International Journal of Language & Communication Disorders

RESEARCH REPORT OPEN ACCESS

JACDI: WS-SF: A Short-Form Version of the Saudi Arabic Communicative Development Inventory

Haifa Alroqi[1] | Michael C. Frank[2] | Khadeejah Alaslani[3] | Aalya Albeeshi[4] | Alaa Almohammadi[1] | Yara Aljahlan[5,6] | Roaa Alsulaiman[7,8] | Abdullah Murad[9] | Fahad Alnemary[10]

[1]Department of Modern Languages and Literatures, King Abdulaziz University, Jeddah, Saudi Arabia | [2]Department of Psychology, Stanford University, Stanford, USA | [3]English Department, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia | [4]Ynmo, Riyadh, Saudi Arabia | [5]College of Medicine, Alfaisal University, Riyadh, Saudi Arabia | [6]King Fisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia | [7]College of Education, King Saud University, Riyadh, Saudi Arabia | [8]Division of Psychology and Language Sciences, University College London, London, UK | [9]College of Computer and Information Systems, Umm Al-Qura University, Mecca, Saudi Arabia | [10]Department of Special Education, Taif University, Taif, Saudi Arabia

Correspondence: Haifa Alroqi (halroqi@kau.edu.sa)

Received: 3 December 2024 | **Revised:** 19 April 2025 | **Accepted:** 20 June 2025

Funding: Open-access publication was supported by an institutional agreement between King Abdulaziz University and Wiley. The authors acknowledge the generous funding provided by Al Muhaidib Social Foundation and Ynmo in support of this study. This research is part of the Faseeh Initiative, which aims to provide free virtual early screening and intervention for language disorders to 10,000 children in Saudi Arabia.

Keywords: CDI | language assessment | language development | Saudi | toddlers | vocabulary development

ABSTRACT

Purpose: This study addresses the lack of a short Communicative Development Inventory (CDI) form tailored to Saudi Arabic, specifically designed based on local norms, for assessing early communicative skills in toddlers. The JISH Arabic CDI: Words and Sentences Short-Form (JACDI: WS-SF) is introduced to meet this need. This tool is particularly useful in contexts of low parental literacy and is designed to offer a culturally relevant and efficient assessment of early language development in Saudi Arabic-speaking children.

Method: The JACDI: WS-SF was developed using a cross-sectional study design. Participants were toddlers aged 16 to 36 months and their parents, who provided data on language development through this newly developed tool.

Results: Preliminary findings indicate that the JACDI: WS-SF effectively captures essential aspects of early communicative development, aligning with local linguistic and cultural norms. The tool proves to be accessible to parents with varying literacy levels and sensitive to the linguistic context of Saudi Arabic. Notably, the study highlighted a significant association between children's ability to combine words and their vocabulary size, emphasising the need for considering language complexity in early interventions. Gender differences in vocabulary size were also noted, calling for equitable support in language assessment and intervention.

Conclusions: The JACDI: WS-SF effectively fills a significant gap in developmental assessment tools in the region, confirming its efficacy in initial testing. The outcomes support the tool's utility and cultural relevance, enhancing its potential to expedite language disorder diagnoses, facilitate rigorous and comparable early language development studies, and broaden accessibility to language assessments. These results underscore the importance of the tool in both research and clinical practice, especially for advancing early intervention strategies.

WHAT THIS PAPER ADDS:

What is already known on this subject

- Previous studies have established the utility of MacArthur-Bates Communicative Development Inventories (CDIs) for assessing early communicative skills in infants and toddlers, with over 100 adaptations in various languages and dialects globally. However, the existing Saudi Arabic version is lengthy, exceeding the length of most long-form CDIs, and includes outdated items. This highlights the need for updated, culturally appropriate tools that accurately reflect current language and cultural experiences.

What this paper adds to the existing knowledge

- This study introduces the JISH Arabic CDI: Words and Sentences Short-Form (JACDI: WS-SF), a culturally and linguistically updated short-form inventory tailored for Saudi Arabic-speaking children. This new tool is designed to be accessible to parents with varying literacy levels and provides a valid and efficient method for assessing early language development. It addresses the need for updated vocabulary reflecting modern Saudi Arabic-speaking children's environments and supports rapid, reliable developmental assessments.

What are the potential or actual clinical implications for this work?

- The JACDI: WS-SF fills a critical void in the assessment toolkit for clinicians working with Saudi Arabic-speaking infants and toddlers. By providing a culturally relevant and efficient means of evaluating early language development, this tool enables more accurate and timely diagnoses of language disorders, ultimately enhancing clinical assessment practices. Its streamlined format is particularly advantageous for use in busy clinical settings and benefits parents with lower literacy levels. By facilitating accurate and timely interventions, the JACDI: WS-SF supports improved developmental outcomes for children across Saudi Arabia.

This paper describes the JISH Arabic CDI: Words and Sentences Short-Form (JACDI: WS-SF) and its development, provides an overview of recently obtained normative data, and discusses the applications of this instrument both in research and clinical settings. JACDI: WS-SF contains a 100-word checklist for vocabulary production for children aged 16 to 36 months. See the appendices for a copy of the short form and the normative tables.

The MacArthur Bates Communicative Development Inventories (CDIs; Fenson et al. 1993, 2007) are parent-report checklists that provide a comprehensive assessment of a wide range of communicative skills in infants and toddlers, including vocabulary, gestures and early language development. They are also cost-efficient and can be easily administered. These inventories supply essential normative data, facilitating the comparison of a child's development against a larger population, thereby aiding in the early identification of potential communication delays or concerns. Nevertheless, the full-length CDI forms sometimes prove extensive and daunting for caregivers, who often contend with demanding schedules and multiple responsibilities (Bello et al. 2018; Fenson et al. 2000; Frota et al. 2016; Kim et al. 2016; Mayor and Mani 2019). The implementation of a shorter CDI form would significantly reduce the time and effort required for completion, enhancing feasibility for caregivers. (Fenson et al. 2000; Frota et al. 2016; Demayo et al. 2021; Mayor and Mani 2019). While we do not claim that shorter forms yield more accurate or reliable responses than long forms, shorter forms may help mitigate issues such as caregiver fatigue, low education levels and time constraints, which can negatively impact response quality, particularly toward the end of longer forms. Transitioning to shorter forms could substantially improve data quality and completion rates (Demayo et al. 2021). Studies have demonstrated that short forms are highly correlated with long forms and can serve as reasonable estimates of vocabulary and grammatical development. For example, Fenson et al. (2000) found that simulated short-form vocabulary scores were highly correlated with vocabulary scores on the full CDI. They also noted that short-form vocabulary scores were highly correlated with grammatical development, suggesting that the short form can effectively capture key aspects of language development. While Fenson et al. (2000) did not explicitly compare caregiver feasibility or response accuracy between short and long forms, their findings support the reliability of short forms as a practical alternative to long forms. Other studies, such as Jackson-Maldonado et al. (2013), have directly compared short and long forms and found strong convergence between the two, further supporting the validity of shorter forms for clinical and research applications. The concise nature of shorter assessments significantly increases caregiver compliance, ultimately leading to larger and more representative sample sizes, thereby enhancing the validity of the data collected (Chai et al. 2020; Demayo et al. 2021; Fenson et al. 2000). Additionally, shorter forms are particularly accessible for caregivers with limited literacy or language skills or from disadvantaged backgrounds. These caregivers may find extensive checklists daunting due to the required time, reading and focus. By reducing cognitive load and completion time, shorter forms promote inclusivity, ensuring that diverse populations can effectively participate in the assessment process.

There are more than 100 adaptations of the original CDI (Fenson et al. 1993, 2007) in various languages and dialects (Demayo et al. 2021). Currently, two CDI tools exist for assessing the communicative skills of Saudi Arabic-speaking children. The first one is the Jeddah Institute for Speech and Hearing (JISH) Arabic CDI (JACDI) long form (550 words in JACDI-Words and Gestures and 898 words in JACDI-Words and Sentences; Dashash and Safi 2014), designed to be used with children aged 8–16 months and 17–36 months respectively. This tool is widely utilised in Saudi speech

and communication centres. However, a drawback of this tool is its length, which exceeds that of most CDIs in other languages. Furthermore, several items in the JACDI have become outdated, reflecting terms and objects that are no longer widely relevant in children's lives (e.g., Gameboy, cassette, cassette player/recorder). As children's language and cultural experiences evolve, it is essential for assessment tools to reflect current vocabulary and concepts to accurately capture early language development. An updated version of the JACDI is needed to address these changes, ensuring that the tool remains relevant, culturally appropriate and effective in evaluating the language skills of today's Saudi children.

