The MacArthur Inventario del Desarrollo de Habilidades Comunicativas III: A measure of language development in Spanish-speaking two- to four-year-olds

Donna Jackson-Maldonado
Universidad Autónoma de Querétaro

Margaret Friend
San Diego State University

Virginia A. Marchman
Stanford University

Adriana Weisleder
Northwestern University

Alejandra Auza
Hospital General Dr. Manuel Gea Gonzalez

Barbara Conboy
University of Redlands

Marta Rubio-Codina
Inter-American Development Bank*

Philip S. Dale
University of New Mexico

Abstract: Parent report measures are reliable, valid, and cost-effective means for obtaining information about early child language development. Adaptations of the MacArthur-Bates Communicative Development Inventories are available in multiple languages for children below the age of three but there is a need for such measures for older children. This study introduces the Spanish adaptation of the MacArthur-Bates Communicative Development Inventory-III, the MacArthur Inventario del Desarrollo de Habilidades Comunicativas III (IDHC-III) designed for children 2;6 to 4 years of age. This form complements the MacArthur Inventario Del Desarrollo de Habilidades Comunicativas Palabras y Gestos and Palabras y Enunciados (IDHC:PG and IDHC:PE) for younger children. A total of 571 families of monolingual Spanish-speaking children from a diverse socio-economic sample in Mexico completed the IDHC-III and comprise the norming sample. Data are presented by age and maternal education level showing developmental growth curves for *Lista de Vocabulario* (Vocabulary) and *Tipos de Palabras y Oraciones* (Grammatical Complexity) along with norming tables showing variability by age. For the *Pronunciación* (Pronunciation) and *Conceptos Generales* (General Concepts) sections, only descriptive data are presented. We provide a parent report measure to support language assessment for preschoolers acquiring Spanish in Mexico and possibly in other Latin American countries as well.

Keywords: parent report, norms, Spanish.

Corresponding author(s): Margaret Friend, Department of Psychology, San Diego State University, San Diego, CA, US. Email: mfriend@sdsu.edu.

ORCID ID(s): https://orcid.org/0000-0001-9678-0571
https://orcid.org/0000-0002-6839-5163
https://orcid.org/0000-0001-7183-6743
https://orcid.org/0000-0001-6094-8424
https://orcid.org/0000-0002-7592-4625
https://orcid.org/0000-0001-6506-3692
https://orcid.org/0000-0002-1286-7918

**Note*: The views presented here do not represent the Inter-American Development Bank, its board of directors, or the countries it represents.

Citation: Jackson-Maldonado, D., Friend, M., Marchman, V. A., Weisleder, A., Auza, A., Conboy, B., ... Dale, P. (2024). The MacArthur Inventario del Desarrollo de Habilidades Comunicativas III. *Language Development Research, 4*(1), 361—398. http://doi.org/10.34842/12fb-qz93

Introduction

Parent report measures are reliable and valid sources of information about young children's language and have proven useful in both clinical and research settings. Caregivers are an invaluable source of information because they have direct and extensive experience with their child in a variety of settings. Information from caregivers is an essential aspect of the process for identifying children with developmental delays or disabilities, as included in U.S. law (IDEA, 2022) and in guidelines from international organizations (WHO, 2012). In addition to their role in clinical assessment and screening, parent report instruments have also been adopted for estimating population-level metrics about the rate of children who are developmentally on track, a need that has grown in response to the United Nation's Sustainable Development Goals McCray et al., 2023). Beyond clinical and educational applications, parent reports have provided critical insights and expanded our knowledge regarding both the consistency and variability that characterizes early child language development (Skarakis-Doyle et al., 2009; Frank, et al., 2021; Fenson et al., 1994).

Two widely used American English parent report measures of language and communication for children up to 30 months of age are the MacArthur-Bates Communicative Development Inventories (Words & Gestures MBCDI:WG, Words & Sentences MBCDI:WS, and the MBCDI-III; Marchman, et al., 2023). The MacArthur Inventarios del Desarrollo de Habilidades Comunicativas (IDHC) are Spanish language adaptations of these measures (Jackson-Maldonado et al., 2003). At the core of these instruments is a vocabulary checklist asking parents to indicate words their child can "understand" or "understand and say", with other sections focused on early gesture use, morphology, word combinations, and sentence complexity. In a recent large, longitudinal study using in a large longitudinal study in Bogota, Colombia, the IDHC:PE predicted both IQ and school achievement (Rubio-Codina & Grantham-McGregor,

2020). Thus, these instruments offer cost-effective means to provide a comprehensive picture of a range of language and communication milestones in children under three years of age, a period when direct testing can be quite challenging.

The present paper reports the development and psychometric properties of an upward extension of the IDHC for measuring language development in children between the ages of 30 and 48 months acquiring Mexican Spanish. Spanish is spoken by substantial numbers of people in more than 20 countries, and is the third most spoken language in the world (Ethnologue, https://www.ethnologue.com/guides/how-many-languages; Simon-Cereijido, Conboy, & Jackson-Maldonado, 2020). In the United States, more than half the population growth between 2000 and 2010 was Hispanic (Passel, Cohen & Lopez, 2011), and approximately 16-18% of the population is Spanish-speaking (Simon-Cereijido, 2015). There is a growing need for valid Spanish-language instruments that can be used with monolingual Spanish speakers in Latin American countries, where there has been an expansion of research, clinical services, and educational programming focused on young children (Minto-Garcia et al., 2019; Rosemberg et al., 2022; Rubio-Codina et al., 2016; Verdisco et al., 2009). There is a particular need for standardized language assessments that can be used with three-year-old Spanish-speaking children across a broad socioeconomic spectrum for several reasons (Rubio-Codina, et al., 2015). First, the availability of standardized language assessment tools for evaluating Spanish-speaking children, particularly those under the age of 4, is limited. Second, this is an important age as it is often when some important educational management decisions are made. Third, parental input has proven effective in providing indirect assessment of the language of younger Spanish-speaking children and is crucial in understanding language skills even beyond the toddler years since it may reflect parent expectations for acquisition that vary with culture (Auza et al., 2023).

Short Forms and Upward Extensions of the CDI Instruments

The original versions of the MBCDI instruments are quite lengthy, with more than several hundred items, and time consuming for caregivers to complete. To overcome these limitations, short form versions have been developed which typically include a short vocabulary checklist (e.g., 100 items) and only one or two additional questions. While short form versions are less comprehensive than the original long forms, their length is likely to increase their feasibility for use in many contexts, for example, enabling face-to-face oral presentation (Rubio-Codina et al., 2016). Although they do not provide comprehensive information of the type needed for studies of vocabulary composition, short forms have strong correlations with longer versions demonstrating their validity as measures of children's relative status (Fenson et al., 2000; Mokhtari et al., 2022; Urm & Tulviste, 2021). This is especially relevant for the development of measures for somewhat older children, whose full range of language skills is

growing rapidly and cannot be assessed comprehensively.

The original long- and short forms of the CDIs were designed and normed for use for children under 30 months (Fenson et al., 2007), thus there remained a need for a form developed and normed for older children. Dale and colleagues developed an upward extension of the American English CDI, the MBCDI-III, for use with children through the age of 37 months (Marchman et al., 2023). The American English MBCDI-III has three Sections: Vocabulary Checklist (100 items), Using Language (12 yes-no questions concerning semantics and pragmatics) and Grammatical Complexity (1 question about word combinations and 12 sentence pairs for assessing morphology and syntax). The Vocabulary Checklist is necessarily brief, given the typical size of children's vocabulary at this age (see Marchman et al., 2023 for norms).

Evidence for the concurrent validity of the MBCDI-III has been reported in several studies and is summarized in Marchman et al. (2023). Vocabulary scores correlate with the Bayley Scales of Infant and Toddler Development-III (Perra et al., 2015), the McCarthy Verbal Scale (Feldman et al., 2005), the Peabody Picture Vocabulary Test (Feldman et al., 2005; Mercure, 1999), and Number of Different Words in language samples (Feldman et al., 2005). Grammatical complexity scores are correlated with MLU in language samples (Feldman et al., 2005).

Moreover, the MBCDI-III has been utilized in diverse research with typically developing children. For example, MBCDI-III scores have been used to estimate genetic influence on vocabulary, grammar, and their relationship (e.g., Dale et al., 2015). In addition, the MBCDI-III has been used in studies of children with language disorders (Feldman et al., 2003; Skaradis-Doyle et al., 2009), otitis media (Feldman et al 2003, 2005), Autism Spectrum Disorder (ASD, Tek et al., 2008), and children born preterm (Perra et al., 2015). The MBCDI-III has adequate discriminant classification validity (Skarakis-Doyle et al., 2009; Ukoumunne et al., 2012) and has been used to help parents identify children with language disorders (Skeat et at., 2010). Nevertheless, a ceiling effect was identified that limited the usefulness of the CDI-III to children at or below 37 months, rather than up to 48 months as originally intended.

Adaptations of the MBCDI-III into Non-English Languages

The success of the MBCDI-III has led to the development of adaptations for several other languages. As always in the adaptation of CDI instruments to new languages, substantial linguistic and cultural adaptation is needed (http://mb-cdi.stanford.edu/documents/ adaptationsnottranslations2015.pdf). Languages differ not only in their vocabulary and syntax, but also in the stages of acquisition of culturally relevant words, morphosyntactic forms and functions. Even the acquisition of translation equivalents may not be developmentally equivalent across languages. Consequently,

it is important to take into consideration language specific acquisition data when developing measures (Peña, 2007).

Two main categories of adaptations have been developed (see Brieković & Kraljević, 2023, for an overview of most current adaptations). The first category, including e.g., Basque and Hungarian, have been created based fairly directly on the original MBCDI-III, in that the vocabulary list has been very broad with respect to categories. The emphasis has been to find individual words appropriately difficult for the target age range. The Basque adaptation of the MBCDI-III, the KGNZ-3, for example, was extensively adapted and modified to reflect both the cultural context and structure of this non-Indo-European, ergative language with agglutinative morphology (Ezeizabarrena et al., 2013; Barens & García, 2013). Along with changes in the vocabulary list motivated by linguistic and cultural differences, the sections on grammar were expanded to include nominal case inflections, intransitive and transitive auxiliaries, and inflections to determine subject-object relations. New sections were added to assess pronunciation, pre-reading and school abilities, narrative questions and grammatical markers. Many of these changes were motivated by the goal of developing an instrument appropriate for children up through 50 months. Ezeizabarrena et al (2013) reported steady increments through 42 months for vocabulary production and through 50 months of age for sentence complexity with the KGNZ-3. Thus the ceiling effect for vocabulary was partially resolved. Similarly, the Hungarian adaptation (Kas & Lőrik, 2022) showed a ceiling effect for vocabulary at around 39-42 months.

