendorf's alpha interrater reliability coefficient (Hayes & Krippendorf, 2007) was calculated, and the values ranged between .994 and .998, with a mean value of .974.

Coding

We collected all instances of finite verbs uttered by the child participants in each interview. Then, each finite verb was evaluated for inclusion in or exclusion from the study. Included tokens consisted of all instances of finite verbs that either did occur with an overt subject or could have occurred with an overt subject but did not. Children in the DLD group produced an average of 38.58 finite utterances that were included (*SD* = 15.35; range: 18–70), whereas children in the age-matched TD sample produced 67.84 finite utterances that were included (*SD* = 39.0; range: 14–170). Excluded tokens included those that did not occur with an overt subject and could not have. For example, (6) and (7) occur with an overt subject, (8) does not but could have, and (9) does not and could not have.

(6) Elisa quiere pan. *included*
Elisa want-PRES.3SG bread.
"Elisa wants bread."

(7) Yo quiero pan. *included*
I want-PRES.1SG bread.
"I want bread."

(8) Ø Quiero pan. *included*
want-PRES.1SG bread
"I want bread."

(9) Hace frío. *excluded*
make-PRES.3SG cold
"It is cold."

Then, each included token was coded for predictor variables following methods used in previous studies of probabilistic constraints on SPP expression in adult Spanish (see Otheguy & Zentella 2012). Our coding procedures showed high rates of interrater reliability. A sample of 10% of children in the study was taken and coded independently by two raters. Then, Cohen's kappa (Cohen, 1960) was used to determine the degree of agreement between the two raters' determinations of inclusion of tokens in the study, and linguistic predictors. Raters showed perfect agreement regarding the inclusion versus exclusion of tokens in the study, kappa = 1 (95% CI, .100), $p < .001$. With regard to predictor variables, raters showed substantial or near-perfect agreement, kappa = .758–.1 (95% CI [.843–.100]), $p < .001$. Predictor variables considered in this study are as follows.

Verb Person and Number

Verb person and number have been found to be important predictors of SPP expression across dialects of Spanish, though some previous work has favored categorizing person and number based on pronominal referent and not verbal morphology (see Shin, 2016). The difference between pronominal and verbal referent in Spanish is notably with the second-person singular pronoun "usted," which takes verbs with a third-person singular morphology. In this analysis, we considered both verbal and pronominal person and number, but found that they were identical in our data set. In our data, verb person and number included five factor levels: first-, second-, and third-person singulars, and first- and third-person plurals. Items (9) and (10) in Table 2 exemplify the morphosyntactic dimensions coded for this factor.

Verb TMA

In this analysis, verbal TMA included 14 possible levels: present indicative, preterit indicative, imperfect indicative, periphrastic future, future indicative, conditional, present subjunctive, imperfect subjunctive, present perfect, pluperfect, pluperfect subjunctive, future perfect, conditional perfect, and imperative. Items (11) and (12) in Table 2 illustrate the coding for this factor.

Table 2. Overt subject personal pronoun predictor variables with examples.

Coding factor		Example	Coding value
Verbal person and number	(9)	Yo me dormí. I REFL.1SG sleep-PST.1SG "I fell asleep."	first-person singular
	(10)	Él come pasto. He eat-PRES.3SG grass. "He eats grass."	third-person singular
Verbal tense-mood-aspect	(11)	Yo me dormí. I REFL.1SG sleep-PST.1SG "I fell asleep."	preterite
	(12)	Yo tengo un juguete. I have.PRES.1SG a toy "I have a toy."	present indicative
Discourse continuity (switch reference)	(13)	Interviewer: ¿Cuántos años cumple Mariana? How many years reach.PRES.3SG Mariana? "How old is Mariana turning?" Child: No me acuerdo. No REFL.1SG remember.PRES.1SG "I don't remember."	different referent
	(14)	Child: Nosotros jugamos y cantamos We play.PRES.1PL and sing.PRES.1PL "We play and we sing."	same referent
	(15)	Interviewer: ¿Con quién vives? With whomlive.PRES.2SG "Who do you live with?" Child: Vivo con mi mamá y mi papá. live.PRES.1SG with my mom and my dad "I live with my mom and my dad."	same referent
Coreferential pronoun	(17)	Yo me vi en el espejo. I REFL.1SG saw-PST.1SG in the mirror "I saw myself in the mirror."	co-referential pronoun
	(18)	Yo tengo un jugete. I have.PRES.1SG a toy "I have a toy."	no co-referential pronoun

Switch Reference

This measure of discourse continuity was operationalized as two possible levels: same referent and different referent. A token would be coded as having a different referent (i.e., a switch in reference) when the subject of the preceding finite verb was different from the subject of the token in question. Conversely, a token would be considered to have the same referent when the subject of the preceding finite verb was the same as the subject of the verb in question. The underlined verbs in (13), (14), and (15) in Table 2 illustrate coding for this factor.

