

# A Cypriot Greek Adaptation of the CDI: Early Production of *Translation Equivalents* in a Bi-(dia)lectal Context

*Loukia Taxitari*

Cyprus Acquisition Team, University of Cyprus,  
Cyprus University of Technology  
*taxitari@ucy.ac.cy*

*Maria Kambanaros*

Cyprus Acquisition Team, Cyprus University of Technology

*Kleanthes K. Grohmann*

Cyprus Acquisition Team, University of Cyprus

## Abstract

The MacArthur-Bates Communicative Development Inventory (CDI) has been widely used to study children's word production in both monolingual and bilingual contexts, in typical and atypical populations, and for the study of different aspects of language development, such as the use of mutual exclusivity. In this study, an adaptation of the CDI in Cypriot Greek is used to collect production data for post-vocabulary spurt children growing up in a bilectal community, where two different varieties of a language are used. Parents report that their children use translation equivalents for a single concept, and these increase as their total word production increases. Also girls seem to produce more translation equivalents than boys overall. This suggests that lexical development in bilectal communities might be more similar to bilingual rather than monolingual development, and that mutual exclusivity does not constrain word usage in such populations even during early word production.

## Keywords

language acquisition – vocabulary development – bilingual word learning – mutual exclusivity

## 1 Introduction

Lexical development has been studied extensively for the past 30 years or so (Clark 2009; Golinkoff, Hirsh-Pasek, Bloom, Smith, Woodward, Akhtar & Hollich 2000; Hall & Waxman 2004), using both experimental methods as well as questionnaires such as the MacArthur-Bates Communicative Development Inventory (CDI) (Fenson, Dale, Reznick, Bates, Thal & Pethick 1994). Besides the fact that the CDI has received some criticism for its validity, especially with comprehension data in younger infants (Houston-Price, Mather & Sakkalou 2007; Killing & Bishop 2008; Yoder, Warren & Biggar 1997), its concurrent validity as well as its long-term predictive power of later performance and development have been shown with both typical and atypical populations (Heilman, Ellis, Evans & Hollar 2005; Jackson-Maldonad 2012; Luyster, Qiu, Lopez & Lord 2007). It has also been adapted to a wide and growing range of languages, dialects, and varieties—63 to date (March 2014). A constantly updated list can be found on the CDI website ([http://www.sci.sdsu.edu/cdi/adaptations\\_ol.htm](http://www.sci.sdsu.edu/cdi/adaptations_ol.htm)). However, the majority of these adaptations are designed for monolingual children, in a single language/variety.

Although research with monolingual infants has made great advancements in understanding language development, bilingual lexical development has received some attention more recently. A number of CDI adaptations have been developed for children growing up bilingually and an increasing number of researchers have turned research attention towards bilingual word learning and lexical development (Bialystok 2001; Fennell, Byers-Heinlein & Werker 2009; Genesee & Nicoladis 2007; Goldstein 2011; Werker & Byers-Heinlein 2008), looking at both typical (De Houwer, Bornstein & De Coster 2006; Deuchar & Quay 2001) and atypical language development (Luyster et al. 2007; Petitto, Katerelos, Levy, Gauna, Treault & Ferraro 2001). CDI data from bilinguals has shown similarities and differences in development as compared to monolinguals. For example, Hoff & Elledge (2005) compared English-speaking monolingual children to English-speaking bilingual children of 16 to 30 months. They found that bilingual children were slightly behind monolinguals in their expressive vocabulary, but no differences were found between the two groups in measures of sentence complexity (i.e. grammatical development). Other researchers, however, suggest that bilingual children exhibit developmental patterns similar to monolinguals, and that vocabulary estimates of these children should be made with reference to both languages, which brings them closer to their monolingual peers' performance (Pearson, Fernández & Oller 1993; Poulin-Dubois, Bialystok, Blaye, Polonia & Yott 2013).

The CDI has been used in the past to investigate specific behaviors dur-

ing lexical development which allow comparisons across different linguistic situations. It has, for example, been used to shed light on the use of *mutual exclusivity* by young bilingual infants. Experimental research with monolinguals has shown that children between 2 and 3 years of age are reluctant to attach two labels to the same object, and, when presented with a new label in the presence of an unnamed object, they prefer to attach the label to the unnamed, rather than the named, object (Markman, Wasow & Hansen 2003). This bias to avoid two labels for the same object and to look for a new object to attach a new label to has been called the 'mutual exclusivity constraint' (Markman 1992; Markman & Wachtel 1988; Merriman & Bowman 1989), the 'N3C (name–nameless category) principle' (Golinkoff, Hirsh-Pasek, Bailey & Wenger 1992), or the 'principle of contrast' (Clark 1993). Markman et al. (2003) showed that even pre-vocabulary spurt infants exhibit mutual exclusivity, even in the presence of only a single familiar object and no novel objects to attach the new label to. Mutual exclusivity has been shown for young children and adults, and also for monolingual and bilingual children alike, provided that the labels referred to categories of different hierarchy (Au & Glusman 1990).

Several studies used observational or questionnaire data from various languages to investigate the *production* of words with the same meaning but different form, i.e. *translation equivalents* (henceforth, TES). These studies reported that children in their first and second years of life used words in the two languages to refer to the same concept/referent from the very early stages of word learning and with several language pairs such as English–French, English–Spanish, or English–German (Holowka, Brosseau-Lapre & Petitto 2002; Junker & Stockman 2002; Pearson, Fernández & Oller 1995; Poulin-Dubois et al. 2013). These studies document a low percentage of TES in children's word production before 18 months of age, and a steady increase to about 30% by the age of 2 years, as well as high variability in TES across children. De Houwer et al. (2006) used the CDI to investigate bilingual *comprehension* in pre-vocabulary spurt infants (13 months) growing up in Belgium learning French and Dutch. They asked parents and caregivers to fill in the two versions of the CDI and checked for TE pairs in the two languages; that is, words which have the same meaning but appear in different forms in the two languages. Although there was large variation across children in the number of TES they understood, all of these young children were reported to understand at least some *doublets*, that is, different words for the same concept in both French and Dutch. These studies suggest that bilingual infants, even at the very early stages of word learning, can overcome mutual exclusivity and learn two words for the same concept, albeit in different languages.

