Speech, Language and Hearing

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/yslh20

# Assessing vocabulary and grammar development in New Zealand Mandarin- and Cantonese-speaking children: a validation study

**Yuxin Zhang, Elaine Ballard, Taiying Lee, Henrietta Lee, Johanna Schmidt & Elaine Reese**

**To cite this article:** Yuxin Zhang, Elaine Ballard, Taiying Lee, Henrietta Lee, Johanna Schmidt & Elaine Reese (18 Jan 2024): Assessing vocabulary and grammar development in New Zealand Mandarin- and Cantonese-speaking children: a validation study, Speech, Language and Hearing, DOI: 10.1080/2050571X.2024.2303825

**To link to this article:** https://doi.org/10.1080/2050571X.2024.2303825

SPEECH, LANGUAGE AND HEARING
https://doi.org/10.1080/2050571X.2024.2303825

Taylor & Francis
Taylor & Francis Group



# Assessing vocabulary and grammar development in New Zealand Mandarin- and Cantonese-speaking children: a validation study

Yuxin Zhang[a], Elaine Ballard[b], Taiying Lee[c], Henrietta Lee[d], Johanna Schmidt[e] and Elaine Reese[a]

[a]Department of Psychology, University of Otago, Dunedin, New Zealand; [b]Faculty of Science, Psychology, University of Auckland, Auckland, New Zealand; [c]Academic Unit of Human Communication, Development, and Information Sciences, Faculty of Education, The University of Hong Kong, Hong Kong; [d]Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand; [e]Te Kura Aronui – School of Social Sciences, University of Waikato, Hamilton, New Zealand

**ABSTRACT**

New Zealand has a growing population of Mandarin and Cantonese speakers, including children with speech, language, and communication disorders. However, the absence of language development profiles poses challenges in effectively identifying and addressing these concerns. This study aimed to capture the current state of monolingual, bilingual, and multilingual Mandarin- and Cantonese-speaking children living in New Zealand. Over 200 mothers in the *Growing Up in New Zealand* longitudinal cohort study identified their children as understanding Mandarin and/or Cantonese. Mothers completed newly adapted vocabulary and grammar checklists for New Zealand Mandarin and Cantonese speakers at the Age 2 data wave. Both of the adapted New Zealand Mandarin and Cantonese versions showed high reliabilities and validities. Unique demographic predictors of children's vocabulary and grammar were mothers' education, household deprivation level, and children's birth order. Language status and maternal concerns were also unique predictors of children's vocabulary development in Mandarin and Cantonese, with monolingual children whose mothers reported no concerns having higher Mandarin and Cantonese vocabularies. These results have rich implications for researchers, clinicians, and practitioners working with Mandarin- and Cantonese-speaking children in New Zealand and worldwide.

**ARTICLE HISTORY**
Received 17 August 2023
Accepted 7 January 2024

**KEYWORDS**
Bilingualism; multilingualism; Chinese language acquisition; assessment

## Introduction

Oral language development in early childhood is a key focus for professionals and researchers as it serves as a strong predictor of later language and literacy development (Greenwood et al., 2017; Harris, Law, & Roy, 2005). The considerable impact of lexical and grammatical development in the early years on later language growth generates great interest among researchers and clinicians. For instance, children who experience language difficulties in later childhood often exhibited risk factors as toddlers (Horwitz et al., 2003; Reilly et al., 2009). Conversely, children who possess advanced oral language skills during their early years demonstrate better early literacy abilities when they enter school and enhanced reading comprehension later in primary school (Suggate, Schaughency, McAnally, & Reese, 2018). Therefore, understanding key factors influencing early language development is crucial for effective interventions and addressing long-term challenges associated with oral language issues for literacy, academic performance, and social interactions (Simonsen, Kristoffersen, Bleses, Wehberg, & Jørgensen, 2014).

In addition to the significant impact of adult input (Hart & Risley, 1995), various factors such as gender, socioeconomic status (SES), birth order, maternal education, and parental concerns have been identified as influential in shaping early language development (Horwitz et al., 2003; Reese et al., 2018; Reese & Read, 2000; Reilly et al., 2007). Furthermore, the role of bilingualism and multilingualism, the ability to speak two or more languages, has also been recognized as a significant factor in early language growth (Hoff et al., 2012; Pearson, Fernández, & Oller, 1993; Reese et al., 2018). Although a considerable body of research on Chinese children's language development has primarily been conducted in China or Chinese-dominant countries (Hao, Shu, Xing, & Li, 2008; Hua & Wei, 1999; Tardif, Fletcher, Zhang, Liang, & Zuo, 2008b), there has been an increase in studies focusing on this population in English-dominant societies in recent years (Hao, Bedore, Sheng, & Peña, 2019; Jia & Paradis, 2015; Sun, Low, & Chua, 2022; Sun, Yin, Amsah, & O'Brien, 2018). In New Zealand, there has been a noticeable increase in the population of Mandarin and Cantonese speakers, including children

**CONTACT** Yuxin Zhang yuxin.zhang@postgrad.otago.ac.nz Department of Psychology, University of Otago, William James Building, 275 Leith Walk, Dunedin 9016, New Zealand

with speech, language, and communication disorders (Ballard & Lee, 2016). Despite this growing demographic, there is a critical gap in the identification and assessment of these children's language development within the specific context of New Zealand. The primary objective of this study was to provide a current overview of Mandarin and Cantonese acquisition among children participating in the *Growing Up in New Zealand* study. Specifically, we investigated the predictors of variation in language proficiency among Mandarin and Cantonese-speaking children in New Zealand with a newly adapted version of the MacArthur-Bates Communicative Development Inventory (CDI): short form (Fenson et al., 2000), with permission from the CDI board. To achieve this, we assessed 236 Mandarin and Cantonese speakers from the *Growing Up in New Zealand* study, a longitudinal birth cohort study that included over 6,500 diverse families representing the New Zealand population (Morton et al., 2012, 2013).

## Measuring early language in bilingual children

Accurately assessing the linguistic development of young bilingual children in child language research is challenging due to the risk of over- and under-identifying speech, language, and communication disorders (Aldawood, Hand, & Ballard, 2023). Collecting, transcribing, and analysing naturalistic language samples requires extensive time and effort, which is further complicated when dealing with bilingual speakers. To overcome this challenge, parent-report inventories, such as the MacArthur-Bates Communicative Development Inventories (CDIs), have emerged as reliable and valid alternatives for assessing language proficiency in children aged 8–37 months (Fenson et al., 2000; Fenson et al., 2007; Fernald, Perfors, & Marchman, 2006). Originally developed for American English (Fenson et al., 1994, 2000), the CDI has since been adapted to over fifty languages, including different variations of English (e.g., British English: Dionne, Dale, Boivin, & Plomin, 2003; New Zealand English: Reese & Read, 2000; see https://mb-cdi.stanford.edu/adaptations.html for a current list) as well as numerous other languages (e.g., Spanish: Jackson-Maldonado, Marchman, & Fernald, 2013; Danish: Bleses et al., 2008; Canadian French: Poulin-Dubois, Graham, & Sippola, 1995). These adaptations allowed for broader applicability and enabled researchers to assess language development in diverse linguistic contexts.

The CDI:II for toddlers focuses on assessing vocabulary and grammar. Parents or caregivers complete the questionnaire by indicating the words and/or sentences their children are producing (Fenson et al., 1994). They also answer a question about children's acquisition of word combination skills, which is commonly used as an indicator of grammatical development (Bates & Goodman, 1997). The CDI:II long form provides a valid and reliable measure of language development for children aged 16–30 months but can be time-consuming, taking 30–45 min to complete. In contrast, the CDI:II short form offers a more concise version with a 100-word vocabulary checklist and one grammar question, which is also valid and reliable (Fenson et al., 2000, 2007). The short form serves as a practical alternative to the long form, particularly in situations where time and resources are limited. It has gained wide acceptance and is frequently used when administering the long form is challenging or time-consuming (Fenson et al., 2000). Therefore, the CDI short form is well-suited for comprehensive health and development studies like *Growing Up in New Zealand* (Reese et al., 2015, 2018).