Coreferential Reflexive Pronoun

In this analysis, the coreferential reflexive pronoun variable had two factor levels: verbs that occurred with a reflexive pronoun and those that did not. Items (17) and (18) in Table 2 illustrate this distinction. Finally, we also considered several individuals factors for each participant, including continuous measures of age and MLUw, as well as group (DLD vs. TD) and sex (male vs. female).

In summary, we coded the spontaneous production transcripts of a group of preschool-age children identified using Leonard's (1997, 2014) conventional diagnosis of DLD, and an age-matched control group, for overt subject pronoun expression. Together with our conventional DLD diagnosis, we have followed an additional set of conventions from the field of morphosyntactic variation, which has a rich tradition of studying the factors that act as probabilistic predictors of SPP occurrence in Spanish. This approach to studying the language of children with DLD is novel, inasmuch as it takes into account factors that do not have the "grammatical vs. ungrammatical" nature of much of morphosyntax. Rather, the probabilistic constraints we have coded for exist very much at the core interface between

morphosyntax and discourse pragmatics that the Interface Deficit Hypothesis claims to be relevant for understanding an important part of developmental language disorder.

Results

Overview

In the entire sample, children produced 2,021 instances of finite verbs that either did appear or could have appeared with an overt subject. Overall, 70% (n = 1415) were null subjects, and 29.3% (n = 606) were overt. However, rates of overt subjects significantly differed between TD and DLD groups, $t(36) = -3.475$, $p = .001$, as shown in Table 3.

However, despite differences in overall rates of subject expression, the distribution of grammatical subject types was generally consistent across DLD and TD groups. That is to say, the vast majority of tokens for both TD and DLD groups were null subjects, full lexical NPs, and subject pronouns for both groups, as in Serratrice (2005) for child Italian.[5] None of the differences in mean rates of expression of grammatical subject type were statistically significant ($p > .05$).

Inferential Analysis

In order to conduct an analysis that is directly comparable to previous research on adult and child SPP expression in Spanish, we created a subset of our data that included only verb tokens that either did or could have occurred with overt SPPs. That is to say, subject types labeled "null" and "subject pronoun" in Table 4 above were included in this subset (N = 1,421). In this subset, 93% (n = 1,321) of the finite verbs occurred with a null SPP, and 7% (n = 100) occurred with an overt SPP.

Following convention, we then performed a mixed-effects logistic regression, using the lme4 package in R (Bates et al., 2015), on the occurrence of overt pronominal subjects in each child's production, accounting for nesting of subject occurrence within individual child participants with a random effect for participant. We used stepwise methods to build these models with SPP expression as the dependent variable and compared them using a likelihood-ratio test to determine the best-fit model for the data. The output of the best-fit mixed-effects logistic regression model, which shows that switch reference (same vs. different), group (TD vs. DLD), and verb person and number are significant main effects on the probability of SPP usage in the sample, is shown in Table 5. Additionally, this output also illustrates the significant interaction between switch reference and group. None of the other potential predictor variables or their interactions statistically significantly predicted SPP expression in this sample.

Table 3. Rates of subject pronoun use between DLD and control groups.

Subject expression	Control (N = 1288)	DLD (N = 733)
Overt	32.3% (n = 416)	25.9% (n = 190)
Null	67.7% (n = 872)	74.1% (n = 543)

Note. DLD = developmental language disorder.

First, this model illustrates the role of continuity of reference in overt subject expression among children in this sample. In other words, when the subject of the verb token in question was the same as that of the finite verb directly preceding it (n = 711), the combined group of DLD and TD children in this sample used overt SPPs 4.5% (n = 34), but when the subject of the verb token in question was different than that of the preceding finite verb (n = 710), they used overt subjects at a rate of 9.3% (n = 66), consistent with the trend in adult Mexican Spanish to use more overt subjects in switch reference contexts. Second, this model shows that the verbal utterances of children diagnosed with DLD (n = 538) occurred with overt SPPs at a rate of 5.2% (n = 28) overall, whereas TD children's utterances (n = 883) used overt SPPs at a rate of 8.2% (n = 72) overall.

Third, this model shows that only third-person plural referents occur with overt subject pronouns significantly less than first-person plural referents. However, when binned, singular and plurals together were significantly different from each other by chi-squared test (χ^2 = 15.079, degrees of freedom = 1, p value < .001). The overall rates of subject use with singular versus plural referents are summarized in Table 6.

Finally, the model output in Table 5 also reveals a significant interaction, which is illustrated in Figure 1. Here, we can see that TD children approximate adultlike patterns of SPP use based on switch reference, albeit at lower overall rates. In same reference contexts, TD

Table 4. Distribution of subject types for control and developmental language disorder (DLD) groups.