In the present study, the MacArthur Bates CDI was used to investigate a yet more complex linguistic situation, between the two discrete classifications of monolingualism and bilingualism. This 'grey area' of development, which is only now beginning to receive much needed attention and to become understood, has been given many names over the past 30 years, from *diglossia* (Ferguson 1959) to the most recent *bilectalism* (Rowe & Grohmann 2013): language development in children who are exposed to two (or more) varieties of the *same language*. Such situations are not rare in different countries, such as Austria, for which a specially adapted version of the CDI exists that differs from Standard German (Marschik, Einspieler, Vollmann & Einspieler 2005; Vollmann, Marschik & Einspieler 2000) and Cyprus (see Grohmann & Leivada 2012, and references cited). In the latter, the local variety, Cypriot Greek (CG), co-exists with another, the standard and high variety, Standard Modern Greek (SMG). Children who grow up on the island are exposed to CG at home, but they also hear SMG through more formal lines of communication such as television and radio. The question which arises is whether these children grow up more like monolinguals or like bilinguals in linguistic and cognitive terms (as also raised in Kambanaros, Grohmann and Michaelides 2013), or whether there is an intermediate situation with special characteristics in development. Current research suggests that Greek Cypriot children do not acquire morphosyntax the same way that monolingual children from Greece do (see Grohmann 2011; Grohmann & Leivada 2012; Kambanaros, Grohmann, Michaelides and Theodorou 2013; Rowe & Grohmann 2013, and references cited), and this could indeed be related to the close language distance between CG and SMG. However, so far we know nothing about the very early language development in young children, such as around the time of pre- and post-vocabulary spurt (before the age of 3 years). This said, there are no published comparable data available from SMG, i.e. monolingual children from Greece, either.

In the current study, the Cypriot Greek bi(dia)lectal adaptation of the MacArthur-Bates Communicative Development Inventory (henceforth, CG-CDI) was used in order to test whether bilectal post-vocabulary spurt children exhibit behavior consistent with mutual exclusivity or whether they are more flexible in accepting multiple names for objects, as bilingual children do, but coming from two varieties of the same language. For this purpose, we asked parents of children in their third year of life (mean age 31 months) to indicate which words their children produce in the vocabulary section of the CG-CDI, containing both CG and SMG words.

## 2 Methodology

### 2.1 *Participants*

Twenty-two children participated in this study, with a mean age of around 31 months (aged from 27 to 36 months), ten girls and twelve boys. One of the girls was later excluded from the study because she behaved as an outlier in the analysis, leaving a total of 21 children.

All children were recruited for this specific study of the CG-CDI, through advertisements in nurseries and children's clinics in the area of Limassol. Children from a single city were recruited in order to avoid additional dialectal variation. All of those children were exposed to some form of the CG *koiné* as explained in detail in section 2.2 below. Some parents were approached directly by the researcher (first-named author) and others volunteered by contacting the researcher themselves. In order to have a controlled sample, the Language and Background Development Questionnaire (Paradis, Emmerzael & Sorenson Duncan 2010; Paradis, Genesee & Crago 2011) was given to parents to complete, along with the CG-CDI. This provided information about the physical development of children, their language environment, as well as any history of language problems in the family. A more detailed description is provided in section 2.3.2 below.

All children who participated were also exposed to only CG and SMG from birth and on a daily basis. None of the children was systematically exposed to any other language. All children were full-term and had normal weight at birth, and had no history of hearing problems or ear infections; 45% of the children were first-borns, the others ranged from second to fourth children in their families.

Each CDI was completed by the mother, even though parents were given the choice of either the mother or the father completing it. All parents were high school graduates, and all mothers additionally had completed some professional training or university degree by the time of the study. All but two mothers were full-time professionals at the time of the study.

### 2.2 *CG and SMG*

Over the past decades there has been considerable discussion in the literature regarding an exact definition of the linguistic situation on the island. Recently, Rowe & Grohmann (2013) suggest that Cyprus is currently transitioning through a state of diglossia and Tsiplakou (2014) argues for a partial convergence (of the Cypriot *koiné*) to Standard Modern Greek through *innovative, structurally mixed forms together with systematic language alternation in the form of code-switching, code-mixing, and register shifting*. This Cypriot *koiné*

is the variety used in urban centres on the island, retaining many of the characteristics of CG, but also leaving behind many of the features of the geographical sub-varieties of CG and replacing them with more standard-like features. For the purposes of this paper, by CG we will refer to the CG *koiné*.

Differences between CG and SMG can be traced in many linguistic levels.

1. Concerning phonology, CG and SMG mainly differ in terms of certain consonants (consonant germination and no voiced stops in CG), which make the *koiné* sound distinctly different from SMG.
2. CG and SMG differ in several aspects of their inflectional morphology. However, within the *koiné*, CG and SMG often become mixed up with features from either being used with structures from the other variety.
3. In terms of the lexicon, CG and SMG share a large proportion of their vocabulary, with certain lexical tokens existing only in one or the other variety, and others having different meaning across varieties.
4. CG and SMG share most of the Modern Greek syntax, but there are also certain Cypriot-specific structures (such as enclisis in indicative declaratives, wh-question formation, or the syntactic expression of focus).

At all linguistic levels, there are similarities and differences between the two varieties, with some levels more closely related than others. Phonology and syntax seem to remain quite distinct in the two varieties, while morphological features tend to be more mixed in the *koiné*. There is still considerable debate in the literature whether CG and SMG form part of a continuum or not, and the question which arises is when exactly during language development these different features are acquired and when they become separated (or even merged).

Although CG and SMG differ at several levels, for the purposes of this paper, we focus on the lexicons of these varieties which are largely common.

### 2.3 *Instruments*

Two questionnaires were given to parents: the CG adaptation of the MacArthur Bates Communicative Development Inventory *Words and Sentences* (Fenson et al. 1994) and an adaptation for Greek Cypriot parents of the *Developmental and Language Background Questionnaire* (Paradis et al. 2010; Paradis et al. 2011).

#### 2.3.1 CG-CDI

The CDI is a long list of words and parents are asked to mark if their child produces the items on the list. The questionnaire includes a total of 685 words divided in 23 categories (Animal and Other Sounds, Animals, Vehicles, Toys,

Food and Drink, Clothes, Body Parts, Home Objects, Furniture and Rooms, Outside Things, Places to Visit, People, Routines, Actions, Descriptive Words, Words for Time, Pronouns, Question Words, Propositions and Words for Place, Quantitatives and Articles, Colours and Shapes, Modal and Auxiliary Verbs, and Connectives). Although words are presented in isolation, some context is provided to parents because words are divided in different semantic categories. 'Grammatical' categories also exist, such as modal and auxiliary verbs or question words.

The CG-CDI is an adaptation of the CDI in CG containing both SMG and CG forms. Unlike bilingual CDIs, this adaptation includes words from both varieties in one long list. Often words for a concept in the two varieties coincide, with slight phonological differences or in exactly the same form, and in a few cases a completely different word form exists for CG. This results in very few CG-only words in the CG-CDI, and a large number of shared words.

As far as the lexicons of the two varieties are concerned, differences between CG and SMG might be found both lexically and phonologically. So there are three ways a concept might behave across the two varieties:

1. a concept might be lexically the same (for example, the words for *hand* or *mouth*. In that case the word could further be phonologically different (*hand*—['çeri] in SMG and in ['jeri] CG) or the same (*mouth*—['stomɐ] in both varieties).
2. a single concept might be lexically different in CG and SMG (for example, the word for *head*, [çefɐ'li] in SMG and [c<sup>h</sup>:elle] in CG), or
3. a concept could exist in only one of the two varieties (for example, [t<sup>h</sup>:oros] in CG is equivalent to *bath towel*, which does not exist as a single word in SMG, but instead the word for *towel* in general is used, i.e. [pe'fsetɐ]).