Researchers in other countries have created several parental questionnaires tailored to assess the early language development of Mandarin- and Cantonese-speaking children (Hao et al., 2008; Hua & Wei, 1999; Tardif et al., 2008b). Tardif and colleagues adapted the CDI short forms to Mandarin and Cantonese languages to examine vocabulary and grammatical development in children aged 8–30 months from Beijing and Hong Kong, providing valuable insights into language development in Mandarin- and Cantonese-speaking populations (Tardif et al., 2008b; Tardif, Fletcher, Liang, & Kaciroti, 2009). However, Hao et al. (2008) raised concerns about the suitability of certain words on the Mandarin CDI short form, as they were direct translations from English and not colloquial to native Mandarin speakers. To address this limitation, Hao et al. (2008) developed their own Chinese vocabulary inventory and reported its reliability. Indeed, these adapted language measures were conducted in China, where most participants were monolingual and from single-child households. Most New Zealand Chinese families, however, are bilingual or multilingual with multiple children. Therefore, there is a need for adapted Mandarin and Cantonese CDIs for the New Zealand population. In this study, two novel NZ CDI:II short forms were developed exclusively for NZ Mandarin and Cantonese speakers, adapted from Tardif et al. (2008b) and Fenson et al. (2000).

## Demographic factors in children's early language development

When assessing language development in Mandarin or Cantonese speaking children in an English-speaking country, it is fundamental to consider existing research on language development in English-speaking children. Previous studies have suggested that gender can be a predictor of early lexical and grammatical development, with young girls often exhibiting more advanced linguistic skills compared to boys (Fenson

et al., 2000; Reese & Read, 2000; Reilly et al., 2007). However, it should be mentioned that this association does not consistently hold true in all cases (see exceptions in Horwitz et al., 2003). More robust predictors of language development are family SES and maternal education; children from low-SES families and with mothers who have lower levels of education are more likely to have lower language skills and experience language delay (Hart & Risley, 1995; Horwitz et al., 2003; Reilly et al., 2007). Several studies have indicated that birth order can also influence language development. First-born children tend to have larger vocabularies and more advanced syntax compared to later-born children (Hoff-Ginsberg, 1998; Reese et al., 2018; Reilly et al., 2009). On the other hand, later-born children often demonstrate more advanced pronoun use and conversational skills (Hoff-Ginsberg, 1998; Oshima-Takane, Goodz, & Derevensky, 1996). These birth order effects have been linked to differences in maternal language input, with mothers using longer utterances with first-born children and employing more questioning techniques with later-born children (Hoff-Ginsberg, 1998). Finally, several studies have consistently demonstrated that parental concern (see Appendix 3 for an example), measured by a question asking mothers whether they have concerns about their children's speech or hearing, serves as a reliable predictor of children's language development (Klee, Pearce, & Carson, 2000; Reese et al., 2018).

Research on Chinese children's language acquisition has found that girls, first-born children, those from higher socioeconomic backgrounds, and those with highly educated parents tend to show language advantages (Tardif et al., 2009; Zhang, Jin, Shen, Zhang, & Hoff, 2008). In the current study, we aimed to investigate whether similar language disparities associated with gender, socioeconomic status (area-level deprivation), language status (monolingual/bilingual/multilingual), birth order, maternal education, and maternal concern can be observed among Mandarin- and Cantonese-speaking children in New Zealand. Given that the majority number of children in each sample used English as an additional language (67% in Mandarin, 71% in Cantonese), we also examined language differences in English vocabulary and word combinations within the English-Mandarin[1] and English–Cantonese[2] subsamples as a function of these demographic predictors.

### *Bilingual and multilingual language development*

Recent studies focusing on bilingual and multilingual children's language development have contributed to a better understanding of their developmental patterns and trajectories (De Houwer, Bornstein, & Putnick, 2014; Grech & McLeod, 2012). Previous research has yielded mixed findings when comparing bilingual and monolingual children in terms of vocabulary and grammatical development pace (Hoff et al., 2012; Pearson et al., 1993). It is now established that bilingual children may take longer to acquire each language compared to monolingual children (Hoff et al., 2012). However, this does not necessarily imply that bilingual or multilingual children experience language delays compared to monolingual children (Grech & McLeod, 2012). Although bilingual children may have smaller vocabularies in each language individually, their combined vocabularies (De Houwer et al., 2014) and grammatical development in any language are comparable (Hoff et al., 2012). Some researchers argue that the total number of concepts for bilingual children is similar to monolingual children when considering their languages together (Pearson et al., 1993). This conceptual vocabulary approach combines bilingual children's vocabulary knowledge across languages, taking advantage of their ability to respond in either language (Peña, Bedore, & Kester, 2016). Bilingual children are estimated to have a comparable 'conceptual vocabulary' when considering words across both languages without counting cross-language synonyms twice (Reese et al., 2018). However, it remains unclear how conceptual vocabulary differs from word forms. This could be because acquiring concepts and word forms varies among languages (Matthews, 1991). For example, English word form changes for tense, whereas Mandarin maintains consistent word forms regardless of grammar. Interestingly, Australian Mandarin-English bilingual children aged 3–7 showed balanced development in both languages, unaffected by language input and dominance (Yang, Kim, Tuomainen, & Xu Rattanasone, 2022).

Although there is limited research on multilingual children's speech and language development, it follows similar patterns to that of bilingual children (Grech & McLeod, 2012). It is suggested that multilingual children may take longer to reach language milestones compared to bilingual children, as their language input is distributed across three or more languages (Mieszkowska et al., 2017). To gain a comprehensive understanding of bilingual/multilingual children's development, it is crucial to compare measures of conceptual vocabulary and expressive vocabulary, along with grammatical measures, in each of their languages whenever feasible.

## The present study

The emergence of Chinese immigrants in New Zealand can be historically situated in the mid-1860s. Since then, there has always been a consistent number of Chinese populations residing in New Zealand (Ballard & Lee, 2016). The 2018 census reported that a total of 247,700 Chinese individuals currently inhabit in

New Zealand, constituting approximately 4.9% of the overall populace (Statistics New Zealand, 2018).

Early Chinese immigrants predominantly originated from the Canton regions of China, where Cantonese serves as their primary language. However, with the implementation of the Immigration Act 1987, there has been a 23.6% increase in Mandarin speakers over the past few decades. Many subsequent immigrants from various parts of China exhibit bilingual or multilingual abilities (Ballard & Lee, 2016). Indeed, Mandarin and Cantonese are distinct languages, lacking mutual intelligibility due to differences in their phonological systems and lexical structures (Tardif et al., 2009). One phonological contrast lies in the number of tones, with Mandarin having four and Cantonese having six. Additionally, Mandarin and Cantonese share less than half of their vocabulary, indicating significant lexical divergence (Snow, 2004). Findings from the nine-month datawave in *Growing Up in New Zealand study* indicated that among 6,383 mothers, 3.3% spoke Mandarin and 1.3% spoke Cantonese to their infants. By age two, 40% of the children in the full sample demonstrated comprehension of two or more languages, and 7% exhibited comprehension of three or more languages (Morton et al., 2014). Hence, the prevalence of bilingualism and multilingualism among the present generation of children is a prominent feature, highlighting the importance for professionals to update their knowledge about these children to effectively support their needs (Aldawood et al., 2023).

To the best of our knowledge, there has been no research on Chinese children's language development from birth to two years old in a longitudinal cohort sample in New Zealand. This study focuses on Mandarin- and Cantonese-speaking children recruited through the *Growing Up in New Zealand* study cohort at age two. The study aimed to (1) validate the adapted New Zealand Mandarin and Cantonese versions of the CDI, and (2) explore the predictors of language acquisition in a representative group of toddlers. We hypothesized that factors such as gender, socioeconomic status (area-level deprivation), maternal education, birth order, language status, and maternal concerns would influence children's language performance. Consistent with previous research (Bleses et al., 2008; Harris et al., 2005; Horwitz et al., 2003; Jackson-Maldonado et al., 2013; Reese et al., 2015, 2018; Reese & Read, 2000; Reilly et al., 2009), we hypothesized that children from less deprived areas, with more educated parents, female, first-born, monolingual, and with no maternal concerns about speech and hearing would exhibit higher scores on the Mandarin and/or Cantonese CDI inventories. For the children who also spoke New Zealand English, we had no specific hypotheses since this study was the first to assess New Zealand English language acquisition for Chinese speakers.

**Table 1.** Demographics for children and families responding to the NZ Mandarin CDI-sf and the Cantonese CDI-sf.

| | Mandarin speakers $N = 177$ | Cantonese speakers $N = 59$ |
|---|---|---|
| Maternal education (highest level) | | |
| High school or lower[a] (%) | 10 | 13 |
| Diploma/Trade certificate (%) | 23 | 30 |
| Bachelor's degree (%) | 46 | 38 |
| Postgraduate degree (%) | 21 | 17 |
| Maternal birthplace (% of mothers born outside NZ) | 98 | 90 |
| Maternal self-prioritised ethnicity (% Asian) | 98 | 95 |
| Area-level deprivation | | |
| Most deprived[b] (%) | 25 | 15 |
| Moderately deprived (%) | 51 | 53 |
| Least deprived (%) | 24 | 30 |
| Child gender (% male) | 49 | 52 |
| Birth order (% first born) | 61 | 55 |
| Child language status | | |
| Monolingual (%) | 31 | 23 |
| Bilingual[c] (%) | 52 | 42 |
| Multilingual[d] (%) | 18 | 35 |

Notes: [a]The value of this variable for the Cantonese sample is fewer than 10.
[b]The value of this variable for the Cantonese sample is fewer than 10.
[c,d]Among the bilingual and multilingual children, 67% of the Mandarin speakers and 71% of the Cantonese speakers used English as another language.