Subject type	Control (N = 1288)	DLD (N = 733)
null	67.7% (n = 872)	74.2% (n = 544)
lexical noun phrase	23.1% (n = 297)	18.3% (n = 134)
subject pronoun	5.6% (n = 72)	3.8% (n = 28)
indefinite pronoun	1.2% (n = 15)	1.1% (n = 8)
demonstrative pronoun	1.1% (n = 14)	0.82% (n = 6)
question word	0.69% (n = 9)	0.95% (n = 7)
quantifier	1.1% (n = 14)	0.82% (n = 6)
negative word	0.08% (n = 1)	0% (n = 0)

[5]An anonymous reviewer notes that "child Italian" may not be a conventional expression to some of the journal's readers. This usage stems from the Continuity Hypothesis of Macnamara (1982), Pinker (1984), and Hyams (1986) for similar claims.

Table 5. Output of best-fit mixed-effects logistic regression model for subject pronoun expression.

Group	Variable	Estimate	*SE*	*Z* value	Pr(>|z|)
	Intercept	–3.2349	0.6157	–5.254	1.49e-07
Switch reference (reference level = different)	same	–0.9937	0.2826	–3.516	< 0.001 ***
Group (reference level = TD)	DLD	–1.1851	0.3966	–2.988	0.003 **
Person and number (reference level = 1PL)	1SG	1.8872	0.6121	3.088	0.003 **
	2SG	1.0827	0.7086	1.528	0.126 ns
	3PL	–0.6959	0.8269	–0.841	0.400 ns
	3SG	0.9009	0.6259	1.439	0.150 ns
Switch Reference × Group	Same × DLD	1.0818	0.4868	2.222	0.026 *

Note. DLD = developmental language disorder; TD = typically developing; 1SG = first-person singular; 1PL = first-person plural; 2SG = second-person singular; ns = not significant; 3PL = third-person plural; 3SG = third-person singular.
***< .001. ** < .01. *< .05.

children used overt SPPs just 4.6% (n = 20) of the time but used overt SPPs 11.5% (n = 52) of the time in cases of switch in reference. However, DLD children did not follow these same patterns. Instead, DLD children employed relatively low rates of over SPPs in both same (5%, n = 14) and switch reference (5.5%, n = 14) contexts.

Discriminant Function Analysis

While the mixed-effects model allows us to retain the variance associated with each child's individual expressions of overt subject pronouns, we must average across the utterances produced by each child in order to have an independent variable that can be included in a discriminant function analysis. The distribution of overt subject pronouns in switch-reference contexts, as well as the distributions of our other measures, is nonnormal. Nonnormality can be caused by outliers, but in our case, it is caused by skew. Consequently, a quadratic discriminant function analysis, which is robust to nonnormality of this type, is appropriate. In the following table, we see the results of several discriminant functions, composed of different linguistic measures, including MLUw, the grammaticality choice measure of tense from J. Grinstead, De la Mora, Pratt, and Flores (2009),[6] and our current overt subject pronoun use in switch-reference contexts measure (see Table 7).

Over 80% discriminant accuracy is considered "fair," whereas above 90% discriminant accuracy is considered "good" accuracy by Plante and Vance (1994). The switch-reference subject pronoun + MLUw discriminant function correctly classifies all 19 out of 19 children diagnosed with DLD (100% or good sensitivity) and 16 out of 19 aged-matched controls (84.2% or fair specificity). Similarly, there was good sensitivity for the switch-reference subject pronoun + MLUw + tense function, which applied to a reduced sample of the children, who had taken the tense measure (14 of 19 in the DLD sample and 15 of 19 in the TD aged-matched sample, because some children did not pass filler items on the tense measure). In this case, there was 100% sensitivity for the DLD sample (14 of 14 children) and 100% specificity for the TD age-matched controls (15 of the 15 children). Notice, further, that the discriminant function composed of only pronominal subjects in switch reference contexts provided "fair" (above 80%) sensitivity.

Discussion

With respect to our first research question, regarding the distribution of grammatical subject types between DLD and TD populations, we can say that there do not appear to be significant differences. Children produced statistically nondistinct rates of pronouns, full NPs, and demonstratives, in subject position, regardless of their group.

With respect to our second and third research questions, regarding whether children in both DLD and TD groups appeared to be sensitive to the probabilistic constraints that guide overt subject pronoun use, the answer is mixed. On the one hand, when the groups are pooled, the

Table 6. Rates of subject pronoun expression by number.