The second CDI tool is a short form developed by Abdelwahab et al. (2021) to assess the vocabulary skills of children aged 8 to 30 months across 17 Arabic dialects, including Saudi Arabic. This tool, adapted primarily from Egyptian norms, presents several limitations when applied to Saudi Arabic-speaking children, highlighting the need for a Saudi-specific version. First, the existing tool's sample is heavily skewed, with data drawn primarily from 436 Egyptian children and only 15 Saudi children. This imbalance limits the tool's capacity to accurately represent linguistic and cultural nuances specific to Saudi Arabia. In addition, there are significant socioeconomic differences between the populations of Saudi Arabia and Egypt. Such differences can impact language development patterns uniquely within Saudi contexts. Furthermore, the considerable dialectal and linguistic variability across Arabic-speaking regions—including phonological, morphological and lexical differences—means that dialects unique to Saudi Arabia, such as Najdi, Hijazi and Gulf Arabic, are not adequately represented in the Egyptian-based tool.

We commend Abdelwahab et al. (2021) for their pioneering work in developing a tool intended to serve multiple Arabic dialects. Their efforts to create a resource spanning diverse linguistic contexts represent a significant step forward in Arabic language assessment. However, tools designed for multiple dialects often require refinement by native speakers of each dialect to ensure accuracy and cultural relevance.

The Saudi version of the CDI short form in Abdelwahab et al. (2021) includes several linguistic features that do not reflect the realities of Saudi dialects. For example, 'duck' is represented as بِطّه (/bɪtˤːah/), instead of the Saudi pronunciation بَطّه (/batˤːah/). The word 'hammer' appears as مطْرَجه (/matˤraʤah/), which is inaccurate for Saudi dialects, where the correct form is مطرقة (/mətˤragah/), using a /g/ sound instead of /ʤ/. These discrepancies may stem from differences in the pronunciation of the /q/ sound, which is realised as /q/ or /g/ in Saudi dialects but as a glottal stop /ʔ/ in most Egyptian dialects. The word 'a lot' is represented as وايِض (/waːjɪdˤ/), a form stereotypically used in Egyptian media to mimic Gulf speakers, yet it is not representative of most Saudi dialects. A closer variant, واجِد (/waːʤɪd/) or (/waːjɪd/) with a final /d/ rather than /dˤ/, is used in the Gulf (Sharqi) dialect of the Eastern Region of Saudi Arabia, which shares similarities with other Gulf dialects (e.g., Kuwaiti, Bahraini and Emirati) due to geographical proximity and other factors. Still, it is not representative of Saudi dialects as a whole.

These discrepancies reflect broader differences between Egyptian and Saudi Arabic, including variations in the realisation of sounds such as /q/ and /ʤ/, as well as in vowel usage. Geographical, ethnic and sociolinguistic factors contribute to these differences, making the direct application of Egyptian norms problematic for Saudi populations. The JACDI: WS-SF addresses these concerns by incorporating vocabulary and pronunciation that reflect contemporary Saudi speech, offering multiple word and vowel options to accommodate dialectal variation.

As noted in the literature, Arabic dialects, including Saudi ones, differ more significantly in vowel systems than in consonants (Alorifi 2008; Biadsy et al. 2009), a pattern also observed in languages like English (Cutler et al., 2012; Ladefoged and Johnson 2011). This makes vowel accuracy particularly critical in tools designed for language assessment. By attending to such details, the JACDI: WS-SF ensures both linguistic precision and cultural relevance for Saudi children.

The term 'Saudi Arabic,' as used in the current study, refers to either the White Dialect or the prominent regional dialects spoken across Saudi Arabia (e.g., Al Madani 2022; Albirini 2015; Alyahya and Druks 2015; Al-Rojaie 2020; Rabab'ah and Al-Qarni 2012). The White Dialect denotes a linguistically neutral koiné form that has emerged as a lingua franca among the country's diverse linguistic communities. This dialect is mutually intelligible across different countries and regions and is increasingly used in media alongside Modern Standard Arabic (MSA). The White Dialect serves as a practical linguistic bridge among the diverse Arabic dialects spoken in the country, including Najdi, Hijazi and Gulf Arabic. This dialectal evolution is particularly relevant in urban centres like Riyadh and Jeddah, where significant internal migration has created a melting pot of dialects. The resultant White Dialect is not only widely accessible but also extensively used by both children and adults, facilitating communication across different linguistic backgrounds within the country. In a chapter titled 'The Emergence of a National Koiné in Saudi Arabia: A Perceptual Dialectology Account' in *The Routledge Handbook of Arabic and Identity*, Al-Rojaie (2020) discusses the 'White Dialect' in the context of its perception as a national koiné in Saudi Arabia, noting that this dialect is known by several names, such as 'The White Dialect,' 'The Common Dialect,' 'The Saudi Dialect,' and 'The Standard Dialect,' with 'The White Dialect' often implying a speech form free from distinctive accents or stigmatised features that may reveal a speaker's regional background, thus serving as a neutral, unifying variety for various Saudi populations (Al-Rojaie 2020). This dialect is particularly prevalent among younger Saudis, as observed by Yamani et al. (2024), who noted its growing use among youth in urban settings. This shift reflects a trend towards a more unified and softened form of speech that diverges from traditional regional dialects. This form is increasingly used in both casual interactions and social media, signalling linguistic unification trends driven by internal migration, urban growth, and social media use (Al Madani 2022). Given this scholarly precedent and the sociolinguistic reality in Saudi Arabia, we used 'Saudi Arabic' to denote this broadly understood and used form of Arabic in the country, ensuring our terminology is up-to-date and reflective of current linguistic studies and trends.

Understanding the current linguistic landscape and the widespread use of the White Dialect among Saudi children

is central to developing valid and culturally relevant language assessment tools. Given the limitations of existing CDIs in capturing the specific linguistic features of Saudi Arabic, whether in its regional forms or the emerging national koiné, there is a clear need for an updated, concise tool tailored to the vocabulary and usage patterns of Saudi toddlers today. Consequently, there arises a need to develop a short Saudi Arabic CDI form based on a representative sample of Saudi Arabic-speaking infants and toddlers. Short forms prove essential when rapid assessment of a child's language skills is required or when parental literacy levels are low. To address this need, the present study introduces the JACDI: WS-SF for children aged 16 to 36 months, representing a promising tool for both researchers and clinicians. Given that we derived the vocabulary items used in the current short form from the long form of the JISH CDI, we have decided to name this new tool 'JISH Arabic Communicative Development Inventory: Words & Sentences—Short Form (JACDI: WS-SF)' as a gesture of gratitude towards the JISH and its commendable endeavours in supporting children with language disorders in Saudi Arabia.

1 | Description and Development of the Instrument

JACDI: WS-SF is a short form of the JISH Arabic CDI. It is designed for children between 16 and 36 months, and it contains a 100-word vocabulary production checklist. The directions specify that caregivers should only check the vocabulary items that their child spontaneously says. In addition, the form contains a question asking the caregiver if their child has begun to combine words, with three response options: 'Not Yet,' 'Sometimes,' and 'Often.' The CDI Advisory Board provided authorisation of the JACDI: WS-SF tool as a version of the CDI Words & Sentences form for Saudi Arabic-speaking children.

Items for the short form were drawn from the JACDI: WS long form, which contains 898 words. Our approach was to use item response theory (IRT) to select a subset of items from the original form that would be both semantically diverse and related to overall vocabulary size. IRT describes a family of statistical models that allow the parameters of individual test questions to be estimated from item-level data from a group of test takers; these parameters can then be used to answer questions about test-takers' ability, about individual questions' properties, and to construct related tests (see e.g., Embretson and Reise 2013). To create an item bank, we combined data from the original normative dataset collected by JISH ($N = 643$, ages 8–36 months, mean age = 22.6 months) and another dataset ($N = 199$, ages 12–36 months, mean age = 24.0 months) collected by Alroqi, Almohammadi and Alaslani (available in Wordbank; Frank et al. 2017), with summary scores reported in Kartushina et al. (2022).

We used this combined dataset, which is available for download through Wordbank (Frank et al. 2017), to fit a variety of IRT models, including a Rasch model, a two-parameter logistic model (2PL), and a three-parameter logistic model (3PL). These models all seek to estimate the parameters of individual items based on the performance of participants with a range of different abilities. The Rasch model assumes that every item has a single difficulty parameter, but is otherwise identical. In contrast, the 2PL and 3PL assume that items also vary in their discrimination ability (e.g., that some items are more diagnostic of higher vocabulary children than others). The 3PL additionally includes a parameter that captures the tendency of some words to be known by all children in the sample (e.g., 'mama'). Prior work using IRT models on CDI data has typically adopted a 2PL (e.g., Kachergis et al., 2021; Frank et al. 2021). We compared model fits using two standard indices of model fit, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC); both quantify the tradeoff between model complexity and data likelihood under the model. In our simulations, we found that the 3PL was favoured over the 2PL and Rasch model based on AIC (Rasch AIC = 375986, 2PL AIC = 368675, 3PL AIC = 368383), but the 2PL was favoured based on BIC (Rasch BIC = 379998, 2PL BIC = 376688, 3PL BIC = 380402). Based on these fits and prior work, we decided to adopt the more parsimonious 2PL (see Kachergis et al. 2022 for further details on this comparison).