## Method

### Participants

*Growing Up in New Zealand* is a comprehensive longitudinal pre-birth cohort study that involved 6,853 children and their families, representing a broad and diverse sample in terms of ethnicity and socioeconomic status (Morton et al., 2013). Recruitment of mothers and their partners (if available) occurred between 2009 and 2010, prior to the birth of their children (Morton et al., 2013). When children were approximately 24 months old (Age two data wave), approximately 196 children who spoke Mandarin and/or Cantonese and had mothers capable of reporting on their vocabulary in those languages were included in the sample, out of a total of 6,327 mother respondents. Among the overall mother respondents, 196 (3%) reported that their child could understand Mandarin, and 71 (1%) reported that their child could understand Cantonese. Further analysis was conducted on 177 (90%) of the children whose mothers could complete the vocabulary checklist in Mandarin, excluding three subsequent births of twin pairs, and on 59 (83%) of the children whose mothers could complete the checklist in Cantonese without the need for an interpreter or interviewer. Among the Mandarin-speaking sample, 20 (11%) of the children spoke Cantonese and 119 (67%) of the children also spoke English. Among the Cantonese-

speaking sample, 20 (34%) of the children spoke Mandarin and 42 (71%) of the children also spoke English.

The broader *Growing Up in New Zealand* study collected additional demographic information relevant to language development, including birth order, maternal self-prioritised ethnicity, area-level deprivation, maternal education, and birthplace. These data were gathered during the antenatal period or through a phone call at 6 weeks, along with child gender information. Socioeconomic status was determined using the well-established 2006 New Zealand Index of Deprivation, which utilizes eight socioeconomic dimensions derived from census data developed (Salmond, Crampton, & Atkinson, 2007). Scores on this scale range from 1 (least deprived) to 10 (most deprived), and a tripartite classification of deprivation levels was used: low (≤ 3), medium (4–7), and high (8–10). Maternal education was categorized into five subgroups based on the highest level of education completed: intermediate, high school, trade certificate/diploma, bachelor's degree, or postgraduate degree. Table 1 presents the demographic characteristics of the final sample of Mandarin- and Cantonese-speaking children.

## Measures

### Expressive language

Child vocabulary in Mandarin and Cantonese and word combination skills in any language were measured using the two new adapted versions of the MacArthur-Bates Communicative Development Inventory (CDI) short form in Mandarin and Cantonese for New Zealand children. For children who also spoke English, their English vocabulary was also assessed with the NZ CDI:II short form (Reese et al., 2018). To construct the New Zealand adapted versions of the CDI for Mandarin and Cantonese, the third and fourth authors, who are Mandarin and Cantonese native speakers respectively and had experience working with children, consulted with Chinese families more generally and made recommendations to Tardif's (2008b) existing word list. These were discussed with the second author before items that did not fit the NZ context were removed to produce 100-word versions that were similar in scope to the long-form Chinese versions (Tardif et al., 2008b) regarding lexical categories, such as action words, pronouns etc. Each word was assigned a score of one point, resulting in a scoring range from 1 to 100 for the questionnaire. Additionally, the questionnaire included a question regarding the children's ability to combine words in any language for a measure of grammatical development. The response options for this question were '*not yet*' '*sometimes*' and '*often*'; the scoring range for this question was from 1 to 3.

## Procedure

### Antenatal measures

During the antenatal phase, mothers were interviewed using a 90-minute face-to-face computer-assisted personal interview (CAPI) conducted by trained interviewers at the family's home. The interview covered a range of topics, including questions about the mothers themselves, their families, and other important information such as the household SES relevant to their children.

### Child expressive language at two years

At around 24 months of age, the child expressive language assessment took place during the computer-assisted personal interview (CAPI) with Mandarin- and Cantonese-speaking mothers as part of the *Growing Up in New Zealand* study. Initially, mothers were asked to indicate all the languages that their children could understand, including New Zealand English. Subsequently, for each of these languages, mothers were presented with show cards containing numbered word lists specific to the respective language (see the Appendices for the word lists in Mandarin and Cantonese; for the word list in New Zealand English, please refer to Reese et al., 2018). The standard instructions for the CDI: II short form (Fenson et al., 2000) were slightly adapted for the computer-assisted procedure: *Children understand many more words than they say. With this question, we are particularly interested in the words your child SAYS. Please look at this list and tell me the numbers for the words you have heard [Child's name] use. If [Child's name] uses a different pronunciation of a word (e.g., 'sketti' instead of 'spaghetti'), still tell me the number for that word. Remember that this is a list of all the words that are used by many different children – don't worry if your child only uses a few of these words* (Reese et al., 2018). Interviewers entered the participant's response for each numbered word on each show card.

Following the completion of the vocabulary checklist in all the languages the mother believed her child understood, the word combination question was asked in English (*Has your child begun to combine words yet, such as 'another cookie' or 'doggie bite'?*; Fenson et al., 2000), provided with three options: '*not yet*' '*sometimes*' and '*often*'. For this particular question, mothers were asked to provide their response considering all languages spoken by the child.

After completing the word combination question, mothers were asked if they had any concerns about their child's speech or hearing. If they expressed a concern, they were further prompted to specify the type of concern using the categories provided in Appendix 3. Additionally, mothers were asked to rate the level of their concern, ranging from not concerned at all to very concerned. Furthermore, they were asked whether they had sought professional advice or

treatment regarding their child's speech and/or understanding.

## Results

### Psychometric properties

To evaluate the reliability of each inventory, Cronbach's alpha was employed as a measure of inter-item correlations within each scale. The Mandarin and Cantonese vocabulary scales demonstrated high reliability, with Cronbach's alpha coefficients of .99 for both. These high alpha values are consistent with those reported for the English CDI: II short forms (Fenson et al., 2000).

### Vocabulary development

Table 2 displays the average vocabulary scores for Mandarin and Cantonese speakers as a function of by gender, birth order, language status (monolingual, bilingual, and multilingual), maternal education, area-level deprivation (least, moderately, and most deprived), and maternal concerns regarding the child's speech or hearing. Language status was dummy coded as monolingual, bilingual, and multilingual; area-level deprivation was dummy coded into least, moderately, and most deprived. Prior to regression analysis, the intercorrelations among the predictors were assessed, revealing low correlations among all eight predictors ($r_s < .6$); therefore, all were retained and included in the final regression model. Due to minimal variation in maternal self-prioritised ethnicity and maternal birthplace, these two variables were excluded from the regression analysis. Specifically, 98% of Mandarin and 95% of Cantonese participants identified themselves as Asian, while 98% of Mandarin and 90% of Cantonese participants were born outside of New Zealand. Multiple linear regression analyses were conducted to explore the best predictors for children's vocabulary knowledge. Two separate regression models were built for the Mandarin and Cantonese samples and two separate regression models for the English-Mandarin and English–Cantonese subsamples. Percentile rankings for monolingual, bilingual, and multilingual Mandarin- and Cantonese-speaking children on both inventories can be found in Appendices 4–6.

### Mandarin speakers

For Mandarin speakers, the average total vocabulary was 28.94 words. The most common words were 鞋 (shoe, 65%), 水 (water, 64%), 狗 (dog, 64%), 手 (hand, 61%), 抱 (hug, 53%) (see Appendix 1). When all the predictors were entered simultaneously into a regression analysis, the model accounted for 36% of the variance in Mandarin vocabulary, $R^2 = .36$, $F(8, 168) = 11.62$, $p < .001$, which could be explained by two unique predictors in the final model: monolingual ($\beta = 0.49$, $p < .001$) and maternal concerns ($\beta = -0.16$, $p = .01$). Mandarin monolingual children and those whose mothers had no concerns about children's speech and hearing exhibited larger vocabularies than Mandarin bilingual children or those whose mothers had concerns.

### English-Mandarin speakers

For English-Mandarin speakers, the average total English vocabulary was 19.56 words. When all the predictors were entered simultaneously into a regression analysis, the model accounted for 13% of the variance in English vocabulary, $R^2 = .13$, $F(7, 111) = 2.41$, $p = .02$. Maternal education was the only significant predictor ($\beta = 0.28$, $p = .00$); children whose mothers reported lower education had smaller English vocabularies than children whose mothers reported higher education.