Group	Present (n = 100)	Absent (n = 1321)
Singular	8.6% (n = 95)	91.4% (n = 1061)
Plural	1.9% (n = 6)	98.1% (n = 304)

[6]The Grammaticality Choice Measure of tense, used in J. Grinstead, De la Mora, Pratt, and Flores (2009), asks children to choose between an adultlike tensed verb, for example, *Yo abro la boca.* "I open my mouth." and a child-particular, nontensed form, for example, *Yo abre la boca.* "I open (bare stem verb) my mouth." The goal of the method is to compare non-adultlike forms attested in spontaneous language to the adult form. See J. Grinstead, De la Mora, Pratt, and Flores (2009) for further details.

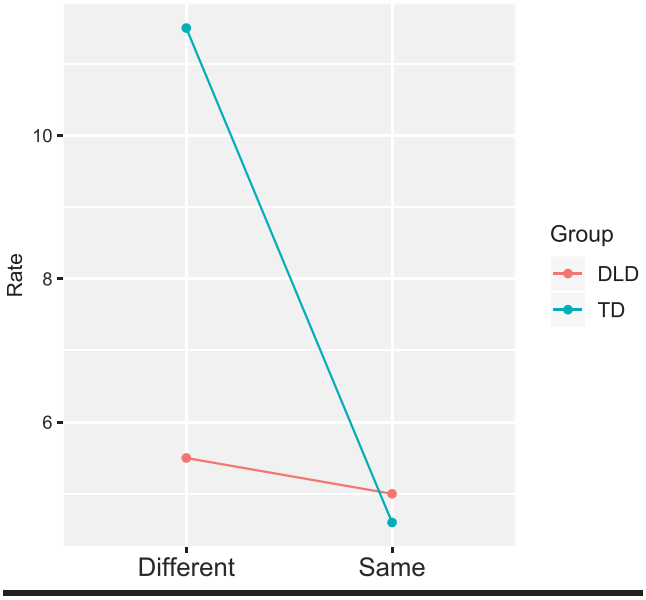
Table 7. Discrimination function analysis of control versus developmental language disorder (DLD) groups.

Classification variables	Control	DLD	Average
Switch reference	52.63% (10/19)	84.21% (16/19)	68.42%
Switch reference, MLUw	84.21% (16/19)	100% (19/19)	92.11%
MLUw, tense	86.67% (13/15)	100% (14/14)	92.86%
Switch reference, MLUw, tense	100% (15/15)	100% (14/14)	100%

Note. MLUw = mean length of utterance in words.

same constraints that have been established for TD, monolingual Spanish-speaking children of this age are predictive, namely, verbal person–number agreement and switch reference. On the other hand, and perhaps of greatest theoretical relevance, child Spanish-speakers with DLD use significantly fewer overt subject pronouns in switch-reference contexts than do their age-matches. This is of theoretical significance because children with DLD appear unable to coordinate discourse-pragmatic information regarding the identity of the subject antecedent with the morphosyntax of choosing a null subject versus an overt subject morpheme from the lexicon. Rather, as with definite articles, direct object clitics and the other cases alluded to children appear to be following the pattern stated in the Interface Deficit Hypothesis of overassuming interlocutor familiarity with the subject antecedent's identity, and consequently producing null subjects, where a TD child or adult would not. We take this as an additional and more specific case of interface deficit. Namely, not only do children with DLD overuse null subjects, in general, as in J. Grinstead et al. (2018), but rather replace a precise grammatical type of subject (tonic pronouns) with null subjects, consistent with overassumed familiarity, in a precise discourse context (switch-reference). We take this finding to be a confirmation of the Interface Deficit Hypothesis.

Figure 1. Rate differences between switch and nonswitch reference contexts between groups. DLD = developmental language disorder; TD = typically developing.



Finally, regarding our fourth research question about whether the SPP in switch-reference context rate could be profitably used in a discriminant function analysis, we see that it can. In particular, it appears useful in the "sensitivity" dimension regarding identification of children with DLD. Our analysis showed that using only this factor, we could successfully identify 16 of 19 DLD children (84.21%), which, as we said, is considered "fair" accuracy. Adding MLUw to the function gave us 100% sensitivity (19 of 19 DLD children correctly identified) and "fair" accuracy with TD children (16 of 19—84.21%). By adding our grammaticality choice measure of verb tense to the function, at least for the children who successfully completed it, we were able to achieve 100% sensitivity (14 of 14 DLD children identified) and 100% specificity (15 of 15 TD children identified). In summary, the use of overt subject pronouns in switch-reference contexts appears to play a useful role in identifying monolingual Spanish-speaking children diagnosed with DLD.