In the CG-CDI, we listed as separate entries *only* items which differed lexically. For this, we included both concepts with different words in the two varieties and concepts which can be found in only one variety. Words which differed phonologically in the two varieties were entered in the CG-CDI as a single entry, for example ['çeri] and ['jeri] for *hand*.

The CG-CDI contains a total of 685 words (100 words found only in CG and 585 found in SMG as well) and a total of 639 concepts. Fewer concepts exist than words because a single concept can correspond to both a CG and an SMG word, as described above. There are 44 such TE pairs, with words from both varieties which correspond to a single meaning.

The CG adaptation of the CDI has not been normed yet, since this is a first pilot study providing data for Greek Cypriot children.

### 2.3.2 Developmental and Language Background Questionnaire

This questionnaire is based on the ALEQ and ALDeQ questionnaires, developed by Paradis et al. (2010, 2011) and was subsequently modified in COST Action IS0804 in order to collect information about the child's development, the language environment she grows up in, and other demographic data. It has been translated into Greek for the purposes of this study, in order to control for the different factors which could affect children's lexical development. It consists of five sections:

1. general information about the child (name, birth date, sex, order of birth in family);
2. child's developmental history (complications at birth, birth weight, health problems, important milestones, exposure to languages);
3. languages used by and with the child (parents, caregivers, grandparents);
4. information about mother and father demographics (education, work);
5. any kind of health problem (e.g., frequent ear infections or other kinds of difficulties, e.g., language or learning difficulties/impairments in family).

All children participating in the study exhibited low risk of language problems on the basis of the questionnaire (no frequent ear infections, no low birth weight, no health problems, no birth problems, no parental concerns for the child's language development, and all children produced at least two-word combinations in their productive language). Participating children were aged between 27 and 36 months, of whom around 40% were first-borns. The remainder ranged from 2nd to 4th children in the family.

In Section 3 of the Questionnaire, parents were asked to rate the frequency of exposure to the two varieties of Modern Greek on a scale from 0 to 4 (0 = never, 1 = rarely, 2 = sometimes, 3 = often, 4 = always) and also the age of first exposure to the two varieties (in months). Apart from a general frequency of exposure rating, separate ratings were provided for the use of the two varieties from main caregivers and family members of the children.

All children were reported to have been exposed to both CG and SMG from birth. They were also reported to be exposed to the two varieties very often (mean rating: CG: 4 (always), SMG: 3.27 (often–always); Std. Dev.: CG: 0, SMG: .63). The difference between CG and SMG probably reflects the fact that parents are aware that their children hear CG at home and everyday talk, while they would hear SMG only through more formal communication, such as TV or radio shows.

#### 2.4 Procedure

Parents were asked to complete the CG-CDI and the Development and Language Background Questionnaire together and at their own time. They were given the choice for either the mother or the father to complete the questionnaire, but not both together. All mothers completed both questionnaires at home and returned both questionnaires to the researchers upon completion.

For the vocabulary checklist of the CG-CDI, parents were instructed to mark the field if their child produces a word, or leave it unmarked otherwise. They were also informed that they would find some words in CG in the word list and that sometimes they would find a word for an object in both CG and SMG. They were instructed to mark the version their child uses or mark both if the child uses both. They were also asked to write down a different pronunciation that they child might be using.

### 3 Analysis and Results

#### 3.1 Scope

The analysis in this study focuses only on the *vocabulary section* of the CG-CDI and the *productive* vocabulary of the children. There are 685 words in the CG-CDI and 639 concepts.

In this study we were especially interested in TE pairs, that is, words in the two varieties which express one core meaning (Pearson et al. 1995). It is assumed that CDI production data are a good estimation of the child's possible use of two different words for a corresponding meaning (and possibly the same object). Such pairs include, for example, the words for the concept *ball* (CG [mɛp<sup>h</sup>:v] and SMG [mbɛlɛ]) or for the concept *ladybird* (CG [pɛpɛ'runɛ] and SMG [pɛsxɛ'lifɛ]).

Out of the 685 words in the CG-CDI, only 88 words are part of TE pairs, that is, only 12.85% of the total word count. Unlike bilingual CDIs, in which the vast majority of words in the two language are members of TE pairs, in the CG-CDI TEs are a small percentage of the total number of words. This is the case because SMG and CG are typically considered varieties of the same language and they actually share a large portion of their vocabularies. Only some concepts are lexicalized differently in the two varieties, making this bilingual situation fall somewhere in between bilingualism and monolingualism. This creates a very interesting case study because for the most part children learning the two varieties are faced with a common vocabulary (albeit with small phonological differences), but in a few cases they are presented with two different words for

the same concept/object. To make things even more interesting, some of those are high-frequency words, so children should be expected to come across them quite often.

Although TES are the focus of this study and can be very informative concerning bilectal development, a number of other variables and factors were taken into account in the analysis in order to better appreciate the status of TES themselves: children's age, birth order, gender, other demographic factors (mother's and father's education), total vocabulary produced, and total conceptual vocabulary were all tested to see whether and how they might affect production of TES.

### 3.2 *Scoring*

For every item in the CG-CDI the parent reported that their child produced, a single point was given. For fields left unmarked, no points were given. Words in the two varieties which correspond to the same concept were marked as TES; for example, [pəsxɛ'lifʃɛ] in SMG and [pɛpɛ'rʉnɛ] in CG for *ladybird*. There were 44 such pairs in the CG-CDI. For these pairs, each word received one point to yield a total vocabulary score for each child. In order to calculate a conceptual vocabulary score, all TES received one point, irrespective of whether the child produced only the CG word, only the SMG word, or both.

Following the terminology from De Houwer et al. (2006), CG and SMG words which make up a TE pair are called *members* of that pair. So, when a child produces only the CG or only the SMG member, she is said to produce a *singlet*; when, on the other hand, the child produces both members of the pairs, she is said to produce a *doublet*.

### 3.3 *Measures*

In order to test children's productive vocabulary in this bilectal CG-CDI, two measures were calculated, *total vocabulary score* (the total number of words the child can say, coming from both SMG and CG) and *total conceptual vocabulary* (by subtracting the number of doublets a child says from her total vocabulary score). Both measures were calculated as percentages of the corresponding totals. Also percentages of singlets and doublets were calculated using the following formula:  $\text{number of singlets (doublets)} * 100 / \text{total number of singlets (doublets)}$ . Also, normality tests were run for all measures and factors tested, and normality was confirmed, after the exclusion of one outlier (from the girls group; see Section 2.1).