Using this model, we then attempted to construct maximally useful short forms of 100 items following the length of prior CDI short forms. In IRT models, it is possible to compute the *item information* for a particular item, that is, how informative it is about the ability of test-takers. Typically, items with higher discrimination in a 2PL will have higher item information. When item information is summed, it yields *test information*, the amount of information yielded by a particular test. We adopted test information as the criterion for the evaluation of different potential form compositions.

We were not only interested in constructing a highly informative test, but we also wanted to preserve some of the semantic diversity on the JACDI form. We hence considered a number of possible strategies for short-form construction, including

1. selecting items randomly,
2. selecting items with the highest discrimination,
3. selecting the five items within each of the 20 semantic categories[1] on the JACDI that showed the highest discrimination,
4. balancing difficulty and discrimination within categories by choosing sets of items that jointly maximised the average discrimination and the standard deviation of difficulty, and
5. selecting four items within each category that showed the highest discrimination and adding the easiest item.

Figure 1 shows the resulting test information curves. The simple option of maximising discrimination resulted in a highly informative test. The scores of children on this subset of items were very highly correlated with their total score ($r = 0.979$). This strategy had two weaknesses, however. First, the items were asymmetrically distributed across semantic categories, resulting in a less holistic assessment of vocabulary. Our goal in constructing the JACDI: WS-SF was to include some coverage of all of the different categories represented in the original JACDI. Representing such a wide range of categories necessarily required decreasing coverage of some categories that were highly represented in the original form (e.g., action words; see Table 1), but we decided that having a balanced and diverse set of words was an important goal. Second, when we pursued the maximal discrimination strategy, we found that within children 18 months and under, the correlation with

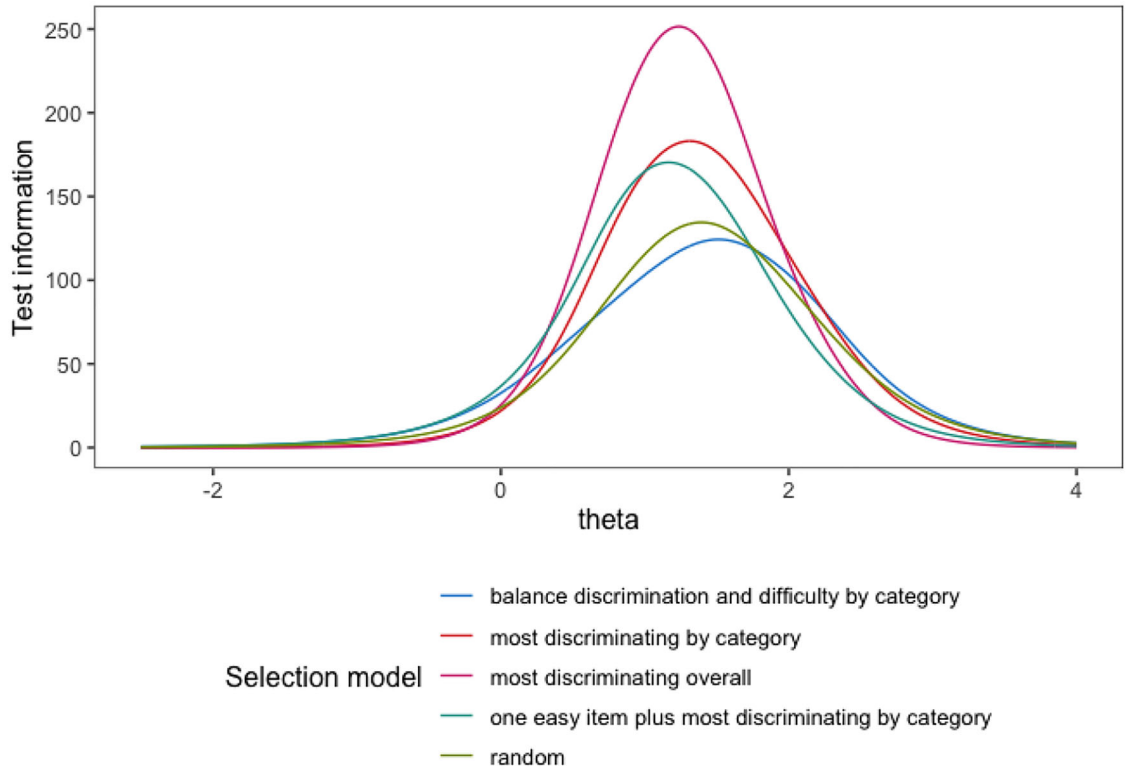

FIGURE 1 | Test Information across a range of ability (theta) values for short forms constructed following different selection models.

TABLE 1 | The representation of semantic categories in the initial JACDI and in JACDI: WS-SF.

Semantic category	# items on JACDI	% of total on JACDI	# items on JACDI: WS-SF	% of total on JACDI: WS-SF
Action words	130	15.4	5	5
Animals (real or toy)	52	6.2	5	5
Body parts	35	4.1	5	5
Clothing	39	4.6	5	5
Connecting words and let	6	0.7	5	5
Descriptive words	84	9.9	5	5
Food and drink	85	10.0	5	5
Furniture and rooms	53	6.3	5	5
Games and routines	49	5.8	5	5
Negation words	11	1.3	5	5
Outside things	46	5.4	5	5
People	32	3.8	5	5
Prepositions and location	21	2.5	5	5
Pronouns	18	2.1	5	5
Quantifiers	28	3.3	5	5
Question words	6	0.7	5	5
Small household items	67	7.9	5	5
Sound effects and animal	26	3.1	0	0
Toys	28	3.3	5	5
Vehicles (real or toy)	18	2.1	5	5
Words about time	12	1.4	5	5

total scores was substantially lower ($r = 0.909$). This weakness is visible on the left side of Figure 1, where test information drops to zero for lower-ability test-takers more quickly for the high-discrimination test than for other alternatives.

We found that strategy #5 – in which we selected a mix of one easy item and four highly discriminating items for each semantic category—helped to overcome these weaknesses. Not only did it yield a test with even semantic category coverage, but the inclusion of an easy item in each category yielded a strong correlation with scores of all children ($r = 0.975$) and a substantially increased correlation with the scores of children 18 months and under ($r = 0.958$). Hence, we adopted this form construction strategy for the final sample.

After revising the final list for cultural appropriateness, six words were replaced with others that belonged to the same semantic categories and had the same approximate discrimination parameter as the ones that were removed. Appendix 1 shows the final production vocabulary list of JACDI: WS-SF with parameters. Table 1 shows the representation of semantic categories in the initial JACDI and in the short form. Note that sound effects showed ceiling effects in this age range and did not contribute to overall test information, so were omitted.

2 | Norming Study

A new norming study was conducted to obtain data on Saudi Arabic-speaking infants and toddlers aged 8 to 36 months. The total number of participants within the 8–15 months age range was insufficient to generate percentiles for the Words and Gestures (WG) form. Therefore, for the current study, only data from children aged 16–36 months whose parents completed the Words and Sentences (WS) form were included. This study obtained ethics approval from the research ethics committee of King Saud University and adhered to the principles outlined in the Helsinki Declaration.

2.1 | Sample

Data was collected in two stages. In Stage 1, participants were recruited via social media platforms (X & WhatsApp) between August and October 2022. They were directed to complete the JACDI: WS-SF via a Google Form link. A total of 361 responses were received, yet 175 were excluded for various reasons: falling outside the original target age range (8–36 months) (88 responses), completion of the Words and Gestures form instead of the Words and Sentences form (77 responses), and caregiver-provided information indicating concerns regarding hearing or language development, or significant medical problems (10 responses). Thus, the total number of valid responses in Stage 1 amounted to 186.

In Stage 2, participants were recruited through the Ynmo Family app, which features the same short form used in Stage 1. The Ynmo Family app is designed to provide parents and professionals in Saudi Arabia with access to a large pool of assessments that enhance clinical evaluation and guide special education and other therapeutic services. This app serves families with young children by offering a platform for language screening and enrollment in a virtual early intervention program if necessary. A total of 301 respondents completed the form via the app. Among these, 167 were excluded from the final sample for similar reasons as in Stage 1: falling outside the target age range (8–36 months) (95 responses), completion of the Words and Gestures form (38 responses), and caregiver-provided information indicating concerns regarding hearing or language development, or significant medical problems (34 responses). Consequently, the total number of valid responses in Stage 2 was 134.