More recent adaptations of the MBCDI-III have generally followed the Swedish adaptation (Eriksson, 2017), which incorporates a different design for the vocabulary list. Here, four specific semantic categories have been selected for more in-depth assessment based on developmental appropriateness and substantial growth during the target age range: food-related words, body-related words, cognitive words and emotion words. The vocabulary section also contains relatively more verbs, adjectives, and adverbs than the Swedish CDIs for younger children. To evaluate morphology and grammar, a section of 10 items asks about the child's use of complex phrases and another section of eight items queries the child's use of grammatical markers. A section of seven items on metalinguistic awareness asks caregivers to comment on their child's phonological and orthographic awareness as well as their awareness of the existence of other languages. Finally, one question asks about whether children pronounce words more like slightly younger children, their peers, or slightly more advanced children. This version has been normed on a nationally representative Swedish sample for children up to 48 months of age. Eriksson (2017) provided an initial evaluation of developmental validity based on correlations with age. As in most MBCDI studies, vocabulary and syntax were correlated. Internal consistency was high for vocabulary and syntax, and somewhat lower for the other components.

The Estonian MBCDI-III (the ECDI-III; Tulviste & Schultz, 2020) was also based to a considerable extent on the Swedish version. The vocabulary section consists of 101 words, mostly verbs and adjectives, with similar categories to the Swedish forms (body words, food words, mental and emotion words). The grammatical complexity section has 10 sentence pairs focusing on the agglutinative nature of the language. There are also sections which assess metalinguistic and general concepts, and pronunciation. Pilot data on the validity of the ECDI-III have been reported for children at 3 years, a sub-sample of the full longitudinal normative sample ranging from 30-48 months, based on correlations with Reynell Developmental Language Scale (Edwards et al., 2011).

Other adaptations which have generally followed the Swedish model, both with respect to the structure of the vocabulary checklist and the incorporation of scales for aspects of language beyond vocabulary and grammar are those for Norwegian (Holm et al., 2023), Finnish (Stolt, 2023), European Portuguese (Cadime et al., 2021), and Croatian (Brieković & Kraljević, 2023). Although the existing reports differ with respect to design, age range, and validation measures, overall the results are positive and similar to those for Swedish.

Evolution of the MBCDI-III in Spanish

Two preliminary Spanish parent report instruments for three-year-olds similar to the English-language MBCDI-III have been developed: the Pilot Inventario-III (INV-III; Guiberson, 2008 1 & b; Guiberson and Rodriguez, 2010, 2014; Guiberson, et al., 2011; and the Spanish Vocabulary Extension (SVE; Mancilla-Martinez, et al., 2016, Mancilla-Martínez, et al., 2011; Mancilla-Martínez et al., 2013). The INV-III is a direct translation of the English MBCDI-III; it includes a vocabulary checklist, a grammatical complexity section, and a request to provide examples of their child's three longest utterances. Scores on the INV-III correlate with the Ages and Stages Questionnaire (ASQ; Squires & Bricker, 2009) ($r_{(46)} = .69$, $p = .01$) and the Preschool Language Scale-4 (PLS-4; Zimmerman, et al., 2002) ($r_{(46)} = .62$, $p = .01$) and there is good classification accuracy of children with and without language delays (sensitivity = .82 and specificity = .81). However, because participants varied in age, the correlations with other measures are likely somewhat inflated. Also, data on this instrument for monolingual or near-monolingual children are limited.

The Spanish Vocabulary Extension (SVE) (Mancilla-Martínez et al., 2016) consists of a 100-word vocabulary checklist, drawn from the IDHC-PE and spontaneous languages samples. Correlations with the Short Form of the Spanish IDHC-PE (IDHC-IISF) for lower-income Spanish-speaking bilingual children in the U.S (N=48) suggest concurrent and discriminant validity and SVE scores also correlate with the full IDHC-IISF, the Woodcock Language Proficiency Battery–Revised (WLPB-R; Woodcock & Muñoz-

Sandoval, 1995), the Picture Vocabulary subtest, and the Test de Vocabulario en Imágenes Peabody (TVIP; Dunn, et al., 1986). The correlations were strongest for the younger children, and for the WLPB.

Although these measures filled a need in language assessment for Spanish-speaking three-year-olds, both have significant limitations. The INV-III is a direct translation of the English form rather than an adaptation, limiting its validity due to the lack of consideration of the cultural and linguistic relevance of specific words and sentence structures. The SVE was developed for a specific project and consists of a word list only. Norms were not obtained for either measure. Therefore, the present research sought to develop a Spanish MBCDI-III with indicators of both vocabulary and grammar, culturally and linguistically relevant items, and norms derived on a monolingual Spanish-speaking sample. Both the INV-III and the SVE were considered, and their authors contacted, before developing the current Spanish IDHC-III.

The Development of the MacArthur Inventario del Desarrollo de Habilidades Comunicativas III (IDHC-III)

The development of the IDHC-III has drawn on the Basque, Estonian, and Swedish adaptations. The development process followed the process for the original IDHC:PG and IDHC:PE adaptations (Palabras y Gestos and Palabras y Enunciados; Jackson-Maldonado et al, 2003) and the Spanish Short Forms (IDHC-ISF & IDHC-IISF; Jackson-Maldonado et al., 2012), with careful consideration of cultural and linguistic relevance and inspection of Spanish language acquisition data. The process of development of this form consisted of a preliminary norming study and this final version.

Pilot Instrument Development

For both Vocabulary and Grammatical Complexity, item selection began with examination of results at 30 months on the IDHC-PE norming study, to identify items selected by less than 30% of parents. This did not yield enough advanced vocabulary items, so additional words were identified by several other means, including narrative language samples from Mexican children (Jackson-Maldonado & Maldonado, 2015) and two Spanish-language corpora of 3- and 4-year-old children from CHILDES (Diez-Itza Corpus, Diez-Itza et al., 1999). Spanish-language acquisition researchers reviewed the preliminary list and the developers of the English form were consulted.

The *Tipos de Palabras y Oraciones* (Grammatical Complexity) section was expanded to increase the ceiling from the IDHC-II for younger children. On these items, parents are asked to identify which sentence of two examples "sounds most like how your child speaks". Each example sentence captures the same basic meaning, but one sentence is morphosyntactically more complex. New phrases were constructed from

CHILDES Spanish databases and narrative samples of preschool monolingual Spanish-speakers and Spanish language acquisition studies (Fernández & Aguado, 2007; Jackson-Maldonado & Maldonado, 2015, 2016; Jackson-Maldonado & Conboy, 2007; Morgan, et al., 2009; Perez-Leroux et al., 2012; Sanz-Torrent et al., 2008; Uccelli, 2009; Uccelli & Pavez, 2007). Note that, in most cases, these forms also convey more semantic information, as is common when children begin to use more complex sentence structures.

Preliminary norms were developed using a 100-word list and 26 sentence pairs completed by caregivers for 579 middle and low SES children in Mexico and 640 low SES children in Colombia between the ages of 30 and 47 months. Validity studies compared scores to the INV-III (Guiberson 2008a & b; Guiberson & Rodríguez, 2010; Guiberson et al., 2011) and sub-sections of the Clinical Evaluation of Language Fundamentals-Preschool (CELF-P; Wiig, Secord, Semel, 2004). Significant positive correlations (.54 in both cases) were found between corresponding sections (Vocabulary or Grammatical Complexity) of the IDHC-III and the INV-III. A discriminant analysis of children with varied language disorders yielded moderate sensitivity, 75%, and high specificity, 92%.

However, developmental trends were limited and there were ceiling effects with most participants producing 50 or more words on the 100-item test. Further, several words were identified as extremely low or extremely high frequency, all indicating a need to revise the vocabulary list. In contrast, the complexity section evinced an expected linear increase with age.

Current IDHC-III

Based on these considerations, the current IDHC-III was developed. Following Eriksson (2017), we included a more focused vocabulary list to include more advanced word classes in food related, body related, cognitive and emotional words categories. Further, pilot data on 108 Guatemalan children from low SES backgrounds, half of whom were monolingual Spanish-speaking and half of whom were Spanish-dominant from Kaqchikel-Spanish bilingual homes (Conboy et al., 2017a & 2017b), motivated the inclusion of additional culturally relevant categories (e.g., nature, health, school, abstract nouns–including culture-religion, action specific verbs, and change of state verbs) to allow the inventory to be used with children from a wider range of cultural backgrounds. A new 140-word list was piloted with 45 participants. Extremely high and extremely low frequency words and words with low correlations with age were deleted to obtain the final 100-word list reported here.

The final word categories, number of items and examples are presented in Table 1 and the examples of the grammatical complexity items are shown in Table 2. The full

form is presented in Appendix A. As can be seen from Table 1, the vocabulary list can be viewed as intermediate between that of the original MBCDI-III and that of the Swedish model. It is more focused than the original MBCDI-III, but is not as focused as the Swedish model.

Two additional sections are included. *Pronunciación* (Pronunciation), as in the Swedish version, consists of one question about how the child pronounces words. In *Conceptos Generales* (General Concepts), what Eriksson (2017) called metalinguistic awareness, parents are asked about school concepts, specifically, writing letters or numbers, counting, and naming shapes; the wording is based on the preschool academic programs for public schools in Mexico (https://www.gob.mx/sep/acciones-y-programas/educacion-preescolar, SEP 2017-ref) and consultation with preschool teachers and early literacy specialists.

Table 1. ***Word categories for the Vocabulary Checklist***

Category	Number of Items	Example
Abstract Noun	8	Accidente -*accident*
Attributes	13	Envidioso- *envious*
Action	11	Aguantar -*stand it, hold out*
Body	5	Cachete -*cheek*
Change of state	12	Desaparecer -*disappear*
Food	1	Postre- *dessert*
Function wds	7	Desde -*from*
Health	6	Calentura -*fever*
Objects	4	Grúa -*tow truck*
Locatives	8	Ciudad- city
Outside-nature	9	Insecto -*insect*
People	3	Mecánico -*mechanic*
Quantifier	10	Cada -*every*
School	3	Cuadrado -*square*
Total	100	

Table 2. ***Example sentence pairs for Grammatical Complexity***

Sentence Pair	Translation
Como pollo Voy a comer pollo con el tenedor	(I) eat chicken (I) am going to eat chicken with the fork
Ma caí y me duele Cuando me caigo, me duele	I fell and it hurts When I fall, it hurts
Se enfermó No pudo porque se enfermó	They got sick They couldn't because they got sick
No lo pongo aquí No creo que pueda ponerlo	I don´t put it here I don´t think I can put it

The Current Study

The goals of this study are to: (1) present developmental norms for vocabulary and grammatical complexity on the newly developed Spanish IDHC-III for monolingual, Spanish-speaking children in Mexico; (2) compare vocabulary and complexity development in children from different socioeconomic backgrounds; and (3) determine the relation between vocabulary and complexity on this instrument. Based on previous findings, we expect a strong relationship between vocabulary and complexity. We also expect that there will be developmental change with age and variation in scores as a function of maternal education.