TABLE 1 *Mean Total Vocabulary Score (TVS) and Conceptual Vocabulary Score (CVS) and corresponding percentages for the entire sample*

	Mean	Std. Deviation	Minimum	Maximum
TVS	523.64	107.38	247.00	675.00
Percentage TVS	76.44	15.68	36.06	98.54
CVS	487.05	100.72	231.00	631.00
Percentage CVS	76.22	15.76	36.15	98.75

### 3.4 Results

#### 3.4.1 General Language Abilities

A first analysis of the data involved the total word production in the CG-CDI irrespective of language variety and whether words were TES or not. Children who participated in our study exhibited huge variation in word production, as shown in Table 1. They produced a mean of 537 words, which constitutes approximately 78% of the total word count in the CG-CDI, and also a mean of 499 concepts, which again corresponds to 78% of the total CG-CDI concepts (Table 1). Children produced on average 79.57% (range: 39.83–99.66%) of the SMG words and 60.92% of the CG words. A paired-samples t-test confirmed that children in our sample produced a significantly higher percentage of SMG than CG words ( $t(21) = 5.64, p < .01$ ). Pearson R correlations showed no meaningful relationship between language input (frequency of exposure to SMG) as reported in the Development and Language Background Questionnaire and either total word production ( $r(21) = -.06, p = .81$ ), total concept production ( $r(21) = -.06, p = .79$ ), or percentage of SMG ( $r(21) = -.02, p = .92$ ) and CG words ( $r(21) = -.12, p = .46$ ) produced by children.

In order to test whether the large variation in word production was due to the age range of our sample (27–36 months), bivariate Pearson R correlations with age and TVS/CVS tested whether word production increased with age, as children progressed in their third year. No such correlation was found for either TVS ( $r(21) = .34, p = .13$ ) or TCS ( $r(21) = .34, p = .13$ ), suggesting that in the third year of life, post-vocabulary spurt language production can present huge differences across children, irrespective of exact age in months.

Additionally, we tested to see whether other factors controlled for in this study might have affected the word production as reported by parents. Univariate Analyses of Variance (ANOVAs) were run for both TVS and CVS with Birth Order (first or second+), Gender (female or male), Mother's Education, or Father's Education. All factors were shown not to have an effect on children's

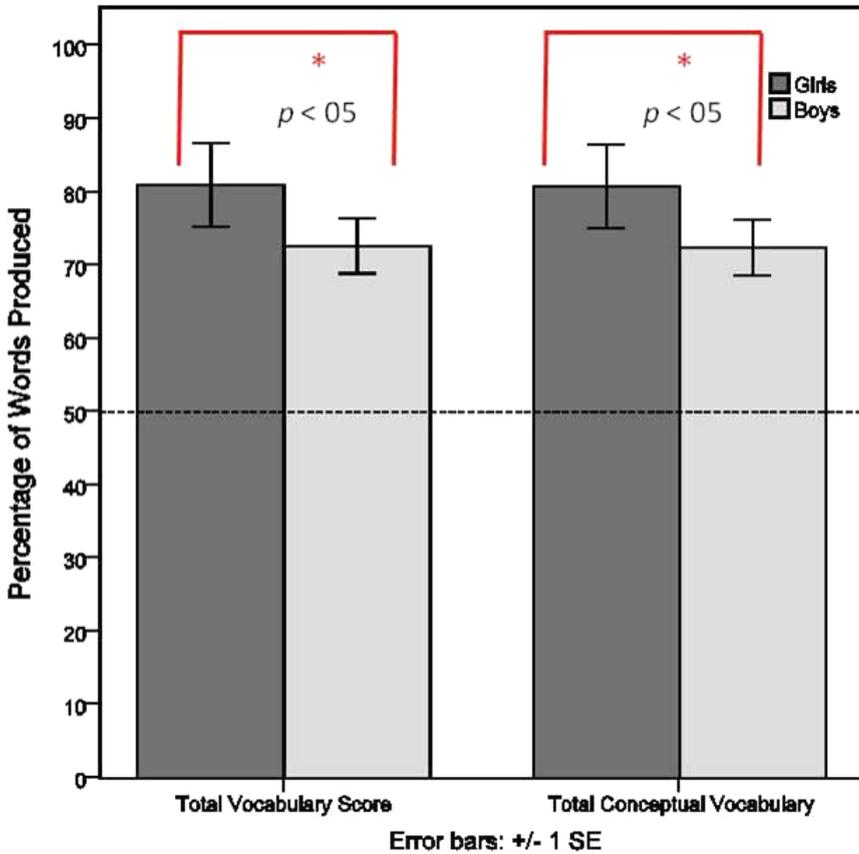


FIGURE 1 Percentages of the TVIs and CVIs produced for each gender separately. The difference between boys and girls is significant at the .05 level for both measures.

word production (Birth Order: TVIs:  $F(1, 19) = .44, p = .52$  / CVIs:  $F(1, 19) = .45, p = .51$ —Mother's Education: TVIs:  $F(3, 17) = 2.07, p = .14$  / CVIs:  $F(3, 17) = 1.98, p = .16$ —Father's Education: TVIs:  $F(3, 16) = 1.57, p = .24$  / CVIs:  $F(3, 16) = 1.58, p = .23$ ), except for Gender (TVIs:  $F(1, 19) = 6.81, p < .05$  / CVIs:  $F(1, 19) = 6.66, p < .05$ ). As shown in Figure 1, girls produced significantly more words as well as concepts than boys, suggesting an advantage for girls at these early stages of lexical development.

### 3.4.2 Translation Equivalents (TES)

As mentioned before, of special interest in this study were TES, words that have the same meaning but different phonological forms in the two varieties of Greek. These are expected to be quite informative in terms of the way bilingual children develop, and whether their language development is closer to

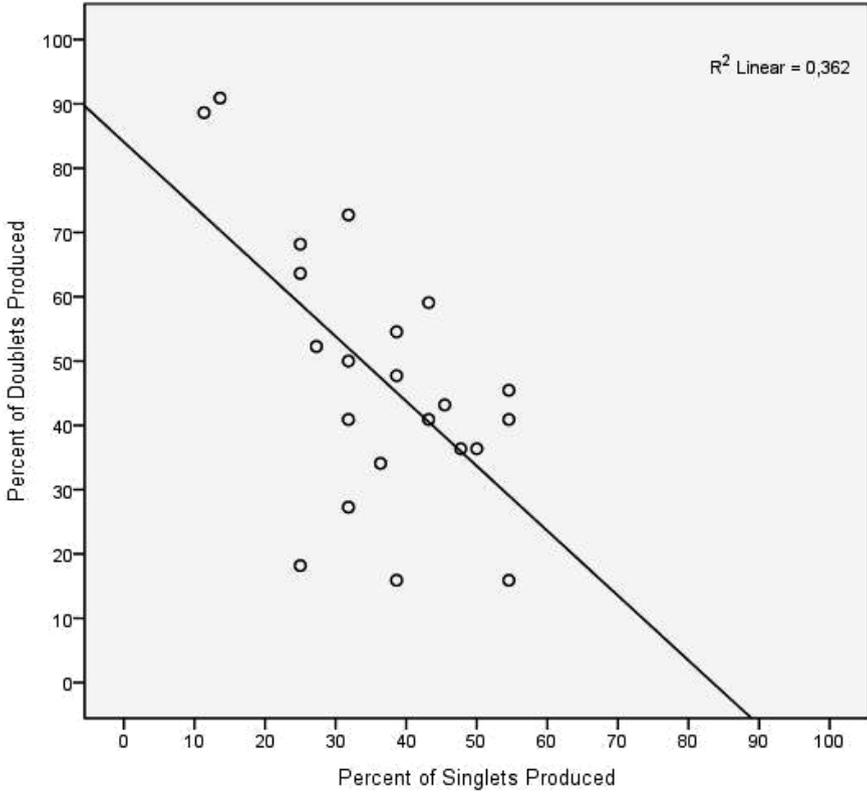


FIGURE 2 Scatter-plot showing the correlation between the percentage of singlets and the percentage of doublets produced. As the number of doublets increases, the number of singlets decreases ( $r(21) = -.60, p < .01$ ).