In total, 662 families completed the questionnaire across two recruitment stages. After applying exclusion criteria (e.g., out-of-range child age, completion of the wrong form, and reported developmental or medical concerns), 342 responses were excluded, resulting in a final valid sample of 320. Thus, the response rate (valid responses/total received) was 48.3% (320/662), and the attrition rate was 51.7% (342 / 662).

Upon combining data from both stages, the final sample comprised 320 toddlers (Mean age in months: 27.26; $SD = 6$). Table 2 summarises the socioeconomic characteristics of the final sample. As shown in the table, in our norming sample, the sex distribution reflects a relatively balanced representation. Specifically, 54% of the participants were male, while 46% were female. This sex balance ensures that our assessment tool captures vocabulary size across both boys and girls, contributing to its validity and reliability.

Children and parents in the norming study identify themselves as speaking Saudi Arabic, which has been discussed earlier. The JACDI: WS-SF captures the linguistic diversity within Saudi Arabia by providing alternative vocabulary items for most listed words, reflecting variations in word choice and pronunciation across the most commonly spoken colloquial dialects in Saudi Arabia. This inclusiveness makes the tool highly relevant to a broad demographic. For instance, for the term 'now,' variations like الحين (alˈħiːn), الآن (alˈʔaːn), دحين (daˈħiːn) and ذلحين (ðalˈħiːn) are included to match different regional uses within Saudi Arabia.

Our norming sample includes respondents from the three most populated regions of the country: the Central Region (43%), the Western Region (31%) and the Eastern Region (18%), with other regions each comprising less than 1% of the sample. These areas represent the primary linguistic landscapes where Saudi citizens predominantly reside, as per the latest census data from the Saudi General Authority for Statistics (2022). In these regions, Najdi Arabic is primarily spoken in the Central Region, Hijazi Arabic in the Western Region, and Gulf Arabic, or Shargi, in the Eastern Region. Each region has its unique linguistic traits, all well represented in the JACDI: WS-SF, ensuring that the tool is both relevant and representative of the linguistic profiles across Saudi Arabia.

It was challenging to ensure that the children in the sample were exclusively monolingual. This difficulty stemmed from many households where children are in regular contact at home with domestic workers or nannies who are non-native speakers of Arabic, or they are being exposed to screen media in another language, thereby influencing the linguistic environment in which

TABLE 2 | Socioeconomic characteristics of the sample.

Characteristic	*n*	%
Child sex		
Male	172	53.75
Female	148	46.25
Monthly household income		
SAR 4999 or less	34	10.62
SAR 5000 to 9999	56	17.5
SAR 10,000 to 14,999	68	21.25
SAR 15,000 to 19,999	44	13.75
SAR 20,000 to 24,999	32	10
SAR 30,000 to 34,999	11	3.43
SAR 35,000 to 39,999	8	2.5
SAR 40,000 or more	20	6.25
NC	47	14.68
Arabic only	82	25.6
Hears another language	238	74.37
Hears another language from screens	162	68.06
Hears another language from family members	43	18.06
Hears another language from different sources	21	8.82
Hears another language from school	12	5.04
Parental education	**Mothers**	
	n	**%**
Doctoral degree	11	3.43
Master's degree	40	12.5
Bachelor's degree	194	60.62
Diploma	26	8.12
High school	43	13.43
Intermediate	3	0.93
Elementary	2	0.63
Others	1	0.31

Note: $N = 320$.

they are raised. Parental reports indicate that a significant portion of the children, approximately 74%, are exposed to another language, primarily English. This exposure predominantly occurs through watching shows in English, which accounts for 68% of the reported exposure to another language. Such a multilingual context is common for most Saudi children. Instead of considering it as an interfering factor, it should be regarded as a constant element inherent to the language acquisition process in the Saudi context. Consequently, the participants constitute a relatively homogeneous group; all have occasional exposure to other languages, though not necessarily systematically (such as through attending an English-speaking school or receiving explicit foreign language instruction). This exposure does not reach the extent of being classified as a bilingual situation.

The parental education levels in our sample show notable differences between mothers and fathers. Specifically, 77% of mothers have attained a bachelor's degree or higher, while only 65% of fathers have achieved the same level of education. Conversely, 22% of mothers have completed high school or obtained a diploma after high school, compared to 31% of fathers. Additionally, a smaller proportion of mothers (2%) have low education levels (elementary or intermediate school) compared to fathers (3%). This disparity suggests that mothers in our sample are more highly educated than fathers, reflecting broader demographic trends observed in the national population. National statistics from the Saudi General Authority for Statistics (2020) indicate a higher proportion of female college graduates compared to male graduates. Therefore, the higher education levels among mothers in our sample mirror these national trends.

For the purpose of this study, household income was categorised into three groups based on monthly earnings: lower-income households, defined as those earning less than SAR 10,000 a month; middle-income households, categorised as earning between SAR 10,000 and SAR 19,999 a month; and higher-income households, comprising those earning over SAR 20,000 a month (1 SAR = 0.27 USD as of 24 April 2024). As seen in Table 2, the distribution of income levels within our sample indicates that the middle-income category was the most represented (35%), followed by the lower-income category (28%) and the higher-income category (22%). Additionally, 15% of participants chose not to disclose their monthly income. According to the Saudi General Authority for Statistics (2018), the mean average monthly income for Saudi households is SAR 14,823, while the household median income is SAR 11,494. This comparison provides context for understanding the income distribution in our sample relative to national averages.

3 | Results

Following the approach of Marchman et al. (2023), we estimated developmental curves using generalised additive models. We modelled vocabulary as a beta-distributed proportion of the total form score, using smooth P-splines over both the effect of age on vocabulary and the standard deviation of the age effect (allowing variance to change smoothly with age). Figure 2 shows estimated developmental curves alongside the raw data. All code and data are available at https://github.com/langcog/arabicCAT.

Females tend to have larger vocabulary than males (Frank et al. 2021), and this trend was found in the norming sample as well. We fit a generalised additive model as above, but including a main effect of sex and found that sex was a highly significant predictor ($\beta = -0.43$, $t = -3.35$, $p < 0.001$). Figure 3 shows predictions from this model.

The growth of vocabulary is typically tightly coupled to the growth of word combination, such that children who produce more words tend to combine them more and make more complex sentences (Frank et al. 2021). In JACDI: WS-SF, parents could check whether their child was combining words 'not yet,' 'sometimes,' or 'often.' Table 3 shows the percentages of children in the sample with 'not yet' and affirmative ('sometimes' or 'often') responses in combining words. The proportion of children combining words grew with age, with 'often' showing the most pronounced developmental growth (Figure 4). In addition,

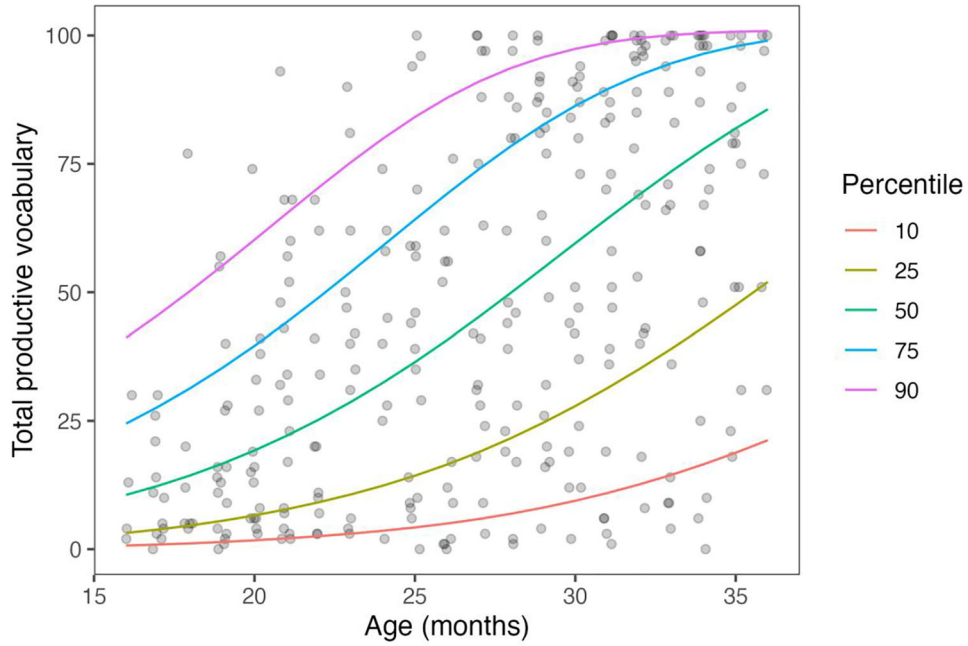

FIGURE 2 | Productive vocabulary plotted by age. Each dot shows an individual child, with points slightly jittered on the horizontal axis to avoid overplotting. Coloured lines show percentile curves derived from the generalised additive model (see text).