Method

Participants

Data were originally compiled from $n = 577$ caregivers across multiple data collection sites. A total of 6 children were excluded because they were older than the target age range when the forms were completed. The final sample consisted of $n = 571$ caregivers, mostly mothers, who completed the IDHC-III and had children between

30 and 48 months of age (290 M, 281 F). Vocabulary checklist data were available for the full sample, however, not all sites chose to administer the form in its entirety and so data were available for only a subset of the children for *Grammatical Complexity* (n = 502) and *Pronunciación* and *Conceptos Generales* (n = 542).

Participants were recruited by multiple means to ensure a diverse sociodemographic sample from urban and rural areas of central Mexico. Caregivers were contacted through day care centers, preschools, recreation centers and personal contacts. An additional sample was also obtained as part of the piloting of the first child development module of the 2018-19 Mexican National Health and Nutrition Survey (ENSANUT 2018-19). This subset of caregivers, all beneficiaries of the government Conditional Cash Transfer program *Prospera*, were invited to attend a special session in which the goals of the project were explained, and they were offered a nonobligatory opportunity to fill out the forms with the interviewers. All caregivers completed a consent document fulfilling the first author's university Bio-Ethics committee requirements prior to the study. Caregivers then completed a Basic Information Questionnaire that included questions about the child's gestational age, birth weight, health issues, languages spoken in the home, as well as each caregivers' education and occupation.

For descriptive purposes, participants were divided into six age groups: 30-32 months, 33-35 months, 36-38 months, 39-41 months, 42-44 months, and 45-48 months. Participants were also divided into groups based on maternal education level: Middle School or less (MS), some High School (SHS), Completed High School (HS), and More than High School (MHS). The sample is described by age, child sex, and maternal education in Table 3. The sample is relatively evenly distributed over age, with a balance of females vs. males in each age group. Levels of maternal education were not evenly distributed, as the majority of the sample consisted of caregivers in the lower two groups. Just under 1/3 of the sample had more than a high school education across all age groups. This sample consists of a large and relatively representative sample of the Mexican population, as determined by educational attainment (OECD, 2023).

Procedure

Caregivers completed the forms following two administration formats. Some caregivers filled out in person with the help of linguistics and psychology students and teachers. This method was used most often at day care centers or the government health facility where *Prospera* program activities were carried out. Other caregivers received the forms in person. Parents could complete the forms on site, or if desired, they could take them home, and the forms were picked up no longer than 2 weeks later.

Written instructions appear at the beginning of each section, but to ensure understanding, full instructions with examples were always first explained verbally.

Table 3. ***Number of participants (%) by child age, child sex, and level of maternal education in full sample (n = 571)***

				Level of Maternal Education			
Age Group	Total	Female	Male	Middle School or less	Some High School	High School Graduate	More than High School
30-32 mos	96	46 (47.9)	50 (52.1)	10 (10.4)	46 (47.9)	11 (11.5)	29 (30.2)
33-35 mos	90	46 (51.1)	44 (48.9)	12 (133)	34 (37.8)	10 (11.1)	34 (37.8)
36-38 mos	103	54 (52.4)	49 (47.6)	4 (3.9)	52 (50.5)	11 (10.7)	36 (35.0)
39-41 mos	95	45 (47.4)	50 (52.6)	9 (9.5)	49 (51.6)	9 (9.5)	28 (29.5)
42-44 mos	82	38 (46.3)	44 (53.7)	6 (7.3)	41 (50.0)	12 (14.6)	23 (28.0)
45-48 mos	105	52 (49.5)	53 (50.5)	14 (13.3)	47 (44.9)	16 (15.2)	28 (26.7)
TOTAL	571	281 (49.2)	290 (50.8)	55 (9.6)	269 (47.1)	69 (12.1)	178 (31.2)

Measures

In the Lista de Vocabulario (Vocabulary) section, parents are asked to indicate the words that their child "comprende y dice" 'understands and says', yielding a maximum production vocabulary score of 100 words. Caregivers are told that the child should be able to produce the word spontaneously (i.e., repetitions are not allowed), but the words can be pronounced in a "childlike" manner (e.g., *momingo* for "domingo" 'Sunday'). Similarly, the child may use a different grammatical form that is equivalent to the one listed on the form. For example, if a child says *sabo* for "saber," an overgeneralization of the regular first-person singular form applied to the irregular verb 'to know', or *pesada* for "pesado," a feminine-marked form for the adjective 'heavy,' these would be counted as correct. The caregiver may also substitute synonyms for words that are used in their own family or dialect (e.g., *cavar* instead of *excavar* or *bonito* instead of *hermoso*).

The *Tipos de Palabras y Oraciones* (Word and Sentence Types or, as we refer to it in this paper, Grammatical Complexity) section is used to evaluate emerging morphology and grammar. On each item, the caregiver is asked to indicate which sentence, of each pair, *sounds most* like how their child currently speaks. They are told that the child does not have to produce the same sentence exactly, but rather the caregiver should reflect on which sentence sounds the most like something their child might say. Thus, in the pair, "Me caí y me duele" / "Cuando me caigo, me duele" '*I fell and it hurts'* / '*When I fall, it hurts,*" the child is given a score of 0 if the parent chooses the first phrase and a score of 1 if the parent chooses the second phrase. The maximum score is 15, reflecting the number of times the parent chose the second, more complex, sentence in the pair across 15 items.

For the *Pronunciación* section, caregivers were asked if it was difficult to understand their child´s speech. The score reflects *difficulty of understanding*, that is, when caregivers answered "sí (yes)", a score of 1 was recorded. This was the only item in which a higher score was indicative of less sophisticated development and a lower score was indicative of more advanced development. For *Conceptos Generales*, each question about academic concepts received a score of 1 if the parent answered "sí (yes)." Scores were summed to provide a total score (out of 3 possible responses). Scoring instructions for each section can be found in Appendix B.

Analysis Plan.

We first report descriptive statistics for vocabulary and grammatical complexity for children in each age group. We next report developmental patterns using modeling techniques that allow us to estimate age-related changes for each measure at monthly intervals. We chose to apply generalized additive models in the beta distribution family using non-parametric monotonic P-splines with GAMLSS (Stasinopoulos et al., 2017) in the R statistical package (Version 4.0.3; R_Core_Team, 2020). GAMLSS is a general framework for modeling a range of functions within a regression framework. This approach was recently applied in the 3rd Edition of the American English norms (Marchman et al., 2023) and has advantages over other techniques because it allows fit to a range of possible functions, and provides fit estimates of standard deviation, as well as the central tendency. Based on earlier work (Fenson et al., 2007; Frank et al., 2021), we assumed that the distributions were best captured within the family of Beta distributions, i.e., limited by 0 and 1. We modeled age and child sex, as well as interactions between age and sex as fixed effects. Following Smithson and Verkuilen (2006), scores were first converted to a proportion out of possible responses and extreme scores were imputed as 0.001 and 0.999 so that we could include all observations in the models. Based on our expectations of developmental change, all models used a very high value of lambda (10^4) and set the number of knots at 20. These parameters resulted in estimates of development that approached linear and were

smooth over age, while nevertheless being constrained at the higher and lower values. We report unstandardized beta coefficients (B) for all fixed effects. To test the significance of goodness of fit between nested models (e.g., those with and without an interaction term), we applied likelihood ratio tests (LRT, df = 1). Alpha was set at $p < .05$, two tailed, for all analyses. To generate the normative values, we extracted the percentile ranks for developmental trajectories estimated in the GAMLSS models by 5-percentile intervals from the 5^{th} – 95^{th} percentiles and the 99^{th} percentile over age in months. Plots present the values for the quintiles (10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th}). Normative values were also generated separately by child sex.

We next conducted several exploratory analyses examining the intercorrelation between scores on the vocabulary and grammatical complexity sections. We anticipated that the indices would be highly correlated, consistent with earlier reports (Bates & Goodman, 1994). To analyze effects of socioeconomic status, we again modeled age-related changes introducing maternal education level as a potential moderator for each measure. Finally, we present descriptive statistics for the remaining measures of *Pronunciación,* and *Conceptos Generale*s in Table 7; however, no further analyses of those measures are presented here.

Results

Descriptives

Mean scores and standard deviations on the Vocabulary and Grammatical Complexity sections are presented in Table 4 by child age group and sex. For vocabulary, note that even the children in the youngest group were reported to know just under half of the items on the form. For complexity, the children in the youngest group were reported to say the second, more complex, example, only about $1/5^{th}$ of the time. For both measures, there were substantial increases over age group in children's performance suggesting developmental changes in these critical language abilities over this important period.

Age-related trends.

Vocabulary Size.

To explore these developmental effects more fully, we conducted models that allowed us to capture age-related changes in vocabulary score, as shown in Table 5. As expected, Model 1, the unconditional model, shows a significant main effect of age, reflecting developmental change in vocabulary score from 30 to 48 months. Model 2 adds the factor of child sex. Again, results revealed a significant main effect of age, but the main effect of child sex was not statistically significant, $p = 0.35$. Thus, unlike

previous studies of vocabulary development using parent report (Frank et al., 2021; Marchman et al., 2023), the evidence for sex-related differences in vocabulary size was not statistically reliable. Moreover, adding the interaction term in Model 3 did not significantly increase overall model fit, $LRT(1) = 0.22$, $p = 0.64$, suggesting no differences in the magnitude of any sex differences across the age period.

Figure 1 illustrates the developmental effects from the unconditional model in terms of the fitted quantile estimates for all children in the sample and Figures 2 and 3 for boys and girls separately. Full values for percentile levels for both vocabulary and complexity scores, in 5-percentile increments, are presented in Tables 1 – 3 in Appendix C for all children and for girls and boys separately. Even though the main effect of sex and the sex by age interaction terms were not statistically reliable, we nevertheless provide norming tables separately for girls and boys to be consistent with earlier studies and to conform with some requirements for clinical reporting.

Table 4. Means and (SD) of scores for Vocabulary (n = 571) and Grammatical Complexity by age group for all children and by child sex

	Vocabulary[a]			Grammatical Complexity[b]		
Age Group	All	Female	Male	All	Female	Male
30-32 mos	46.2 (23.8)	47.8 (21.8)	44.8 (25.7)	3.2 (3.8)	3.2 (3.7)	3.2 (3.9)
33-35 mos	49.7 (22.6)	47.7 (23.4)	51.8 (21.9)	3.7 (3.6)	4.2 (3.6)	3.3 (3.5)
36-38 mos	56.8 (21.4)	63.1 (21.1)	49.8 (19.7)	5.5 (4.2)	5.7 (4.6)	5.2 (3.7)
39-41 mos	64.7 (21.7)	64.1 (23.1)	65.3 (20.5)	6.2 (4.1)	5.6 (4.0)	6.7 (4.2)
42-44 mos	64.4 (23.3)	62.2 (24.8)	66.3 (22.0)	6.5 (3.8)	6.1 (3.8)	6.9 (3.9)
45-48 mos	67.4 (20.9)	68.2 (21.8)	66.6 (20.4)	6.5 (4.0)	6.2 (4.0)	7.0 (4.0)
All children	58.3 (23.6)	59.1 (23.8)	57.5 (23.4)	5.3 (4.2)	5.2 (4.1)	5.4 (4.2)

Note: [a]Production vocabulary reflects the number of items caregivers selected on the vocabulary checklist (max = 100); [b]Grammatical complexity reflects the number of times the parent chose the more complex answer of two choices (max = 15).