Determining how to best diagnose and treat DLD depends upon what dimensions of cognition are most affected by it. Though we do not presume to have a theory of all the dimensions of this disorder, we believe that we have provided some evidence that its linguistic dimensions relate to the interaction of syntax and discourse pragmatics described above. On this account, whatever else may occur with the language of DLD children, the flavor of their language is that, in both the nominal and verbal domains, they overassume familiarity of referents for their interlocutors. This appears true in English for overuse of definite articles (e.g., Maratsos 1974, 1976) and pronouns (e.g., Avrutin, 1994), and it also appears true with null versus overt subjects in Spanish. The grammatical nature of the null subject phenomenon is not common to all languages. Rather, there seem to be different grammatical versions of null subjects across even languages as closely related as Spanish and Italian.[7] It may turn out that even across languages in which null subjects are licensed distinctly, the discourse-pragmatics

[7]See Cantú-Sánchez et al. (2021) for a formal account of the different grammatical relations underlying the null subject nature of Spanish versus Italian.

motivating overt subject pronoun use is substantially the same. Investigations of subject pronoun use in the language of children with DLD in an array of such languages could better help us understand whether what we have struck on here is limited to languages like Spanish and perhaps Arabic, Mohawk, and other languages with these type of null subject, or whether it might hold of all null subject languages, regardless of the grammatical basis for null subjects.

Having said that, we should also recognize the limitations of what we have shown here. Our sample was limited by the labor-intensive nature of both diagnosing children with DLD and of transcribing and coding these data. The diagnostic test we used was locally normed, which allowed us to calculate 1.25 *SD*s below the mean, but it did not include published sensitivity and specificity rates. Furthermore, not all children were able to complete the tense measure, reducing the scope of our diagnostic ability. Greater statistical power could reveal presently invisible differences in our sample. For example, other constraints might be operating, but a low enough level to be difficult to detect in this sized sample. Furthermore, we have considered only one dialectal variant of Spanish. There are many, many more, including some that have been studied carefully, at least in their adult forms, from the variationist perspective. What, for example, will be the role played by switch reference in Caribbean dialects of Spanish that tend to use more overt subjects across the board? Will such a difference limit or enhance the sensitivity and specificity of our measures? Only future work will tell.

Morphosyntactic Variation and Language Disorders

Valuable work has been done to understand dialectal differences, without which many inappropriate diagnoses could occur. This is an important contribution of sociolinguistics to creating a scientifically accurate understanding of language differences versus language disorders. An additional contribution of sociolinguistics to the study of language disorders is the one we have attempted to pursue here. Namely, how grammatical and discourse factors influence the use of overt subject pronouns. This kind of approach to child language differs from what has come before inasmuch as we are not trying to simply identify a set of different grammatical rules that apply to a different dialect, but rather we aspire to identify factors that increase the probability of a morpheme being expressed. To those of us accustomed to working with morphosyntactic rules that obey more categorical, inviolable rules of grammar, this idea is somewhat foreign. However, the field of variationist sociolinguistics has considered a wide array of construction types in adult languages that could be susceptible to the same type of analysis we have deployed here. What determines the use of any sociolinguistically conditioned optionality in language, if that optionality is discourse-sensitive, could be helpful in deepening our understanding of DLD. It is our hope that the approach adopted here to the analysis of sociolinguistically conditioned variation may prove valuable in future studies of language disorders.

Data Availability Statement

The data sets analyzed during the current study are not publicly available due to the age of the research subjects and the manner in which the data were collected. Please contact the corresponding author for any data-related inquiries (kendra.dickinson@rutgers.edu).

Acknowledgments

This project was supported by a National Science Foundation (NSF) Grant to the second author (NSF-1551903). The authors are grateful to lab collaborators Myriam Cantú-Sánchez, Juliana De la Mora, Mariana Vega-Mendoza, Ximena Carreto-Guadarrama, and Amy Pratt for their invaluable comments and support.

References

Anderson, R. T., & Souto, S. M. (2005). The use of articles by monolingual Puerto Rican Spanish-speaking children with specific language impairment. *Applied Psycholinguistics, 26*(4), 621–647. https://doi.org/10.1017/S0142716405050332

Avrutin, S. (1994). Psycholinguistic investigations in the theory of reference [Doctoral dissertation, Massachusetts Institute of Technology].

Austin, J. R., Blume, M., Parkinson, D., Nuñez del Prado, Z., & Lust, B. (1998). Interactions between pragmatic and syntactic knowledge in the first language acquisition of Spanish null and overt pronominals. In J. Lema & E. Treviño (Eds.), *Theoretical analyses on romance languages: Selected papers from the 26th Linguistic Symposium on Romance Languages (LSRL XXVI), Mexico City, 28–30 March, 1996* (pp. 35–52). John Benjamins. https://doi.org/10.1075/cilt.157.04aus

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baugh, J. (1995). The law, linguistics, and education: Educational reform for African American language minority students. *Linguistics and Education, 7*(2), 87–105. https://doi.org/10.1016/0898-5898(95)90002-0

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

Cameron, R. (1993). Ambiguous agreement, functional compensation, and nonspecifictúin the Spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language Variation and Change, 5*(3), 305–334. https://doi.org/10.1017/S0954394500001526

Cantú-Sánchez, M., Ordóñez, F., Silva-Pereyra, J., & Grinstead, J. (2021). *Cousins do not always agree: Unagreement in Spanish*

and Italian. Hispanic Linguistics Symposium 2021, Wake Forest University, Winston-Salem, NC, United States.