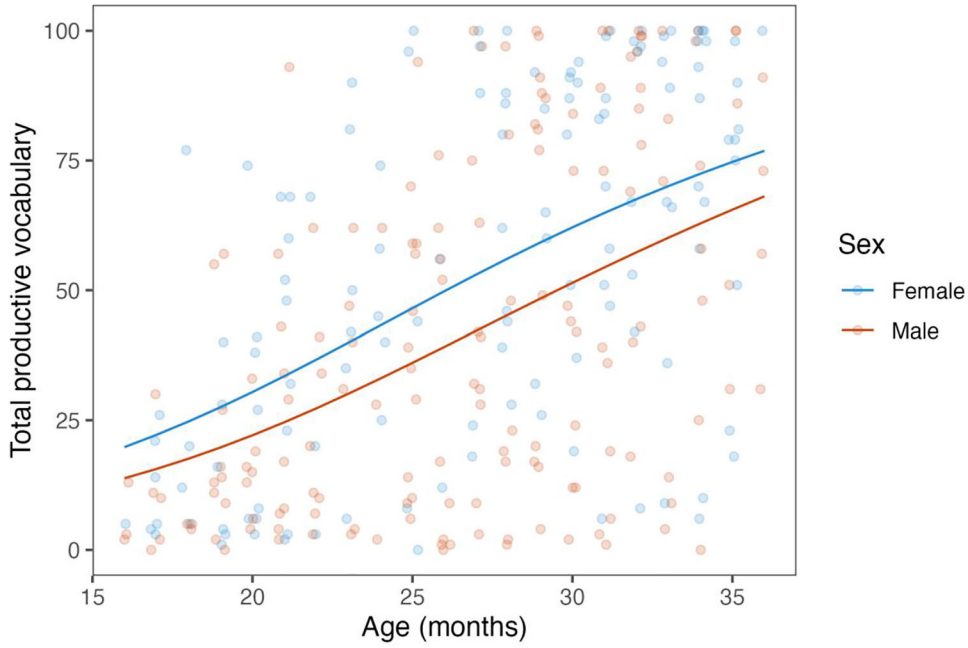

FIGURE 3 | Sex effects on productive vocabulary. Plotting conventions and models are as in Figure 2.

children combining words more frequently had substantially higher vocabulary, even controlling for age (Figure 5). To quantify this pattern, we fit the same generalised additive models as in our prior analyses, but added coefficients for combining words 'sometimes' and 'often.' Both of these were highly significant predictors of vocabulary size over and above age ($\beta = 1.24$, $t = 8.34$, $p < 0.0001$ and $\beta = 2.51$, $t = 16.969$, $p < 0.0001$, respectively).

3.1 | Normative Tables

Appendices 2, 3 and 4 provide separate normative percentile tables for both sexes combined, for females, and for males, presenting raw score values at intervals of every 5th percentile from the 5th percentile to the 99th percentile. These tables enable the evaluation of a child's level in comparison to other children of the same age and sex.

TABLE 3 | Percentages of children with not yet and affirmative (sometimes or often) responses in combining (JACDI: WS-SF).

	Male		Female	
Age in months	***not yet***	***sometimes or often***	***not yet***	***sometimes or often***
16	1.00	0.00	0.33	0.67
17	0.80	0.20	0.67	0.33
18	1.00	0.00	0.50	0.50
19	0.70	0.30	0.50	0.50
20	0.71	0.29	0.63	0.38
21	0.40	0.60	0.22	0.78
22	0.25	0.75	0.67	0.33
23	0.50	0.50	0.17	0.83
24	0.33	0.67	0.00	1.00
25	0.31	0.69	0.20	0.80
26	0.50	0.50	0.50	0.50
27	0.27	0.73	0.00	1.00
28	0.25	0.75	0.00	1.00
29	0.23	0.77	0.00	1.00
30	0.33	0.67	0.00	1.00
31	0.30	0.70	0.00	1.00
32	0.00	1.00	0.13	0.88
33	0.17	0.83	0.13	0.88
34	0.14	0.86	0.17	0.83
35	0.00	1.00	0.00	1.00
36	0.00	1.00	0.00	1.00

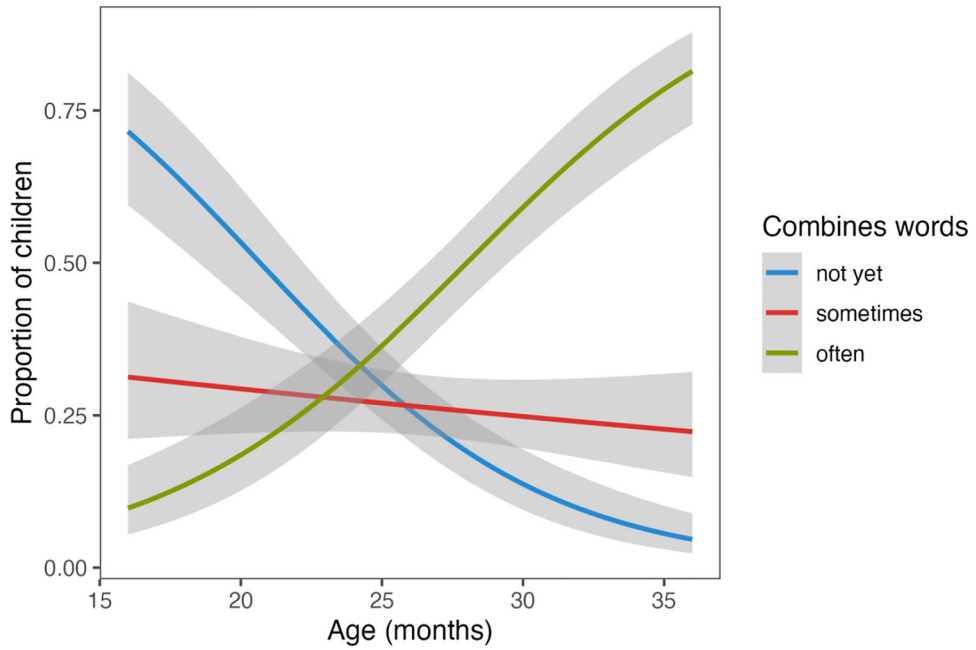

FIGURE 4 | Proportion of children combining words 'Not Yet,' 'Sometimes,' or 'Often' by age, as predicted by a generalised linear model with a logistic link function.

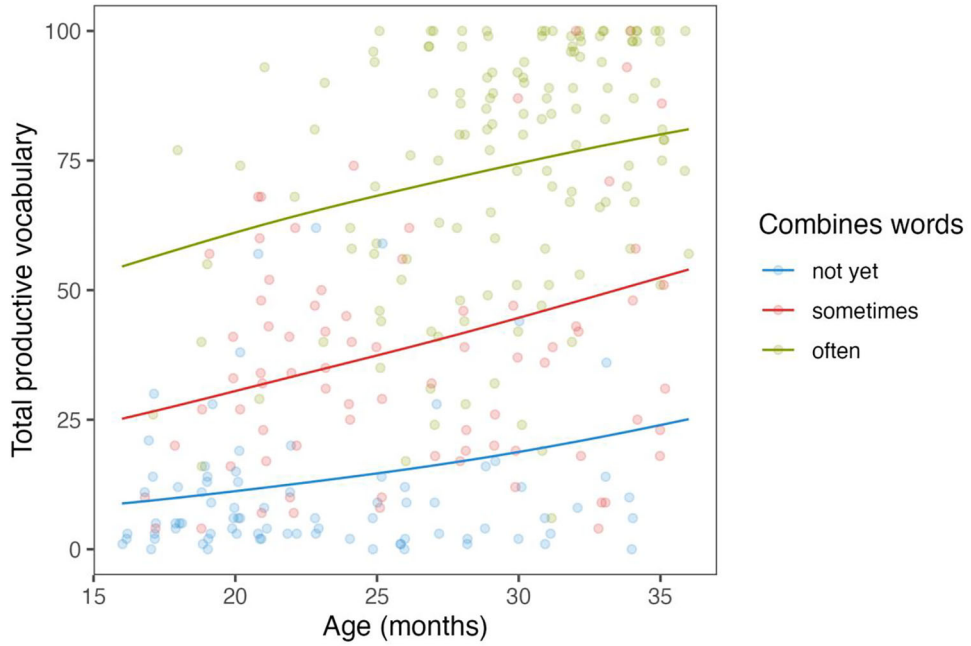

FIGURE 5 | Total productive vocabulary plotted by age and split by the child's level of word combination.