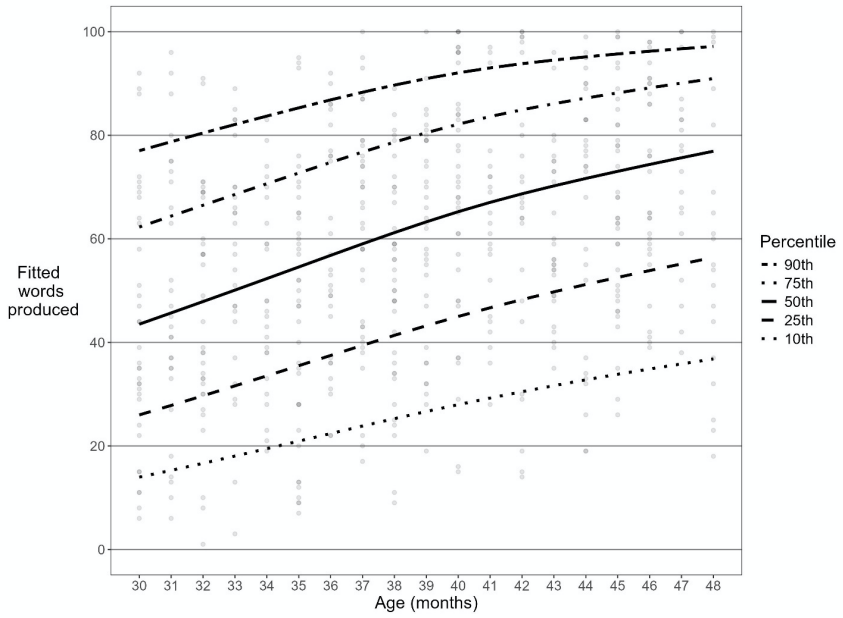

Figure 1. Fitted percentile scores by quintile (10th, 25th, 50th, 75th, and 90th) for Total Words Produced as a function of age group (months), both sexes combined; dots represent individual data points (n = 571).

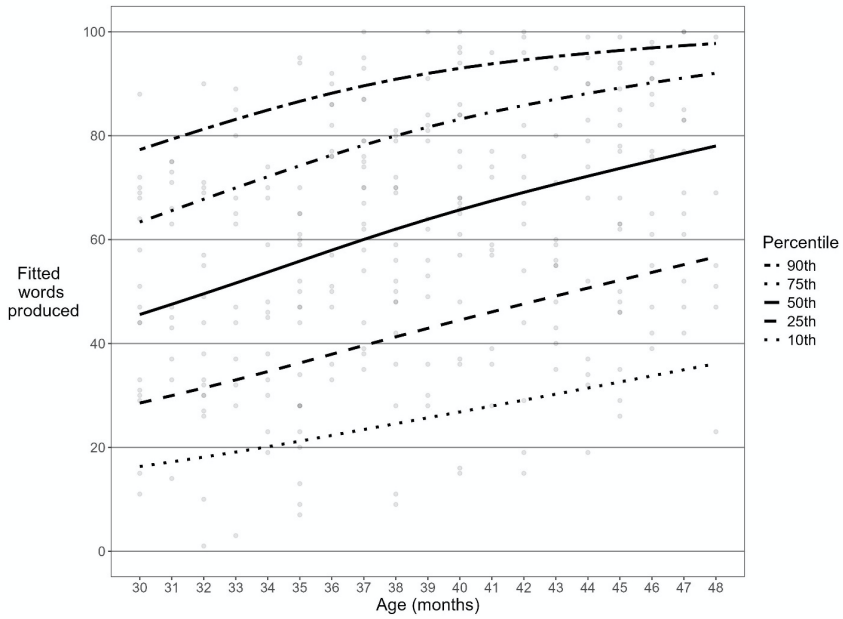

Figure 2. ***Fitted percentile scores by quintile (10th, 25th, 50th, 75th, and 90th) for vocabulary production as a function of age group (months) – girls; dots represent individual data points (n = 281).***

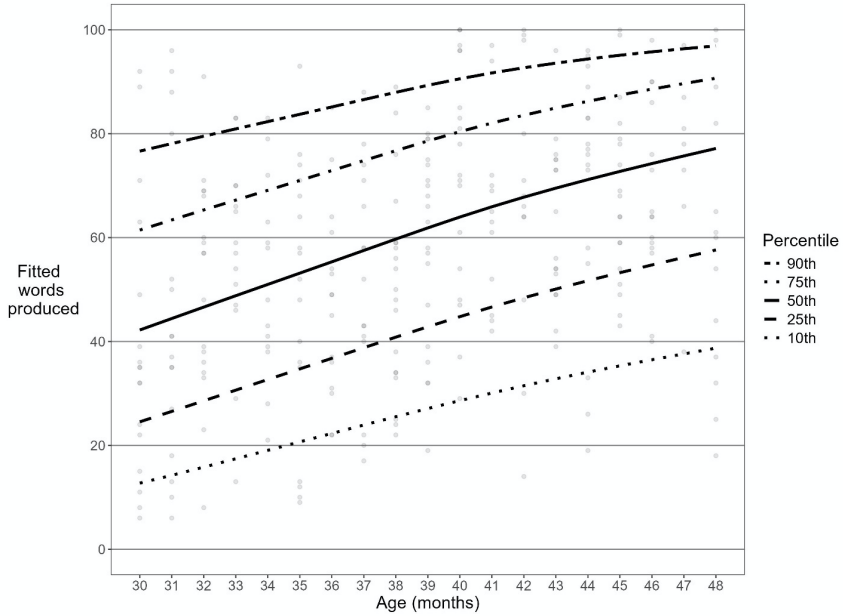

Figure 3. ***Fitted percentile scores by quintile (10th, 25th, 50th, 75th, and 90th) for vocabulary production as a function of age group (months) – boys; dots represent individual data points (n = 290).***

Grammatical Complexity.

Results for models exploring developmental trends in grammatical complexity score are presented in Table 6. Model 4 again revealed a substantial main effect of age on children's scores. Model 3 also revealed a main effect for age, but no main effect of child sex. Adding the interaction term did not increase overall model fit, LRT(1) = 1.5, p = 0.22, again suggesting no advantages for girls over boys in grammatical complexity scores at any developmental level. See Figure 4 for developmental effects from the unconditional model in terms of the fitted quantile estimates for all children in the sample and Figures 5 and 6 for boys and girls separately. Full values for all percentile levels are presented for all children and for girls and boys separately in Tables 4 – 6 in Appendix C.

Interrelation between Vocabulary and Grammatical Complexity.

We next explored the association between scores on the vocabulary and grammatical complexity subsections. Scores on the Vocabulary and the Grammatical Complexity sections were moderately intercorrelated ($r(500) = 0.43$, $p < 0.001$), reflecting that children who scored higher in vocabulary were also scoring higher on the grammatical complexity scale. This correlation remained significant after controlling for age,

$r(499) = 0.37$, $p < 0.001$, suggesting that this association is not due to each measure being individually associated with age.

Maternal Education.

To explore the impact of maternal education on patterns of age-related changes, we added maternal education, as well as the age x maternal education interaction to Models 1 and 3. Looking first at vocabulary production, we again see a significant main effect of age, $B = 0.07$ (0.01), $p < .001$, however, there were no effects of maternal education, such that scores were similar across all 4 groups, as illustrated in Figure 7. Note that children in families with more than high school education had the lowest scores overall when compared to children with less than high school education, $B = -0.29$ (0.16), $p = .07$. This difference did not reach statistical significance and must therefore, be interpreted with caution. Finally, the addition of the interaction term did not increase overall model fit, $LRT(3) = 1.7$, $p = 0.65$, suggesting that the patterns of relations among the caregiver education groups were parallel across age.

To examine whether this pattern was consistent in those sub-samples of families in which the caregivers received additional support in completing the forms, we reanalyzed the effect of maternal education in only those families in which caregivers were not likely to have been given verbal support during administration ($n = 502$). Again, there were no statistically significant group differences on vocabulary scores as a function of maternal education group (all $p > .08$), and adding maternal education to the model did not significantly increase overall model fit, $LRT(1) = 0.96$, $p = .33$.

For Grammatical Complexity, adding maternal education level to Model 4 showed a significant effect of age, $B = 0.07$ (0.01), $p < 0.001$. Importantly, in contrast to the results for vocabulary, children of caregivers with higher education levels were reported to produce more complex sentences than children of caregivers with less than middle school education, $B = 0.66$ (0.19), $p < .001$, as illustrated in Figure 8. Scores for children in all of the other groups were not statistically different from those of children who had caregivers with less than middle school education. Adding the interactiton term did not increase overall model fit, $LRT(4) = 2.83$, $p = 0.63$, suggesting that the advantage for children of caregivers with higher educational levels was consistent across the age period.