Carvalho, A. M., & Child, M. (2011). Subject pronoun expression in a variety of Spanish in contact with Portuguese. In J. Michnowicz & R. Dodsworth (Eds.), *Selected proceedings of the 5th workshop on Spanish sociolinguistics* (pp. 14–25). Cascadilla Proceedings Project.

Castilla-Earls, A., Auza, A., Pérez-Leroux, A. T., Fulcher-Rood, K., & Barr, C. (2020). Morphological errors in monolingual Spanish-speaking children with and without developmental language disorders. *Language, Speech, and Hearing Services in Schools, 51*(2), 270–281. https://doi.org/10.1044/2019_LSHSS-19-00022

Castilla-Earls, A., Pérez-Leroux, A. T., Fulcher-Rood, K., & Barr, C. (2021). Morphological errors in Spanish-speaking bilingual children with and without developmental language disorders. *Language, Speech, and Hearing Services in Schools, 52*(2), 497–511. https://doi.org/10.1044/2020_LSHSS-20-00017

Castilla-Earls, A., Pérez-Leroux, A. T., Restrepo, M. A., Gaile, D., & Chen, Z. (2018). The complexity of the Spanish subjunctive in bilingual children with SLI. *Language Acquisition, 25*(1), 72–84. https://doi.org/10.1080/10489223.2016.1192636

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46. https://doi.org/10.1177/001316446002000104

De la Mora, J. (2004). *Direct object clitics and determiners in Spanish specific language impairment* [Unpublished master's thesis]. University of Alberta.

Erker, D., & Guy, G. R. (2012). The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language, 88*(3), 526–557. https://doi.org/10.1353/lan.2012.0050

Erker, D., & Otheguy, R. (2016). Contact and coherence: Dialectal leveling and structural convergence in NYC Spanish. *Lingua, 172-173,* 131–146. https://doi.org/10.1016/j.lingua.2015.10.011

Flores-Ferrán, N. (2002). *Subject personal pronouns in Spanish narratives of Puerto Ricans in New York City: A sociolinguistic perspective* (Vol. 2). Lincom Europa.

Forsythe, H. (2018). *Discourse and grammatical cues in the acquisition of Spanish pronouns* [Unpublished dissertation]. Michigan State University.

Grindal, T., Schifter, L. A., Schwartz, G., & Hehir, T. (2019). Racial differences in special education identification and placement: Evidence across three states. *Harvard Educational Review, 89*(4), 525–553. https://doi.org/10.17763/1943-5045-89.4.525

Grinstead, J. (2004). Subjects and interface delay in child Spanish and Catalan. *Language, 80*(1), 40–72. https://doi.org/10.1353/lan.2004.0024

Grinstead, J., Baron, A., Vega-Mendoza, M., De la Mora, J., Cantú-Sánchez, M., & Flores, B. (2013). Tense marking and spontaneous speech measures in Spanish specific language impairment: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research, 56*(1), 352–363. https://doi.org/10.1044/1092-4388(2012/11-0289)

Grinstead, J., Cantú-Sánchez, M., & Flores-Ávalos, B. (2008). Canonical and epenthetic plural marking in Spanish-speaking children with specific language impairment. *Language Acquisition, 15*(4), 329–349. https://doi.org/10.1080/10489220802349996

Grinstead, J., De la Mora, J., Vega-Mendoza, M., & Flores, B. (2009). An elicited production test of the optional infinitive stage in child Spanish. In J. Crawford, K. Otaki, & M. Takahashi (Eds.), *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 2008)* (pp. 36–45). Cascadilla Proceedings Project.

Grinstead, J., De la Mora, J., Pratt, A., & Flores, B. (2009). Temporal interface delay and root nonfinite verbs in Spanish-speaking children with specific language impairment: Evidence from the grammaticality choice task. In J. Grinstead (Ed.), *Hispanic child languages: Typical and impaired development* (pp. 239–263). John Benjamins.

Grinstead, J., Lintz, P., Pratt, A., Vega-Mendoza, M., De la Mora, J., Cantú-Sánchez, M., & Flores-Avalos, B. (2018). Null subject occurrence in monolingual Spanish SLI: A discriminant function analysis. *Language, 3*(2), 17. https://doi.org/10.3390/languages3020017

Grinstead, J., Lintz, P., Vega-Mendoza, M., De la Mora, J., Cantú-Sánchez, M., & Flores-Avalos, B. (2014). Evidence of optional infinitive verbs in the spontaneous speech of Spanish-speaking children with SLI. *Lingua, 140,* 52–66. https://doi.org/10.1016/j.lingua.2013.11.004

Grinstead, J., & Spinner, P. (2009). The clausal left periphery in child Spanish and German. *Probus, 21*(1), 51–82. https://doi.org/10.1515/prbs.2009.003

Grinstead, J. A. R. (1998). *Subjects, sentential negation and imperatives in child Spanish and Catalan* [Unpublished dissertation]. University of California.