The development of a new CDI tool, such as the JACDI: WS-SF, necessitates a robust psychometric design prior to its deployment. The utility of the CDI as an evaluative instrument hinges on its ability to sensitively detect deviations in early language development among children (Rinaldi et al. 2019). Consequently, ensuring high levels of reliability and validity is crucial. In the following sections, we detail the methods employed to examine the psychometric properties of the new JACDI: WS-SF tool and present the results obtained.

A reliability analysis was performed on the 100 items of the JACDI: WS-SF using the normative dataset of 320 children. We first computed Cronbach's alpha (α) to assess internal consistency. A Cronbach's α value of 0.7 or higher is generally accepted as indicating strong reliability (Kline 2015). As is common for CDI forms—since they include so many items—α was extremely high (.99), demonstrating consistent measurement of early language. However, α is known to overestimate reliability. Following the IRT-based approach we used in the construction of the short form, we fit a 2-parameter logistic model and computed marginal reliability, which was 0.96. These findings suggest that the 100 items of the JACDI: WS-SF offer a reliable index of individual children's vocabulary size.

High correlations between vocabulary and grammar sections (combining words) were also reported, confirming the tool's reliability. These findings align with previous research confirming the reliability of tools used to measure early language development. Mariscal and Gallego (2012) summarise numerous studies demonstrating a strong correlation between lexical and grammatical development. For instance, Bates et al. (1988) found correlations ranging from 0.54 to 0.83 in children aged 20–28 months, while Fenson et al. (1994) reported a correlation of 0.84 using the MacArthur-Bates Communicative Development Inventories (MBCDI). Similar findings were observed in adaptations of these inventories across Romance languages (Caselli et al. 1999; Devescovi et al. 2005; Jackson-Maldonado et al. 2003; López-Ornat et al. 2005; Pérez-Pereira and García Soto 2003) and non-Romance languages (Barreña et al. 2008). These findings also align with previous research on similar CDI tools, such as the New Zealand CDI-WS (Reese and Read 2000), which reported a long-term reliability of 0.61 for the grammatical complexity subsection and 0.75 for the vocabulary subsection over approximately 6 months. High correlations between vocabulary and grammar sections further confirm the reliability of these tools in capturing early language development, despite differences in methodologies, such as longitudinal versus cross-sectional designs.

As highlighted by Valentini and Serratrice (2021), extensive research demonstrates a strong relationship between vocabulary and grammar skills. This positive correlation has been widely documented across various languages (Devescovi et al. 2005; Mariscal and Gallego 2012; Stolt et al. 2009; Thordardottir et al. 2002) and is evident in both monolingual and bilingual children (Marchman, et al. 2004; Parra et al. 2011; Simon-Cereijido and Gutiérrez-Clellen 2009). In their own study, Valentini and Serratrice (2021) found that all tested models showed a significant correlation between vocabulary and grammar skills in a sample of 100 bilingual children, further reinforcing the interconnectedness of these linguistic domains.

In terms of validity, several key aspects were considered. JACDI: WS-SF was perceived by participating parents as a valuable resource for assessing children's early language development. This perception, supported by a moderate-to-high response rate of 48.3%, indicates that parents took the tool seriously and were willing to engage with it meaningfully. Face validity is further enhanced as, in Stage 2, participants were recruited via the Ynmo app, a trusted platform for language development. The use of this respected resource assured that parents trusted the platform and were more likely to provide honest and accurate responses.

The establishment of content validity was supported by the careful construction and evaluation of the tool by language development and assessment experts, who confirmed the tool's appropriateness and relevance for the linguistic needs of the Saudi Arabic-speaking population. The final short form, constructed using a balanced approach of selecting one easy item and four high-discrimination items from each of the 20 semantic categories, demonstrated strong overall correlation with total scores ($r = 0.975$). Notably, this approach significantly improved correlation scores for children aged 18 months and younger ($r = 0.958$ compared to $r = 0.909$ with other strategies). In addition, revisions for cultural appropriateness further strengthened the tool's ability to assess vocabulary development.

Convergent validity was assessed by comparing vocabulary size with grammatical complexity. Consistent with previous research (e.g., Frank et al. 2021), children with larger vocabulary sizes were found to produce more complex grammatical structures. Additionally, parents reported on their children's ability to combine words, which, as anticipated, became more frequent as vocabulary size and child age increased. This relationship between vocabulary growth and grammar development further validates the JACDI-WS-SF as an effective instrument for assessing early language development milestones.

The tool was initially piloted to ensure that it was appropriate for the Saudi Arabic-speaking context. This pilot phase allowed for necessary adjustments and provided preliminary evidence of the tool's reliability and validity. Before full deployment, the JACDI: WS-SF was piloted on a smaller sample to assess its reliability and functionality. This pilot stage allowed for the refinement of the test items and provided preliminary evidence of reliability through internal consistency and early feedback from participants.

Overall, the development of the Saudi Arabic JACDI: WS-SF was grounded in theoretical models of language acquisition, ensuring that it aligned with current literature on early childhood language development. The psychometric properties of the JACDI: WS-SF, including its internal consistency, face, content and convergent validity, demonstrate that it is a reliable and valid tool for assessing early language development in Saudi children.

4 | Discussion

The primary objective of this study has been successfully achieved through the development of a new and up-to-date norm-referenced language development assessment tool tailored for Saudi Arabic-speaking children aged 16–36 months. This new tool holds significant value, serving not only as a crucial asset for future research endeavors but also as a valuable tool in clinical applications.

In our study, we found that females in the norming sample tend to have a larger vocabulary than males. This aligns with established patterns found in other languages. Gender differences in language development have been extensively researched and documented in the literature (Frank et al. 2021; Talbot 2020; Wallentin 2020). Consistently, girls tend to begin to speak at an earlier age compared to boys (Frank et al. 2021). Research indicates that from early childhood, girls typically outperform boys in both receptive and expressive vocabulary measures (Bleses et al., 2018; Frank et al. 2021). This finding underscores a robust gender-based distinction in language acquisition processes that transcends linguistic boundaries.

Our study revealed a notable association between children's ability to combine words and their vocabulary size, with this relationship becoming more pronounced as children grow older. This finding aligns with previous research indicating that the growth of vocabulary is closely intertwined with the development of word combination skills, suggesting a natural progression in language acquisition (Frank et al. 2021). Specifically, we observed that children who frequently combined words demonstrated a significantly larger vocabulary size, even after accounting for age differences. This relationship underscores the interconnectedness between different aspects of language development, where the ability to combine words reflects not only syntactic complexity but also lexical richness. These findings have important implications for understanding the dynamics of language development in young children. They highlight the importance of assessing not only vocabulary size but also syntactic abilities, such as word combination, in evaluating language skills comprehensively. Clinically, interventions targeting language development may benefit from focusing on promoting both vocabulary expansion and syntactic complexity concurrently, as these aspects of language development appear to be mutually reinforcing.

The introduction of the JISH Arabic Communicative Development Inventory: Words & Sentences—Short Form (JACDI: WS-SF) holds significant implications for both research and clinical practice, offering a valuable tool for assessing and promoting child development in the Saudi Arabic-speaking population. In clinical settings, the JACDI: WS-SF fills a critical void in the assessment toolkit for clinicians working with Saudi Arabic-speaking infants and toddlers. By providing a culturally relevant and efficient means of evaluating early language development, this tool facilitates more accurate and timely diagnoses of language disorders, ultimately enhancing clinical assessment practices. Moreover, in research contexts, the availability of a standardised assessment tool tailored to this population enables more rigorous and comparable studies on early language development. Researchers can now collect data more quickly, assess a larger sample size within a limited time frame, and utilise the short form in longitudinal research to track children's progress over time. Additionally, it aids in identifying children at risk for language delays or disorders and evaluating the effectiveness of early intervention programs in improving communication skills.

The development of a short-form assessment tool promotes accessibility to language assessment for a broader range of caregivers, improving the inclusivity and reach of early intervention efforts. By providing a concise and user-friendly instrument, the JACDI: WS-SF serves as an efficient screening tool during routine check-ups or initial assessments, enables remote administration for telehealth sessions, and supports ongoing monitoring of a child's progress. Additionally, the observed gender differences in vocabulary size underscore the importance of considering gender-related factors in early language assessment and intervention, ensuring equitable support for all children. Furthermore, understanding the relationship between vocabulary size and the ability to

combine words in utterances informs intervention strategies aimed at promoting language fluency and complexity in early childhood. In conclusion, the introduction of the JACDI: WS-SF offers various benefits for enhancing clinical assessment practices, improving research methodologies, promoting accessibility to language assessment, and informing intervention strategies, ultimately contributing to the advancement of early childhood development efforts in the Saudi Arabian context.