TABLE 2 Number of singlets and doublets produced from the 44 TE pairs in the CG-CDI and the corresponding percentages for the entire sample

	Mean	Std. deviation	Minimum	Maximum
<b>Singlets</b>	16.24	5.44	5.00	24.00
<b>(Percentage)</b>	(36.90)	(12.36)	(11.36)	(54.55)
<b>Doublets</b>	21.48	8.84	7.00	40.00
<b>(Percentage)</b>	(48.81)	(20.08)	(15.91)	(90.91)

their mono- or bilingual peers. As is the case with the two measures above, there was huge variation in the number of singlets and doublets these chil-

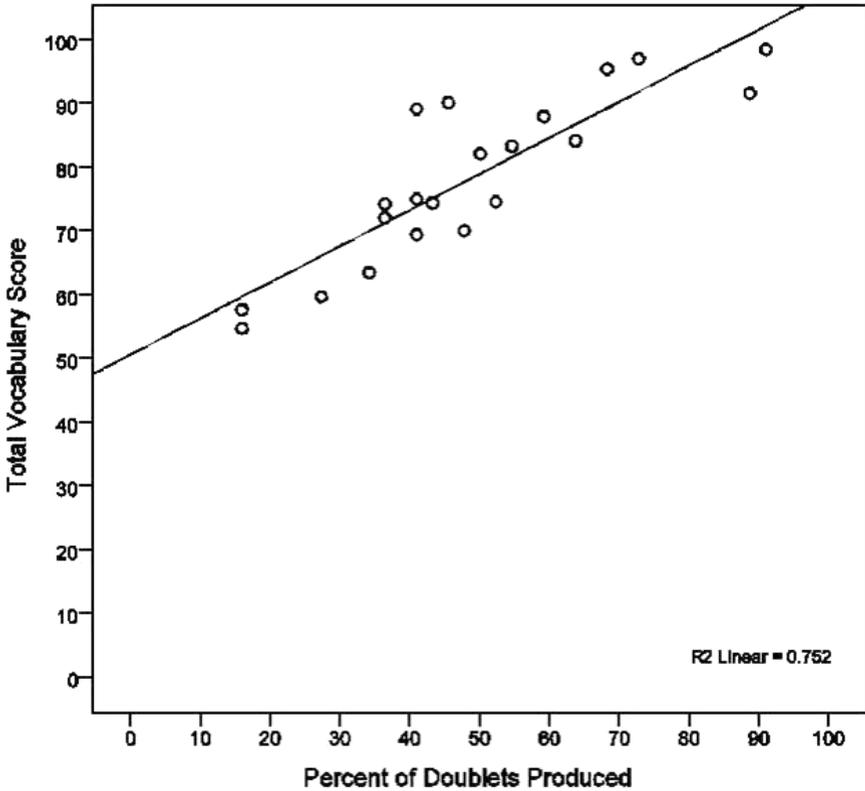


FIGURE 3 Scatter-plot showing the correlation between TVS and the percentage of doublets produced. As TVS increases, the number of doublets increases as well ( $r(21) = .87, p < .01$ ).

dren understand (Table 2), but all children in the sample produced at least a few doublets at the same time as they produced singlets. Children produced on average 36.9% singlets (range: 11.36–54.55%) and 21.48% doublets (range: 7–40%) (Table 2). The relationship in the production of singlets and doublets was inversely related; that is, as the number of doublets increased, the number of singlets decreased,  $r(21) = -.60, p < .01$ .

We also found that the number of doublets increased as both TVS and CVS increased (TVS:  $r(21) = .87, p < .01$  / CVS  $r(21) = .87, p < .01$ ). In other words, the more words and concepts a child knew, the more likely she was to produce more doublets (Figure 3). For the singlets, no such correlation exists (TVS:  $r(21) = -.35, p = .12$  / CVS  $r(21) = -.36, p = .12$ ). As the total word production increases, children thus tend to learn their concepts in both varieties. This is a behavior similar to that of bilingual children producing words in their two different

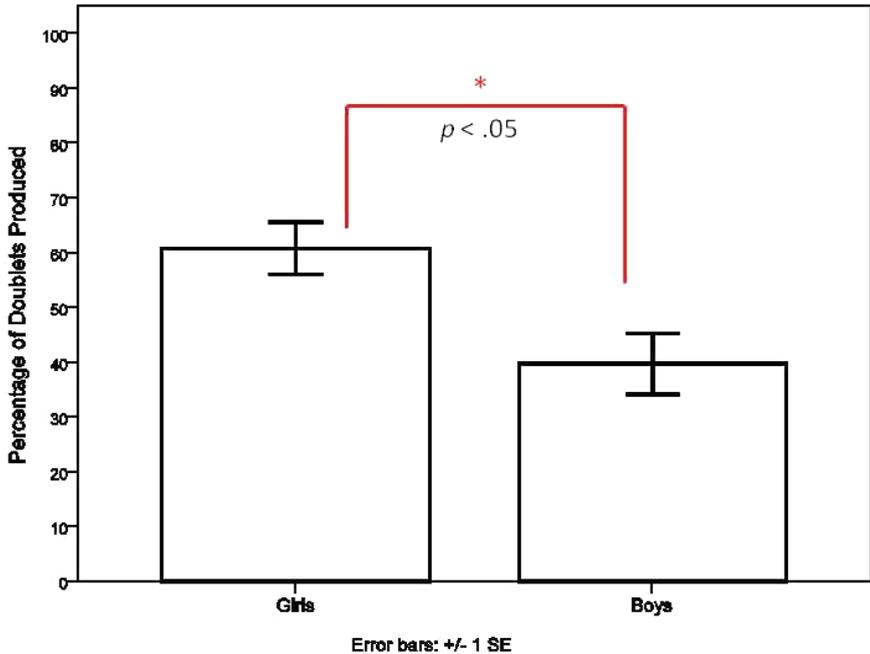


FIGURE 4 Mean percentage of doublets produced by each gender separately, showing a significant advantage for girls ( $F(1, 19) = 7.52, p < .05$ )

languages, and becoming better as their vocabulary grows (De Houwer et al. 2006).

As with TVS and CVS, we tested several other factors to see if and how they affected the production of doublets. ANOVAs were run for both TVS and CVS with Birth Order (first or second+), Gender (female or male), Mother's Education, or Father's Education. As with the previous measures, no factor was shown to have an effect on the number of doublets produced (Birth Order:  $F(1, 19) = .28, p = .61$ —Mother's Education:  $F(3, 17) = 1.68, p = .21$ —Father's Education:  $F(3, 16) = 1.10, p = .38$ ), except Gender ( $F(1, 19) = 7.52, p < .05$ ). As before, girls produced a higher number of doublets ( $M = 60.86, SD = 14.41$ ) than boys ( $M = 39.77, SD = 19.35$ ) (Figure 4).