### Cantonese speakers

For Cantonese speakers, the average total vocabulary was 28.13 words. The most common words were 狗 (dog, 58.3%), 車 (car, 55%), 打 (hit, 53.3%), 頭髮 (hair, 50%), 襪 (sock, 48.3%), 帽 (hat, 48.3%) (see Appendix 2). When all the predictors were entered simultaneously into a regression analysis, the model accounted for 47% of the variance in the Cantonese vocabulary checklist, $R^2 = .47$, $F(8, 50) = 5.49$, $p < .001$, which could be explained by three unique predictors in the best fitting model: maternal education ($\beta = -0.29$, $p = .01$), monolingual status ($\beta = 0.44$, $p < .001$), and maternal concerns ($\beta = -0.24$, $p = .03$). Cantonese monolingual children, those whose mothers reported less education, and those whose mothers reported no concerns exhibited larger vocabularies than Cantonese bilingual children, those whose mothers reported higher education, and those whose mothers had concerns about the child's speech or language.

### English–Cantonese speakers

For English–Cantonese speakers, the average total English vocabulary was 20.26 words. When all predictors were entered simultaneously into a regression analysis, the model was not significant and there were no significant differences in English vocabulary by gender, birth order, language status (bilingual and multilingual), maternal education, area-level deprivation (least, moderately, and most deprived), or maternal concerns, $R^2 = .20$, $F(7, 34) = 1.25$, $p = .31$.

Overall, the patterns for vocabulary size for the Mandarin and Cantonese full samples as a function of demographics were similar for language status and maternal concerns: Monolingual Mandarin and Cantonese speakers and those whose mothers reported no concerns had larger vocabularies

**Table 2.** Mean vocabulary (and SD) and word combination scores on the NZ Mandarin and Cantonese CDI-sf as a function of demographics and maternal concerns.

| | Mandarin vocabulary (total words) | Word combinations (% not yet combining) | Cantonese vocabulary (total words) | Word combinations (% not yet combining) |
|---|---|---|---|---|
| **Child gender** | | | | |
| Girls | 29.69 (29.15) | 10 | 32.79 (32.47) | 3 |
| Boys | 28.16 (27.85) | 20 | 24.20 (24.21) | 10 |
| **Birth order** | | | | |
| First-borns | 31.88 (30.53) | 11 | 26.40 (27.62) | 6 |
| Later-borns | 24.44 (24.47) | 20 | 31.00 (30.25) | 8 |
| **Language status** | | | | |
| Monolinguals | 52.65 (31.77) | 13 | 59.23 (30.40) | 0 |
| Bilinguals | 20.80 (19.61) | 15 | 22.92 (21.53) | 8 |
| Multilinguals | 11.77 (16.71) | 16 | 15.90 (21.22) | 10 |
| **Maternal education**[a] | | | | |
| Less educated | 32.17 (27.93) | 12 | 38.73 (30.98) | 9 |
| More educated | 27.32 (28.68) | 16 | 20.30 (24.16) | 4 |
| **Area-level deprivation** | | | | |
| Most deprived | 30.25 (28.72) | 16 | 32.78 (33.35) | 11 |
| Moderately deprived | 26.81 (26.84) | 13 | 28.44 (28.38) | 10 |
| Least deprived | 32.05 (31.60) | 16 | 26.22 (28.17) | 0 |
| **Maternal concern** | | | | |
| Not concerned | 31.38 (29.62) | 11 | 31.04 (29.00) | 6 |
| Concerned | 18.68 (20.14) | 30 | 11.75 (20.67) | 13 |

Notes: [a]Less educated = highest level of education intermediate or high school qualification; more educated = highest level of education trade certificate/diploma or bachelor's or postgraduate degree.

compared to bilingual children or those whose mother reported concerns in that language. However, for Cantonese speakers only, vocabulary differed as a function of maternal education in the opposite direction to that predicted, with children whose mothers were less educated having a larger Cantonese vocabulary than children whose mothers were more educated. This pattern was the opposite for the English vocabularies in English-Mandarin subsamples as a function of maternal education, with children whose mothers reported lower education having smaller English vocabularies than children whose mothers reported higher education.

## Grammatical development

Table 2 also displays the average word combination scores for Mandarin and Cantonese speakers as a function of gender, birth order, language status (monolingual, bilingual, and multilingual), maternal education, area-level deprivation (least, moderately, and most deprived), and maternal concerns regarding the child's speech or hearing. Again, language status was dummy coded as monolingual, bilingual, and multilingual, while area-level deprivation was dummy coded into least, moderately, and most deprived. Before conducting the regression analysis, the intercorrelations among the predictors were examined to avoid multicollinearity and found to be low ($r_s < .6$) for all eight predictors. Consequently, all predictors were retained and included in the final regression model. Maternal self-prioritised ethnicity and maternal birthplace were excluded from the regression analysis as most mothers of Mandarin and Cantonese speakers self-prioritised as Asian (98% of Mandarin and 95% of Cantonese) and were born outside New Zealand (98% of Mandarin and 90% of Cantonese). Since the dependent variable is discrete, multivariable ordinal regression models were used to estimate the cumulative odds ratio (OR) for word combinations with all predictor variables. Two separate regression models were built for the Mandarin and Cantonese samples and two separate regression models were built for the English-Mandarin and English–Cantonese subsamples.

### Mandarin speakers

A chi-square test of model fit demonstrated that the model was a good fit to the data, $\chi^2$ (8) = 25.64, $p < .001$. Significant variations existed in Mandarin speakers' word combinations as a function of birth order, area-level deprivation (moderately and most deprived), and maternal concerns. The odds of children combining words in any language were 2.40 times higher (95% CI [1.27, 4.51]) for later-born children compared to first-born children, Wald $\chi^2$ (1) = 7.29, $p < .01$. Additionally, the odds of children combining words in any language were 3.62 times lower (95% CI [1.71, 7.67]) for mothers who reported concerns compared to those who did not, Wald $\chi^2$ (1) = 11.31, $p < .001$. Furthermore, the odds of children combining words in any language were 2.36 times higher (95% CI [1.13, 4.91]) for moderately deprived households compared to the most deprived households, Wald $\chi^2$ (1) = 5.22, $p = .02$.

### English-Mandarin speakers

A chi-square test of model fit demonstrated that the model was a good fit for the English-Mandarin subsample, $\chi^2$ (7) = 26.23, $p < .001$. Significant variations existed in English-Mandarin speakers' word combinations as a function of birth order, maternal education, area-level deprivation (moderately and most

deprived), and maternal concerns. The odds of children combining words in any language were 2.37 times higher (95% CI [1.06, 5.28]) for later-born children compared to first-born children, Wald $\chi^2$ (1) = 4.46, $p$ = .04. The odds of children combining words in any language were 3.13 times higher (95% CI [1.26, 7.75]) for children whose mothers reported higher education compared to children whose mothers reported lower education, Wald $\chi^2$ (1) = 6.07, $p$ = .01. The odds of children combining words in any language were 4.31 times higher (95% CI [1.66, 11.19]) for children in moderately deprived household compared to children in most deprived households, Wald $\chi^2$ (1) = 9.00, $p$ <.001. The odds of children combining words in any language were 4.27 times higher (95% CI [1.68, 10.88]) for children whose mothers reported no concerns compared to children whose mothers reported concerns, Wald $\chi^2$ (1) = 9.26, $p$ <.001.

### *Cantonese speakers*

For Cantonese speakers, there were no significant differences in word combinations by gender, birth order, language status (monolingual, bilingual, and multilingual), maternal education, area-level deprivation (least, moderately, and most deprived), or maternal concerns, Wald $\chi^2$ (8) = 7.56, $p$ = .48.

### *English–Cantonese speakers*

For English–Cantonese speakers, there were no significant differences in word combinations by gender, birth order, language status (bilingual and multilingual), maternal education, area-level deprivation (least, moderately, and most deprived), or maternal concerns, Wald $\chi^2$ (7) = 6.02, $p$ = .54. Overall, in both the full Mandarin samples and the English-Mandarin subsample, word combinations were predicted by similar demographic factors such as birth order, maternal concerns, and area-level deprivation (moderately and most deprived). However, maternal education was a significant predictor only in the English-Mandarin subsample, and not in the full Mandarin sample.

## Discussion

This study had two aims: (1) to assess the reliability of the New Zealand-adapted short forms of the Mandarin and Cantonese versions of the Communicative Development Inventories (CDI), and (2) to investigate the language development of monolingual, bilingual, and multilingual Mandarin- and Cantonese-speaking children in the context of the *Growing Up in New Zealand* study. We found that both of the NZ adapted Mandarin and Cantonese CDI short forms demonstrated high reliability. Furthermore, for vocabulary development, language status and maternal concerns emerged as strong predictors of vocabulary in Mandarin and Cantonese-speaking children, with monolingual children whose mothers reported no concerns exhibiting larger vocabularies compared to bilingual speakers whose mothers showed concerns. For Cantonese speakers only, children of mothers having lower education displayed larger Cantonese vocabularies.