Table 5. ***Fitted estimates (unstandardized B (SE)) for vocabulary production by age and child sex, both sexes combined (n = 571)***

	Model 1	Model 2	Model 3
Intercept	-2.13 (0.30)***	-2.10 (0.30)***	-2.01 (0.43)***
Age	0.06 (0.01)***	0.07 (0.01)***	0.06 (0.01)***
Sigma Intercept	-0.46 (0.29)	-0.46 (0.29)	-0.46 (0.29)
Sigma Age	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
Sex	--	-0.08 (0.08)	-0.22 (0.59)
Age x Sex	--	--	0.01 (0.02)
Number of Observations	571	571	571
R^2	0.12	0.12	0.12
Generalized AIC	-158.22	-157.98	-156.04

Table 6. ***Fitted estimates (unstandardized B (SE)) for grammatical complexity by age and child sex, both sexes combined (n = 502)***

	Model 4	Model 5	Model 6
Intercept	-2.84 (0.43)***	-3.02 (0.43)***	-2.57 (0.57)***
Age	0.06 (0.01)***	0.06 (0.01)***	0.05 (0.01)***
Sigma Intercept	0.59 (0.33)	0.56 (0.33)	0.56 (0.33)
Sigma Age	-0.01 (0.01)	-0.01 (0.01)	0.01 (0.01)
Sex	--	-0.07 (0.10)	-0.87 (0.77)
Age x Sex	--	--	0.02 (0.02)
Number of Observations	502	502	502
R^2	0.10	0.10	0.10
Generalized AIC	-268.02	-266.51	-266.01

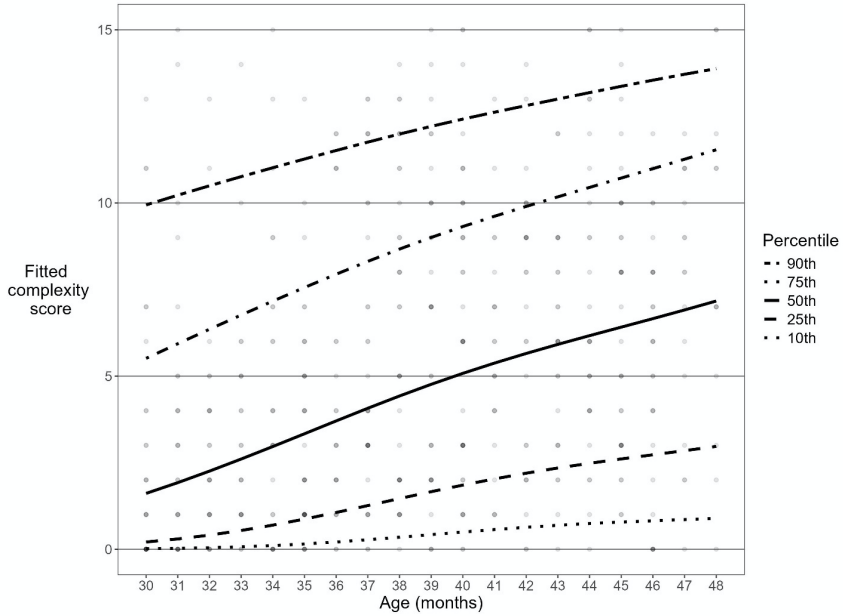

Figure 4. Fitted percentile scores by quintile (10th, 25th, 50th, 75th, and 90th) for grammatical complexity as a function of age group (months) – both sexes combined; dots represent individual data points (n = 502).

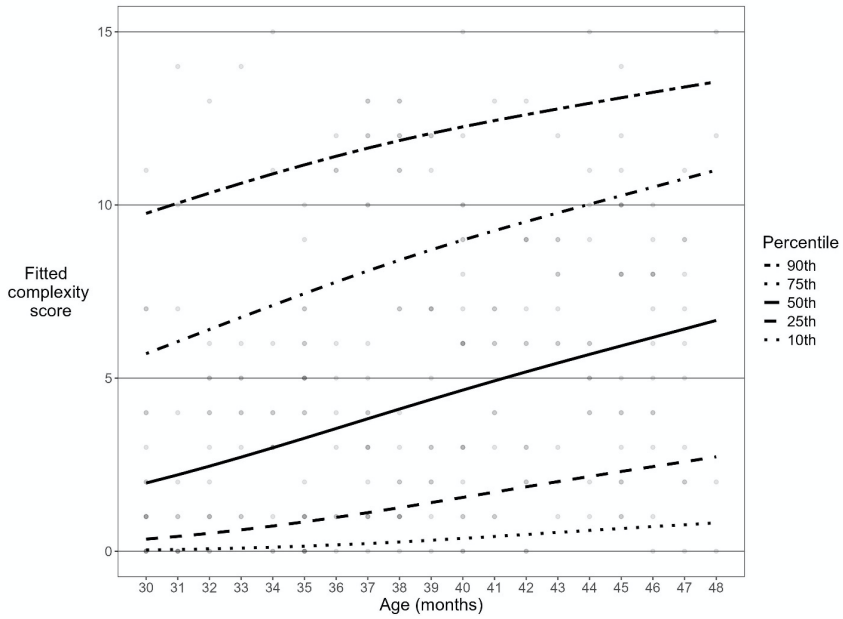

Figure 5. Fitted percentile scores by quintile (10th, 25th, 50th, 75th, and 90th) for grammatical complexity as a function of age group (months) – girls; dots represent individual data points (n = 251).

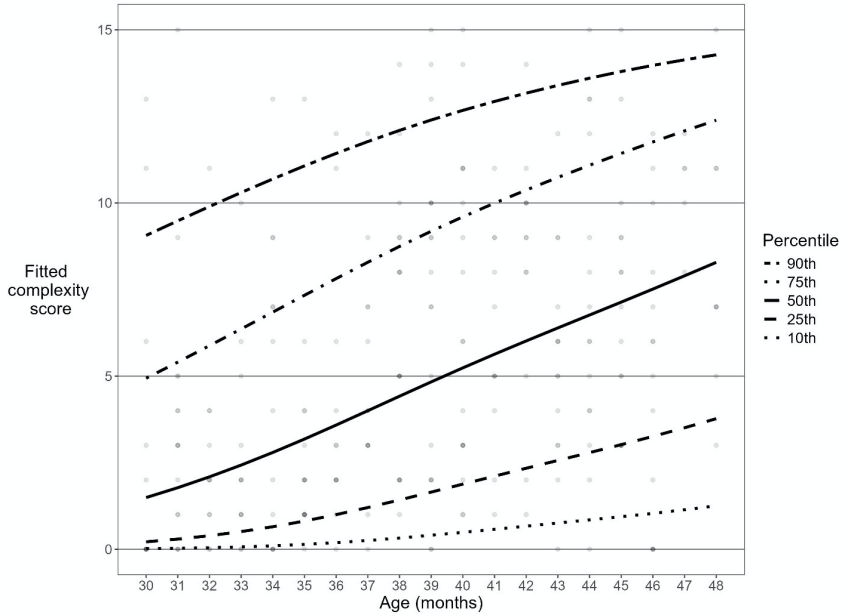

Figure 6. Fitted percentile scores by quintile (10th, 25th, 50th, 75th, and 90th) for grammatical complexity as a function of age group (months) – boys; dots represent individual data points (n = 251).

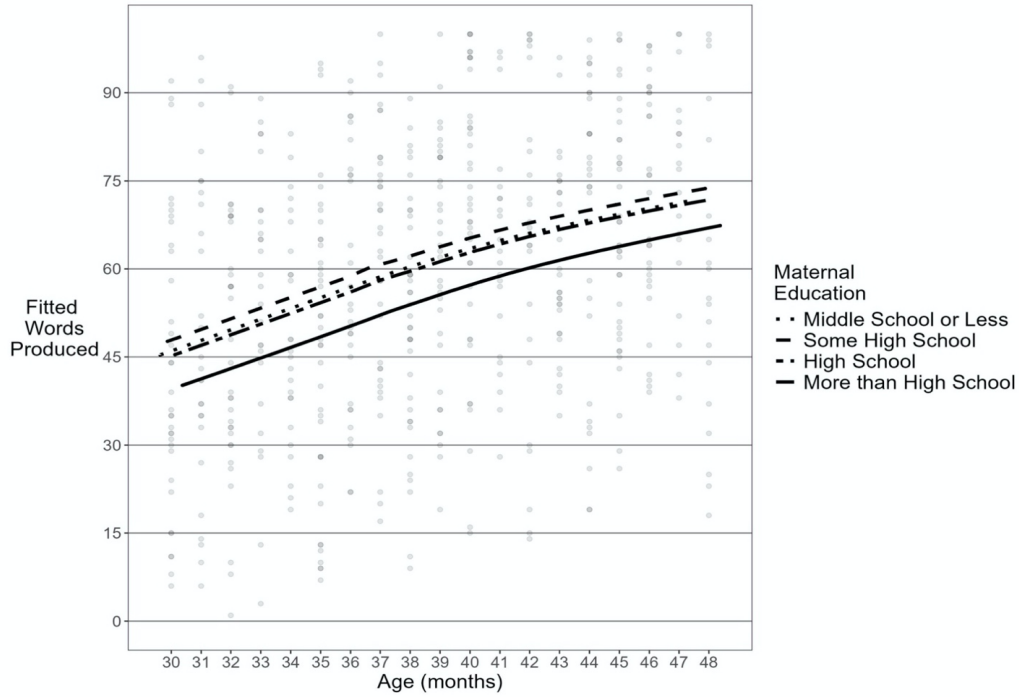

Figure 7. Modeled estimates for words produced as a function of child age and maternal education level; dots represent individual data points (n = 571).

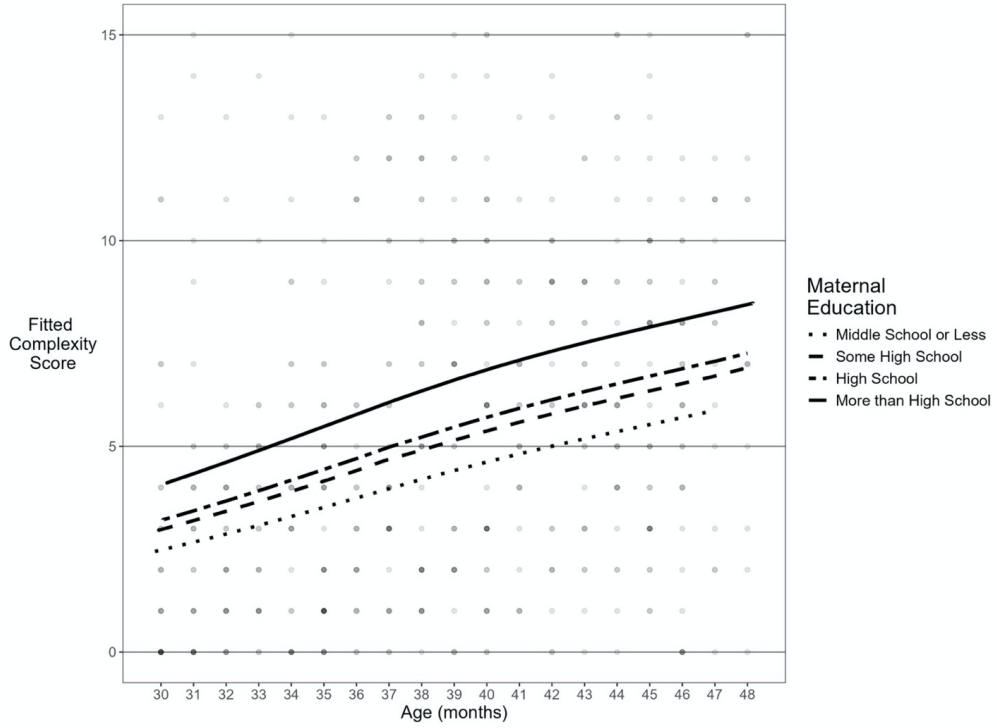

Figure 8. Modeled estimates for grammatical complexity as a function of child age and maternal education level; dots represent individual data points (*n* = 502).