Hayes, A. F., & Krippendorf, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*(1), 77–89. https://doi.org/10.1080/19312450709336664

Hyams, N. M. (1986). Discontinuous models of linguistic development. In *Language acquisition and the theory of parameters* (pp. 128–147). https://doi.org/10.1007/978-94-009-4638-5_5

Jackson-Maldonado, D., & Maldonado, R. (2017). Grammaticality differences between Spanish-speaking children with specific language impairment and their typically developing peers. *International Journal of Language & Communication Disorders, 52*(6), 750–765. https://doi.org/10.1111/1460-6984.12312

Labov, W. (1973). *Sociolinguistic patterns*. University of Pennsylvania Press.

Lastra, Y., & Martín Butragueño, P. (2015). Subject pronoun expression in oral Mexican Spanish. In A. Carvalho, R. Orozco, & N. L. Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective* (pp. 39–58). Georgetown University Press.

Lange, B. P., & Zaretsky, E. (2021). Sex differences in language competence of 4-year-old children: Female advantages are mediated by phonological short-term memory. *Applied Psycholinguistics, 42*(6), 1503–1522. https://doi.org/10.1017/S0142716421000333

Leonard, L. B. (1997). *Children with specific language impairment*. MIT Press.

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

Macnamara, J. (1982). *Names for things: A study of human learning*. MIT Press.

Maratsos, M. P. (1974). Preschool children's use of definite and indefinite articles. *Child Development, 45*(2), 446–455. https://doi.org/10.2307/1127967

Maratsos, M. P. (1976). *The use of definite and indefinite reference in young children: An experimental study of semantic acquisition*. Cambridge University Press.

Merino, B. J. (1983). Language loss in bilingual Chicano children. *Journal of Applied Developmental Psychology, 4*(3), 277–294. https://doi.org/10.1016/0193-3973(83)90023-0

Michnowicz, J. (2015). Subject pronoun expression in contact with Maya in Yucatan Spanish. In A. Carvalho, R. Orozco, & N. L. Shin (Eds.), *Subject pronoun expression in Spanish: Across-dialectal perspective* (pp. 101–120). Georgetown University Press.

Morgan, G., Restrepo, M. A., & Auza, A. (2009). Variability in the grammatical profiles of Spanish-speaking children with

specific language impairment. In J. Grinstead (Ed.), *Hispanic child languages: Typical and impaired development* (pp. 283–302). John Benjamins.

Morgan, P. L. (2021). Unmeasured confounding and racial or ethnic disparities in disability identification. *Educational Evaluation and Policy Analysis, 43*(2), 351–361. https://doi.org/10.3102/0162373721991575

Nilsson, K. K., & de López, K. J. (2016). Theory of mind in children with specific language impairment: A systematic review and meta-analysis. *Child Development, 87*(1), 145–153. https://doi.org/10.1111/cdev.12462

Oetting, J. B. (1992). *Language-impaired and normally developing children's acquisition of English plural* [Doctoral dissertation]. University of Kansas.

Oetting, J. B. (2019). Variability within varieties of English: Profiles of typicality and impairment. In T. Ionin & M. Rispoli (Eds.), *Three Streams of Generative Language Acquisition Research. Selected papers from the 7th Meeting of Generative Approaches to Language Acquisition – North America* (pp. 59–82). University of Illinois at Urbana-Champaign.

Oetting, J. B., Berry, J. R., Gregory, K. D., Rivière, A. M., & McDonald, J. (2019). Specific language impairment in African American English and Southern White English: Measures of tense and agreement with dialect-informed probes and strategic scoring. *Journal of Speech, Language, and Hearing Research, 62*(9), 3443–3461. https://doi.org/10.1044/2019_JSLHR-L-19-0089

Oetting, J. B., Rivière, A. M., Berry, J. R., Gregory, K. D., Villa, T. M., & McDonald, J. (2021). Marking of tense and agreement in language samples by children with and without specific language impairment in African American English and Southern White English: Evaluation of scoring approaches and cut scores across structures. *Journal of Speech, Language, and Hearing Research, 64*(2), 491–509. https://doi.org/10.1044/2020_JSLHR-20-00243

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*(5719), 255–258. https://doi.org/10.1126/science.1107621

Orozco, R., & Guy, G. (2008). El uso variable de los pronombres sujetos:¿ Qué pasa en la costa Caribe colombiana. In M. Westmoreland & J. A. Thomas (Eds.), *Selected proceedings of the 4th International Workshop on Spanish Sociolinguistics* (pp. 70–80). Cascadilla Proceedings Project.