### 3.5 Recapitulation of Main Findings

In this study, children in their third year of life growing up in a diglossic speech community (CG and SMG), all had several *doublets* in their productive vocabulary, that is, words in CG and SMG which have the same meaning. Singlets were also produced by all children, and the relationship between singlet and doublet production was inversely proportional, suggesting that

at the initial stages of word learning children might have a single word for a specific concept and as their vocabularies grow and they add on words from both varieties, some of those words could be referring to the same concepts.

Doublet production was also correlated with children's total word and concept production (TVS and CVS). In other words, the more words or concepts children learn, the more flexible they are in producing them in two different varieties (and also languages, as has been reported in research with bilingual children; cf. De Houwer et al., 2006). Finally, total word/concept production and doublet production exhibited *gender effects*, with girls producing significantly more words/concepts as well as doublets than boys.

#### 4 Discussion

This is the first study that aims at investigating language development in a *bi(dia)lectal* linguistic community where two varieties of the same language co-exist (in the form of a standard/high and an everyday/low variety). Language data from such situations are not available at the moment, and they constitute a 'grey area' between monolingualism and bilingualism, which is only now beginning to be studied and understood. In our study we have found that children in their third year use both CG and SMG words in their speech, and all of them also use singlets as well as doublets for the same concept, besides large variation across children.

Variation in CDI studies is not unexpected for two reasons. First, although several language milestones have been established and CDI norms have been standardized in many languages, the *modal child* has turned out to be a *mythical creature* in the study of language development as Fenson et al. (1994) suggest. Children follow their own course in development showing large variation, especially in the production of language in the first years of life (Huttenlocher 1998; Pearson et al. 1993), within the same and across different countries (Papaeliou & Rescorla 2011). Second, different reporters (mothers, fathers, caregivers, etc.) tend to report differently on their child's language (De Houwer et al. 2006; De Houwer, Bornstein & Leach 2005; O'Toole 2013), especially when they come from different cultures (Hamilton, Plunkett & Graham 2000). This, however, is mediated in our study due to the fact that we have used production and not comprehension data, which is arguably more prone to subjective judgment (Houston-Price et al. 2007; Killing & Bishop 2008). We expect that parental reports at age 2–3 are quite reliable, as these parents can observe their children's language production before filling in our questionnaire. This has been

shown in other studies with children of this age as well (Heilman et al. 2005; O'Toole & Fletcher 2010).

Of special interest in our study were TES, which all children produced, consistent with previous findings from studies with bilingual children of the same age (De Houwer et al. 2006; Poulin-Dubois et al. 2013). Mutual exclusivity predicts that 2- to 3-year-olds should avoid the use of two labels for the same object. However, research with bilingual children has shown this not to be true, from the early stages of word learning. In this study we have the first evidence of mutual exclusivity not being respected within the same language. Children are willing to use two different words, from two different varieties of the same language, to refer to a single concept/referent and this becomes more robust as their vocabularies grow, similar to bilingual lexical development (Poulin-Dubois et al. 2013; Pearson et al. 1995). These words sometimes have distinct phonological markers from one or the other variety (e.g. the presence of the post-alveolar fricative which exists only in Cypriot Greek, as in the word for *slide*, SMG [ʃsu'liθrɛ]—CG [kɛtrɛ'ʃʎili]), but often they are unmarked (*swing*: SMG /kunjə/—CG /susa/), and children can only decide which variety they belong to on the basis of context. The important finding in this case is that, just like bilingual children, toddlers in their third year of life growing in the bilectal community of Cyprus do not obey to mutual exclusivity, suggesting that these children present behavior which departs from that of their monolingual peers.

There is also an interesting negative correlation between the number of singlets and doublets produced by children, which suggests the existence of a period when children have a word for a single concept in one language for most words followed by a period when they learn words from both languages for a single concept, just like in bilingual language development (Poulin-Dubois 2013). This relates to a discussion about the underlying system or systems governing the acquisition of TES: people have suggested that at the beginning of vocabulary development, bilingual children are willing to learn only a single word for a concept from whichever language they happen to pick it up from, and only later are they willing to accept two words for a concept from two different languages when they have appreciated the fact that they belong to two different linguistic systems (Volterra & Teachner 1978; Clark 1987). This was used to suggest that during early vocabulary learning children build a single lexicon with words from both languages. However, Pearson et al. (1995) showed that children do not actually go through an initial stage in lexical development when they have no doublets at all in their vocabularies, casting doubt on the hypothesis of a single lexicon as well as on the functioning of Clark's (1987) Principle of Contrast in bilingual children. Indeed, even though bilectal

children in the current study showed more doublets with larger vocabularies, they still all had at least some doublets in their vocabularies. This, however, cannot be taken as evidence for or against a single lexicon, especially since children in the current study are already in their third year of life, but it does suggest that bilingual children's lexical development follows patterns similar to those of their bilingual peers.

The absence of evidence for mutual exclusivity in bilingual as well as bilingual children, even during early word production, could indicate that such constraints are not innately specified, as suggested several times in the past (for a discussion see Markman 1992), but are heavily influenced by the linguistic environment children grow up in. Just like their bilingual peers, children who are exposed to two varieties of the same language are willing to accept two labels for the same thing in an age when their monolingual peers are reluctant to follow such behavior. A question which arises is how early bilingual children exhibit this kind of behavior and whether evidence from comprehension can be obtained even earlier. An interesting finding in our study concerns the apparent advantage of female participants in both their total word and conceptual production as well as their doublet production. A heads-up for girls in their productive language has been reported in the past (Dale & Fenson 1996; Huttenlocher, Haight, Bryk, Seltzer & Lyons 1991; Papaeliou & Rescorla 2011). For monolingual children acquiring American English, Dale & Fenson (1996) report differences in productive vocabulary between boys and girls, since the beginning of word production with an average of 10.5 words at 12 months and 104.8 words at 18 months for girls, and 9.8 words at 12 months and 63.5 words at 18 months for boys. However, at the age of 30 months this difference becomes smaller for girls (mean: 526.7 words; boys' mean: 510.2). This advantage is less pronounced in comprehension data, both in Dale & Fenson (1996) and in most experimental studies of word learning/comprehension which do not find any gender effects (for example, Mani, Johnson, McQueen & Huettig 2013; Werker, Fennell, Corcoran & Stager 2002). This suggests that the advantage in production may be due to girls' tendency to be more verbal than boys, possibly as a result of early gender stereotypes (for example, mothers talking to girls more than to boys; Huttenlocher et al. 1991).