Discussion

This paper presents initial norming data for Spanish adaptation of the upward extension of the MBCDIs (the IDHC-III), for children 2.5 through 4 years of age. This adaptation consists of a 100-word checklist for word production and a complexity section consisting of 15 sentence pairs to identify word-level and sentence-level grammatical complexity. The other sections of the IDHC-III are not analyzed in depth here. For the Vocabulary and Grammatical Complexity sections, the results included developmental trends by age groups and present differences by maternal education. Importantly, the data set includes a large and relatively representative sample of the population of Spanish-speaking, Mexican population (OECD, 2023).

Our analytic models for vocabulary production revealed developmental changes across age groups, but also substantial individual variation from 30 months through 4 years of age. This checklist, like the IDHC Short Forms, was only 100 words long and yet captured a wide range of levels of vocabulary knowledge. The results for the Grammatical Complexity section also revealed steep age-related changes as well as substantial variation, consistent with findings from the Basque form (Ezeizabarrena et al., 2013; Barnes & Garcia, 2013). The strength of the relation between Vocabulary

and Grammatical Complexity was similar to what has been reported for younger children (e.g., Frank et al., 2021), which may suggest that these two abilities are driven by a common set of learning mechanisms.

In addition, we observed no reliable effects of maternal education on vocabulary production scores. Prior studies with English-speaking children have found that caregivers with lower levels of education report higher vocabulary comprehension scores in very young children (8-12 months) but not older children (13-18 months, e.g., Dollaghan et al., 1999). In older English-speaking (21-30 months; Fenson et al., 2007), and Spanish-speaking children (26-30 months; Jackson-Maldonado et al., 2003) a positive relation between SES and vocabulary production has been reported and several studies show a positive correlation between maternal education and child vocabulary through the preschool years (e.g., Hoff, 2006; 2013). At the same time, other studies report no relation between maternal education and vocabulary in monolingual and bilingual (DeAnda et al., 2015; Friend et al., 2022; Montanari et al., 2020) Spanish-speaking children. Still others report that the relation is mediated by parent literacy behaviors (e.g., Gonzalez et al., 2017).

One possibility is that differences in administration methods for caregivers with varying education levels may have masked our ability to detect any effects of maternal education on vocabulary production in the current study. In accordance with recommendations for low literacy contexts (Alcock et al., 2015), some caregivers with lower maternal education received assistance from a researcher when completing the IDHC-III forms, while caregivers with higher maternal education typically completed the forms independently. This aligns with previous pilot research on the IDHC-III (Conboy et al., 2017b; Conboy, 2019) conducted in Guatemala and Mexico. This format may draw parent attention to how children use words in a way that written checklists do not, facilitating attention to the content of the vocabulary checklist among caregivers with lower levels of education. Indeed, at the ages covered by the IDHC-III, vocabulary can be quite large, and the great majority of words will have low frequency. Based on the enormous growth of vocabulary during the previous year or two in preschoolers, deciding if a given word has been produced by a child is a challenging task. There are two possible unintended consequences of providing additional support to caregivers with lower education levels. Caregivers completing the inventory on their own may not have attended to the checklist in the same way and underreported the words their children knew or, on the other hand, mothers who received additional support may have overreported their children's vocabulary. In an effort to further understand this finding, we reanalyzed the subset of the data in which no systematic assistance in completing the forms was offered to caregivers with low education. Patterns mirrored the effects for the full dataset; administration differences within sub-groups do not appear to have masked effects of maternal education on vocabulary production in this study.

In contrast to the results with Vocabulary, Grammatical Complexity scores showed the expected tendency: caregivers with higher education levels reported that their child produced more complex forms than those with lower education levels. These results are generally consistent with those reported in the literature. Based on a series of studies including older children, Hoff (2013) reported that higher SES children outperform lower SES peers on most tests of grammatical development and produce more complex sentences with a larger variety of structures (Dollaghan et al., 1999; Huttenlocher et al., 2010; Vasilyeva, Waterfall, & Huttenlocher, 2008). In the norming studies for the original English and Spanish long form instruments, caregivers of older children (21-30 months) from mid-SES families report longer mean length of utterance and higher grammatical complexity scores than caregivers from lower-SES families (Fenson et al., 2007; Jackson-Maldonado et al., 2003).

It is important to speak to the appropriateness of parent report in low SES samples. Without question, adapting methods for socio-cultural context is paramount to acquiring valid and reliable indicators of language development. For example, parents with lower levels of education and literacy may need additional support to complete the forms (e.g., Rubio-Codina et al., 2016) or may have different values and motivations than middle class families (Roseberry-McKibbin, 2013; Gonzalez et al., 2018) that may be reflected in their responses. Nevertheless, parental report measures have been used successfully in diverse socio-cultural contexts. For example, Weber et al. (2018) has shown the validity of parental report measures in Wolof communities in Africa, and Alcock et al. (2015) have noted that traditional written formats may need to be modified for face-to-face interviews in Kenya (Alcock et al., 2015). Most studies have shown that low SES parents are valid reporters (Alcock et al 2015, Dar et al, 2015; DeAnda et al., 2015; Hamadani et al 2010; Prado et al, 2016). Indeed, comparisons of parent reports with child speech samples (Dollaghan et al., 1999; Feldman et al, 2000) and with a behavioral comprehension measure (DeAnda et al., 2015) have revealed comparable accuracy across SES in reporting on vocabulary production on the MBCDI. Others, assessing vocabulary by means of parent report, language samples, and language tests, also showed comparable accuracy across SES (Furey, 2011; Sachse & Suchodoletz, 2008) in reporting vocabulary.

Recent efforts in developing vocabulary assessments for young children have used statistical techniques such as Item Response Theory (IRT; Embretson & Reise, 2013) to select a set of items that are most efficient for discriminating children with different ability levels (Bohn et al., 2023; Kachergis et al., 2022; Chai et al., 2020). However, data-driven methods such as IRT require a large set of data to be able to calculate each item's difficulty and discrimination power. Thus, these methods are most useful when one is selecting from a larger population of items with substantial data on each, but do not offer help in identifying items for a new instrument. Given that the goal of

the present paper was to develop these initial set of items, we focused on manual iterations to arrive at the set of items in the current S-CDI III. Once sufficient data has been collected using this instrument, future studies could use IRT to further refine it and improve its psychometric properties.

Limitations

This study had several limitations. First, there is some evidence of a ceiling effect for vocabulary in high-performing children as they approach 48 months of age. This suggests that the instrument may not reflect the full range of variability for these children. Nevertheless, this is quite modest for the IDHC-III compared to the pilot version and even the original MBCDI-III and other language adaptations. Further, the instrument does well at differentiating the lowest performing children from those in the mid and upper ranges. We also observed a floor effect on the sentence complexity scale for younger children. This likely reflects the more complex mature grammatical forms chosen for this instrument relative to other adaptations of the MBCDI-III and may have implications for assessment of grammatical development in children up to about 36 months of age with shorter MLUs.

Second, whereas the maternal education levels of the present sample approximate the larger Mexican population, this led to unique challenges and findings: some caregivers with lower education levels required assistance to complete the form. Our analyses suggest that this difference in administration did not alter the reported effects. Nevertheless, we found no effect of maternal education on vocabulary in contrast to our expectation (e.g., Hoff, 2006; 2013). We note however that other recent studies report similar null effects in monolingual and bilingual Spanish-speaking samples (DeAnda et al., 2015; Friend et al., 2022; Montanari et al., 2020) and encourage further research on the influence of administration practices and on the role of maternal education in early Spanish vocabulary.

Finally, early validity studies (Guiberson 2008a & b; Guiberson & Rodríguez, 2010; Guiberson et al., 201) suggest that the IDHC-III may be useful in clinical settings. In particular, in our pilot research, both Vocabulary and Grammatical Complexity correlated positively with scores on the Clinical Evaluation of Language Fundamentals-Preschool (CELF-P; Wiig, Secord, Semel, 2004) and an assessment of discriminant validity yielded moderate specificity and high sensitivity in the detection of language disorders. These findings require confirmation by additional studies.

Conclusion

Our results suggest that parent report is a useful means to obtain language development information in preschoolers in this language setting, just as it is for toddlers.

Thus, it can contribute to meeting the need for cost-effective, valid language assessment instruments as part of the global effort to bring assessment to scale, that is, making it broadly accessible to diverse populations. There are few measures available for Spanish-speaking children and we find that parent report may be a viable means to fill this need, obtaining information about language development for preschoolers acquiring Spanish in Mexico and possibly in other Latin American countries as well. As expected, we found age-related changes in vocabulary and grammatical complexity scores and a strong relationship between the two measures. We also observed expected differences in grammatical complexity scores with maternal education. Together, these metrics provide preliminary indicators of validity: expected relations of age with both vocabulary and grammatical complexity, between vocabulary and grammatical complexity, and between maternal education and grammatical complexity. Given reported variability in parent reports of Spanish vocabulary and our observations, the relation between maternal education and vocabulary in preschool children merits further research.

This study provides a new parental report instrument, the IDHC-III for Spanish-speaking children, normed on a representative sample in a monolingual setting in Mexico. As data from other research has shown (Mancilla-Martínez et al., 2016, 2011, 2013; Marchman & Martínez-Sussman, 2002) it may also be useful in the assessment of bilingual children, but specific studies need to be developed to analyze its use in this case. Norms are now available for this measure that are appropriate for assessing language development in preschool Spanish-speakers from diverse socio-economic backgrounds.

References

Alcock, K.J., Rimba, K., Holding, P., Kitsao-Wekulo, P., Abubakar, A. & Newton, C.R.J.C. (2015). Developmental inventories using illiterate parents as informants: Communicative Development Inventory (CDI) adaptation for two Kenyan languages. *Journal of Child Language*, *42*(4), 763-785.

Auza, A., Murata, C. & Peñaloza, C. (2023) Predictive validity of a parental questionnaire for identifying children with Developmental Language Disorders. *Frontiers in Psychology 14*: 1110449. doi:10.3389/fpsyg.2023.1110449

Barnes, J., & Garcia, I. (2013). Vocabulary growth and composition in monolingual and bilingual Basque infants and toddlers. *International Journal of Bilingualism*, *17*, 357-374. https://doi.org/10.1177/1367006912438992

Brieković, L. Š., & Kraljević, J. K. (2023). Parental reports on language development in toddlers and preschoolers based on the Croatian version of Communicative

Development Inventories III. *Frontiers in Psychology*. *14*. https/doi.org/10.3389/fpsyg.2023.1188550

Cadime I., Santos A.L., Ribeiro I., Viana F.L. (2021). Parental reports of preschoolers' lexical and syntactic development: Validation of the CDI-III for European Portuguese. *Frontiers in Psychology, 12*. doi: 10.3389/fpsyg.2021.677575.

Carrow-Woolfolk, E. (1999). Test for Auditory Comprehension of Language (3rd Ed.). Austin, TX: ProEd.