4.1 | Limitations

Although our work here establishes a new instrument for use in research and clinical applications, there are some important limitations. First and foremost, we have not validated the JACDI: WS-SF as a diagnostic instrument to be used in making a diagnosis of conditions related to language delay; this limitation is shared with other CDI forms (Marchman et al. 2023). CDI forms, especially short forms, should be used for screening and research, rather than diagnosis. Second, we suggest some caution in the interpretation of percentile scores on the JACDI: WS-SF given the relatively small sample sizes in each age group of our study. We used generalised additive models to smooth out variation across ages following current best practices (Marchman et al. 2023), but a larger sample size would be helpful in more precisely estimating percentile bands.

Additionally, although the JACDI: WS-SF was derived from the JACDI long form and showed high statistical correlation with total scores ($r = 0.975$), a direct empirical within-subject comparison between the long and short forms was not conducted in the current study. Initially, such a comparison was considered; however, it was beyond the scope of this phase of research and thus not completed. Therefore, the strong correlations presented here are based solely on statistical modelling and item-selection processes. Future research is planned to explicitly collect data from parents completing both versions, thus directly evaluating the reliability, validity and practical feasibility of the JACDI: WS-SF in relation to the original long form.

5 | Conclusion

The collaboration of the co-authors in this study exemplifies successful interdisciplinary efforts in advancing global understanding of children's language development, particularly in addressing the linguistic needs of diverse backgrounds and understudied languages, reflecting our commitment to inclusivity and diversity in research. For future research directions, we recommend further collaboration between researchers and clinicians to collect additional data aimed at enhancing the norms of our assessment tool, thereby increasing its validity and reliability. Efforts should also focus on gathering more data among younger children aged 8–16 months in order to develop a Words and Gestures WG short form for this age range. However, it is worth noting that our attempts to collect data on younger children have encountered challenges, notably a low response rate from parents of infants under 1 year old. Our observation suggests that Saudi parents of children in this age group may be hesitant to participate in studies on infants' language development. This hesitancy could stem from the demands of caring for young infants, coupled with the perception that language development is not a pressing concern at such an early stage. Thus, raising awareness about the importance of early childhood research and the potential benefits of collaboration between families and researchers is essential. Parents need to understand that assessing children's comprehension skills, which can precede their production skills, is valuable even at a young age. By emphasising the significance of early language development and the role of families in research endeavours, we can foster greater engagement and participation from parents of infants and toddlers, ultimately enriching our understanding of early language acquisition processes.

Acknowledgement

We gratefully acknowledge King Abdulaziz University for facilitating open-access publication through its institutional agreement with Wiley. We extend our sincere gratitude to all participants for their invaluable contribution to the various stages of this study, aiding in the development of a much-needed assessment tool for evaluating children's vocabulary development. Additionally, we express our thanks to the Jeddah Institute for Speech and Hearing (JISH) for providing the foundation for our study. The naming of our new tool, 'JISH Arabic Communicative Development Inventory: Words & Sentences—Short Form (JACDI: WS-SF)', reflects our deep appreciation for JISH's commitment to aiding children with language disorders and supporting families in nurturing their children's language development.

Conflicts of Interest

The authors declare that this study received funding from Al Muhaidib Social Foundation and Ynmo. Open-access publication of this article was made possible through an institutional agreement between King Abdulaziz University and Wiley. The authors declare no relevant financial or non-financial conflicts of interest with respect to the research, authorship and/or publication of this article.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Note

[1]We excluded the "animal sounds" category from this analysis as items were found to have low discrimination parameters.

References

Abdelwahab, A. G. S., S. Forbes, A. Cattani, J. Goslin, and C. Floccia. 2021. "An Adaptation of the MacArthur-Bates CDI in 17 Arabic Dialects for Children Aged 8 to 30 Months." *Language Learning and Development* 17, no. 4: 425–446. https://doi.org/10.1080/15475441.2021.1916502.

Albirini, A. 2015. *Modern Arabic Sociolinguistics: Diglossia, Variation, Codeswitching, Attitudes, and Identity*. 1st ed. Routledge. https://doi.org/10.4324/9781315683737.

Al Madani, R. 2022. "The Role of Digital Multilingual Patterns and Functions in Branding and Self-Branding: An Investigation of Practices by Saudi Users on Twitter." Doctoral dissertation, Cardiff University Repository.

Alorifi, F. S. 2008. "Automatic Identification of Arabic Dialects Using Hidden Markov Models." Doctoral dissertation, University of Pittsburgh. https://d-scholarship.pitt.edu/8131/.

Al-Rojaie, Y. 2020. "The Emergence of a National Koiné in Saudi Arabia: A Perceptual Dialectology Account." In *The Routledge Handbook of Arabic and Identity*, edited by R. Bassiouney and K. Walters, 1st ed. Routledge. https://doi.org/10.4324/9780203730515.

Alyahya, R. S. W., and J. Druks. 2015. "The Adaptation of the Object and Action Naming Battery Into Saudi Arabic." *Aphasiology* 30, no. 4: 463–482. https://doi.org/10.1080/02687038.2015.1070947.

Barreña A., I. García, M. J. Ezeizabarrena, et al. 2008. *MacArthur-Bates komunikazio garapena neurtzeko zerrenda. Euskerara egokitua. Erabiltzaileentzako gida Eta eskuliburu teknikoa*. UEU.

Bates, E., I. Bretherton, and L. Snyder. 1988. *From First Words to Grammar: Individual Differences and Dissociable Mechanisms*. Cambridge University Press.

Bello, A., D. Onofrio, L. Remi, and C. Caselli. 2018. "Prediction and Persistence of Late Talking: A Study of Italian Toddlers at 29 and 34 Months." *Research in Developmental Disabilities* 75: 40–48. https://doi.org/10.1016/j.ridd.2018.02.006.

Biadsy, F., J. Hirschberg, and N. Habash. 2009. "Spoken Arabic Dialect Identification Using Phonotactic Modeling." In Proceedings of the EACL 2009 Workshop on Computational Approaches to Semitic Languages, 53–61. Association for Computational Linguistics.

Bleses, D., W. Vach, and P. S. Dale. 2018. "Self-Reported Parental Vocabulary Input Frequency for Young Children." *Journal of Child Language* 45: 1073–1090, https://doi.org/10.1017/S0305000918000089.

Caselli, C., P. Casadio, and E. Bates. 1999. "A Comparison of the Transition From First Words to Grammar in English and Italian." *Journal of Child Language* 26, no. 1: 69–111. https://doi.org/10.1017/S0305000998003687.

Chai, J. H., C. H. Lo, and J. Mayor. 2020. "A Bayesian-Inspired Item Response Theory-Based Framework to Produce Very Short Versions of MacArthur-Bates Communicative Development Inventories." *Journal of Speech, Language, and Hearing Research: JSLHR* 63, no. 10: 3488–3500. https://doi.org/10.1044/2020_JSLHR-20-00361.

Cutler, A., T. Otake, and L. Bruggeman. 2012. "Phonologically Determined Asymmetries in Vocabulary Structure Across Languages." *The Journal of the Acoustical Society of America* 132, no. 2: EL155–EL160. https://doi.org/10.1121/1.4737596.

Dashash, N., and S. Safi. 2014. *JISH Arabic Communicative Development Inventory: Saudi Population JACDI: User's Guide and Technical Manual*. Jeddah Institute for Speech and Hearing.

deMayo, B., D. Kellier, and M. Braginsky, et al. 2021. "Web-CDI: A system for online administration of the MacArthur-Bates Communicative Development Inventories." *Language Development Research* 1, no: 1, 55–98. https://doi.org/10.34758/kr8e-w591.

Devescovi, A., M. C. Caselli, D. Marchione, P. Pasqualetti, J. Reilly, and E. Bates. 2005. "A Crosslinguistic Study of the Relationship Between Grammar and Lexical Development." *Journal of Child Language* 32, no. 4: 759–786. https://doi.org/10.1017/s0305000905007105.

Embretson, S. E., and S. P. Reise. 2013. *Item Response Theory*. Psychology Press.

Fenson, L., E. Bates, P. S. Dale, V. A. Marchman, J. S. Reznick, and D. J. Thal. 2007. *The MacArthur–Bates Communicative Development Inventories*. Brookes.

Fenson, L., P. S. Dale, J. S. Reznick, et al. 1993. *The MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual*. Singular.