The picture was different for Mandarin-speaking children who also spoke English. We found that English-Mandarin speaking children whose mothers reported lower education had smaller English vocabularies compared to children whose mothers reported higher education. Additionally, maternal education was a predictor for English-Mandarin speakers' word combinations; English-Mandarin speaking children whose mothers reported lower education were less likely to combine words than children whose mothers reported higher education.

For grammatical skills across all children's languages, similar predictors identified were birth order, SES, and maternal concerns of Mandarin-speaking and English-Mandarin speaking children's grammar, where children from households with high deprivation and whose mothers expressed concerns about their speech or hearing exhibited lower word combination scores. Contrary to expectations, later-born Mandarin speakers exhibited more advanced grammatical skills across their languages than first-borns.

### *Vocabulary development*

The findings revealed that children in both language groups produced less than 30% of the total vocabulary listed in the inventories. Specifically, the most frequently used words among Mandarin-speaking children were 'shoe,' 'water,' 'dog,' 'hand,' and 'hug,' whereas Tardif et al.'s (2008a) study identified 'mommy,' 'daddy,' 'grandma-paternal,' 'grandpa-paternal,' and 'hello/wei' as the top words in their sample. For Cantonese speakers, the prominent words in our study were 'dog,' 'car,' 'hit,' 'hair,' 'sock,' and 'hat,' whereas Tardif et al.'s (2008a) study found 'daddy,' 'aah,' 'mommy,' 'YumYum,' and 'older sister' as the most common words. These variations in word types and categories can be attributed to multiple factors. Firstly, the children in our study were assessed at the age of two, which is older than the children in Tardif et al.'s (2008a) study, who ranged from 8 to 16 months old. Additionally, the differences in word selection may arise from disparities in the inventories themselves. It is interesting that the word 'dog' emerged as a top word for both Mandarin and Cantonese-speaking children in our study but not in Tardif et al.'s (2008a) study. This dissimilarity may be due to age differences or possibly influenced by the prevalence of dog ownership within New Zealand Chinese families. Moreover, the greater exposure to dogs in public spaces in New

Zealand compared to mainland China or Hong Kong could contribute to this discrepancy. Furthermore, the frequent depiction of dogs in children's books in New Zealand, which we call the '*Hairy Maclary effect*' after the popular book series by Lynley Dodd, could influence children's word preferences.

Our hypothesis was that children from less deprived areas, with more educated parents, female, first-born, monolingualism, and no maternal concerns about speech and hearing would score higher on the Mandarin and/or Cantonese CDI inventories. Partially aligned with this hypothesis, our result was consistent with previous research suggesting that bilingual children had lower vocabularies in one of their languages than did monolingual speakers (see also Hoff et al., 2012; Reese et al., 2015). When comparing the language performance of children in our Mandarin ($N = 54$) and Cantonese ($N = 13$) samples to other Chinese samples (i.e., see Hao et al., 2008; Tardif et al., 2009), we found that monolingual children in our study used a similar higher number of words from the two lists (Mandarin: *Mean* = 52.65, *SD* = 31.77; Cantonese: *Mean* = 59.23, *SD* = 30.40). However, this does not imply that bilingual children experience delayed language learning. In fact, bilingual children exhibit comparable language development across both languages compared to monolingual individuals (see Hoff et al., 2012).

It is important to note that the majority of children in both samples had English as an additional language (67% of Mandarin speakers and 71% of Cantonese speakers). For English-Mandarin speakers ($N = 119$), the total vocabularies in English (*Mean* = 19.56, *SD* = 18.42) and Mandarin (*Mean* = 18.21, *SD* = 17.94) were similar. This finding aligns with a recent study by Yang et al. (2022), which suggested no difference in language dominance and the amount of language input on bilingual children's narrative skills in each language between age 3 and 7. It is possible that the bilingual English speakers in our samples performed similarly across both of their languages because they were younger, and therefore had more exposure to all of the languages compared to Yang et al. (2022). However, we did not assess language dominance as perceived by mothers, which is a limitation of our design. Future investigations should explore the relative use and proficiency of each language within home and school environments to obtain a more comprehensive understanding of bilingual children's language acquisition.

Contrary to our hypothesis, Cantonese-speaking children whose mothers had lower education levels demonstrated larger vocabularies in Cantonese. This finding may be attributed to the tendency of mothers with lower education to overestimate their children's language skills using CDIs (Jackson-Maldonado et al., 2013). Alternatively, less educated mothers might predominantly use Cantonese as the primary language with their children due to their limited exposure to alternative languages, thus positively impacting their Cantonese vocabulary development. An additional possible explanation for this phenomenon is the increased influx of Mandarin speakers to New Zealand. Mothers with higher education likely have more language skills across Mandarin/Cantonese and English (and/or other languages) than mothers with less education, leading to higher language fluency and competence in their other languages (Ballard & Lee, 2016). This possibility is supported by the results from our English-Mandarin sample where children whose mothers reported lower education had smaller English vocabularies and word combination scores compared to children whose mothers reported higher education. Additionally, mothers with higher education may be more likely to have employment where they primarily utilize English and/or both Mandarin/Cantonese and English (Jia & Paradis, 2015). This circumstance could result in more bilingual/multilingual interactions with their children, potentially leading to a more balanced split in language skills across the child's languages. These factors could influence bilingual mothers with higher levels of education to opt to speak Mandarin/English rather than Cantonese to their children, perceiving it as a more practical and advantageous language (Ballard & Lee, 2016). Using a non-parametric test, a statistically significant disparity in the median highest maternal education level was detected when comparing the Mandarin and Cantonese samples (excluding 20 participants who spoke both languages). Specifically, mothers of Cantonese-speaking children exhibited lower educational attainment on average in comparison to mothers of Mandarin-speaking children. To better understand the complex relationship between maternal education and child language outcomes in bilingual/multilingual samples, future studies should consider asking mothers what language their education is based on at the point of data collection (see Sorenson Duncan & Paradis, 2020).

### *Grammatical development*

For word combination scores, only 15% of Mandarin speakers and 7% of Cantonese speakers had not yet begun combining words in any language at age two. These results align with word combination scores for New Zealand English speakers (10% not yet combining; Reese et al., 2018) in the full *Growing Up in New Zealand* cohort. In contrast, lower rates of word combinations were observed among NZ Samoan (17% not yet combining; Reese et al., 2015) and Tongan speakers (16% not yet combining; Reese et al., 2015) in the same cohort.

In the Mandarin sample, including English-Mandarin speakers, we observed that children from the most deprived households showed lower rates of word combinations compared to children from moderately deprived households. This finding supported our hypothesis and is consistent with the word combination patterns seen in English-speaking children from the entire group (Reese et al., 2018). It also aligns with previous research indicating that children in less deprived households tend to have language advantages (Horwitz et al., 2003; Jackson-Maldonado et al., 2013; Reese et al., 2015). However, children from the most deprived households did not also have smaller Mandarin vocabularies as predicted. It is possible that these children may receive similar input in Mandarin as children in families from moderately deprived areas. However, their exposure to languages outside of Mandarin, particularly in early childhood education and other social settings, may be more limited.

Contrary to hypotheses, we found that later-born children exhibited more advanced grammar skills compared to first-born children in the Mandarin sample, including the English-Mandarin sample. This finding contradicts previous research indicating the opposite pattern (Hoff-Ginsberg, 1998; Horwitz et al., 2003; Reilly et al., 2009). One possible explanation is that later-born children benefited from increased exposure to grammatical cues through interactions with parents and with older siblings (Oshima-Takane et al., 1996). Given the number of English-Mandarin speakers, it is plausible that later-born children may exhibit a higher frequency of word combinations in English due to their exposure to older siblings who speak English, which could contribute to their advanced grammar skills compared to first-born children.