Conboy, B.T. (2019). Spanish language parent report language inventory data from low- and middle-SES families in Guatemala and Mexico. Paper presented as part of the symposium, The Effect of Socioeconomic Level on Language Development in Diverse Populations, at the IX Congress of the Asociación Para el Estudio de la Adquisición del Lenguaje [Spanish Association for the Study of Language Acquisition], UNED, Madrid, September 4 – 6, 2019.

Conboy, B.T., Stansbury, A., Sanchez, E. & Jackson-Maldonado, C. (2017a). Language acquisition and development in bilingual (Kaqchikel-Spanish) Preschool children in Guatemala. Poster presentation. *American Speech, Language, and Hearing Association (ASHA) Convention*, Los Angeles, CA.

Conboy, B.T., Stansbury, A. & Jackson-Maldonado, D. (2017b). Language Development in Bilingual Kaqchikel-Spanish Speaking Contexts. Poster presentation. *International Symposium on Bilingualism*, Ireland.

Dale, P.S., Grazia Tosto, M., Hayiou-Thomas, M.E., Plomin, R. (2015). Why does parental language input style predict child language development? A twin study of gene–environment correlation. *Journal of Communication Disorders, 57*, 106-117, https://doi.org/10.1016/j.jcomdis.2015.07.004.

Dar, M., Anwaar, H., Vihman, M., & Keren-Portnoy, T. (2015). Developing an Urdu CDI for early language acquisition. *York Papers in Linguistics, 14*, 1-14.

DeAnda, S., Arias-Trejo, N., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2015). Effects of minimal L2 exposure and SES on early word comprehension in English and Spanish: new evidence from a direct assessment. *Journal of Bilingualism: Language and Cognition*.

Díez Itza, E., Snow, C. E., & MacWhinney, B. (1999). La metodología RETAMHE y el proyecto CHILDES: breviario para la codificación y análisis del lenguaje infantil. *Psicothema, 11(3), 517-530*.

Dollaghan, C. A., Campbell, T. F., Paradise, J. L., Feldman, H. M., Janosky, J. E., Pitcairn, D. N., & Kurs-Lasky, M. (1999). Maternal education and measures of early speech and language. *Journal of Speech, Language, and Hearing Research, 42*(6), 1432-1443.

Dunn, L. M., & Dunn, L. M. (1997). Peabody Picture Vocabulary Test—III. Circle Pines, MN: American Guidance Service.

Edwards, S., Letts, C., & Sinka, I. (2011). *The New Reynell Developmental Language Scales*. London: GL Assessment Ltd.

Embretson, S. E., & Reise, S. P. (2013). *Item response theory*. Psychology Press.
Eriksson, M. (2017). The Swedish Communicative Development Inventory III. Parent reports on language in preschool children. *International Journal of Behavioral Development,* 41(5), 647-654.

Ezeizabarrena, M.J., Barnes, J.; García, I., Barreña, A. & Almgren, M. (2013). Using Parental Report Assessment for Bilingual Preschoolers: the Basque experience. In V. C. Mueller Gathercole (ed.), *Solutions for the assessment of bilinguals* (57-80). Clevedon: Multilingual Matters.

Feldman, H. M., Dale, P. S., Campbell, T. F., Colborn, D. K., Kurs-Lasky, M., Rockette, H. E., & Paradise, J. L. (2005). Concurrent and predictive validity of parent reports of child language at ages 2 and 3 years. *Child Development, 76*, 856-868.

Feldman, H. M., Dollaghan, C. A., Campbell, T. F., Kurs-Lasky, M., Janosky, J. E., & Paradise, J. L. (2000). Measurement properties of the MacArthur Communicative Development Inventories at ages one and two years. *Child Development, 71*, 310-322.

Feldman, H. M., Dollaghan C. A., Campbell T. F., Colborn D. K., Janosky J., Kurs-Lasky, M., Rockette, H. E., Dale, P. S. & Paradise, J.L. (2003). Parent-reported language skills in relation to otitis media during the first 3 years of life. *Journal of Speech, Language & Hearing Research, 46*.

Fenson, L., Bates, E., Dale, P., Goodman, J., Reznick, J. S., & Thal, D. (2000). Reply: Measuring variability in early child language: Don't shoot the messenger. *Child Development, 71*(2), 323-328.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J. & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development,* 59 (5, Serial 242).

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories (2nd Ed.).* Baltimore, MD: Brookes Publishing.

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur communicative development inventories. *Applied Psycholinguistics, 21,* 95-116.

Fernández Vázquez, F. & Aguado, G. A. (2007). Medidas del desarrollo típico de la morfosintaxis para la evaluación del lenguaje espontáneo de niños hispanohablantes. *Revista de Logopedia, Foniatría y Audiología, 27,* 140-152.

Frank, M. C., Marchman, V., Yurovsky, D., & Braginsky, M. (2021). The Wordbank Project: Variability and Consistency in Children's Language Learning Across Languages. Cambridge, MA: MIT Press.

Friend, M., Lopez, O., De Anda, S., Abreu-Mendoza, R. A., & Arias-Trejo, N. (2022). Maternal education revisited: Vocabulary growth in English and Spanish from 16 to 30 months of age. *Infant Behavior and Development, 66,* https://doi.org/10.1016/j.infbeh.2021.101685

Furey, J. E. (2011). Production and maternal report of 16-and 18-month-olds' vocabulary in low-and middle-income families. *American Journal of Speech-Language Pathology,* 20, 38-46.

Gonzalez, J.E., Bengochea, A., Justice, L., Yeomans-Maldonado, G., & McCormick, A. (2019). Native Mexican Parents' Beliefs About Children's Literacy and Language Development: A Mixed-Methods Study, *Early Education and Development, 30,* 259-279, 10.1080/10409289.2018.1542889

Guiberson, M. (2008a). Concurrent validity of a parent survey measuring communication skills of Spanish-speaking preschoolers with and without delayed language. *Perspectives on Communication Disorders in Culturally and Linguistically Diverse Populations, 15*, 73-81.

Guiberson, M. (2008b). Validity of a parent vocabulary checklist for young Spanish-speaking children of Mexican immigrants. *International Journal of Speech-Language Pathology, 10,* 279–285.

Guiberson, M., & Rodríguez, B. L. (2010). Measurement properties and classification accuracy of two Spanish parent surveys of language development for preschool-age

children. *American Journal of Speech-Language Pathology, 19*, 225-237.

Guiberson, M., Rodriguez, B. L., & Dale, P. S. (2011). Classification accuracy of brief parent report measures of language development in Spanish-speaking toddlers. *Language, Speech, and Hearing Services in Schools, 42*, 536-549.

Hamadani, J. D., Baker-Henningham, H., Tofail, F., Mehrin, F., Huda, S. N., & Grantham-McGregor, S. M. (2010). Validity and reliability of mothers' reports of language development in 1-year-old children in a large-scale survey in Bangladesh. *Food and Nutrition Bulletin, 31*(2_suppl2), S198-S206.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26*, 55-88.

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: implications for closing achievement gaps. *Developmental Psychology, 49*, 4-14.

Hoff, E., Burridge, A., Ribot, K. M., & Giguere, D. (2018). Language specificity in the relation of maternal education to bilingual children's vocabulary growth. *Developmental Psychology, 54*, 1011–1019. https://doi.org/10.1037/dev0000492

Holm, E., Hansen, P. B., Romøren, A. S. H., & Garmann, N. G. (2023). The Norwegian CDI-III as an assessment tool for lexical and grammatical development in preschoolers. *Frontiers in Psychology, 14*, 1175658.

Individuals with Disabilities Act (2022) Annual Report to Congress. https://sites.ed.gov/idea/2022-individuals-with-disabilities-education-act-annual-report-to-congress/

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology, 61*, 343-365.

Jackson-Maldonado, D. (2012). Verb morphology and vocabulary in monolinguals, emerging bilinguals, and monolingual children with Primary Language Impairment. In B. Goldstein (Ed.), *Bilingual Language Development and Disorders in Spanish-English Speakers. 2nd edition*. Baltimore: Brookes, pp 153-173.

Jackson-Maldonado, D. & Conboy, B. T. (2007). Utterance length measures for Spanish-speaking toddlers: the morpheme vs word issue revisited. En J.G. Centeno, L.K. Obler y R. Anderson (Eds) *Studying Communication Disorders in Spanish Speakers:*

Theoretical, research & clinical aspects. Multilingual Matters: North Somerset, England.

Jackson-Maldonado, D. & Maldonado, R. (2017). Grammaticality differences between Spanish-speaking children with Specific/Primary Language Impairment (SLI/PLI) and their typically developing peers. *International Journal of Language and Communication Disorders 52,* 750-765.

Jackson-Maldonado, D. & Maldonado, R. (2015). La complejidad sintáctica en niños con y sin Trastorno Primario de Lenguaje. In I. Rodríguez Sánchez y E.Vázquez (Eds.) *Lingüística Funcional.* Querétaro: Universidad Autónoma de Querétaro, pp. 253-301.

Jackson-Maldonado, D. & Maldonado, R. (2016). El uso de conectores en niños con y sin Trastorno del Lenguaje. *Lingüística Mexicana, 8,* 33-50.

Jackson-Maldonado, D., Marchman, V. A., & Fernald, L. C. (2013). Short-form versions of the Spanish MacArthur–Bates Communicative Development Inventories. *Applied Psycholinguistics, 34,* 837-868.

Kas, B., Jakab, Z., & Lőrik, J. (2022). Development and norming of the Hungarian CDI-III: A screening tool for language delay. *International Journal of Language & Communication Disorders*. https://doi-org.libproxy.unm.edu/10.1111/1460-6984.12686

McCray, G., McCoy, D., Kariger, P., Janus, M., Black, M. M., Chang, S. M., Tofail, F., Eekhout, I., Waldman, M., Buuren, S. van, Khanam, R., Sazawal, S., Nizar, A., Schönbeck, Y., Zongo, A., Brentani, A., Zhang, Y., Dua, T., Cavallera, V., … Gladstone, M. (2023). The creation of the Global Scales for Early Development (GSED) for children aged 0–3 years: Combining subject matter expert judgements with big data. BMJ Global Health, *8,* e009827. https://doi.org/10.1136/bmjgh-2022-009827

Mancilla-Martinez, J., Gámez, P. B., Vagh, S. B., & Lesaux, N. K. (2016). Parent reports of young Spanish–English bilingual children's productive vocabulary: A development and validation study. *Language, Speech, and Hearing Services in Schools, 47,* 1-15.

Mancilla-Martinez, J., Pan, B. A., & Vagh, S. B. (2011). Assessing the productive vocabulary of Spanish–English bilingual toddlers from low-income families. *Applied Psycholinguistics, 32*(2), 333-357.

Mancilla-Martinez, J., & Vagh, S. B. (2013). Growth in toddlers' Spanish, English, and conceptual vocabulary knowledge. *Early Childhood Research Quarterly, 28,* 555-567.

Marchman, V. A., Dale, P. S., & Fenson, L. (2023). MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual, 3rd Ed., Brookes: Baltimore.