Fenson, L., P. S. Dale, J. S. Reznick, E. Bates, D. J. Thal, and S. J. Pethick. 1994. "Variability in Early Communicative Development." *Monographs of the Society for Research in Child Development* 59, no. 5: 1–185.

Fenson, L., S. Pethick, C. Renda, J. L. Cox, P. S. Dale, and J. S. Reznick. 2000. "Short-Form Versions of the MacArthur Communicative Development Inventories." *Applied Psycholinguistics* 21, no. 1: 95–116. https://doi.org/10.1017/S0142716400001053.

Frank, M. C., M. Braginsky, D. Yurovsky, and V. A. Marchman. 2017. "Wordbank: An Open Repository for Developmental Vocabulary Data." *Journal of Child Language* 44, no. 3: 677–694. https://doi.org/10.1017/S0305000916000209.

Frank, M. C., M. Braginsky, D. Yurovsky, and V. A. Marchman. 2021. *Variability and Consistency in Early Language Learning: The Wordbank Project*. MIT Press.

Frota, S., J. Butler, S. Correia, C. Severino, S. Vicente, and M. Vigário. 2016. "Infant Communicative Development Assessed With the European Portuguese MacArthur–Bates Communicative Development Inventories Short Forms." *First Language* 36, no. 5: 525–545. https://doi.org/10.1177/0142723716648867.

Jackson-Maldonado, D., V. A. Marchman, and L. C. H. Fernald. 2013. "Short-Form Versions of the Spanish MacArthur–Bates Communicative Development Inventories." *Applied Psycholinguistics* 34, no. 4: 837–868. https://doi.org/10.1017/S0142716412000045.

Jackson-Maldonado, D., D. Thal, V. Marchman, L. Fenson, T. Newton, and B. Conboy. 2003. *CDI Inventarios MacArthurBates Del Desarrollo de Habilidades Comunicativas [CDI MacArthur-Bates Communicative Development Inventories]*. Brookes Publishing.

Kachergis, G., V. A. Marchman, and M. C. Frank. 2021. "Toward a "Standard Model" of Early Language Learning." *Current Directions in Psychological Science*, 31, no. 1: 20–27. https://doi.org/10.1177/09637214211057836.

Kachergis, G., V. A. Marchman, P. S. Dale, J. Mankewitz, and M. C. Frank. 2022. "Online Computerized Adaptive Tests of Children's Vocabulary Development in English and Mexican Spanish." *Journal of Speech, Language, and Hearing Research* 65, no. 6: 2288–2308. https://doi.org/10.1044/2022_JSLHR-21-00372.

Kartushina, N., N. Mani, A. Aktan-Erciyes, et al. 2022. "COVID-19 First Lockdown as a Window Into Language Acquisition: Associations Between Caregiver-Child Activities and Vocabulary Gains." *Language Development Research* 2, no. 1:. https://doi.org/10.34842/abym-xv34.

Kim, S. W., J. Y. Kim, S. Y. Lee, and H. R. Jeon. 2016. "The Comparison of M-B CDI-K Short Form and K-ASQ as Screening Test for Language Development." *Annals of Rehabilitation Medicine* 40, no. 6: 1108–1113. https://doi.org/10.5535/arm.2016.40.6.1108.

Kline, P. 2015. *A Handbook of Test Construction (Psychology revivals): Introduction to Psychometric Design*. Routledge.

Ladefoged, P., and K. Johnson. 2011. *A Course in Phonetics*, 6th ed. Wadsworth.

López-Ornat, S., C. Gallego, P. Gallo, A. Karousou, S. Mariscal, and M. Martínez. 2005. Inventarios de desarrollo comunicativo MacArthur. Manual Técnico [MacArthur Communicative Development Inventories. Technical manual]. Madrid, Spain: TEA, Ediciones.

Marchman, V. A., P. S. Dale, and L. Fenson. 2023. *MacArthur-Bates Communicative Development Inventories User's Guide and Technical Manual*, 3rd ed. Brookes Publishing.

Marchman, V. A., C. Martínez-Sussmann, and P. S. Dale. 2004. "The Language-Specific Nature of Grammatical Development: Evidence From Bilingual Language Learners." *Developmental Science* 7, no. 2: 212–224. https://doi.org/10.1111/j.1467-7687.2004.00340.x.

Mariscal, S., and C. Gallego. 2012. "The Relationship Between Early Lexical and Grammatical Development in Spanish: Evidence in Children With Different Linguistic Levels." *Spanish Journal of Psychology* 15, no. 1: 112–123. https://doi.org/10.5209/rev_SJOP.2012.v15.n1.37293.

Mayor, J., and N. Mani. 2019. "A Short Version of the MacArthur-Bates Communicative Development Inventories With High Validity." *Behavior Research Methods* 51, no. 5: 2248–2255. https://doi.org/10.3758/s13428-018-1146-0.

Parra, M., E. Hoff, and C. Core. 2011. "Relations Among Language Exposure, Phonological Memory, and Language Development in Spanish–English Bilingually Developing 2-Year-Olds." *Journal of Experimental*

Child Psychology 108, no. 1: 113–125. https://doi.org/10.1016/j.jecp.2010.07.011.

Pérez-Pereira, M., and X. R. García-Soto. 2003. "El Diagnóstico del Desarrollo Comunicativo en la Primera Infancia: Adaptación de las Escalas MacArthur al Gallego [The diagnosis of communicative development in early childhood: adaptation of the MacArthur inventories Into Galician]." *Psicothema* 15, no. 3: 352–361.

Rabab'ah, G., and A. M. Al-Qarni 2012. "Euphemism in Saudi Arabic and British English." *Journal of Pragmatics* 44, no. 6-7: 730–743. https://doi.org/10.1016/j.pragma.2012.02.008.

Reese, E., and S. Read 2000. "Predictive Validity of the New Zealand MacArthur Communicative Development Inventory: Words and Sentences." *Journal of Child Language* 27, no. 2: 255–266. https://doi.org/10.1017/S0305000900004098.

Rinaldi, P., P. Pasqualetti, S. Stefanini, A. Bello, and M. C. Caselli 2019. "The Italian Words and Sentences MB-CDI: Normative Data and Concordance Between Complete and Short Forms." *Journal of Child Language* 46, no. 3: 546–566. https://doi.org/10.1017/S0305000919000016.

Saudi General Authority for Statistics. 2018. Household Income and Expenditure Survey. https://www.stats.gov.sa/sites/default/files/household_income_and_expenditure_survey_2018_en_27-6-2019.pdf.

Saudi General Authority for Statistics. 2020. Saudi Women: The Partners of Success. https://www.stats.gov.sa/sites/default/files/woman_international_day_2020EN.pdf.

Saudi General Authority for Statistics. 2022. Population by Region, Nationality and Gender. https://portal.saudicensus.sa/portal/public/1/15/45?type=DASHBOARD.

Simon-Cereijido, G., and V. F. Gutiérrez-Clellen. 2009. "A Cross-Linguistic and Bilingual Evaluation of the Interdependence Between Lexical and Grammatical Domains." *Applied Psycholinguistics* 30, no. 2: 315–337. https://doi.org/10.1017/S0142716409090134.

Stolt, S., L. Haataja, H. Lapinleimu, and L. Lehtonen. 2009. "Associations Between Lexicon and Grammar at the End of the Second Year in Finnish Children." *Journal of Child Language* 36, no. 4: 779–806. https://doi.org/10.1017/S0305000908009161.

Talbot, M. 2020. *Language and Gender*, 3rd ed. Polity Press.

Thordardottir, E. T., S. E. Weismer, and J. L. Evans. 2002. "Continuity in Lexical and Morphological Development in Icelandic and English-speaking 2-Year-Olds." *First Language* 22, no. 1: 3–28. https://doi.org/10.1177/014272370202206401.

Valentini, A., and L. Serratrice. 2021. "What Can Bilingual Children Tell Us About the Developmental Relationship Between Vocabulary and Grammar?" *Cognitive Science* 45, no. 3: e13062. https://doi.org/10.1111/cogs.13062.

Wallentin, M. 2020. "Gender Differences in Language Are Small but Matter for disorders." In *Handbook of Clinical Neurology*, edited by R. Lanzenberger, G. S. Kranz, and I. Savic, Vol. 175, 81–102. Elsevier. https://doi.org/10.1016/B978-0-444-64123-6.00007-2.

Yamani, A., R. Alziyady, R. AlYami, et al. 2024. "The KIND Dataset: A Social Collaboration Approach for Nuanced Dialect Data Collection." In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: Student Research Workshop*, 32–43. Association for Computational Linguistics.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supplementary File: jlcd70108-sup-0001-Appendix.pdf