Otheguy, R., & Zentella, A. C. (2012). *Spanish in New York: Language contact, dialectal leveling, and structural continuity*. Oxford University Press.

Otheguy, R., Zentella, A. C., & Livert, D. (2007). Language and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language,* (4), 83, 770–802. https://doi.org/10.1353/lan.2008.0019

Pinker, S. (1984). *Language development and learnability*. Harvard University Press.

Plante, E., & Vance, R. (1994). Selection of preschool language tests. *Language, Speech, and Hearing Services in Schools, 25*(1), 15–24. https://doi.org/10.1044/0161-1461.2501.15

Pray, L. (2003). *An analysis of language assessments used in the referral and placement of language minority students into special education*. Arizona State University.

Rangel, E., Gómez-Palacio, M., & Romero, S. (1988). *Batería de Evaluación de la Lengua Española (BELE) para niños mexicanos de 3 a 11 años: Manual de aplicación, calificación e interpretación [Assessment Battery of Spanish Language for Mexican children ages 3–11]*. Secretaría de Educación Pública, Dirección General de Educación Especial.

Reichenbach, H. (1947). The tenses of verbs. In J. Meinster & W. Schernus (Eds.), *Time: From concept to narrative construct: A reader*. De Gruyter. https://doi.org/10.1515/9783110227185.1

Restrepo, M. A., & Gutierrez-Clellen, V. F. (2001). Article use in Spanish-speaking children with specific language impairment. *Journal of Child Language, 28*(2), 433–452. https://doi.org/10.1017/S0305000901004706

Rice, M. L., & Oetting, J. B. (1993). Morphological deficits of children with SLI: Evaluation of number marking and agreement. *Journal of Speech and Hearing Research, 36*(6), 1249–1257. https://doi.org/10.1044/jshr.3606.1249

Serratrice, L. (2005). The role of discourse pragmatics in the acquisition of subjects in Italian. *Applied Psycholinguistics, 26*(3), 437–462. https://doi.org/10.1017/S0142716405050241

Shin, N. L. (2016). Acquiring constraints on morphosyntactic variation: Children's Spanish subject pronoun expression. *Journal of Child Language, 43*(4), 914–947. https://doi.org/10.1017/S0305000915000380

Shin, N. L., & Cairns, H. S. (2012). The development of NP selection in school-age children: Reference and Spanish subject pronouns. *Language Acquisition, 19*(1), 3–38. https://doi.org/10.1080/10489223.2012.633846

Shin, N. L., & Erker, D. (2015). The emergence of structured variability in morphosyntax: Childhood acquisition of Spanish subject pronouns. In A. Carvalho, R. Orozco, & N. L. Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective* (pp. 169–189). Georgetown University Press.

Shin, N. L., & Montes-Alcalá, C. (2014). El uso contextual del pronombre sujeto como factor predictivo de la influencia del inglés en el español de Nueva York [English influence on Spanish in New York: Evidence from subject pronouns in context]. *Sociolinguistic Studies, 8*(1), 85–110. https://doi.org/10.1558/sols.v8i1.85

Silva-Corvalán, C. (1982). Subject expression and placement in Mexican-American Spanish. In J. Amaste & L. Elías-Olivares (Eds.), *Spanish in the United States: Sociolinguistic aspects* (pp. 93–120). Cambridge University Press.

Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles*. Clarendon Press.

Sossaman, E. C. (2017). *Relationship between subject pronoun and verb finite marking in AAE-speaking children* [Unpublished master's thesis]. Louisiana State University.

Stalnaker, R. (1974). Pragmatic presuppositions. In M. Munitz & P. Unger (Eds.), *Semantics and philosophy*. New York University Press.

Torres Cacoullos, R., & Travis, C. E. (2011). Testing convergence via code-switching: Priming and the structure of variable subject expression. *International Journal of Bilingualism, 15*(3), 241–267. https://doi.org/10.1177/1367006910371025

Travis, C. E. (2007). Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change, 19*(2), 101–135. https://doi.org/10.1017/S0954394507070081

Valian, V., & Eisenberg, Z. (1996). The development of syntactic subjects in Portuguese-speaking children. *Journal of Child Language, 23*(1), 103–128. https://doi.org/10.1017/S0305000900010114

Woodward, A. L., Sommerville, J. A., Gerson, S., Henderson, A. M., & Buresh, J. (2009). The emergence of intention attribution in infancy. *Psychology of Learning and Motivation, 51,* 187–222. https://doi.org/10.1016/S0079-7421(09)51006-7