Marchman, V. A., & Martínez-Sussmann, C. (2002). Concurrent validity of caregiver/parent report measures of language for children who are learning both English and Spanish. *Journal of Speech, Language, and Hearing Research, 45,* 983-997.

Minto-García, A., Canto, E. A. A., & Arias-Trejo, N. (2020). Mothers' Use of Gestures and their Relationship to Children's Lexical Production. *Psychology of Language and Communication, 24*, 175–200. https://doi.org/10.2478/plc-2020-0010

Mokhtari, F., Kazemi, Y., Feizi, A., & Dale, P. (2022). Psychometric Properties of the MacArthur-Bates Communicative Development Inventories-III (CDI-III) in 30 to 37 Months Old Persian-Speaking Children. *Archives of Rehabilitation, 23*, 372-391.

Morgan, G., Restrepo, M. A., & Auza, A. (2009). Variability in the grammatical profiles of Spanish-speaking children with specific language impairment. *Hispanic child languages: Typical and impaired development, 50*, 283-303.

Montanari, S., Mayr, R., & Subrahmanyam, K. (2020). Speech and language outcomes in Spanish-English preschoolers: the role of maternal education. *International Journal of Bilingual Education and Biilngualism*, 1–19. https://doi.org/10.1080/13670050.2020.1781780

Muggeo, V. (2013). quantregGrowth: growth charts via regression quantiles. *R package version 0.1*, *1*.

Organization for Economic Cooperation and Development (2023), Education at a Glance 2023: OECD Indicators, OECD Publishing, Paris, https://doi.org/10.1787/e13bef63-en.

Passel, J. S., Cohn, C., & Lopez, M. H. (2011). Hispanics Account for more than Half of Nation's Growth in Past Decade. Census 2010: 50 Million Latinos. Washington, DC: Pew Hispanic Center.

Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child Development, 78*, 1255-1264.

Pérez-Leroux, A. T., Castilla-Earls, A. P., & Brunner, J. (2012). General and specific effects of lexicon in grammar: Determiner and object pronoun omissions in child Spanish. *Journal of Speech, Language, and Hearing Research, 55*, 313–327.

Perra, O., McGowan, J. E., Grunau, R. E., Doran, J. B., Craig, S., Johnston, L., Jenkins, J., Holmes, V. & Alderdice, F. A. (2015). Parent ratings of child cognition and language compared with Bayley-III in preterm 3-year-olds. *Early Human Development, 91*, 211-216.

Prado, E. L., Adu-Afarwuah, S., Lartey, A., Ocansey, M., Ashorn, P., Vosti, S. A., & Dewey, K. G. (2016). Effects of pre-and post-natal lipid-based nutrient supplements on infant development in a randomized trial in Ghana. *Early Human Development, 99*, 43-51.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Roseberry-McKibbin, C. (2013). Increasing language skills of students from low-income backgrounds: Practical strategies for professionals. Plural Publishing.

Rosemberg, C. R., & Barreiro, A. (2022). *Interacción social, interaccionales y discursivos en el uso infantil de lenguaje mentalista.* XVIII REUNIÓN NACIONAL - VII ENCUENTRO INTERNACIONAL Asociación Argentina de Ciencias del Comportamiento, 6-7. https://dialnet.unirioja.es/descarga/articulo/9066708.pdf
.

Rubio-Codina, M., Attanasio, O., Meghir, C., Varela, N., & Grantham-McGregor, S. (2015). The socioeconomic gradient of child development: Cross-sectional evidence from children 6–42 months in Bogota. *Journal of Human Resources, 50*, 464-483.

Rubio-Codina M., Araujo M. C., Attanasio O., Muñoz P., Grantham-McGregor S. (2016). Concurrent validity and feasibility of short tests currently used to measure early childhood development in large scale studies. *PLoS ONE,* 11: e0160962. https://doi.org/10.1371/journal.pone.0160962

Rubio-Codina M., Grantham-McGregor S. (2020). Predictive validity in middle childhood of short tests of early childhood development used in large scale studies compared to the Bayley-III, the Family Care Indicators, height-for-age, and stunting: A longitudinal study in Bogota, Colombia. *PLoS One*. doi: 10.1371/journal.pone.0231317.

Sachse, S., & Von Suchodoletz, W. (2008). Early identification of language delay by direct language assessment or parent report? *Journal of Developmental & Behavioral Pediatrics, 29*(1), 34-41.

Sanz-Torrent, M., Serrat, E., Andreu, L., & Serra, M. (2008). Verb morphology in Catalan and Spanish in children with specific language impairment: a developmental study. *Clinical Linguistics & Phonetics, 22*, 459-474.

Secretaría de Educación Pública. (2017). *Aprendizajes Clave para la Educación Integral. Educación Preescolar*. México: Secretaría de Educación Pública.

Simon-Cereijido, G. (2015). Preschool language interventions for Latino dual language learners with language disorders: What, in what language, and how. *Seminars in Speech and* Language, *36*, 154-164.

Simon-Cereijido, G., Conboy, B.T.**,** & Jackson-Maldonado, D. (2020). El derecho humano de ser multilingüe: recomendaciones para logopedas [The human right of being multilingual: recommendations for speech-language therapists]. *Revista de Logopedia, Foniatría y Audiología*, *40*, 178-186.

Skarakis-Doyle, E., Campbell, W., & Dempsey, L. (2009). Identification of children with language impairment: Investigating the classification accuracy of the MacArthur–Bates Communicative Development Inventories, Level III. *American Journal of Speech-Language Pathology, 18*, 277-288.

Skeat, J., Eadie, P., Ukoumunne, O., & Reilly, S. (2010). Predictors of parents seeking help or advice about children's communication development in the early years. *Child: Care, health and development, 36*, 878-887.

Smithson, Michael and Verkuilen, Jay, (2006). A Better Lemon Squeezer? Maximum-Likelihood Regression with Beta-Distributed Dependent Variables. Psychological Methods, *11*, 54-71.

Squires, J. & Bricker, D. (2009). *The Ages and Stages Questionnaire (ASQ) in Spanish*. Baltimore: Brookes Publishing Co.

Stasinopoulos, M. D., Rigby, R. A., Heller, G. Z., Voudouris, V., & De Bastiani, F. (2017). *Flexible regression and smoothing: using GAMLSS in R*. CRC Press.

Stolt, S. (2023). Internal consistency and concurrent validity of the parental report instrument on language in pre-school-aged children–The Finnish Communicative Development Inventory III. *First Language, 43*, 492-515.

Tek, S., Jaffery, G., Fein, D., & Naigles, L. R. (2008). Do children with Autism Spectrum Disorders show a shape bias in word learning? *Autism Research, 1*, 208-222.

Tulviste, T., & Schultz, A. (2020). Parental reports of communicative development at the age of 36 months: The Estonian CDI-III. *First Language, 40,* 64-83.

Uccelli, P. (2009). Emerging temporality: past tense and temporal/aspectual markers in Spanish-speaking children's intra-conversational narratives. *Journal of Child Language, 36,* 929-966.

Uccelli, P., & Páez, M. M. (2007). Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associations among English and Spanish skills. *Language, Speech, and Hearing Services in Schools, 38,* 225–236.

Ukoumunne, O. C., Wake, M., Carlin, J., Bavin, E. L., Lum, J., Skeat, J. & Reilly, S. (2012). Profiles of language development in pre-school children: a longitudinal latent class analysis of data from the Early Language in Victoria Study. *Child: Care, Health and Development, 38,* 341-349.

Urm, A., & Tulviste, T. (2021). Toddlers' Early Communicative Skills as Assessed by the Short Form Version of the Estonian MacArthur-Bates Communicative Development Inventory II. *Journal of Speech, Language, and Hearing Research, 64,* 1303-1315.
Vasilyeva, M., Waterfall, H., & Huttenlocher, J. (2008). Emergence of syntax: Commonalities and differences across children. *Developmental Science, 11,* 84-97.

Verdisco, A., Cueto, S., Thompson, J., Engle, P., Neuschmidt, O., Meyer, S., González, E., Oré, B., Hepworth, K., & Miranda, A. (2009). Urgency and possibility results of PRIDI A first initiative to create regionally comparative data on child development in four Latin American countries technical annex. Technical Annex. Inter-American Development Bank, Washington DC.

Weber, A. M., Marchman, V. A., Yatma, D. I. O. P., & Fernald, A. (2018). Validity of caregiver-report measures of language skill for Wolof-learning infants and toddlers living in rural African villages. *Journal of Child Language, 45,* 939-958.

Wiig, E. H., Secord, W. A., & Semel, E. (2004). *Clinical Evaluation of Language Fundamentals—Preschool (2nd ed.).* (CELF Preschool-2). Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.

World Health Organization & UNICEF. (2012). Early childhood development and disability: A discussion paper. https://apps.who.int/iris/handle/10665/75355Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2002). *Preschool Language Scale, Fourth Edition: Spanish Edition*. San Antonio, TX: Harcourt Assessment.

Data, Code and Materials Availability statement

The IDHC-III form is non-commercially distributed by the MacArthur-Bates CDI Advisory Board, Larry Fenson, Chair, and is available free-of-charge from https://mb-cdi.stanford.edu. The form is also available, along with all data, code, and materials (including Appendices) on OSF at https://osf.io/r6fep/, DOI 10.17605/OSF.IO/R6FEP

Ethics statement

Ethics approval was obtained from the Universidad Autónoma de Querétaro by the first author. All participants gave informed written consent before taking part in the study.

Authorship and Contributorship Statement

DJM developed the measure, data collection and analysis, and prepared the first draft of this manuscript, MF contributed refinements to the measure, data interpretation, organization, and writing the new manuscript, VM contributed refinements to the measure, data analyses, interpretation, and writing of the new manuscript, AW contributed refinements to the measure, data interpretation, and writing of the new manuscript, AA contributed refinements to the measure, acquisition of original data files, data interpretation and writing of the new manuscript, BC contributed to the interpretation of the data and writing of the new manuscript, MRC contributed to the acquisition of original data files and interpretation of data, and PD contributed refinements to the measure, data interpretation, and writing the new manuscript. All authors reviewed and approved the final manuscript.

Declaration of conflict of interests

Authors Friend, Weisleder, Marchman, and Dale are members of the MacArthur-Bates CDI Advisory Board.

Dedication and acknowledgement

This paper is dedicated to the memory of our colleague and good friend, Donna Jackson-Maldonado, whose untimely passing occurred on November 29, 2021. We miss her infectious good spirit, the stories, jokes, and laughter that made collaboration with her such a pleasure. Underpinning that friendship was her intense commitment to the study of language development and disorders, to training new clinicians and researchers, and to helping young children grow. We will miss her tremendously.

We gratefully acknowledge Marta Rubio-Codina, Ph.D., at Inter-American Development Bank, for her assistance in locating many of the files that facilitated our work on this manuscript.

License