This advantage of girls is also reflected in their use of doublets. One explanation, which goes together with the girls' tendency to be more verbal than boys, could relate to women being known to use more standard-like language (Labov 2006; Trudgill 1972). That is, girls could be using more standard variants than boys, therefore increasing their use of doublets to agree with the linguistic conventions in their community. This would mean that when they use CG grammar they employ the CG word for a concept and when they use SMG grammar they

employ the corresponding SMG word. This should not imply that boys do not follow or ignore these conventions, but girls might be more sensitive in opting for the “right” option. Another explanation for the higher number of doublets for girls could be that they might have greater flexibility in their use of language in different situations from a very early age. However, the fact that boys also use doublets in their productive vocabulary suggests that the higher number of doublets for girls could actually be a statistical artefact of their larger productive vocabularies. In any case, both boys and girls appear not to be driven by mutual exclusivity in their production of words and share the common flexibility of using two labels for the same concept/referent. In the bilectal language community of Cyprus, the two varieties, SMG and CG, share the vast majority of their vocabulary, either identical or with mere phonological differences. In order for children to acquire their lexicon, they need to learn a single set of words which can be used across two varieties. At the same time, a small percentage of concepts is lexicalized differently in the two varieties, with those words forming parts of TE pairs (e.g., the concept *slide*: CG [kɛtrɛ'ʃili] and SMG [ʃsu'liðrɛ]). This creates a special language situation, in which children can use most of their words in both varieties, but reserve some of those words for use in only their everyday or their more formal speech.

For TEs, it is unknown what the exact meaning of the two words is for young children and whether they are real equivalents in the two varieties. When using questionnaires, such as the CDI, to measure language abilities of children, there is no way to have a qualitative measure of what is the representational status of those words (De Houwer et al. 2006; Deuchar & Quay 2001), either in comprehension or in production, and either in one or in two languages. Additionally, in the first years of language development, children often use their words differently from adults, either underextending or overextending (Bloom 2004; Clark 2009), so knowing if those TEs, in both bilingual and bilectal situations, are exactly equivalent in meaning or not is an issue which has not been addressed or resolved adequately to date.

It is also unknown whether the words which exist in only one of the two varieties are recognized as belonging to CG or SMG by children. Many of those words use sounds which are common across the two varieties, so children cannot use phonology to help them separate the two (e.g., *run*: CG [vu'ro] and SMG ['tɹexo]). However, other cues are available for children to use in order to classify those words appropriately; different syntactic structures (e.g., clitics) and grammatical markers (e.g., verb endings) result in clearly different constructions in the two varieties, for example:

- (1) μεν βουράς, έσυρες το κάτω (CG)  
 NEG run.IMP dropped.2SG CL-it down  
 'Don't run, you dropped it.'
- (2) μην τρέχεις, το έριξες κάτω (SMG)  
 NEG run.IMP CL-it dropped.2SG down  
 'Don't run, you dropped it.'

Using the context in which these words are produced, children in bilingual communities can begin to classify their words as belonging to one or the other variety, or words which can be used in both. However, when and how exactly this takes place will require additional and systematic work in several directions before we begin to better understand it.

## 5 Conclusion

Bilingual development is an area of language and cognitive development which has received minimal attention to date. This is the first study investigating language in children growing up in a community where two varieties of the same language, CG and SMG, co-exist, with the former being used in everyday interactions and the latter being reserved for formal communication and education. The CDI is proving a useful tool for investigating this area of development and is giving us interesting insights into the manner in which these children develop. The very first evidence puts them closer to bilingual children concerning a behavior very well-studied in monolingual development, namely, the use of mutual exclusivity. The adaptation of the CDI for CG opens up a world of new and exciting research possibilities for understanding this grey area in development better, for better designing language studies in such bilingual communities and ultimately for obtaining a better understanding of language development as a whole.

## References

- Au, Terry Kit-Fong and Mariana Glusman. 1990. The principle of mutual exclusivity in word learning: To honor or not to honor? *Child Development* 61(5): 1474–1490.
- Bialystok, Ellen. 2001. *Bilingualism in Development: Language, Literacy, and Cognition*. Cambridge: Cambridge University Press.

- Bloom, Paul. 2004. Myths of word learning. In *Weaving a Lexicon*, D.G. Hall & R.S. Waxman (eds.), 205–224. Cambridge, MA: MIT Press.
- Clark, Eve. 1987. The principle of contrast: A constraint in language acquisition. In *Mechanisms of language acquisition*, B. MacWhinney (ed.), 1–33. Hillsdale, NJ: Erlbaum.
- Clark, Eve. 1993. *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.
- Clark, Eve. 2009. *First Language Acquisition* (2nd edn.). Cambridge: Cambridge University Press.
- Dale, Paul and Larry Fenson. 1996. Lexical development norms for young children. *Behavior Research Methods, Instruments and Computers* 28(1): 125–127.
- De Houwer, Annick, Marc Bornstein and Sandrine De Coster. 2006. Early understanding of two words for the same thing: A CDI study of lexical comprehension in infant bilinguals. *International Journal of Bilingualism* 10(3): 331–347.
- De Houwer, Annick, Marc Bornstein and Diane Leach. 2005. Assessing early communicative ability: A cross-reporter cumulative score for the MacArthur CDI. *Journal of Child Language* 32(4): 735–758.
- Deuchar, Margaret and Suzanne Quay. 2001. *Bilingual Acquisition: Theoretical Implications of a Case Study*. Oxford University Press.
- Fennell, Christopher, Krista Byers-Heinlein and Janet Werker. 2009. Using speech sounds to guide word learning: The case of bilingual infants. *Child Development* 78(5): 1510–1525.
- Fenson, Larry, Paul Dale, Steven Reznick, Elizabeth Bates, Donna Thal and Stephen Pethick. 1994. Variability in early communicative development. *Monographs of the Society for Research in Child Development* 59(5): 1–173.
- Ferguson, Charles. (1959). Diglossia. *Word* 15: 325–340.
- Goldstein, Brian (ed.). 2011. *Bilingual Language Development and Disorders in Spanish-English Speakers*. Baltimore, MD: Brookes Publishing.
- Golinkoff, Roberta, Kathy Hirsh-Pasek, Lois Bloom, Linda Smith, Amanda Woodward, Nameera Akhtar, Michael Tomassello and George Hollich (eds.). 2000. *Becoming a Word Learner: A Debate on Lexical Acquisition*. Oxford: Oxford University Press.
- Golinkoff, Roberta, Kathy Hirsh-Pasek, Leslie Bailey and Neill Wenger. 1992. Young children and adults use lexical principles to learn new nouns. *Developmental Psychology* 28(1): 99–108.
- Grohmann, Kleanthes K. 2011. Some directions for the systematic investigation of the acquisition of Cypriot Greek: A new perspective on production abilities from object clitic placement. In *The Development of Grammar: Language Acquisition and Diachronic Change*, Esther Rinke & Tanja Kupisch (eds.), 179–203. Amsterdam: John Benjamins.
- Grohmann, Kleanthes K. and Evelina Leivada. 2012. Interface ingredients of dialect design: Bi-x, socio-syntax of development, and the grammar of Cypriot Greek. In