### *Maternal concerns*

We predicted that children whose mothers had concerns with their children's speech or hearing would be more likely to have lower language levels. Aligned with our hypothesis, maternal concerns were a unique predictor of Mandarin- and Cantonese-speaking children's language development for vocabulary skills; those whose mothers displayed concerns about children's speech and hearing had less vocabulary skills than those who showed no concerns. Furthermore, we found Mandarin-speakers (including English-Mandarin speakers) whose mothers expressed no concerns were more likely to combine words than children whose mothers expressed concerns. Our findings align with existing research that highlights parental concern as a significant predictor for children's language delay in New Zealand and in other countries (Harris et al., 2005; Horwitz et al., 2003; Klee et al., 2000; Reese et al., 2018), despite some previous studies failing to establish its predictive value (e.g., Dale, Price, Bishop, & Plomin, 2003). It is important to consider the potential influence of shared reporting bias, as mothers who had concerns could have underestimated their children's linguistic abilities. Additionally, it is possible that mothers with concerns intentionally used less sophisticated vocabulary and grammar when interacting with their children, thereby limiting their expressive language skills. However, this effect was not observed in our Cantonese sample, possibly due to limited statistical power. Further investigation is required to explore the quantity and quality of maternal input in cases where concerns are present, as well as their associations with child language development.

## Strength, limitations, and future directions

The present study encompasses several strengths and limitations. Notably, the inventories employed demonstrated high reliability in both languages and exhibited predictive relationships with sociodemographic factors as anticipated. However, certain limitations need to be addressed. First, since we used adapted Mandarin and Cantonese CDIs to evaluate bilingual language acquisition, direct comparisons between the results obtained from the new adapted inventories and other versions for assessing children's overall vocabulary across both languages are not feasible. The two wordlists are not aligned since they were selected to include common words used in Mandarin and Cantonese speaking families (see Peña, 2007). As a result, we were unable to calculate composite scores based on the number of concepts across all languages of the children (see Pearson et al., 1993). Moreover, the study did not assess other language varieties, such as Shanghainese or Fukien, which may limit the ability to accurately determine vocabulary scores for bilingual and multilingual children who can speak Mandarin or Cantonese in addition to these other languages.

Second, we recognize that our study could have benefited from a more comprehensive collection of background information about the families, especially those where fathers or grandparents were the main caregivers. Caution should be exercised when generalizing our findings given our focus only on mothers as respondents. Since many young Chinese children are raised primarily by their grandparents, parental report measures may not reflect language acquisition accurately (Reese et al., 2015; Tardif et al., 2009). According to Morton et al. (2014), 27% of Asian children in the *Growing Up* cohort lived with extended families at age two. This suggests that the family dynamics in immigrant Chinese families can be vastly different from other families in the cohort, with grandparents often interacting with their young

grandchildren in their native dialect rather than English. In future research, reports from all caregivers should be included for better estimates of Mandarin and Cantonese proficiency in children, and conducting ethnographic interviews will provide valuable insights into the family dynamics. Language samples of children interacting with grandparents and siblings would also round out the picture to provide a more holistic understanding of bilingual language acquisition in young Chinese-speaking children.

The findings of our study hold significant implications for clinicians in assessing the language development of Chinese-speaking children in New Zealand. Clinicians often have robust resources for assessment/intervention in English, but face barriers when working with children who speak a language other than English and/or in addition to English, particularly non-European languages. Given that more Chinese-speaking children may be turning up in practice, it is of great benefit to the profession to have assessments in Mandarin/Cantonese that have been validated for the New Zealand context. We would like to temper that by reminding our readers of the importance of collecting detailed information about language exposure/environment and taking caregiver concerns into consideration. Specifically, clinicians should take into account the language environment of the child, whether monolingual or bilingual, and to assess all of the child's languages whenever possible with multiple respondents for a more complete picture. Ideally, bilingual and multilingual children should also be assessed with separate questions about word combinations in each of their languages. These word combination questions are vital for identifying language delays, especially in children under 24 months old (Klee et al., 2000). Finally, given our findings that maternal concerns regarding children's speech or hearing were robust predictors of both total vocabulary and word combinations among Mandarin speakers, and of vocabulary among Cantonese speakers, caregivers' concerns need to be taken seriously. It is recommended that general practitioners, teachers, and clinicians in New Zealand inquire about caregivers' concerns regarding speech or hearing as a preliminary screening tool prior to conducting further assessments or making referrals for Mandarin- and Cantonese-speaking children.

## Conclusions

In summary, this study validated versions of Mandarin and Cantonese CDIs specifically tailored for the New Zealand context, and taking into account their proficiency in New Zealand English. The study also examined significant predictors of language acquisition among 24-month-old Mandarin- and Cantonese-speaking children residing in New Zealand. Notably, the adapted versions of the Mandarin and Cantonese CDI short forms demonstrated high levels of reliability and validity. Furthermore, the findings underscored the role of demographic factors on the language acquisition of Chinese children. It is expected that these findings provide a practical resource with insights on the development of bilingual and multilingual children, as well as for clinicians and practitioners engaged in supporting Mandarin- and Cantonese-speaking children both in New Zealand and on a global scale.

## Notes

1. For this study, English-Mandarin speakers refers to bilingual or multilingual Mandarin speakers who used English as an additional language.
2. For this study, English-Cantonese speakers refers to bilingual or multilingual Cantonese speakers who used English as an additional language.

## Acknowledgements

We express our gratitude to the whānau participating in the *Growing Up in New Zealand* study. We would also like to thank Philip Dale and the MacArthur-Bates Advisory Board for granting permission to adapt the Communication Development Inventory. The study was funded by various organizations, including the New Zealand Ministry of Social Development, Ministry of Health, University of Auckland, Social Policy and Evaluation Research Unit (Superu, formerly Families Commission), Ministry of Business, Innovation and Employment, Ministry of Education, Ministry of Justice, Sport New Zealand (previously Sport and Recreation New Zealand), New Zealand Police, Te Puni Kokiri, Housing New Zealand, Ministry of Pacific Island Affairs, Ministry of Women's Affairs, Department of Corrections, and the Mental Health Commission.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work was supported by Ministry of Business, Innovation and Employment; Ministry of Education- New Zealand; Mental Health Commission - New Zealand; Ministry of Social Development - New Zealand; Ministry of Health - New Zealand; Department of Corrections - New Zealand; Ministry of Justice - New Zealand; Sport New Zealand; New Zealand Police; Te Puni Kokiri; Housing New Zealand; Ministry of Pacific Island Affairs - New Zealand; Ministry of Women's Affairs - New Zealand; Social Policy Evaluation and Research Unit; University of Auckland.

## References

Aldawood, Z., Hand, L., & Ballard, E. (2023). Language learning environments for Arabic-speaking children in New Zealand: Family demographics and children's Arabic

language exposure. *Speech, Language and Hearing*, *26*, 1–12. doi:10.1080/2050571X.2023.2212537

Ballard, E., & Lee, T. Y. (2016). Language maintenance and attitudes of Mandarin speaking families in Auckland. In R. Muhr, K. E. Fonyuy, Z. Ibrahim, & C. Miller (Eds.), *Pluricentric languages and non-dominant varieties worldwide: Pluricentric languages across continents – features and usage* (Vol. 1, pp. 223–235). Vienna: Peter Lang Verlag.

Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia, and real-time processing. *Language and Cognitive Processes*, *12*(5–6), 507–584. doi:10.1080/016909697386628

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. O., & Basbøll, H. (2008). Early vocabulary development in Danish and other languages: A CDI-based comparison. *Journal of Child Language*, *35*(3), 619–650. doi:10.1017/S0305000908008714

Dale, P. S., Price, T. S., Bishop, D. V., & Plomin, R. (2003). Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech, Language, and Hearing Research: JSLHR*, *46*(3), 544–560. doi:10.1044/1092-4388(2003/044)

De Houwer, A., Bornstein, M. H., & Putnick, D. L. (2014). A bilingual-monolingual comparison of young children's vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics*, *35*(6), 1189–1211. doi:10.1017/S0142716412000744

Dionne, G., Dale, P. S., Boivin, M., & Plomin, R. (2003). Genetic evidence for bidirectional effects of early lexical and grammatical development. *Child Development*, *74*, 394–412. doi:10.1111/1467-8624.7402005

Fenson, L., Bates, E., Dale, P. S., Marchman, V. A., Reznick, J. S., & Thal, D. J. (2007). *MacArthur–Bates communicative development inventories*. Baltimore, MD: Brookes.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., … Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*(5), 1–185. doi:10.2307/1166093

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*(1), 95–115. doi:10.1017/S0142716400001053

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*(1), 98–116. doi:10.1037/0012-1649.42.1.98

Grech, H., & McLeod, S. (2012). Multilingual speech and language development and disorders. In D. E. Battle (Ed.), *Communication disorders in multicultural and international populations* (4th ed., pp. 120–147). St Louis, MI: Elsevier.