- Towards a Bilingual Understanding of Grammar: Essays on Interfaces*, Anna Maria Di Sciullo (ed.), 239–262. Amsterdam: John Benjamins.
- Hall, George and Sandra Waxman (eds.). 2004. *Weaving a Lexicon*. Cambridge, MA: MIT Press.
- Hamilton, Antonia, Kim Plunkett and Graham Schafer. 2000. Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language* 27(3): 689–705.
- Heilman, John, Susan Weismer, Julia Evans and Christine Hollar. 2005. Utility of the MacArthur–Bates Communicative Development Inventory in identifying language abilities of late-talking and typically developing toddlers. *American Journal of Speech-Language Pathology* 14(1): 40–51.
- Holowka, Siobhan, Francoise Brosseau-Lapre and Laura Ann Petitto. 2002. Semantic and conceptual knowledge underlying bilingual babies' first signs and words. *Language Learning* 52(2): 205–262.
- Houston-Price, Carmel, Emily Mather and Elena Sakkalou. 2007. Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *Journal of Child Language* 34(4): 701–724.
- Huttenlocher, Johanna. 1998. Language input and language growth. *Preventive Medicine* 27(2): 195–199.
- Huttenlocher, Janellen, Wendy Haight, Anthony Bryk, Michael Seltzer and Thomas Lyons. 1991. Early vocabulary growth: Relation to language input and gender. *Developmental Psychology* 27(2): 236–248.
- Jackson-Maldonado, Donna. 2012. *MacArthur-Bates Communicative Development Inventories The Encyclopedia of Applied Linguistics*, New York, NY: Springer New York. doi:10.1002/9781405198431.wbeal0745
- Junker, Doerte and Ida Stockman. 2002. Expressive vocabulary of German–English bilingual toddlers. *American Journal of Speech-Language Pathology* 11: 381–394.
- Kampanaros, Maria, Kleanthes K. Grohmann and Michalis M. Michaelides. 2013. Lexical retrieval for nouns and verbs in typically developing bilingual children. *First Language* 33: 182–199.
- Kampanaros, Maria, Kleanthes K. Grohmann, Michalis Michaelides and Elena Theodorou. 2013. On the nature of verb–noun dissociations in bilingual SLI: A psycholinguistic perspective from Greek. *Bilingualism: Language and Cognition* 11(7): 169–188.
- Killing, Sarah and Dorothy Bishop. 2008. Move it! Visual feedback enhances validity of preferential looking as a measure of individual differences in vocabulary in toddlers. *Developmental Science* 11(4): 525–530.
- Labov, William. 2006. *The Social Stratification of English in New York*. Cambridge: Cambridge University Press.
- Luyster, Rhiannon, Shanping Qiu, Kristina Lopez and Catherine Lord. 2007. Predicting outcomes of children referred for autism using the MacArthur–Bates Commu-

- nicative Development Inventory. *Journal of Speech, Language and Hearing Research* 50(3): 667–681.
- Mani, Nivedita, Elizabeth Johnson, James M. McQueen and Falk Huettig. 2013. How yellow is your banana? Toddlers' language-mediated visual search in referent-present tasks. *Developmental Psychology* 49(6): 1036–1044.
- Markman, Ellen. 1992. Constraints on word learning: Speculations about their nature, origins, and domain specificity. In *Modularity and Constraints in Language and Cognition*, Megan R. Gunnar and Michael Maratsos (eds.), 59–101. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Markman, Ellen and Gwyn Wachtel. 1988. Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology* 20: 120–157.
- Markman, Ellen, Judith Wasow and Mikkel Hansen. 2003. Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology* 47: 241–275.
- Marschik, Peter, Hannes Einspieler, Ralf Vollmann and Christa Einspieler. 2005. Lexikonerwerb im zweiten und dritten Lebensjahr: Wie viel erzählen uns die ersten Wörter? *LOGOS Interdisziplinär* 13: 8–14.
- Merriman, William and Laura Bowman. 1989. The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development* 54(3–4): 1–132.
- O'Toole, Clara. 2013. Using parent report to assess bilingual vocabulary acquisition: A model from Irish. In *Bilinguals and Assessments: Solutions*, Virginia Mueller Gathercole (ed.), 81–102. Bristol: Multilingual Matters.
- O'Toole, Clara and Paul Fletcher. 2010. Validity of a parent report for Irish-speaking toddlers. *First Language* 30: 199–217.
- Papaliou, Christina and Leslie Rescorla. 2011. Vocabulary development in Greek children: A cross-linguistic comparison using the Language Development Survey. *Journal of Child Language* 38: 861–887.
- Paradis, Johanne, Kristyn Emmerzael and Tamara Sorenson Duncan. 2010. Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders* 43: 474–497.
- Paradis, Johanne, Fred Genesee and Martha Crago. 2011. *Dual Language Development and Disorders: A Handbook on Bilingualism and Second Language Learning* (2nd ed.). Baltimore, MD: Brookes.
- Pearson, Barbara, Sylvia Fernández and Kimbrough Oller. 1993. Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning* 43(1): 93–120.
- Pearson, Barbara, Sylvia Fernández and Kimbrough Oller. 1995. Cross-language synonyms in the lexicons of bilingual infants: one language or two? *Journal of Child Language* 22(2): 345–368.
- Petitto, Laura Ann, Marina Katerelos, Bronna Levy, Kristine Gauna, Karine Tetreault

- and Vittoria V. Ferraro 2001. Bilingual signed and spoken language acquisition from birth: implications for the mechanisms underlying early bilingual language acquisition. *Journal of Child Language* 28(2): 453–496.
- Poulin-Dubois, Diane, Ellen Bialystok, Agnes Blaye, Alexandra Polonia and Jessica Yott. 2013. Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism* 17(1): 57–70.
- Rowe, Charley and Kleantes K. Grohmann. 2013. Discrete bilingualism: Towards covert prestige and diglossic shift in Cyprus. *International Journal of the Sociology of Language* 224: 119–142.
- Trudgill, Peter. 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society* 1(2): 175–195.
- Tsiplakou, Stavroula. 2014. How mixed is a “mixed” system? The case of the Cypriot Greek Koine. *Linguistic Variation* 14(1): 161–178.
- Vollmann, Ralf, Peter Marschik and Christa Einspieler. 2000. Elternfragebogen für die Erfassung der frühen Sprachentwicklung für (österreichisches) Deutsch. *Grazer Linguistische Studien* 54: 123–144.
- Volterra, Virginia and Traute Taeschner. 1978. The acquisition and development of language by bilingual children. *Journal of Child Language* 5: 311–320.
- Werker, Janet and Krista Byers-Heinlein. 2008. Bilingualism in infancy: First steps in perception and comprehension. *Trends in Cognitive Sciences* 12(4): 144–151.
- Werker, Janet, Christopher T. Fennell, Kathleen M. Corcoran and Christine Stager. 2002. Infants’ ability to learn phonetically similar words: effects of age and vocabulary size. *Infancy* 3(1): 1–30.
- Yoder, Paul, Steven Warren and Heather Biggar. 1997. Stability of maternal reports of lexical comprehension in very young children with developmental delays. *American Journal of Speech-Language Pathology* 6: 59–64.