Greenwood, C. R., Carta, J. J., Walker, D., Watson-Thompson, J., Gilkerson, J., Larson, A. L., & Schnitz, A. (2017). Conceptualizing a public health prevention intervention for bridging the 30 million word gap. *Clinical Child and Family Psychology Review*, *20*(1), 3–24. doi:10.1007/s10567-017-0223-8

Hao, M., Shu, H., Xing, A., & Li, P. (2008). Early vocabulary inventory for Mandarin Chinese. *Behavior Research Methods*, *40*(3), 728–733. doi:10.3758/BRM.40.3.728

Hao, Y., Bedore, M., Sheng, L., & Peña, E. D. (2019). Narrative skills in two languages of Mandarin–English bilingual children. *International Journal of Speech-Language Pathology*, *21*(4), 325–335. doi:10.1080/17549507.2018.1444092

Harris, F., Law, J., & Roy, P. (2005). *The third implementation of the sure start language measure*. Nottingham, UK: DfES.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experiences of young American children*. Baltimore, MD: Paul H. Brookes.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*(1), 1–27. doi:10.1017/S0305000910000759

Hoff-Ginsberg, E. (1998). The relation of birth order and socioeconomic status to children's language experience and language development. *Applied Psycholinguistics*, *19*(4), 603–629. doi:10.1017/S0142716400010389

Horwitz, S. M., Irwin, J. R., Briggs-Gowan, M. J., Bosson Heenan, J. M., Mendoza, J., & Carter, A. S. (2003). Language delay in a community cohort of young children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*(8), 932–940. doi:10.1097/01.CHI.0000046889.27264.5E

Hua, Z., & Wei, L. (1999). Stylistic variations in the early lexical development of young Putonghua-speaking children. *Asia Pacific Journal of Speech, Language and Hearing*, *4*(1), 39–51. doi:10.1179/136132899805577114

Jackson-Maldonado, D., Marchman, V. A., & Fernald, L. C. (2013). Short-form versions of the Spanish MacArthur–Bates communicative development inventories. *Applied Psycholinguistics*, *34*(4), 837–868. doi:10.1017/S0142716412000045

Jia, R., & Paradis, J. (2015). The use of referring expressions in narratives by Mandarin heritage language children and the role of language environment factors in predicting individual differences. *Bilingualism: Language and Cognition*, *18*(4), 737–752. doi:10.1017/S1366728914000728

Klee, T., Pearce, K., & Carson, D. K. (2000). Improving the positive value of screening for developmental language disorder. *Journal of Speech, Language, and Hearing Research*, *43*(4), 821–833. doi:10.1044/jslhr.4304.821

Matthews, P. (1991). Word, word-form and lexeme. In *Morphology* (Cambridge Textbooks in Linguistics, pp. 24–41). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139166485.005.

Mieszkowska, K., Łuniewska, M., Kołak, J., Kacprzak, A., Wodniecka, Z., & Haman, E. (2017). Home language will not take care of itself: Vocabulary knowledge in trilingual children in the United Kingdom. *Frontiers in Psychology*, *8*, 1358. doi:10.3389/fpsyg.2017.01358

Morton, S. M. B., Atatoa-Carr, P. E., Grant, C. C., Berry, S. D., Bandera, D. K., Mohal, J., … Wall, C. R. (2014). *Growing up in New Zealand: A longitudinal study of New Zealand children and their families: Report 3: Now we are two: Describing our first 1000 days*. Auckland: University of Auckland, Growing up in New Zealand.

Morton, S. M. B., Atatoa-Carr, P. E., Grant, C. C., Lee, A. C., Bandara, D. K., Mohal, J., … Wall, C. R. (2012). *Growing Up in New Zealand: A longitudinal study of New Zealand children and their families. Report 2: Now we are born*. Auckland: Growing Up in New Zealand.

Morton, S. M. B., Atatoa-Carr, P. E., Grant, C. C., Robinson, E. M., Bandara, D. K., Bird, A., … Wall, C. (2013). Cohort profile: Growing up in New Zealand. *International Journal of Epidemiology*, *42*, 65–75. doi:10.1093/ije/dyr206

Oshima-Takane, Y., Goodz, E., & Derevensky, J. (1996). Birth order effects on early language development: Do second-born children learn from overheard

speech? *Child Development*, *67*(2), 621–634. doi:10.2307/1131836

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning*, *43*(1), 93–120. doi:10.1111/j.1467-1770.1993.tb00174.x

Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child Development*, *78*(4), 1255–1264. doi:10.1111/j.1467-8624.2007.01064.x

Peña, E. D., Bedore, L. M., & Kester, E. S. (2016). Assessment of language impairment in bilingual children using semantic tasks: Two languages classify better than one. *International Journal of Language & Communication Disorders*, *51*, 192–202. doi:10.1111/1460-6984.12199

Poulin-Dubois, D., Graham, S., & Sippola, L. (1995). Early lexical development: The contribution of parental labeling and infants categorization abilities. *Journal of Child Language*, *22*(2), 325–343. doi:10.1017/S0305000900009818

Reese, E., Ballard, E., Taumoepeau, M., Taumoefolau, M., Morton, S. B., Grant, C., … Perese, L. (2015). Estimating language skills in Samoan- and Tongan-speaking children growing up in New Zealand. *First Language*, *35*(4–5), 407–427. doi:10.1177/0142723715596099

Reese, E., Keegan, P., McNaughton, S., Kingi, T. K., Carr, P. A., Schmidt, J., … Morton, S. (2018). Te Reo Māori: Indigenous language acquisition in the context of New Zealand English. *Journal of Child Language*, *45*(2), 340–367. doi:10.1017/S0305000917000241

Reese, E., & Read, S. (2000). Predictive validity of the New Zealand MacArthur communicative development inventory: Words and sentences. *Journal of Child Language*, *27*(2), 255–266. doi:10.1017/S0305000900004098

Reilly, S., Bavin, E. O., Bretherton, L., Conway, L., Eadie, P., Cini, E., & Wake, M. (2009). The Early Language in Victoria Study (ELVS): A prospective, longitudinal study of communication skills and expressive vocabulary development at 8, 12 and 24 months. *International Journal of Speech-Language Pathology*, *11*, 344–357. doi:10.1080/17549500903147560

Reilly, S., Wake, M., Bavin, E. L., Prior, M., Williams, J., Bretherton, L., … Ukoumunne, O. C. (2007). Predicting language at 2 years of age: a prospective community study. *Pediatrics*, *120*(6), e1441–e1449. doi:10.1542/peds.2007-0045

Salmond, C., Crampton, P., & Atkinson, J. (2007). *New Zealand 2006 index of deprivation*. Wellington, NZ: Department of Public Health, University of Otago.

Simonsen, H. G., Kristoffersen, K. E., Bleses, D., Wehberg, S., & Jørgensen, R. N. (2014). The Norwegian communicative development inventories: Reliability, main developmental trends and gender differences. *First Language*, *34*(1), 3–23. doi:10.1177/0142723713510997

Snow, D. (2004). *Cantonese as written language: The growth of a written Chinese vernacular* (Vol. 1). Hong Kong: Hong Kong University Press.

Sorenson Duncan, T., & Paradis, J. (2020). How does maternal education influence the linguistic environment supporting bilingual language development in child second language learners of English? *International Journal of Bilingualism*, *24*(1), 46–61. doi:10.1177/1367006918768366

Statistics New Zealand. (2018). *Chinese ethnic group*. Retrieved from https://www.stats.govt.nz/tools/2018-census-ethnic-group-summaries/chinese

Suggate, S., Schaughency, E., McAnally, H., & Reese, E. (2018). From infancy to adolescence: The longitudinal links between vocabulary, early literacy skills, oral narrative, and reading comprehension. *Cognitive Development*, *47*, 82–95. doi:10.1016/j.cogdev.2018.04.005

Sun, H., Low, J. M., & Chua, C. (2022). Maternal language proficiency and Bilingual children's heritage language learning. *International Journal of Bilingual Education and Bilingualism*, *26*, 861–875. doi:10.1080/13670050.2022.2130153

Sun, H., Yin, B., Amsah, F., & O'Brien, B. A. (2018). Differential effects of internal and external factors in early bilingual vocabulary learning: the case of Singapore. *Applied Psycholinguistics*, *39*(2), 383–411. doi:10.1017/S014271641700039X

Tardif, T., Fletcher, P., Liang, W., & Kaciroti, N. (2009). Early vocabulary development in Mandarin (Putonghua) and Cantonese. *Journal of Child Language*, *36*(5), 1115–1144. doi:10.1017/S0305000908009185

Tardif, T., Fletcher, P., Liang, W., Zhang, Z., Kaciroti, N., & Marchman, V. A. (2008a). Baby's first 10 words. *Developmental Psychology*, *44*(4), 929–938. doi:10.1037/0012-1649.44.4.929

Tardif, T., Fletcher, P., Zhang, Z. X., Liang, W. L., & Zuo, Q. H. (2008b). *The Chinese communicative development inventory (Putonghua and Cantonese versions): Manual, forms, and norms*. Beijing: Peking University Medical Press.

Yang, J., Kim, J.-H., Tuomainen, O., & Xu Rattanasone, N. (2022). Bilingual Mandarin-English preschoolers' spoken narrative skills and contributing factors: A remote online story-retell study. *Frontiers in Psychology*, *13*, 797602. doi:10.3389/fpsyg.2022.797602

Zhang, Y., Jin, X., Shen, X., Zhang, J., & Hoff, E. (2008). Correlates of early language development in Chinese children. *International Journal of Behavioral Development*, *32*(2), 145–151. doi:10.1177/0165025407087213

# Appendices

## Appendix 1. Short form of the NZ Mandarin CDI:II (NZM CDI-sf).

1 喵 (猫叫) Meow
2 哎哟 Oh
3 宝宝 Baby
4 阿姨 Aunt
5 弟弟 Younger brother
6 老师 Teacher
7 要 Want
8 你好 Hello
9 请 Please
10 (真) 棒! (So) Great!
11 坐 Sit
12 有 Have
13 玩 Play
14 给 Give
15 送 Send
16 掉 Drop
17 拍 Pat/Slap
18 等 Wait
19 揉 Rub
20 翻 Flip/Turn over
21 摘(楸) Pick
22 记(得) Remember
23 用 Use
24 游泳 Swim
25 套 Cover
26 躲 Hide
27 问 Ask
28 假装 Pretend
29 希望 Hope
30 水 Water
31 粥 Porridge
32 饺子 Dumplings
33 肉 Meat
34 豆芽 Sprouts
35 水果 Fruit
36 梨 Pear
37 辣椒 Chili pepper
38 头发 Hair
39 舌头 Tongue
40 手 Hand
41 膝盖 Knee
42 狗 Dog
43 动物 Animal
44 鸟 Bird
45 蚊子 Mosquito
46 恐龙 Dinosaur
47 好 Good
48 香 Fragrant
49 烫 Hot (to the touch)
50 可怜 Pitiful
51 生气 Angry
52 够了 Enough
53 小心 Be careful
54 碗 Bowl
55 电扇 Electric fan
56 东西 Things/Stuff
57 球 Ball
58 棋子 Chess piece
59 背心 Vest
60 上衣 Top (clothing)
61 鞋 Shoes
62 床 Bed
63 抽屉 Drawer
64 花 Flower
65 树枝 Branch
66 棍子 Stick
67 云 Cloud
68 车 Car
69 卡车 Truck
70 动物园 Zoo
71 幼儿园/托儿所 Kindergarten/Nursery
72 中国 China
73 低下 Lower down/Below
74 旁边 Next to
75 二 Two
76 全部 All
77 自己 Self
78 那个 That
79 我们的 Ours
80 人家 Others/They
81 条 Classifier for long or thin objects
82 块 Classifier for piece or chunk
83 位 Classifier for people
84 些 Some
85 在哪儿/哪儿呢? Where is it?
86 哪个 Which
87 多少 How many
88 嘛 Modal particle used to express obviousness
89 噢 Okay
90 今天 Today
91 已经 Already
92 以前 Before
93 能 Can
94 肯 Willing to
95 不许 Not allowed
96 跟 With/Follow
97 因为 Because
98 抱 Hug
99 看 Look/See/Watch
100 凳子/椅子 Stool/Chair

## Appendix 2. Short form of the NZ Cantonese CDI:II (NZC CDI-sf).

1 喵 (貓叫聲) Meow
2 呷 (鴨叫聲) Quack
3 媽媽/媽咪Mom/Mommy
4 爸爸/爹地 Dad/Daddy
5 嬸嬸 Aunt
6 妹妹 Younger sister
7 哥哥 Older brother
8 司機 Driver
9 水 Water
10 粥 Porridge
11 腸粉 Rice noodle roll
12 蝦 Shrimp
13 豆腐 Tofu
14 蘋果 Apple
15 雪糕 Ice cream
16 吐 (luer) Spit
17 打 Hit
18 俾 Give
19 攞 Take
20 抱 Hold
21 行(路) Walk
22 學 Learn
23 熄(燈) Turn off (light)
24 用 Use
25 喊 Shout
26 問 Ask
27 唱 Sing
28 睇 Watch
29 聽 Listen
30 梳(頭) Comb (hair)
31 摸 Touch
32 拋 Throw
33 著(衫) Wear (clothes)
34 知道 Know
35 試 Try
36 送 Send
37 揭(開) Open
38 多謝/謝謝 Thank you
39 唔該 Please
40 刷牙 Brush teeth
41 洗手 Wash hands
42 食飯 Eat meal
43 打電話 Make phone calls
44 波 Ball
45 積木/Lego Building blocks
46 故仔/故事 Story
47 冇 Don't have
48 得意/可愛 Cute
49 肉酸/核突 Sour
50 凍 Cold
51 顎渴/口渴 Thirsty
52 滑 Slip
53 快 Fast
54 黑(色) Black
55 狗 Dog
56 蛇 Snake
57 熊仔/啤啤 Teddy bear
58 蝴蝶 Butterfly
59 車 Car
60 飛機 Airplane
61 碗 Bowl
62 番揀 Choose
63 鎖匙 Key
64 垃圾 Garbage
65 樹 Tree
66 遮(雨傘) Umbrella
67 天(空) Sky
68 海 Sea
69 襪 Socks
70 帽 Hat
71 (衫) 鈕/扣 Button (on a shirt)
72 拉鍊 Zipper
73 頭髮 Hair
74 腩 Tongue
75 背脊 Back
76 床 Bed
77 櫃捅 Cabinet
78 洗衣機 Washing machine
79 屋企 Home
80 多 More
81 一樣 Same
82 嗰邊/嗰度 There
83 前面 Front
84 我嘅 My
85 佢 He/She
86 我地 We
87 (一) 個 Classifier for object
88 (一) 隻 Classifier for animals
89 (一) 本 Classifier for books
90 (一) 張 Classifier for flat objects
91 乜嘢 What
92 得唔得? Is it okay?
93 啦 Final particle to soften the tone
94 喎 Final particle to express agreement
95 想 Want
96 鍾意 Like
97 依家 Now
98 先 First
99 同(埋) With
100 咁樣 Like this

## Appendix 3. Parental concern question

Parents may have a range of concerns about their children's speech or hearing, such as those listed on this showcard. Do you feel that [child's name] has any of these issues?

(1) no concerns
(2) reluctant to speak
(3) speech not clear to family
(4) speech not clear to others
(5) difficulty finding words
(6) difficulty putting words together
(7) doesn't understand you when you speak
(8) doesn't understand others when they speak
(9) voice sounds unusual
(10) stutters, stammers, or lisps
(11) other

## Appendix 4. Percentile rankings for vocabulary scores on the NZ Mandarin monolingual and Cantonese monolingual CDI-sf.

| Percentile ranking | Mandarin $N = 54$ | Cantonese[a] $N = 13$ |
|---|---|---|
| 10th | 12 | 11 |
| 20th | 16 | 26 |
| 30th | 26 | 34 |
| 40th | 41 | 54 |
| 50th | 58 | 67 |
| 60th | 63 | 76 |
| 70th | 78 | 84 |
| 80th | 88 | 88 |
| 90th | 96 | 95 |

Notes. [a]Given the small sample size for the Cantonese sample, this percentile ranking should be used with caution.

## Appendix 5. Percentile rankings for bilingual children's vocabulary scores on the NZ Mandarin and Cantonese CDI-sf.

| Percentile ranking | Mandarin $N = 92$ | Cantonese[a] $N = 25$ |
|---|---|---|
| 10th | 2 | 0 |
| 20th | 6 | 2 |
| 30th | 8 | 8 |
| 40th | 12 | 12 |
| 50th | 16 | 18 |
| 60th | 19 | 24 |
| 70th | 23 | 29 |
| 80th | 33 | 46 |
| 90th | 48 | 60 |

Notes. [a]Given the small sample size for the Cantonese sample, this percentile ranking should be used with caution.

## Appendix 6. Percentile rankings for multilingual children's vocabulary scores on the NZ Mandarin and Cantonese CDI-sf.

| Percentile ranking | Mandarin $N = 31$ | Cantonese[a] $N = 21$ |
|---|---|---|
| 10th | 0 | 0 |
| 20th | 0 | 0 |
| 30th | 1 | 2 |
| 40th | 2 | 8 |
| 50th | 4 | 12 |
| 60th | 6 | 19 |
| 70th | 12 | 27 |
| 80th | 19 | 40 |
| 90th | 46 | 52 |

Notes: [a]Given the small sample size for the Cantonese sample, this percentile ranking should be used with caution.